On the feasibility of HyMap imagery to investigate hydrocarbon polluted areas over time for remediation purposes

Laura Annika Hakkarainen
March, 2010

Course Title: Geo-Information Science and Earth Observation for Environmental Modeling and Management

Level: Master of Science (M.Sc)

Course Duration: September 2008 - March 2010

Consortium partners: University of Southampton (United Kingdom)
Lund University (Sweden)
University of Warsaw (Poland)
University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC) (The Netherlands)
GEM thesis number: 2010 - 13

On the feasibility of HyMap imagery to investigate hydrocarbon polluted areas over time for remediation purposes

by

Laura Annika Hakkarainen

Thesis submitted to the Faculty of Geo-information Science and Earth Observation (ITC), University of Twente, in a partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation for Environmental Modelling and Management

Thesis Assessment Board

Chairman: Prof. Freek van der Meer
External Examiner: Prof. Dr. Hab. Katarzyna Dabrowska
First Supervisor: Dr. Marleen F. Noomen
Second Supervisor: Dr. Mark van der Meijde
Course coordinator: M.Sc. Andre Kooiman

ITC

Faculty of Geo-information Science and Earth Observation (University of Twente)
Enschede, the Netherlands

Disclaimer

Abstract

Hydrocarbons (HC) have become a severe environmental and economical problem due to increasing use of petroleum products causing vast number of HC contaminated sites all over the world. As organic pollutants, HCs have negative effect on humans and the environment. HC polluted sites should therefore be remediated as soon as possible. This unfortunately is not always possible, since clean up of such sites is expensive and may take several years to accomplish. In order to optimize the cleanup of such areas, non-destructive methods such as hyperspectral remote sensing (RS) have been recognized to be a potential tool for detecting and monitoring HC polluted areas.

In the current study, the aim was to investigate the applicability of hyperspectral (RS) techniques to detect the change over HC polluted areas on two HyMap scenes near a leaking gas pipeline in the Northern Netherlands. Different spectral techniques, such as absorption feature analysis, continuum removal, first and second derivative analysis, existing HC and vegetation health indices, were investigated on the spectra collected from polluted and clean sites. A novel change detection method based on an image processing technique developed by van der Werff et al. (2008) was further applied, in order to detect the change over HC polluted sites throughout the whole study area. This technique allowed the comparison of two images automatically with normalized data, highlighting the variation of vegetation health status within the agricultural fields. The use of spatial factors such as pipeline maps further assisted on investigating vegetation stress related to only HCs.

The results suggested that direct detection of HCs over non-vegetated areas was not possible for a number of reasons: for example the HCs from the leaking pipeline were either not situated on the surface or the remaining vegetation within the pixels hampered the detection of diagnostic HC absorption features. In contrast, using vegetation stress as indirect indicator of HC pollution, improved results were acquired: it was found that the Red Edge Position (REP) - index from several tested indices was the only index able to detect vegetation stress related to HC pollution, especially in densely vegetated areas with Normalized Difference Vegetation Index (NDVI) > 0.6. Lack of drilling data did not allow establishing spectra techniques for detection of HC pollution over partially vegetated sites (NDVI < 0.6). It was therefore recommended that vegetation indices taking into account the soil background reflectance (e.g. OSAVI, Optimized Soil Adjusted Vegetation Index), should be further investigated in detection of HC pollution over areas with less vegetation cover, in order not to leave any HC contaminated areas unnoticed.

In the change detection analysis, several HC polluted locations were identified in which vegetation had become healthier between years 2005 and 2008. This indicated that HC pollution level had decreased on these sites, allowing improved growth of the vegetation. An overall accuracy of 58.8 % and Kappa coefficient of 0.13 was acquired for the change detection image. The accuracy assessment was however affected by the nature of the ground truth data (e.g. sampling depth) and therefore additional validation was recommended. This would not just verify the performance of the change detection method but also provide information if less cleanup actions are needed, further resulting in decrease of remediation costs in the study area.

Acknowledgements

Life is full of surprises and I suppose this is one of them.

"I thought that if we looked for this Pit, we'd
be sure not to find it, which would be a Good
Thing because then we might find something
that we weren't looking for, which might be
just what we were looking for, really!"

(A. A. Milne)

First of all, I would like to thank my supervisors, Dr. Marleen Noomen and Dr. Mark van der Meijde, for providing me good ideas and supporting me with my own ones throughout the whole six months. My sincere acknowledgements go also to Dr. Harald van der Werff for giving tools to perform my analysis and DHV&NAM for providing the data I could use in my work. In addition, thank you for all the other people at ITC for helping me with various things related to my work; especially Dr. Freek van der Meer, Dr. David Rossiter and Mila Luleva and all the others not listed in here.

My sincere acknowledgements goes also to the GEM coordinators, professors and lecturers; Andrew, Petter, Terry, Kasia, Andre, Ted, Angela, Ian, Craig, Samantha, Karin, Jonas, Thomas, Michael, Gosia and all the others whose name has not been mentioned on this list. Thank you for all the new experiences, knowledge and help throughout the whole GEM course. Despite of hard work, I also had a lot of fun.

Now it is time to praise my girls; Nami, Olena and Zory. Without you by my side I wouldn't have made this far. You have really taught me what life is about: NO STRESS and spending time with good friends. I do not even have words of how much I will miss you after the GEM. Thank you for all the other GEMs as well, for being the family for me for the past 18 months. You are all special people with hearts at their right position. Let the Jungle song play and keep on rocking!

I would also like to thank all of my lovely friends back home, who encouraged me with my decision to go abroad and explore the world. In addition, without the additional financial support from Niska & Nyyssönen Oy, this 18 month of studies would have not been as pleasant as it was.

Finally, to my lovely family back home: mom, dad, Simo, Lilli, Saija, Sebu, little Sophie and the yet unborn family member. You are my rock! Without doubt you all have helped me through this time. I liked being abroad but home is still home, always! Then the last but definitely not the least: Paavo. Without you, my life would not have any meaning. I have no other words than I love you so much and thank you for your patience and support! Let our life begin now.

Table of contents

List of figures

Chapter 2: Literature review

Chapter 3: Materials and methods

Chapter 4: HC pollution detection over non-vegetated areas from 2005 the image

Chapter 5: HC pollution detection over vegetated areas from the 2005 image

Chapter 6: HC pollution detection over time

List of tables and appendices

Tables

Chapter 3: Materials and methods

Chapter 4: HC pollution detection over non-vegetated areas from the 2005 image

Chapter 5: HC pollution detection over vegetated areas from the 2005 image

Chapter 6: HC pollution detection over time

Appendices

List of Abbreviations

A = *Area of the absorption feature*
BR = *Band Ratio*
AS = *Asymmetry of the absorption feature*
C = *Carbon*
CR = *Continuum Removal (Removed)*
D = *Depth of the absorption feature*
DHV **=** *Consultancy and Engineering Company who has investigated the study area*
DLR = *German Aerospace Centre who had pre-processed the HyMap images*
EEA = *European Environment Agency*
EEA = *European Economic Area*
EIA = *Energy Information Administration*
EMS = *Electromagnetic spectrum*
ESI = *Environmental Simulations International*
GPR = *Ground Penetrating Radar*
H = *Hydrogen*
HC = *Hydrocarbon(s)*
IS = *Imaging Spectroscopy or Spectrometry*
NAM = *Nederlandse Aardolie Maatschappij (Dutch Oil Company owning the pipeline)*
NIR = *Near-infrared part of the EMS (~ 0.7 – 1.0 µm)*
$\boldsymbol{NP_{mean}}$ = *Non-Polluted average spectra*
$\boldsymbol{NP_{+/-std}}$ = *Non-Polluted + or - 1σ from the average spectra*
NRCAN = *Natural Resources Canada*
OPEC = *Organization of the Petroleum Exporting Countries*
P = *Position of the absorption feature*
$\boldsymbol{P_{mean}}$ = *Polluted average spectra*
$\boldsymbol{P_{+/-std}}$ = *Polluted + or - 1σ from the average spectra*
RS = *Remote Sensing*
QAF = *Quantitative Absorption Feature analysis*
SWIR = *Short-wave infrared part of the EMS (~ 1.0 – 2.5 µm)*
U.S. EPA = *United States Environmental Protection Agency*
VI = *Vegetation index or indices*
VIS = *Visible part of the EMS (~ 0.4 – 0.7 µm)*
VNIR = *Visible and NIR part of the EMS (~ 0.4 – 1.0 µm)*
W = *Width of the absorption feature*
WHO = *World Health Organization*

1. Introduction

1.1. Background and significance

According to the European Environment Agency (2007), approximately 250 000 contaminated sites in EEA (European Economic Area) countries are in a need for cleanup. It has been predicted that the amount will increase with up to 50 % by 2025 (EEA, 2007). The main sources of contamination are municipal and industrial waste disposals, mining and military sites no longer in operation, as well as former and present industrial plants such as metal, chemical, oil and wood industries (Agostini et al., 2007). It has been identified that in Europe, the primary pollutants are heavy metals and mineral oil causing soil contamination in 37.3 % and 33.7 % cases, respectively (EEA, 2007).

Contaminated locations may cause a significant health and ecological risk limiting the use or re-use of the sites (Carlon et al., 2008). A shortage of space for residential areas and fear of litigation further forces owners to clean up (=‘remediate’) polluted locations (Gay and Korre, 2006). The remediation of contaminated sites is not that straightforward, especially when considering the costs of these projects. For instance, countries such as the United Kingdom and the Netherlands spend tens of millions of Euros for remediation annually (EEA, 2003). As suggested by Sorvari et al. (2009), one possible reason for such elevated prices is the ineffective evaluation of remediation requirements at the initial stage of each venture. This results in a need for proper investigations before any remediation actions should take place (Broos et al., 1999).

For remediation efforts, the first step is to detect and characterize the spatial distribution of contaminated sites. Conventional ways to approach this issue are methods such as soil sampling, drilling and geochemical analysis (van der Meijde., 2009). However, these techniques are time-consuming, costly and destructive in nature (van der Werff et al., 2008). In many cases the contamination may occur over large areas or in remote locations. For example, the heavy metal polluted river floodplains cover vast territories in the Netherlands (Clevers et al., 2004 and Kooistra et al., 2003) and leaking oil and gas pipelines may be situated in far-off locations (Degermenci, 2001). It is obvious that in such cases it is difficult or even impossible to use traditional ways to detect and characterize the spatial extent and nature of contaminated sites.

Quite recently, hyperspectral remote sensing (RS), also known as imaging spectroscopy (IS), has become of an interest in investigating polluted sites (e.g. van der Werff et al., 2008 and Winkelmann, 2005). It has shown a great potential to overcome the limitations of conventional techniques since it covers wide areas and it is non-destructive in nature (Kemper and Sommer, 2002). The high spectral and spatial resolution of the imagery may further allow the differentiation of even weak spectral signatures of various components on the Earth's surface (van der Meer, 2001). To provide few examples of such studies, hyperspectral RS has been implemented to detect heavy metal contaminated sites by Kooistra (2004) and Choe et al. (2008). In addition, many researchers have focused on hydrocarbon (HC) polluted areas either with direct detection (e.g. Bihong et al., 2007, Hörig et al., 2001, Kühn et al., 2004, Malley et al., 1999, van der Werff, 2006 and Winkelmann, 2005) or by indirect means through vegetation stress (e.g. Noomen et al., 2008, van der Meijde et al., 2009 and van der Werff et al., 2008). According to all of the above mentioned studies, hyperspectral RS is a promising tool to investigate contaminated areas.

1.2. Research problem

In developing further insight on the matter of HC polluted areas, a few important issues should be mentioned. HC pollution may occur either naturally through gas seepages escaping from underlying gas-bearing rocks (Noomen et al., 2006) or by man-made leakages such as the ones from oil and gas pipelines (van der Meijde et al., 2009). If released to the ground, HCs may travel to the groundwater causing a risk of contaminants to spread over larger areas (Kim and Corapcioglu, 2003). This is problematic since HCs, such as benzene and hexane, are organic chemical components (Chemistry Glossary, 2009) from which many are toxic causing cancer and other health related problems (U.S. EPA, 2007). The rapid detection and remediation of HC polluted sites is therefore essential in order to prevent the risk for human-health and the environment.

It is known, that HCs have unique absorption features for example in 1.73 µm and 2.13 µm of the electromagnetic spectrum (EMS) (Cloutis, 1989), which theoretically enables their detection through hyperspectral RS. However, many other objects reflect from the Earth's surface within the same spectral range, making it complicated to differentiate these elements from each other (Campbell, 2006). For example, when investigating the spectral profile of heterogeneous soils, properties such as organic matter, soil moisture and color affect the spectral signature (Baumgardner et al., 1985), possibly hampering the differentiation of distinct HC absorption features. In laboratory conditions the explicit absorption features may be separable but when attempted to upscale to airborne hyperspectral data, spectral and spatial resolution may yet be too low to directly detect HC polluted soils with different spectral techniques (Winkelmann, 2005).

HC pollution can further be detected by indirect means (Noomen et al., 2005). This is due to the fact, that the oxidation processes of HCs may cause geochemical and botanical anomalies that are detectable through RS (van der Werff et al., 2007). However, some of the observed anomalies (e.g. vegetation stress), can also be caused by other factors such as water deficiency or plant diseases (Noomen et al., 2005 and van der Werff et al., 2006). This becomes problematic especially when applying automated imaging processes as false anomalies are detected by the software hampering the classification accuracy of the final products (van der Werff et al., 2008). Hence, it is of utmost importance to find distinct indicators of HC pollution from the reflectance curve of stressed vegetation. In some of the cases the developed algorithms for indirect detection of HCs are tested only in densely vegetated areas (e.g. van der Werff et al., 2008). It can therefore be that different spectral techniques are required for HC polluted areas with partial vegetation cover due to disturbing background reflectance from the soil (Noomen et al., 2008).

Different types of HCs deplete over time varying greatly depending on the chemical and biological properties of the soil and the form of HCs (Luis and Zemba, 1993). This can result in decrease of number and the size of HC polluted areas through time (Osuji et al., 2006). To date, it has not been studied if, and to what extent, the depletion of HCs over time could be detected by remotely sensed means. For example, land cover may change over time, having an effect on the detection of polluted locations in different time periods.

To summarize, it is important to accurately detect and characterize the spatial distribution of pollutants for remediation purposes of HC contaminated sites. Hyperspectral RS has shown great potential to address this issue. Difficulties have however been found in differentiating HC polluted areas over various land cover types and the traditional image processing methods are still producing false anomalies. The amount and spatial extent of HC polluted areas may further change over time due to natural attenuation of HCs. In order to avoid additional remediation costs, this study investigates if hyperspectral RS techniques could be used to distinguish the changes occurring over HC polluted areas in different time steps.

1.3. Research objectives and questions

The main objective of this study was to investigate ***the applicability of hyperspectral RS techniques to detect the change over HC polluted*** areas on ***two HyMap*** scenes near a leaking gas pipeline in the Northern Netherlands. The following specific objectives were set for the study:

Objective 1

To ***detect*** HC pollution over ***non-vegetated*** areas near a leaking gas pipeline in the Northern Netherlands using HyMap imagery from year 2005.

1. Can HC polluted sites be identified over ***non-vegetated*** areas using ***absorption feature analysis***?
2. Can HC polluted sites be identified over ***non-vegetated*** areas using ***existing HC indices***?

Objective 2a

To ***detect*** HC pollution over ***densely vegetated*** areas near a leaking gas pipeline in the Northern Netherlands using HyMap imagery from year 2005.

1. Can HC polluted sites be identified over ***densely vegetated*** areas using ***absorption feature analysis***?
2. Can HC polluted sites be identified over ***densely vegetated*** areas using ***existing vegetation health indices***?

Objective 2b

To ***detect*** HC pollution over ***partially vegetated*** areas near a leaking gas pipeline in the Northern Netherlands using HyMap imagery from year 2005.

3. Can the ***spectral techniques*** that were ***used for non-vegetated vegetated*** areas also be used to identify HC polluted sites in ***partially vegetated*** sites?
4. Can the ***spectral techniques*** that were ***used for densely vegetated areas*** also be used to identify HC polluted sites in ***partially vegetated*** sites?

Objective 3

To ***detect*** the change over HC polluted areas near a leaking gas pipeline in the Northern Netherlands ***using two HyMap images*** from year 2005 and 2008

1. Can the change related to HC pollution ***be identified over HC polluted sites in the 2008 image*** using the ***same spectral techniques applied on the scene from 2005***?
2. Are there ***any changes in the spatial extent of HC polluted sites*** between years 2005 and 2008 that ***can be detected using hyperspectral imagery***?

3. Is the ***identification of the natural depletion of HC pollution*** in 2008 image ***affected by the variation in land cover on the polluted sites*** between years 2005 and 2008?

2. Literature review

This chapter presents the review of the literature, in terms of investigating HC contaminated sites, their environmental effects and remediation. Since the cleanup of HC contaminated sites can be problematic, the potential of hyperspectral RS in alleviating these difficulties will be discussed. Different RS techniques identified in HC pollution detection are then critically evaluated and their reliability for investigating HC contaminated sites is highlighted in the context of this thesis.

2.1. Hydrocarbon (HC) contaminated sites

2.1.1. HCs as environmental contaminants

HCs are organic components consisting entirely of carbon (C) and hydrogen (H) molecules (Calabrese and Kostecki, 1993). They come in a variety of forms such as gases (methane and propane), liquids (hexane and benzene) and waxes (paraffin wax) (Olah and Molnár, 2003). Based on the configuration of C and H atoms, HCs can be divided into three different groups: aromatic (arenes), saturated (alkanes), and unsaturated HCs (alkenes and alkynes) (Hornback, 2005).

As environmental contaminants, HCs are of particular significance. This is due to the fact, that most of the World's energy demand is fulfilled with crude oil and its refined products (OPEC, 2009). For example, in 2008 approximately 85 thousand barrels of crude oil were produced every day in 16 different countries worldwide (EIA, 2008). In addition, in all countries of the world, oil and its products are being utilized for electricity production and as fuels to propel most of the cars, trucks, trains, planes and ships (Laherrere, 2003). Due to the extensive production and consumption of oil products, human-induced HC pollution may take place in any location at different stages of crude oil recovery, extraction, transportation and consumption (ESI, 2008).

2.1.2. Importance of remediating HC contaminated sites

2.1.2.1. Effect of HCs on humans

The most immediate concern of HC polluted sites is public safety. When released to the environment, HCs may migrate from the soil to the groundwater leading to pollution in wider areas than the point sources might suggest (Davidson, 2000). HCs via contaminated groundwater are considered as one of the main routes for human exposure (Kostecki and Calabrese, 1991). The number of different types of HCs is vast (Winkelmann, 2005): most dangerous for humans being aliphatic HCs (e.g. cyklohexane and hexane) and aromatic benzene-containing constituents (Davidson, 2000). For example, long-term exposure to benzene has been proved to develop cancer in humans (WHO, 1993). Other adverse health effects such as drowsiness, dizziness and unconsciousness have been perceived mainly due to short-term exposure to HCs (WHO, 1993).

2.1.2.2. Effect of HCs on soil environment

When released to the environment, HCs may cause changes in the soil environment. The soil oxygen (O_2) may be replaced by HC gases (Schumacher, 1996). This can be further accelerated by Methanobacteria oxidizing, for example, methane gases into carbon dioxide (CO_2) and water (H_2O) thus creating anaerobic conditions in the soil media (Hanson and Hanson, 1996). This may have an effect on the soil pH and redox potential (Schumacher, 1996) stabilizing both alkaline and acidic soils into neutral (Neue, 1993). The changing acidity conditions in soils may lead to release of toxic trace elements such as manganese (Mn) and ferric iron (Fe_2O_3), increasing their availability to plants (Adams and Ellis, 1960). Some studies further suggest that the amount of nitrogen (N) and phosphorous (P) may decrease due to HC pollution (Obire and Nwaubeta, 2002), further deteriorating the soil quality.

2.1.2.3. Effect of HCs on vegetation

The health and vigour of vegetation can be diminished due to HCs (van der Meijde et al., 2009). Passive or active uptake through roots and direct contact between soil and plant tissues are considered as the main pathways of organic contaminants to enter the plants (Lin et al., 2007). The deposition of gaseous and particulate HCs onto leaves and uptake through stomata may further result in vegetation to be affected by these organic components (Simonich and Hites, 1995). It is not known if HCs themselves are causing negative effects on plants (Noomen, 2007). Therefore, effects occurring in the soil environment such as lack of oxygen (Noomen, 2007, Smith et al., 2004:2005), strong increase in CO_2 (Hoeks, 1972) and mobilization of possibly phytotoxic nutrients (e.g. manganese) (Adams and Ellis, 1960) due to HC pollution, are assumed to be the main factors causing stress for plants.

2.1.3. Remediation of HC contaminated sites

Due to adverse effects on humans and the environment, it is of utmost importance to remediate HC contaminated sites. One of the oldest ways to clean up these areas is to leave them as they are (Friend, 1996), since it is known that HCs deplete over time naturally (Osuji et al., 2006). The term for the depletion is ‘weathering’, caused by different physical, chemical and biological processes such as volatilization, leaching, and biotransformation of HCs (van den Berg et al., 1995).

As remediation method, natural attenuation may sound ideal since it can create fewer costs for stakeholders responsible of HC polluted sites (Palmer et al., 2002). However, serious attention and regulation efforts has been given since 1980, after noticing that HCs may travel from soil to the groundwater and contaminate the drinking water of humans and animals (Calabrese and Kostecki, 1993). This has resulted in conclusions that HC polluted areas should be remediated with methods such as excavation and *ex-situ* treatment in landfill sites (Friend, 1996). Since this may not be possible in all of the cases, *in-situ* remediation methods are still being used as well (Friend, 1996). These processes should however occur in controlled environment with regular monitoring in order to avoid groundwater pollution

(Palmer et al., 2002). Natural depletion of HCs can be further accelerated by adding O_2 and nutrients in to the soil that may increase the amount of bacteria consuming HCs (Shirdam et al., 2008).

Despite of the remediation method used, in most of the cases investigation and remediation of HC contaminated sites is expensive. For example, ENSR International (2005) has estimated that the total cost of remediating all petroleum polluted sites in the United States is 2.0 billion dollars. In addition, many European countries spend more than million Euros per year for cleaning up HC contaminated areas (EEA, 2007). These costs stem mainly from preliminary investigations (Sorvari et al., 2009), which are commonly carried out with not just expensive but also time-consuming and destructive soil sampling, drilling and geochemical analysis (van der Meijde et al., 2009). New methods, such as hyperspectral RS alleviating these difficulties, have therefore become of interest in investigating HC polluted areas (e.g. van der Werff et al., 2008).

2.2. Investigating HC contaminated sites with hyperspectral RS

2.2.1. What is hyperspectral RS?

Figure 2.1 illustrates the general concept of hyperspectral remote sensing (Winkelmann, 2005), also known as IS (‘imaging spectrometry’, ‘imaging spectroscopy’) (van der Meer, 2001). Compared to multispectral remote sensing, hyperspectral RS provides much narrower, continuous spectral bands in the visible (VIS), near-infrared (NIR) and shortwave infrared (SWIR) part of the spectrum (from 0.4 to 2.5 μm) (Curran, 1994). The amount of bands is typically from 100 to 300 with nominal band with of two to ten nanometres and the spatial resolution of hyperspectral sensors varies from 1 to 30 meters (Borengasser et al., 2008). Hyperspectral data can be represented as three-dimensional image cube (Figure 2.1), where x and y axes depict the usual map or image display and the third (z) dimension is illustrating the many spectral bands superimposed on top of each other (Winkelmann, 2005). Every pixel includes data from each of the bands and as the spatial resolution is high, detailed inspection of the electromagnetic spectrum (EMS) is possible.

When solar radiation interacts with the materials on the Earth’s surface, it is either reflected, absorbed or transmitted (van der Meer, 2004). Depending on the wavelength of the incident light and the reflectance and absorption capacities of objects, the reflected radiation varies between different materials over the EMS. For example, in the VIS, NIR and SWIR portion of the EMS (from 0.4 to 2.5 μm), many objects have distinct absorption features that are 0.02 to 0.04 μm wide (Hunt, 1980). As hyperspectral RS samples data with narrow bands at the same wavelength regions where most of the diagnostic absorption features of materials exist, it is considered ideal for the identification and characterization of different objects, both in the laboratory and field conditions (MacDonald et al., 2009).

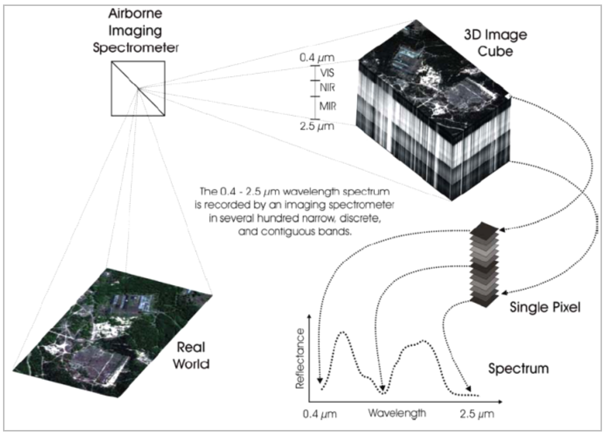


Figure 2.1 *The concept of hyperspectral RS (Winkelmann, 2005).*

Recent applications of hyperspectral RS include for example, mineralogical exploration (e.g. Gingerich et al., 2002), vegetation analysis for precision agriculture (e.g. Rama, 2008), investigation of water bodies and its constituents (e.g. Randolph et al., 2008) and soil, geologic and vegetation mapping (e.g. Demattê et al., 2009, Schmidt and Skidmore 2003 and Smailbegovic et al., 2004). The utility of hyperspectral RS has been recognized in environmental contamination studies as well (e.g. van der Werff et al., 2008 and Winkelmann, 2005). According to Ben-Dor et al. (2009), HC detection through remotely sensed means is commonly divided into direct and indirect detection. Direct detection of HCs with remote sensing refers to the identification of visible oil pools on the Earth's surface (van der Meer et al., 2000a) or inspection of the spectral fingerprints of HC-bearing materials (Ben-Dor et al., 2009). In addition, HC pollution may be detected indirectly using for example botanical anomalies as indicators of HC contamination (Ben-Dor et al., 2009). In the following chapters, the spectral properties of HCs, HC impacted soils and stressed vegetation have been described. These allow the possibility to use RS in direct or indirect detection of HC pollution as will be discussed in the subsequent sections.

2.2.2. Direct detection of HCs

The spectra of organic compounds, such as HCs, are dominated by their C-H stretches, causing distinct absorption features due to their vibrational processes (van der Meer, 2004). One of the pioneer studies of Cloutis (1989) showed that the most prominent HC absorption features were located in the SWIR part of the EMS, particularly at wavelengths 1.73 µm and 2.13 µm (also e.g. Hörig et al., 2001). In addition, other locations for distinct HC absorption features have been found at 0.9 µm, 1.2 µm, 1.38 µm and 2.3 – 2.45 µm (McCoy et al., 2001 and Winkelmann, 2005). Considering some of these features, it has been noted that they may be weak and mixed with absorptions of other natural or man-made materials such as plastics (at 0.9 µm) and water H_2O (at 1.38 µm) (Winkelmann, 2005).

Several studies have been conducted to directly detect HC absorption features using either laboratory or field spectral measurements (e.g. Ellis et al., 2001, Hörig et al., 2001, Kühn et al., 2004, Malley et al., 1999, Orlov et al., 1999 and Winkelmann, 2005). According to laboratory spectroscopy studies made by Malley et al. (1999) and Winkelmann (2005), distinct HC features can be separated from contaminated soils. However, soils have many other properties such as organic matter, moisture, clay content and particle size, which may play dominant role in the reflectance spectrum of soils (e.g. Baumgardner et al., 1985 and Orlov et al., 1999). For example, Winkelmann (2005) stated that these factors may have more effect on the detection limit of HC pollution from soils than the actual spectral characteristics of different HC contaminants and their concentrations. In addition, Malley et al., (1999) found that the absorption features of several HCs in mixtures of diesel fuels may be hard to differentiate from each other.

The promising results from the laboratory experiments have been upscaled to hyperspectral imagery by many researchers (e.g. Ellis et al., 2004, Hörig et al., 2001, Kühn et al., 2001, McCoy et al., 2001, Winkelmann, 2005). The most common methods used have been different band ratios and indices built on the diagnostic HC absorption features. Most of these studies found that the performance of these detection methods varied depending on the site under investigation and the techniques used. For example, Winkelmann (2005) found that the detection of HCs was difficult in study sites where the land was not purely covered with soils. In addition, other authors have concluded that the heterogeneity of the land cover and the amount of vegetation cover within a pixel mostly define the usefulness of hyperspectral images to detect HCs directly (e.g. Hörig et al., 2001 and Taylor, 2000). Ellis et al. (2001) noted that other materials such as roads, parking lots, roofs and oil strains may be misclassified as HC contaminated sites in case automated processed are used. Many researchers stated that the spectral and spatial differences between the laboratory, field data and airborne sensors should be understood in order to optimize the direct detection of HC pollution from hyperspectral imagery (e.g. Ellis et al., 2001).

Apart from using diagnostic HC absorption features for HC detection, mineralogical alterations can be used as indicators of HC presence in soils (Schumacher, 1996). Examples of such alterations are formation of calcite, pyrite and iron sulphides, bleaching of red beds and clay mineral alterations (Schumacher, 1996), formed due to locally anomalous redox zones in the soils (van der Meer et al., 2000a). Considering man-induced HC contamination, it may not be that straightforward to utilize these mineralogical alterations in pollution detection. This is due to the fact that normally the alterations occur over long periods of time due to macro and micro seepages from HC reservoirs (Schumacher, 1996). In contrast, in most of the human-induced contaminated sites, the exposure of HCs to soils can be relatively short compared to the natural leakages. This is most likely the reason why most of the studies in detecting mineralogical alterations have been related to investigating natural gas seepages or exploration of HC deposits (e.g. van der Werff, 2006; Khan and Jacobson, 2008 and Yang et al., 2000). However, as these alterations may cause diagnostic absorption features in the soil spectra and possibly reveal HC pollution, they should be kept in mind when studying man-made HC contamination with RS.

2.2.3. Indirect detection of HCs

HCs may have an effect on vegetation health and vigour due to changes in the soil environment (see section 2.1.2.3.). In remotely sensed images, these adverse health effects can be seen as botanical anomalies (Schumacher, 1996) such as abnormal distribution of plants, presence of indicator species or morphological changes of plants (van der Meijde et al., 2009). The latter is referred to vegetation stress caused by lack of oxygen (O_2) or soil nutrients that induce deficiencies in chlorophyll concentration and other leaf biochemical constituents due to for example structural damages of plant cells (Elvidge, 1990).

In normal conditions, the spectral signature of healthy vegetation is dominated by chlorophyll in the VIS part of the EMS (particularly at 0.43 & 0.66 nm and 0.45 & 0.64 nm for chlorophyll a and b, respectively), by the cell structure in the NIR and by water content in the SWIR regions (Jensen, 2000). When the concentration of chlorophyll and other biochemical constituents of plants declines, the following changes can be detected in the otherwise normal vegetation spectrum: the height of the infrared shoulder decreases, the maximum reflectance in the red part of the spectrum increases and the red edge (steep slope between red and near-infrared wavelength, 0.7µm) is shifted towards shorter wavelengths (e.g. Guyot and Baret, 1988 and Carter, 1993). The plants may further exhibit water stress that results in increased reflectance in the dominant water absorption peaks in the SWIR region (1.4 and 1.9 µm) (Hunt and Rock, 1989). According to many researchers, these changes in the vegetation spectra may be indirectly caused by HC contamination (e.g. van der Meijde et al., 2009, van der Werff et al., 2006 and White, 2007). By investigating the vegetation spectrum and its possible changes over polluted areas, HC pollution may therefore be detected indirectly.

In vegetation studies, possibly the most common way to detect HCs through plants has been to use existing vegetation health indices that may improve the differentiation of stressed and healthy vegetation from each other. For example, Noomen (2007) compared the effect of natural gas, methane and ethane for two plant species, maize and grass, under controlled laboratory and field conditions. The author tested several existing vegetation indices together with continuum removal (i.e. spectral enhancement technique, Kokaly and Clark, 1999), finding that reduced band depths in the water (1.37 – 1.57 µm and 1.87 – 2.17 µm) and chlorophyll (0.55 – 0.75 µm) absorption features were related to an increase in PRI (Photochemical Reflectance Index). This result suggested that HC gasses, especially ethane, decreased the photosynthetic activity of plants (Noomen, 2007).

Furthermore, van der Meijde et al., (2009) tested several vegetation stress indices such as Carter Stress Indices (Carter, 1994) and Red Edge Position Index (Guyot and Baret 1988) to detect vegetation anomalies around a leaking pipeline. The researchers found that many indices showed lower index values above the pipeline than further away, indicating that HC leakage had caused stress for the aboveground vegetation. In order to upscale these field results to airborne imagery, van der Werff et al. (2008) applied a novel automated image processing technique that utilized both the spectral information of stressed vegetation and spatial knowledge of the pipeline location to detect HC pollution over larger areas. Despite of the lack of validation data, the automated procedure was able to identify possibly polluted sites in heterogeneous land cover, which has been seen as challenge by other authors (e.g. White, 2007, Winkelmann, 2005).

The change in ‘Red edge’ position due to HC pollution has been under extensive research. The ‘Red edge’ is known as the rapid change of vegetation reflectance from red to NIR region (Campbell, 2006), which has been found to shift either to shorter (e.g. Bammel and Birnie 1994, Reid et al., 1988 and Smith et al., 2004) or longer (e.g. Yang et al., 1999 and 2000) wavelengths in vegetation above natural HC seepage or simulated gas leakages. The change in ‘Red edge’ position can be very subtle (Campbell, 2006). This may hamper its detection from remotely sensed data with low spatial and spectral resolution. In order to overcome these difficulties, several authors have used first and second derivatives of vegetation spectra to enhance the subtle differences in the ‘Red edge’ of HC affected plants compared to control treatments (e.g. Smith et al., 2004:2005 and White, 2007). Based on the derivative calculations, these authors compiled simple derivative ratios to successfully detect gas leaks from buried pipelines. According to the researchers, band ratios calculated from first derivatives were insensitive to soil background effect (e.g. White, 2007). This could make them useful in detection of pollution especially over partially vegetated areas, where many indices such as NDVI (Normalized Difference Vegetation Index) may not perform properly (Jensen, 2000).

2.2.4. Detection of HCs over time

Change detection in RS has been defined as a process, which aims to identify environmental changes using two or more images taken over a period of time from the same area (NRCAN, 2005). In the change detection analysis, several things should be taken into account in order to acquire a reliable result of the occurred changes over time (Jensen, 1996). The acquisition date and time should be similar in order to eliminate changes in the diurnal and seasonal sun angle effect. The inconsistency in spatial, spectral and radiometric resolution should be accounted for, as these should be comparable between the two or more images used in the analysis. In addition, the changing environmental effects such as differences in soil moisture conditions and vegetation phenology should be taken into consideration.

During the last few decades, RS has been used for example to detect changes in disturbances in natural forests (e.g. Nielsen et al., 2008) or sea surface temperature related issues (e.g. Yang et al., 2004), with promising results. When considering the change detection in relation to HCs, oil spill monitoring in marine environments is probably the most commonly known. Several studies exist, where for example optical (e.g. De Domenico et al., 1994 and Howari et al., 2004) or Synthetic Aperture Radar (SAR) (e.g. Ferraro et al., 2009) have been used in detection and monitoring of oil slicks in surface waters. In addition, thermal RS using Landsat data has been utilized in the multi-temporal investigation of Kuwait's oil lakes formed in the Arabian Gulf War in the 1990s (e.g. Ud Din et al., 2008 and Yaw Kwarteng, 1998).

The common feature for all the above mentioned studies was the utilization of moderately low spatial resolution of remotely sensed data. This manifests from the increased availability of several low resolution sensors and their overpasses above the same areas during the past few decades. In high resolution hyperspectral RS, not many temporal studies have been conducted in general, or especially in relation to man-induced HC contamination. This is due to the lack of multi-temporal imaging spectrometry data caused by rare bypasses over the same areas (Li et al., 2005). The availability of such data has recently increased, drawing the interest of scientists to use multi-temporal hyperspectral data in different applications (Eismann and Meola, 2008). In HC contamination studies, for example Li et al. (2005) used two AVIRIS images to investigate vegetation stress and cover change caused by oil spill in Jornada, New Mexico. The authors found that the vegetation became stressed and the land cover changed over time due to the oil leakage. In addition, hyperspectral RS has been used in off-shore oil spill detection and monitoring by Salem and Kafatos (2001). These authors concluded that difficulties, such as changing weather conditions, change in oil appearance over time, sun angle effects and other natural phenomena having similar features as oil spills can be reduced by using high resolution data (Salem and Kafatos, 2001). Since it is known that HCs deplete over time naturally (see section 2.1.3.), it could be possible that this change occurring over HC polluted areas near a leaking gas pipeline could be detected with hyperspectral RS.

2.3. Concluding remarks

To summarize, the remediation of HC contamination is important but expensive. In order to decrease the cleanup costs, hyperspectral RS techniques have been found to be a potential substitute or addition for traditional investigation methods such as soil sampling and geochemical analysis. Hyperspectral RS allows the investigation of HC pollution either directly or indirectly using diagnostic HC absorption features or subtle changes in the vegetation spectra as indicators of pollution. Various spectral enhancement techniques such as continuum removal, first and second derivative analysis, vegetation stress and HC indices have been used to identify HC pollution from RS data. Furthermore, the recently increased availability of multitemporal hyperspectral data has permitted to conduct change detection analyses that have not yet been utilized to large extent in the detection of man-induced HC pollution. Hence, it can be that natural attenuation of HCs could be detected by remotely sensed means close to the leaking gas pipeline.

3. Materials and methods

This chapter provides a detailed description of the study area, the data and the methodology used for the analysis.

3.1. Study area

The study area was located in the Northern Netherlands along a 21 km long pipeline (Figure 3.1), which has been used to transport HC condensates (mainly benzene) from Grijpskerk to Anjum (located in the provinces of Groningen and Friesland, respectively). The land use in the area was mainly agricultural (van der Werff et al., 2008) and the predominant soil type was clay (Poelman, 1966). The terrain was mostly flat with a few undulations in the form of dikes and canals (Cundill, 2008). After the construction in 1996, the pipeline had started 'sweating' from its connection points, which were approximately nine meters apart from each other (van der Werff et al., 2008). Due to this, the pipeline has been closed around 2000 – 2002 (personal communication with Dr Mark van der Meijde). According to the geochemical drilling data, the leaking of HC condensates has caused HC pollution in several locations in the area. Proper investigations of the extent of contamination were therefore needed, in order to provide information for remediation purposes.

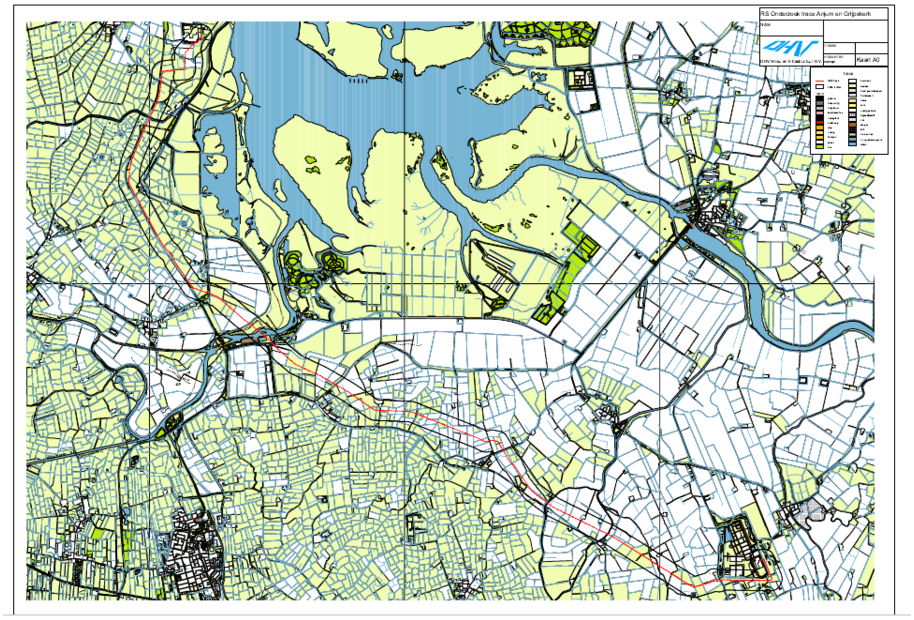


Figure 3.1 *Map of the study area, scale 1:50 000. (Provided by DHV). Pipeline indicated with red line.*

3.2. Data

The following datasets were used for the study:

Analysis data

- Two HyMap scenes acquired in years 2005 and 2008 from the study area (Datum: WGS 1984, Projection: UTM Transverse Mercator, Zone 32 North)
- Drilling data (groundwater samples) with chemical record from 2003/2004 to 2008 (n=10)

Validation data

- Drilling data (groundwater samples) with chemical record from 2003/2004 to 2008 (n=17)
- Drilling data (soil samples) with chemical record from 2003 (n=9) for non-vegetated areas
- Drilling data (soil and groundwater samples) with chemical record from 2002/2003 (n=22) for vegetated areas

Ancillary data

- Vector layer showing the exact location of the pipeline
- Digital soil map (1:50 000)
- Land use map from 2005[1] including classes meadow and arable land
- Vector layer indicating the potato fields in the area in 2005
- Raster layer indicating the anomalous regions (n=28) detected by van der Werff et al. (2008) from the 2005 image (i.e. patches of stressed vegetation possibly due to HC pollution)

The hyperspectral images have been acquired with HyMap airborne imaging spectrometer (Cocks et al., 1998) on the 19th of June in 2005 and on the 24th of July in 2008. The HyMap sensor covered VIS, NIR and SWIR regions (~ 0.45 – 2.48 µm) using 126 bands with nominal bandwidth of 15 – 20 nm (Holzwarth and Bachmann, 2009). The pixel size of the images was 4x4 m. The images had been pre-processed by the German Aerospace Centre (DLR) in Oberpfaffenhofen using ATCOR4 (Richter and Schlaepfer, 2002) for atmospheric correction and PARGE (Schlaepfer and Richter 2002) for geometric correction. In the 2008 image, some clouds were identified, which fortunately did not fall within the pipeline area where the research was focused on. Small discrepancies were found in spatial alignment and in the spectral resolution between the two images, which were taken into account in the analysis, using appropriate correction techniques if seen necessary (see section 3.3.2). The ancillary datasets were provided by the engineering consultancy DHV, who worked for the private oil company NAM (Nederlandse Aardolie Maatschappij) owning the pipeline. These were mainly used for visual display and result interpretation.

---

[1] http://www.agrigeoplaza.com/pirireis.html

The geo-chemical drilling data (groundwater and soil sample information) was provided by DHV/NAM. The groundwater dataset contained information on the contamination concentration (µg/l), location (x and y-coordinates) and the sampling depths (varied from 1.5 m to 4.5 m) in years 2003 and 2008. The soil drilling data included the same information except that in some locations the pollution level was indicated with the soil threshold-values defined by the Dutch law[2] in years 2002/2003. All the drilling data was organized in shape-files according to years, while the assigned geographic coordinates measured with DGPS (personal communication with Dr. Mark van der Meijde) allowed accurate mapping and display of the observed locations. The detailed information for each of the drilling location has been summarized in Appendices 1 and 2.

3.3. Methodology

3.3.1. Methodology overview

The methodology used for the analysis of the HyMap data in detection of HC pollution in non-vegetated and vegetated areas and over time, is outlined in Figure 3.2. Detailed description of the different techniques used can be found in the subsequent sections (see 3.3.2 – 3.3.8).

3.3.2. Data pre-processing

The general differences between the two images were visually inspected. A small non-linear spatial misalignment (2 - 3 pixels) between the two images existed, despite of the fact that they were geometrically corrected by DSL (see section 3.2.). In order to acquire reliable change detection, accurate image registration is necessary (Dai and Khorram, 1999). This can be performed manually or automatically (Dai and Khorram, 1999 and Li et al., 1995). In this research, a manual procedure was applied by selecting tie points from objects such as roads, roofs and field corners to register the 2008 HyMap image to the 2005 HyMap scene (i.e. it was assumed that the 2005 image was accurately registered by DSL). All together 100 tie points were used, which were equally distributed across the image. For image warping, cubic convolution method (Keys, 1981) was chosen since the misalignment between the images was non-linear and cubic convolution has been recommended by many researchers as one of the most accurate interpolation method for resampling images (e.g. Lehmann et al., 1999 and Shi and Reichenbach, 2006). The accuracy of the image-to-image registration was measured with Root Mean Square Error (RMS) and with visual inspection by overlaying the two images.

---

[2] http://www2.minvrom.nl/Docs/internationaal/annexS_I2000.pdf

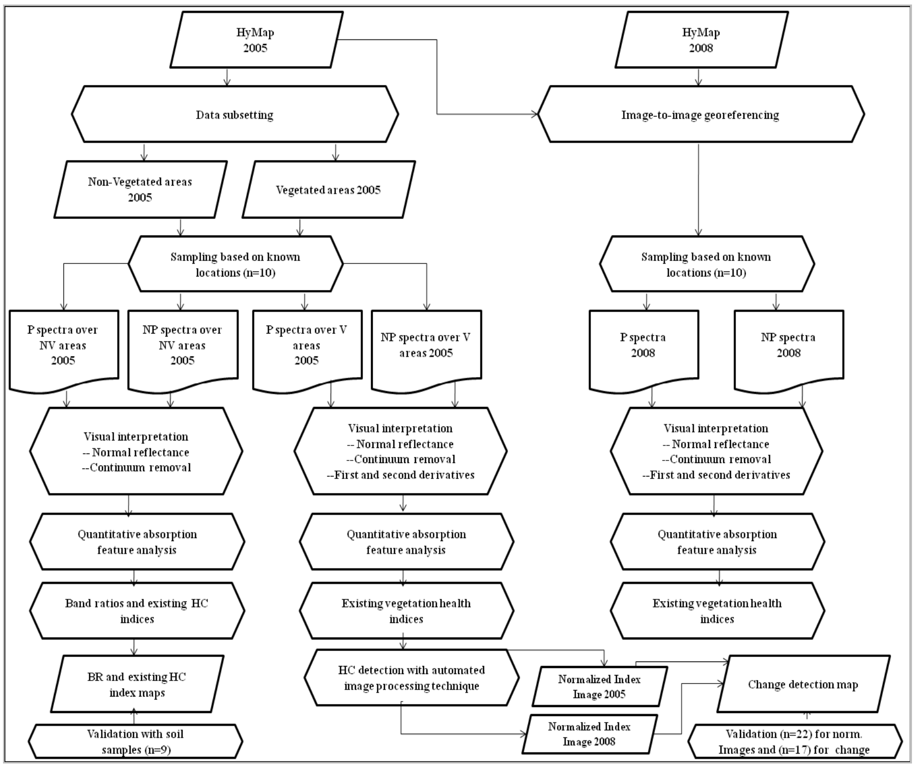


Figure 3.2 *The general overview of the methodology applied in the study. Detailed description of each analysis can be found in sections 3.3.2. – 3.3.8.*

In change detection analysis, the spectral differences between imagery should be taken into consideration when investigating changes over time through remotely sensed data (Jensen, 1996). In this research, the wavelength bands of the 2008 image were located at slightly different wavelengths (i.e. from 0.0003 – 0.02 µm) compared to the bands of the 2005 image. This could be corrected with spectral resampling with different types of cross-calibration methods such as the probably most commonly used empirical line approach (e.g. Furby and Campbell, 2001 and Hall et al., 1991). Spectral resampling may however result in loss of valuable information (Li et al., 2005). A preprocessing test was therefore conducted. The spectra of the 2008 image were first resampled with the empirical line method to correspond to the 2005 image spectra. A sensitivity analysis was then performed by calculating vegetation indices with the original and the resampled spectra. A comparison of the acquired results led to conclusions that the spectral resampling could result in a loss of valuable information in particular in the inflexion points of the spectra. As inflexion points were an important element in this research, it was decided to use the original spectra from both of the images.

For data subsetting (i.e. the delineation of vegetated and non-vegetated areas), the band ratio technique was chosen. It has been used by many researchers to delineate regions of interest for further investigation (e.g. van der Meer et al., 2002 and van der Werff et al., 2008). According to van der Meer et al. (2002), infrared/red (IR/R) band ratios allow effective differentiation of vegetated sites from non-vegetated surfaces (i.e. open water, man-made features, bare soil), based on the specific spectral response of vegetation in the VNIR part of the spectrum (i.e. strong absorption of red light in the VIS and high reflectance in the NIR part of the EMS). The hyperspectral datasets were divided into non-vegetated and vegetated subsets with probably the most commonly known IR/R band ratio NDVI (Normalized Difference Vegetation Index, Rouse et al., 1974). The threshold-value was based on previous studies where NDVI-values < 0.2 have been used for non-vegetated areas (e.g. Momeni and Saradjian, 2007 and Sobrino and Raissouni, 2000). According to these authors, the threshold value may vary from one study site to another depending on the heterogeneity of the land cover. Since previous studies have shown that vegetation may hamper the detection of unique HC absorption features (e.g. Hörig et al., 2001), the original HyMap spectra were carefully inspected with the calculated NDVI-values. The specific spectral response of vegetation in the VNIR part of the EMS (see above) was not visible under NDVI of 0.15 and therefore it was seen as an appropriate threshold-value for the study.

3.3.3. Sampling

The groundwater data (n=10) used for the spectral analysis (Table 3.1) was overlaid with the subsets created in section 3.3.2.

Table 3.1 The HC concentration (μg/l) based on the geochemical measurements and the NDVI-values acquired from the 2005 and 2008 NDVI-images in the known drilling locations used for the spectral analysis.

Location	HC concentration * 2003 (μg/l)	HC concentration * 2008 (μg/l)	NDVI 2005 (NV/V**)	NDVI 2008 (NV/V**)
16	0	0	0.95 (V)	0.46 (V)
28	0	< 0.2	0.12 (NV)	0.91 (V)
11	< 0.2	< 0.2	0.95 (V)	0.76 (V)
21	0.89	< 0.2	0.93 (V)	0.20 (NV)
24	1.1	< 0.2	0.90 (V)	0.15 (NV)
29	150	< 0.2	0.12 (NV)	0.88 (V)
8	320	< 0.2	0.78 (V)	0.76 (V)
10	2 000	11	0.95 (V)	0.76 (V)
1	32 000	< 0.2	0.51 (V)	0.68 (V)
30	94 000	< 0.2	0.11 (NV)	0.84 (V)

*<0.2 μg/l = groundwater is considered clean, >30 μg/l = intervention level (i.e. remediation actions are needed)[3]

**NV=Non-vegetated and V=Vegetated based on the division conducted in section 3.3.2.

---

[3] http://www2.minvrom.nl/Docs/internationaal/annexS_I2000.pdf

The spectra from these known groundwater locations were collected as illustrated in Figure 3.3. The sampling of non-polluted pixels (n=30) was performed from the same fields as where the drilling points were located. The field and road edges were avoided, since these locations may cause unwanted noise on the spectra. For example, shadowing and possible cross-pollution (e.g. due to exhaust fumes) may affect on the growth of vegetation, causing artefacts or false anomalies in the spectra not related to HC pollution. The size of the edges was selected based on the experiment of Harald van der Werff (2009) with hyperspectral edge removal algorithm applied on a subset of the 2005 image, in which it was shown that in most of the sampled fields the edges were approximately 2 – 4 pixels in size. After discarding the edges, it was assumed that the soil quality and vegetation type and distribution were homogeneous in the remaining parts of the fields.

Different studies suggest that HCs may migrate horizontally one or several meters from a point source. For example, Smith et al. (2004) found that HCs have horizontal extent of one meter in thick clayey soils and van der Meijde et al. (2009) indicated that possible indicators of HC pollution were not found further than 10 meters away from a leaking pipeline. DHV (2008) stated that in the current study area, the pollution from one location had spread more than ten meters away from the pipeline. As the spread of HC pollution is dependent on many factors such as soil parent material (Kostecki and Calabrese, 1991), the non-polluted spectra (n=30) were collected beyond 20 meters away from the pipeline, in order to minimize the possibility of cross-pollution. It was assumed that HC pollution had not spread beyond this buffer, as the soils in the study area were mainly clay (Poelman, 1966) and the migration of HCs is decreased by the low porosity of soils, which is the case of clays (Kostecki and Calabrese, 1991).

The polluted spectra above the pipeline were collected from the pixel in which the drilling location fell and the eight locations surrounding it (n=9). As mentioned above, previous studies have indicated that the HC pollution might spread horizontally to some extent from a point source (e.g. Smith et al., 2004, van der Meijde et al., 2009 and DHV, 2008). Using the neighbouring pixels in the analysis, but not exceeding the 20 meter buffer of possibly polluted zone, the collection of spectra from possibly polluted sites was maximized (i.e. due to the limited amount of drilling data available, it could not be established if HC plumes were only located in that particular pixel in which the drilling data fell).

From this point onwards, the spectra collected on top of the pipeline are referred to as ‘polluted’ while the spectra collected further away are denoted as ‘non-polluted’. This applies to all the sampled locations despite of the level of HC concentration.

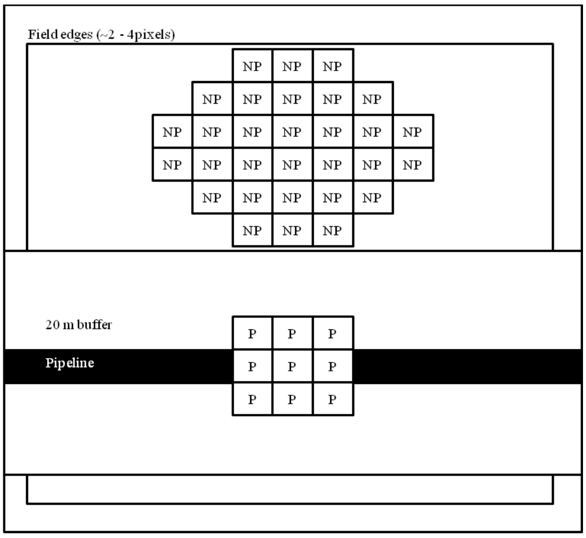


Figure 3.3 *Schematic illustration of the sampling scheme used to collect the spectra from polluted (P=9 pixels) and non-polluted (NP=30 pixels) sites from the drilling locations with known HC concentration.*

3.3.4. Visual interpretation

The average spectra from the polluted and non-polluted sites were plotted in the same graph for each location separately and visually inspected throughout the whole EMS (from 0.4 – 2.5 µm). The focus was on the wavelength regions at which diagnostic HC features have been identified previously (see section 2.2.2) and the main vegetation absorption features that could reveal vegetation stress related to HCs (see section 2.2.3).

In order to enhance the spectral differences and limit the effect of spectral brightness, the reflectance spectra were analyzed with continuum removal (CR). In the CR, the spectra are normalized by applying a convex hull over that part of the spectrum that will be analyzed (Kokaly and Clark, 1999). When the reflectance at a certain wavelength in the absorption feature is divided by the value of the hull at that wavelength, a relative reflectance value of the absorption between 0 – 1 is acquired (Clark and Roush, 1984). CR has been used for example to find relationship between absorption feature depths and biochemical concentrations of dried and fresh plant materials (Kokaly and Clark, 1999 and Mutanga et al., 2004). Furthermore, Noomen (2007) applied CR to correlate gas concentrations (i.e. methane, ethane and natural gas) with the band depths of the main vegetation absorption features, in order to see if the gasses had caused stress symptoms in wheat and maize.

The data collected from vegetated areas was further explored with first and second derivatives, which are known to provide a sensitive analysis of the subtle absorption and reflectance features in the spectra associated with vegetation stress (Filella and Penuelas, 1998). They highlight the wavelength regions where the rate of change in reflectance is high (Smith et al., 2004). The derivatives have been commonly used to investigate for example the subtle changes in the 'Red edge' region at 0.7 µm of the vegetation spectra (e.g. Merton, 1998, Smith et al., 2004 and White, 2007).

3.3.5. Quantitative absorption feature (QAF) analysis

Quantitative absorption feature (QAF) analysis is commonly used in identifying materials from RS data (van der Meer, 2004). For example in soil science it has been used to examine soil constituents such as organic matter (e.g. Ben-Dor et al., 1997) and in vegetation studies to estimate foliar biochemistry (e.g. chlorophyll content, Lichtenthaler et al., 1996). The method is based on the fact that the position (P), depth (D), width (W), area (A) and asymmetry (AS) of the distinct absorption features in different parts of the EMS, are controlled by the crystal configuration of the sample and the chemical structure of the material (van der Meer, 2004). This allows the size and the shape of the absorption features to be used as indicators of materials present in a sample (van der Meer, 2004).

The P of the absorption feature is defined as the minimum CR reflectance (van der Meer, 2004). The W is the range between the left and the right shoulder of the absorption feature (i.e. where the CR reflectance has a value of 1) (van der Meer et al., 2009). The D, A, and AS are calculated with the following equations: (Eq. 3.1.) $D(\lambda) = 1 - CRR(\lambda)$ (Kokaly and Clark, 1999), where CRR is the continuum removed reflectance; (Eq. 3.2.) $A(\lambda) = (L1) + (R1)$ where $L1$ and $R1$ are the area between the maximum wavelength position and the left and right shoulder of the absorption feature, respectively (van der Meer et al., 2009); (Eq. 3.3.) $AS(\lambda) = A_{left} / A_{right}$, where A_{left} is the area of the absorption feature from the left shoulder to the maximum point and A_{right} is the area from the maximum point to right shoulder of the absorption feature. Positive asymmetry values (from 0 to infinity) indicate that the absorption feature is skewed to the longer wavelengths while negative represent the opposite (van der Meer, 2004). In several studies, the D of the absorption feature has been utilized when exploring absorption features of different materials (e.g. Noomen, 2007, Mutanga et al., 2004 and Mutanga et al., 2005). The investigation of other absorption features parameters has not been so profound.

The QAF-analysis varied slightly for the spectra from non-vegetated and vegetated sites, in terms of choosing absorption features for the analysis. For the non-vegetated spectra, all the possibly HC related absorption features identified in the visual interpretation were analyzed, since the subtle differences may not be clear by only visually inspecting the spectra. For the spectra from the vegetated locations, the main vegetation absorption features were further analyzed, in order to investigate if indicators of vegetation stress could be found. For example, the amount of chlorophyll may decrease in stressed vegetation that could be seen as shallower absorption features in the known chlorophyll features at 0.44 µm and 0.67 µm (Carter, 1994). The QAF-analysis for the main vegetation absorption features was applied on the wavelength regions suggested by Noomen (2007) (Figure 3.4). This was done with an adjustment on the available HyMap bands (Table 3.2) where the edges of the absorption feature were situated in such way that the deepest absorption was located approximately in the centre of each of the absorption features (Noomen, 2007).

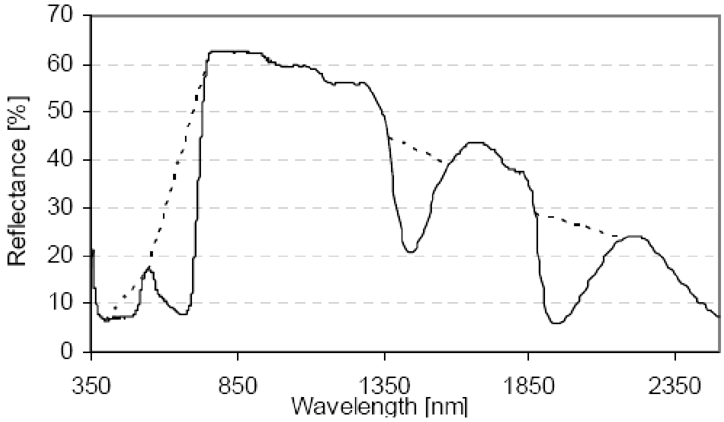


Figure 3.4 *The main wavelength regions in which QAF-analyses were applied (indicated with dotted line, copied with permission from Noomen, 2007). See Table 3.2 for starting and ending wavelength points for each of the regions.*

Table 3.2 The exact HyMap wavelength regions (2005 and 2008) of the main vegetation absorption features in which the QAF-analysis was applied and their function in plants.

Region	HyMap wavelength (nm) 2005	HyMap wavelength (nm) 2008	Function in plants
VIS (blue)	435 – 556	457 – 545	Chlorophyll
VIS (red)	556 – 770	545 – 777	Chlorophyll
SWIR	1340 – 1613	1333 – 1611	Water
SWIR	1811 – 2189	1797 – 2191	Water

The identified subtle differences between the absorption feature parameters of polluted and non-polluted spectra could be also related to other factors such as changes in soil background reflectance (Kokaly and Clark, 1999). A Normalization procedure was therefore performed on the main vegetation absorption features using the following equation: (Eq. 3.4.) $NBD(\lambda) = BD(\lambda) / BD_{max}$ (Kokaly and Clark, 1999), in which the band depth ($BD(\lambda)$) (or any other feature value such as area A) at a certain wavelength was divided by the maximum band depth (BD_{max}) for that absorption feature. This resulted in normalized band depth values ($NBD(\lambda)$) or normalized feature values for any of the measured absorption feature parameters.

The differences in the original and normalized absorption feature parameters between polluted and non-polluted spectra were visually inspected with Box plots. In case there was consistent variation (either negative or positive) in P, D, W, A, or AS between polluted and non-polluted spectra at the clean and polluted locations, the differences in the absorption feature parameters were compared statistically. The absorption feature parameter values were first tested for normality with Shapiro-Wilkins W test (Moore, 2000). Since all the datasets were not normally distributed and the applied normalization procedure (log) did not normalize these datasets, either parametric *Student's t-test* (Moore, 2000) (for normally distributed data) or non-parametric *Mann-Whitney U-test* (Moore, 2000) (for not normally distributed data) was used to identify significant differences (p-level ≤ 0.05).

3.3.6. Band ratios and existing HC indices

Band ratios (BRs) and existing HC indices have been used to detect HC pollution from hyperspectral data with promising results (Kühn et al., 2001, McCoy et al., 2001 and Winkelmann, 2005). Several existing HC indices and BRs built in this study were applied over the non-vegetated areas in the 2005 image. The existing HC indices used are listed in Appendix 3. These indices have been built on the wavelength regions where the known HC absorption features are located in such way, that higher index values indicate the presence of HCs (Kühn et al., 2001, McCoy et al., 2001 and Winkelmann, 2005).

The BRs calculated in this study, were simple mathematical operations in which carefully selected wavelength bands were divided with each other in order to extract information on the presence of HCs in the image. They were built on the wavelength regions in which consistent variation (either negative or positive) was found in all of the contaminated locations in the depth (D) of the absorption feature between polluted and non-polluted spectra in the QAF-analysis. As noted by McCoy et al. (2001) and Penn (2002), the wavelength band where the material was highly reflective (i.e. the CR reflectance was one), was chosen as the numerator and the denominator was the wavelength band where the absorption feature was present. This resulted in index values that were expected to be either higher or lower in the known polluted sites, depending on if the identified absorption feature in the QAF-analysis was deeper or shallower for the polluted spectra (i.e. if the absorption feature was deeper for the polluted than for the non-polluted spectra, lower index were expected at the polluted areas and *vice versa*)

For each of the produced images, the pixel values at the known polluted and non-polluted locations were inspected in detail. The images were further validated with additional drilling data from 2003 (n=9) that had not been included in the analysis (see Appendix 2 for further details).

3.3.7. Vegetation health indices

In remote sensing studies, vegetation indices that are a combination of the VIS and NIR reflectance are often used to assess plant and canopy health (e.g. Carter, 1993, Haboudane et al., 2004, Zarco-Tejada et al., 2005). Most of these indices react to changes in leaf chlorophyll content and canopy structure at wavelengths 0.67 µm and 0.8 µm, respectively (Haboudane et al., 2004). In addition, indices built around the main water absorption features in the SWIR (1.4 and 1.9 µm) have been developed to take into account the possible changes in leaf water content caused by water deficiency and therefore vegetation stress (e.g. Penuelas et al., 1997).

For the current study, 50 indices were compiled from the literature and calculated for the collected spectra from polluted and non-polluted sites over the vegetated areas (see Appendix 4 for the index abbreviations and the equations and Appendix 5 for the exact HyMap wavelengths used for the calculation). The applied indices included for example narrow-band indices developed to detect the subtle changes in the 'Red edge' position (e.g. REP, Guyot and Baret, 1988). Since other factors such as underlying soil may hamper the detection of subtle differences between polluted and non-polluted sites (Darvishzadeh et al., 2006 and Miller et al., 1997), several vegetation indices taking into account the background reflectance of soils, such as OSAVI (Rondeaux et al., 1996) and MSAVI (Qi et al., 1994) were further tested.

The differences in the calculated index values of polluted and non-polluted sites were inspected with Box plots that were created separately for each index at all of the sampled locations. The differences in the index values were compared statistically in the same way as described in section 3.3.5. The aim of this part of the methodology was to identify the index or indices that could differentiate vegetation stress related to HC pollution over partially and densely vegetated areas. This information was further utilized in the detection of HCs with the automated image processing technique as described in the subsequent section.

3.3.8. HC detection with automated image processing technique

A subset from both of the HyMap images was created with 200 m buffer zone from the pipeline in order to decrease computation time. The non-vegetated areas were masked out as the performance of vegetation indices in areas with less vegetation cover was not known (van der Werff et al. 2008). The index or indices that were found to differentiate vegetation stress related to HC pollution in the vegetation index analysis were calculated on the subsets as described in section 3.3.7. The index values were further normalized with a normalization technique developed by van der Werff et al. (2008). This was done in order to acquire normalized datasets that were comparable between the two images taken in different time steps. It further enhanced the variation within instead of between the agricultural fields, allowing vegetation stress detection, independent on vegetation cover or species (van der Werff et al., 2008).

In the automated procedure, the average value of a selected region was subtracted from each image pixel (van der Werff et al., 2008). The average regions, created with region growing procedure, were set to grow in a ring of 5 pixels (20 m) distance from the seed (centre pixel), in order to prevent possible anomalous regions for example along the pipeline to influence the average index values. The regions were growing 40 pixels (160 m) from the seed unless a user-defined index threshold value was exceeded. This threshold-value was chosen based on the variation (σ) of the index value(s) over the polluted locations, found from the box plots created in section 3.3.7. The produced average image was subtracted from the original index image and the pixel values were scaled from – 1 to 1 with respect to the average background value.

The normalization procedure was first applied on the index images from both of the years. The change detection map was then acquired by simply subtracting the produced 2008 image from the 2005 image. The definition for the acquired normalized index values has been given in Table 3.3.

Table 3.3 The definition of the pixel values in the normalized vegetation index-images.

2005	2008	Change * (2005 – 2008)
Healthy (+)	Healthy (+)	Healthy (0) or Healthier (-)
Stressed (-)	Stressed (-)	Stressed (0) or More stressed (+)
Healthy (+)	Stressed (-)	More Stressed (+)
Stressed (-)	Healthy (+)	Healthier (-)

*NOTE! The values represented the opposite in the change detection image than in the normalized 2005 and 2008 index-images.

The pipeline vector layer was used to detect anomalous regions close to pipeline that would be most likely related to only HC pollution. It was assumed that all the anomalous regions further than 20 meters away (see section 3.3.3) were not related to pollution. Spectral and spatial criteria such as anomalies related to shades from high vegetation and traces of agricultural activities were not considered as anomalies related to pollution. Furthermore, van der Werff et al. (2008) detected several (n=28) 'most likely', 'likely' and 'possible' HC related anomalous regions on the 2005 image from the same study area. At these sites, the change was further inspected in detail with the change detection analysis.

The normalized index-image(s) from 2005 was validated with additional drilling data available for the vegetated areas (for further details see section 3.2 and Appendix 2). Two classes were created: 0 = polluted (negative values) and 1 = non-polluted (positive values). The drilling data was classified to polluted and non-polluted based on the given threshold-values in the geochemical data (for further details, see Appendix 2). In the confusion matrix, the following accuracy measures were calculated: overall accuracy and kappa coefficient for the whole map and user and producer accuracies for both of the classes. The definition for each of the accuracy measurements has been given in Table 3.4, which is based on the review of Congalton (1991). Since these drilling data included geochemical record only for 2002/2003 (Appendix 2), the level of accuracy for the normalized 2008 index image could not be reliably calculated with these data points as they were already used for the accuracy assessment of the 2005 image. In addition, more time had passed between drilling measurements and the image acquisition in year 2008. The accuracy of the 2008 image was therefore assumed to be reflected in the validation of the change detection map, which is described in the following paragraph.

The change detection map was validated with confusion matrix using the groundwater data (n=17) having geochemical record from 2003 and 2008 (for further details see section 3.2 and Appendix 1). Two classes (0=no change and 1= change) were created as shown in Table 3.5. The intervention value for benzene polluted groundwater (30µg/l[4]) was used as a threshold value to evaluate if change had occurred in the ground truth data (see Table 3.5 for examples). This value has been derived from the soil/sediment intervention values that are based on the extensive studies of the National Institute for Public health and environmental Protection of both human and ecotoxicological (i.e. animals and plants) effects of soil contaminants (Ministerie van Volkshuisvesting, 2000). It was therefore seen as appropriate threshold-value to indicate in which HC concentration levels the vegetation could yet be affected by the pollution.

In principle, the change detection map was produced in such way that values close to zero indicated little or no change and values close to -1 and 1 stronger negative (healthier) or positive (stressed) change. Different threshold values to define the change were tested: 0.1, 0.2., 0.3, 0.4, 0.5 and 0.6 and similarly -0.1, -0.2, -0.3, -0.4, -0.5 and -0.6 on the negative sides. This resulted in calculating six different confusion matrices that were further compared with each other. This was done in terms of choosing the most appropriate threshold-value to be used in detecting the change over possibly HC polluted areas with the normalized image processing method. The locations, in which changes possibly related to HCs were detected, were also inspected in order to further evaluate the difference when using various threshold-values.

---

[4] http://www2.minvrom.nl/Docs/internationaal/annexS_I2000.pdf

Table 3.4 The definition and the interpretation of the calculated accuracy parameters in the validation of the normalized index images (based on Congalton, 1991).

Accuracy parameters	Definition	Value range	Value Interpretation
Overall accuracy	The probability that a randomly chosen point is correctly classified in a map	0 – 100 %	Higher values indicate improved accuracy
Kappa coefficient	How well the map corresponds to a map produced by chance	– 1 to 1	< 0 = worse than chance; 0 = as good as map produced by chance; > 0 = better than a map produced by chance
User Accuracy	For each class the probability that a classified point actually represents that class in reality	0 – 100 %	Higher values indicate improved accuracy
Producer accuracy	For each class the probability that a randomly chosen point on the field is actually that class in the map	0 – 100 %	Higher values indicate improved accuracy

Table 3.5 The classes used in the validation of the change detection map and examples of how the intervention value of benzene polluted groundwater (30 µg/l[5]) was used to evaluate if a change had occurred in the groundwater data.

Class	Description		Change
0	polluted to polluted	clean to clean	NO
1	polluted to clean	clean to polluted	YES
Example	**HC 2003 (µg/l)**	**HC 2008 (µg/l)**	**Change**
1	< 0.2	< 0.2	NO
2	1900	480	NO
3	2000	11	YES
4	94000	< 0.2	YES
5	1.9	1000	YES

---

[5] http://www2.minvrom.nl/Docs/internationaal/annexS_I2000.pdf

4. HC pollution detection over non-vegetated areas from the 2005 image

As mentioned in the literature review, diagnostic HC absorption features in certain parts of the EMS have been used to investigate HC pollution from RS data (see section 2.2.2). The results of these studies have been promising but yet inconclusive. In order to investigate if HyMap data from 2005 could be used to detect HC pollution, several different spectral techniques (e.g. QAF-analysis, BRs and existing HC indices) were applied over the non-vegetated areas on the 2005 image (for further details see sections 3.3.3 – 3.3.6). In this chapter, the acquired results from the analysis are described and discussed, in relation to the findings of other researchers. The aim of this chapter was to answer the research questions under objective 1 (see section 1.3).

4.1. Results

4.1.1. Visual interpretation

Figure 4.1 depicts the average normal reflectance spectra from polluted and non-polluted sites collected from the three non-vegetated locations along the pipeline. At the two contaminated locations (Locations 29 and 30), the polluted spectra deviated from the standard deviation of the clean spectra only at location 29 (Figure 4.1). At this site, the polluted spectrum was more reflective than the non-polluted spectrum throughout the whole EMS (Figure 4.1). At the clean location 28, the polluted spectra were less reflective compared to the non-polluted spectra (Figure 4.1).

In the CR spectra, variation between polluted and non-polluted spectra was identified at several wavelength regions within the VNIR and SWIR part of the EMS (Figures 4.2 and 4.3, respectively). The detected wavelength regions and their possible function in relation to HCs are summarized in Table 4.1.

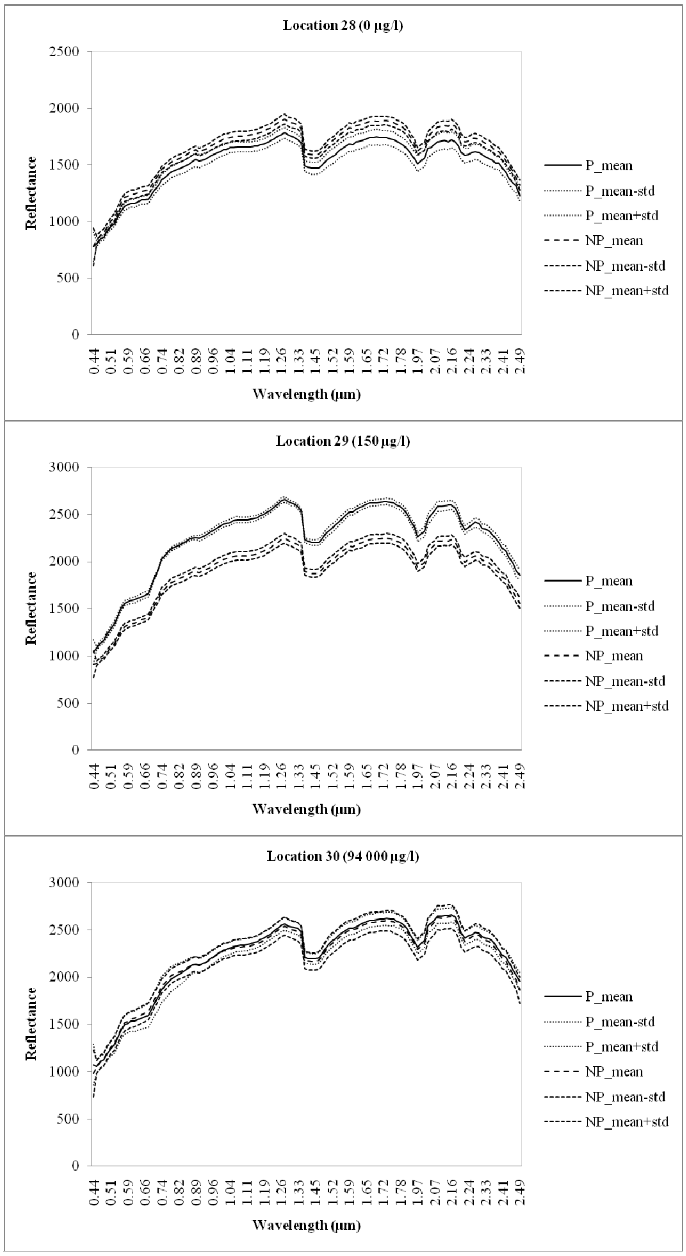


Figure 4.1 *The normal spectra of the sampled non-vegetated locations along the pipeline from the 2005 image analysis: 28 (0 µg/l) (uppermost), 29 (150 µg/l) (middle) and 30 (94 000 µg/l) (lowermost). No consistent differences were identified between the polluted (P) and non-polluted (NP) sites in the normal reflectance spectra of the sampled locations.*

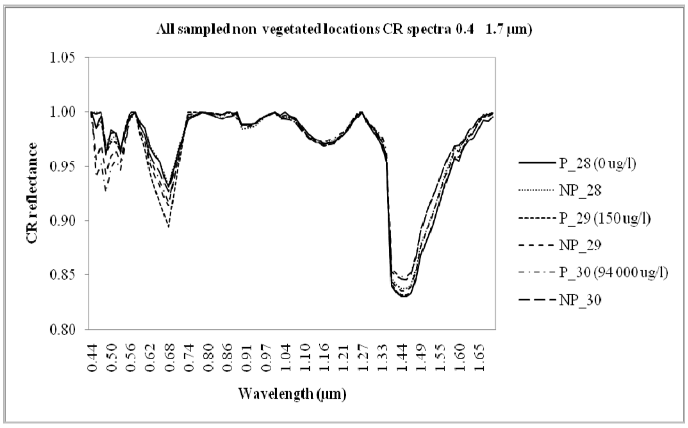


Figure 4.2 *The average continuum removed (CR) spectra from 0.4 – 1.7 µm for all the non-vegetated locations from the 2005 image analysis. The absorption features in which differences were identified between polluted (P) and non-polluted (NP) spectra are listed in table 4.1.*

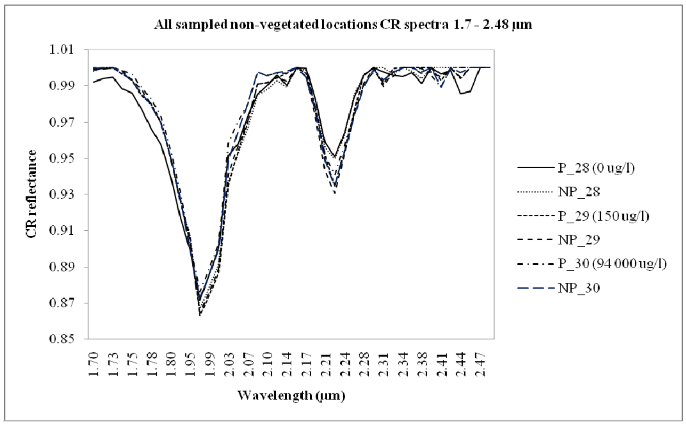


Figure 4.3 *The average continuum removed (CR) spectra from 1.7 – 2.45 µm for all the non-vegetated locations from the 2005 image analysis. The absorption features in which differences were identified between polluted (P) and non-polluted (NP) spectra are listed in table 4.1.*

Table 4.1 The wavelength regions at which differences were identified between polluted and non-polluted CR spectra at the three sampled non-vegetated locations in the 2005 image analysis. The possible function of these absorption features in relation to HCs is indicated.

No	VIS/NIR /SWIR	Wavelength (µm)	Function in relation to HCs	Reference
1	VIS	0.45 - 0.5	Unknown	-
2	VIS	0.67	Unknown	-
3	NIR	0.9	H_2O or HC	(Winkelmann, 2005)
4	NIR	1.15	Unknown	-
5	SWIR	1.4	Water	(Jensen, 2000)
6	SWIR	1.9	Water	(Jensen, 2000)
7	SWIR	2.1	HC	(e.g. Cloutis, 1989, Hörig et al., 2001)
8	SWIR	2.2	HC	(e.g. Cloutis, 1989, van der Meer et al., 2000a)
9	SWIR	2.3	HC or CO_3^{2-}	(e.g. Cloutis, 1989, van der Meer et al., 2000a)
10	SWIR	2.35 – 2.37	HC or CO_3^{2-}	(e.g. Cloutis, 1989, van der Meer et al., 2000a)
11	SWIR	2.4	HC or CO_3^{2-}	(e.g. Cloutis, 1989, van der Meer et al., 2000a)
12	SWIR	2.45	HC	(e.g. Xu et al., 2008)

4.1.2. QAF-analysis

The results of the QAF-analysis of the 2005 image are listed in the absorption feature parameter table in Appendix 6 including the exact HyMap bands used for the QAF-analysis. At all of the calculated wavelengths, most of the absorption features parameter values were not significantly different between polluted and non-polluted spectra at any of the sampled locations (Appendix 6). A diagnostic difference was found only at wavelength 2.2 µm, where the A of the absorption feature was significantly smaller for polluted than for non-polluted spectra at the two contaminated locations (Locations 29 and 30) ($p \leq 0.05$) (Figure 4.4 and also Appendix 6).

The depth (D) of the absorption feature was consistently smaller at wavelengths 2.2 µm and 2.35 – 2.37 µm and higher at wavelengths 0.67 µm, 2.4 µm and 2.45 µm at both of the contaminated locations (Locations 29 and 30, Appendix 6). Even though the differences in the D of these wavelengths were not statistically significant (for more details, see Appendix 6), band ratios were established on these absorption features in order to test if they could detect HCs over the whole study area (for further details, see section 3.3.6).

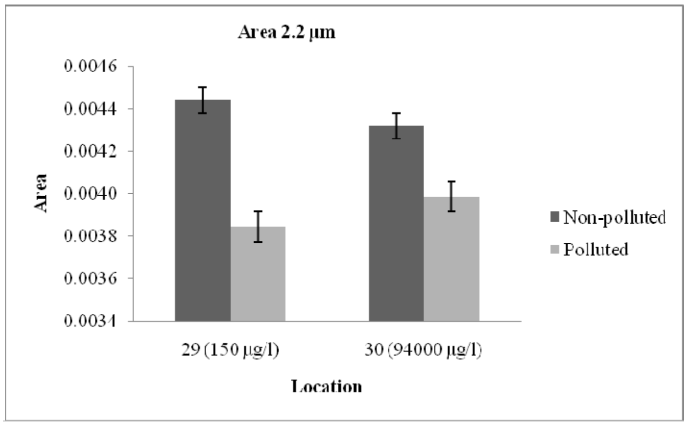


Figure 4.4 *The most diagnostic difference identified in the QAF-analysis at wavelength 2.2 µm at the two contaminated non-vegetated locations (29 and 30) in the 2005 image analysis. The area (A) of the absorption feature was significantly smaller for polluted than for non-polluted spectra ($p \leq 0.05$) at both of the sites.*

It could be the case that the HC plumes had not spread away from the drilling location to the eight surrounding cells as was assumed in section 3.3.3. The same QAF-analysis was therefore conducted for all the non-vegetated locations (Locations 28, 29 and 30, Table 3.1) by comparing the polluted pixel in which the drilling location fell against the thirty non-polluted pixels collected from the same field. As a result, no difference was found between this analysis and the findings of the investigation where the nine polluted pixels were used.

4.1.3. BRs and existing HC indices

The calculated BRs (n=5) built on the wavelengths where the depth (D) of the absorption feature was either smaller or larger for polluted spectra at both of the contaminated locations (for further details, see section 4.1.2), are listed in Appendix 3. The HyMap bands used for the calculation and the expected values for polluted locations are indicated in Appendix 3. All the band ratios represented here were calculated over the non-vegetated areas in the 2005 image as described in section 3.3.6.

No clear difference between any locations along the pipeline and further away was acquired with band ratios 2.39/2.37 µm and 2.45/2.43 µm. The BR image 2.39/2.37 µm was almost fully covered with low index values. In contrast, high index values were acquired across the whole image for the band ratio 2.45/2.43 µm.

The BR image 0.78/0.67 µm showed patches of higher index values within the fields from which not all of them were close to the pipeline and especially in the pixels surrounding the areas that were masked out as vegetated pixels (i.e. the boundary regions between non-vegetated and vegetated areas).

The BR 2.39/2.4 µm was speckled and did not exhibit clear spatial patches close to the pipeline that could be clearly related to HC pollution. Detailed investigation of the index values in the known locations revealed that the average index values were similar but a little higher for the contaminated locations (index value 1.16 for locations 29 and 30) than for the clean site (index value 1.15 for location 28). This result was not as expected, since this band ratio was built in such way that the contaminated locations should exhibit lower index values (see Appendix 3 for further details).

The BR image 2.1/2.2 µm (Figure 4.5) exhibited higher values for the contaminated sites (index values 1.11 and 1.10 for locations 29 and 30, respectively) than for the clean area (index value 1.08 for location 28). In addition, close to location 28 clear patches of higher values adjacent to the pipeline were detected (Figure 4.5). This suggested possible pollution, since higher values were expected on the polluted locations with the BR 2.1/2.2 µm (for further details, see Appendix 3).

The validation of the produced BR images with 9 additional non-vegetated locations did not indicate consistent difference between pixel values of polluted and non-polluted sites in any of the BR images (Figure 4.6).

In the analysis with the existing HC indices, the HI-index image (Kühn et al., 2001) was speckled and the differentiation between possibly polluted areas along the pipeline could not be made. The band ratio images by Winkelmann (2005) and McCoy et al. (2001) provided either similar results as the HI-index or in some cases only the roads or/and roofs were detected as polluted (i.e. higher index values were acquired). No clear spatial patches were found close to the pipeline, indicating that HC pollution due to pipeline leakage was not detected with any of the existing HC indices. Validation with the known drilling locations did not reveal consistently higher index values for the known polluted than for the non-polluted sites.

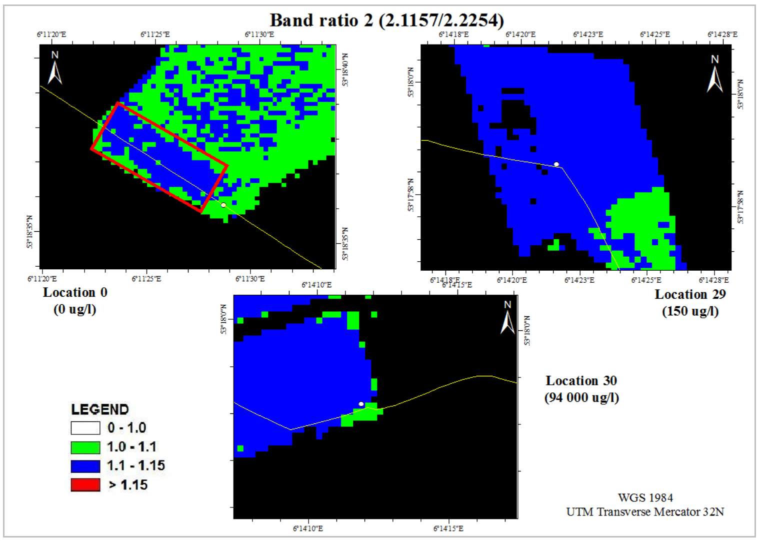


Figure 4.5 *Band ratio image (2.1/2.2 µm) over the non-vegetated areas in the 2005 image. Close to location 28, adjacent to the pipeline, higher pixel values were acquired (indicated with red rectangular) as was expected at polluted sites with this particular BR (for further details, see Appendix 3).*

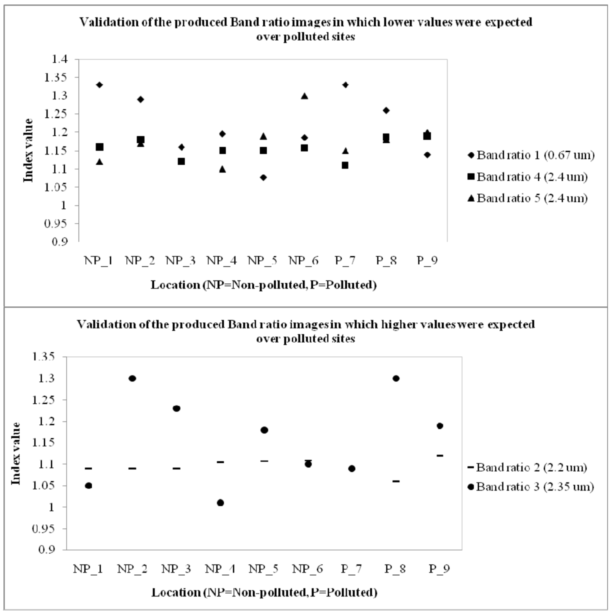


Figure 4.6 *The validation of the band ratio images produced in this study with nine additional drilling locations (NP=non-polluted=6 and P=polluted=3) that were not used for the analysis.*

4.2. Discussion

According to Orlov et al. (1991), the overall reflectance should be lower for polluted than for non-polluted soils. In contrast, Winkelmann (2005) states that the overall reflectance of soil samples is not the best indicator of pollution as soil reflectance is often more affected by other materials such as organic matter and water content of the soil. In this study, only the one clean location (i.e. Location 28) exhibited lower reflectance for the spectra collected on top of the pipeline. The differences in the overall reflectance of the sampled locations are therefore indicating the variation of other substances such as organic matter in soil rather than HC pollution.

Several possibly HC related features were identified in the enhanced CR spectra as indicated in Table 4.1. It is known that wavelengths 1.4 µm and 1.9 µm are due to water as the atmospheric water denies the transmission of infrared radiation in these wavelengths (Jensen, 2000). These features were therefore not diagnostic to HCs. Wavelengths 0.9 µm, 2.1 µm, 2.2 µm, 2.3 µm, 2.35 – 2.37 µm, 2.4 µm and 2.45 µm however are known to be regions where HCs exhibit diagnostic absorption features (e.g. Cloutis, 1989, Winkelmann, 2005 and McCoy et al., 2001, Xu et al., 2008). In addition, the absorption features at 2.3 µm, 2.35 – 2.37 µm and 2.4 µm can be related to the absorption of carbonate (CO_3^{2-}) ion which is an alteration product of HC effect on soils (van der Meer et al., 2000a).

Wavelengths 0.45 µm, 0.67 µm and 1.15 µm have not been associated with HCs in previous research. It is not however fully understood how HCs change the soil environment (Schumacher, 1996). For example, HCs may cause excessive presence of other materials in the soil (van der Meer et al. 2000b) further causing changes in the otherwise normal soil spectra. The detected features could therefore be indirectly related to the presence of HCs in soil.

The detailed investigation of the detected absorption features with QAF-analysis revealed that the differences in the absorption feature parameters between polluted and non-polluted spectra were not consistent at the contaminated or clean sites in almost all of the cases. The area (A) of the absorption feature area at wavelength 2.2 µm was the only feature found to be significantly smaller for polluted spectra at the two sampled contaminated locations ($p \leq 0.05$) (Figure 4.4 and Appendix 6). The feature at wavelength 2.2 µm has been identified to be diagnostic to HCs (e.g. McCoy et al., 2001 and Winkelmann, 2005). In Figure 4.5 it was shown that with the band ratio 2.1/2.2 µm, patches of higher index values were acquired adjacent of the pipeline, suggesting that these areas are polluted. The feature at wavelength 2.2 µm can however be explained by the absorption feature of clays (Mc Coy et al., 2001) which is the main soil parent material in the study area (Poelman, 1996). Hence, the mixing or ploughing of the soil layers during and after the installation of the pipeline (personal communication with Dr. Mark van der Meijde), may have also caused the variation in the soil spectra along the pipeline, further explaining the higher index values detected close to location 28 (see Figure 4.5.). Drilling data was not available for that particular location. In order to ascertain if the identified feature at wavelength 2.2 µm was related to HCs or to that of clay, the site close to location 28 should be further verified with ground truth data.

The detection limit of HCs from clayey soils have been suggested to be 50 000 mg/kg (Winkelmann, 2005). At the location with the highest HC concentration (i.e. 94 000 µg/l), indicators of HC pollution were however not found. Hence, in the study area, the HC pollution may not have migrated to the surface in such large quantities, making the diagnostic HC absorption features to be detectable from the soil spectra. This is supported by Kim and Corapcioglu (2003), who indicated that in many cases HC pollution, is found to move downwards to the groundwater. It should however be noted that this is dependent on many factors such as soil parent material (e.g. soil porosity) and groundwater slope (Kim and Corapcioglu, 2003). Due to this, one cannot conclude for certain if the pollution had moved upwards or downwards in the study area, unless further investigations of the subsurface movement of the HCs are conducted. Nevertheless, benzene condensates, being the main HC components in the study area, are highly volatile (Chemistry Glossary, 2009). This indicates that in case the HCs migrate upwards to the surface instead of down to the groundwater, they are likely to evaporate from the non-vegetated surfaces before they can be directly detected with hyperspectral imagery.

Several other factors may have further affected the acquired results. The data discrepancy between the drilling data and HyMap imagery (i.e. difference in data acquisition time and sampling depth), may be the reason for not detecting diagnostic HC absorption features. The limited amount of drilling data did not permit establishing the exact location of the HC plumes. It could therefore be that the HC leakages were smaller than the sample size used in the study. This was tested by applying a QAF-analysis for the spectra collected from that particular pixel where the drilling data fell at all the sampled locations, which was then compared with the spectra collected from the thirty non-polluted pixels. As a result, no difference was found between this analysis and the findings when using the nine polluted pixels. It can therefore be concluded that the plumes with the highest amount of HC contamination on the surface are either smaller than the pixel size (i.e. which is the smallest identifiable unit in the image) or situated beyond the sampling unit of the possibly polluted non-vegetated pixels that were assumed in this study. With small amount of drilling data, a strong statistical relationship between the detected absorption features and HC contamination could not be established which could have improved and strengthened the results.

Kühn et al. (2001) indicated that vegetation cover may hamper the detection of diagnostic HC absorption features causing 'noisy' images when using for example the HI-index the authors developed. Many other natural and man-made materials also exhibit absorption features at the same wavelengths as HCs, causing false anomalies in the final classification maps (Winkelmann, 2005). The results from the existing HC indices and band ratio images, suggested that the remaining vegetation cover hampered the detection of diagnostic HC absorption features over the non-vegetated areas. This was clearly shown for example in the HI-index image, that was speckled (i.e. ‘noisy’) and clear differentiation between HC polluted and clean sites could not be made. Even though great care was taken to delineate non-vegetated areas (see section 3.3.2), some vegetation signature may have remained on the spectra of the polluted sites. In the current study area, the land use is mainly agricultural and vegetation dominates the spectral signature recorded by the sensor. A reason for not detecting HC features could therefore be that the soil spectra in the imagery still contain a weak vegetation signature, caused by yet too coarse spatial resolution (4x4m) of the HyMap data. In such case, the features identified at wavelengths 0.45 and 0.67 μm would not be indirectly related to HCs as suggested above, but to that of chlorophyll, as it is known that chlorophyll exhibits its two absorptions at these particular wavelengths (Jensen, 2000). Nevertheless, there are several (n ~ 10) of such non-vegetated sites along the pipeline. Different ways to detect HC pollution spectrally from these areas should therefore be explored further in order not to miss any HC contaminated sites.

5. HC pollution detection over vegetated areas from the 2005 image

In chapter 4, the results of the investigation of HC pollution with direct means from the non-vegetated areas on the 2005 image were presented in terms of answering research questions under objective 1 (see section 1.3). HC contamination may be detected indirectly, using for example vegetation stress as an indicator of pollution (see section 2.2.3). Several studies have been conducted to detect HC related vegetation stress with promising results (e.g. van der Meijde et al., 2009, van der Werff et al., 2008 and Noomen, 2007). In order to identify if HC pollution could be detected by investigating the relative differences between polluted and non-polluted spectra, several spectral techniques such as continuum removal, first and second derivative analysis and vegetation health indices were applied on the spectra collected from the vegetated locations on the 2005 image (for further details see section 3.3). In this chapter, the acquired results from the analysis are depicted and discussed, in relation to the findings of other previous studies. The aim of this chapter was to answer research questions under objectives 2 a, and b (see section 1.3).

5.1. Results

5.1.1. Visual interpretation

No differences between polluted and non-polluted spectra were identified in the normal spectra at most of the sampled locations, except at location 24 (HC conc. of 1.1 µg/l), where indicators of vegetation stress were found (Figure 5.1). At this site, the NIR shoulder was shorter and decrease of reflectance was found in the chlorophyll absorption features at the VIS (at 0.44 and 0.67 µm) and the main water absorption features at the SWIR regions (at 1.4 and 1.9 µm) of the polluted spectra (Figure 5.1).

In the CR spectra, no differences were identified between polluted and non-polluted sites in the main vegetation absorption features at most of the locations (Figure 5.2). Only at location 24 the differences between polluted and non-polluted spectra were even more visible in the CR spectra (Figure 5.2) than in the normal spectra of that particular location (Figure 5.1). Difference in the absorption feature at wavelength 2.4 – 2.5 µm was found between polluted and non-polluted spectra in all of the sampled vegetated locations (Figure 5.2).

In the first derivative analysis, no differences were identified between polluted and non-polluted spectra at most of the sampled vegetated locations (Figure 5.3). At location 24, however, the ‘Red edge’ region at 0.7 µm showed lower first derivative values for the polluted spectra compared to the non-polluted one (Figure 5.3). In the second derivative spectra, the results were showing similar differences between polluted and non-polluted spectra at all of the locations, as for the analysis of the first derivative spectra (see examples in Appendix 7).

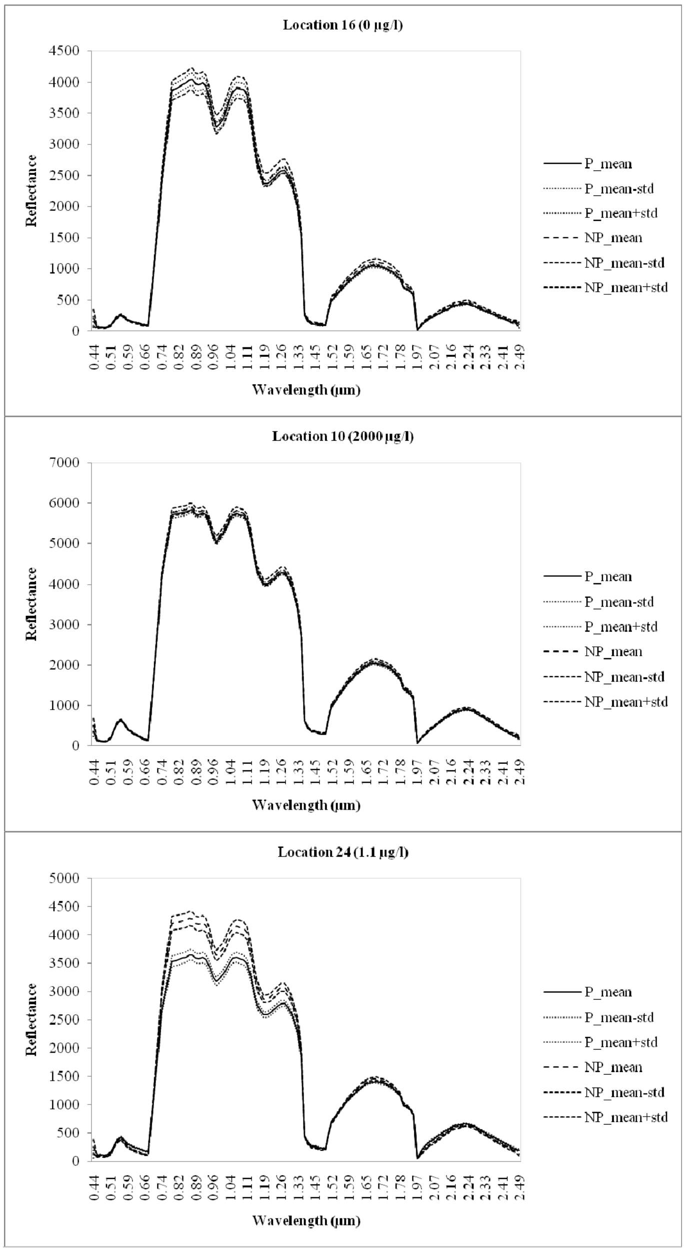


Figure 5.1 *Examples of the differences identified in the normal spectra between the polluted (P) and non-polluted (NP) sites at the sampled vegetated locations in the 2005 image analysis. No difference was identified at most of the locations except at location 24 with HC concentration of 1.1 µg/l (lowest).*

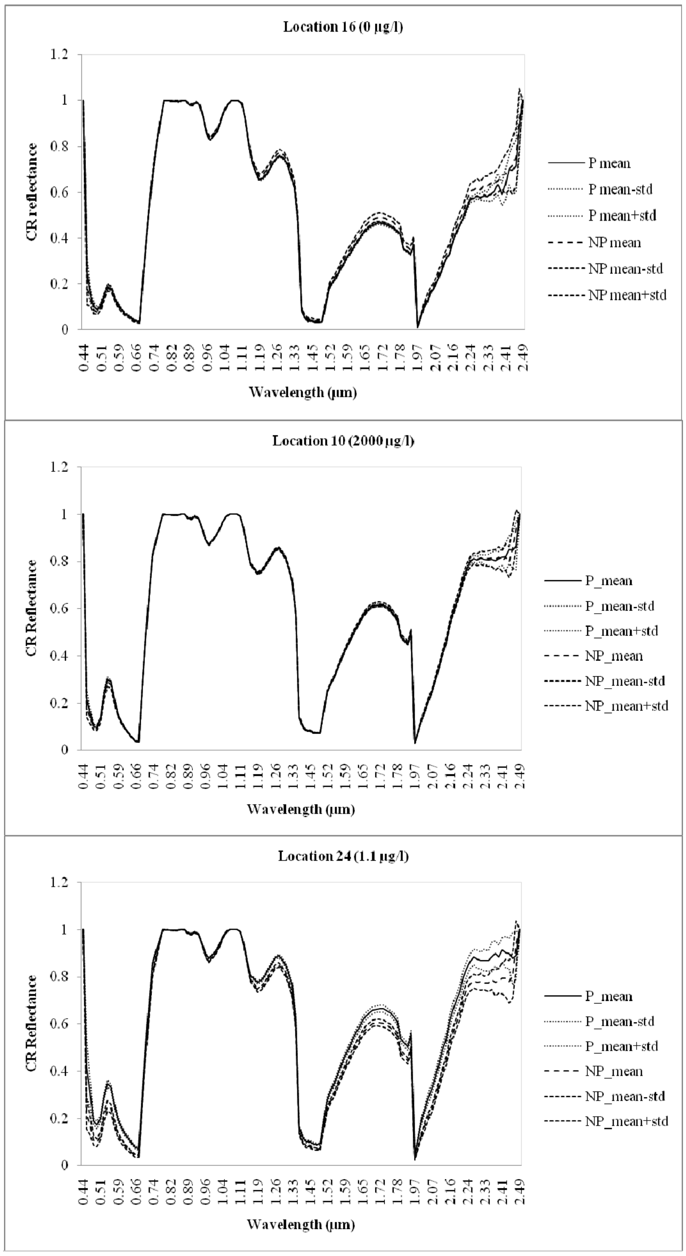


***Figure 5.2** Examples of the differences identified in the continuum removed (CR) spectra between the polluted (P) and non-polluted (NP) sites at the vegetated locations in the 2005 image analysis. No difference was identified at most of the sampled locations except at location 24 with HC concentration of 1.1 µg/l (lowest).*

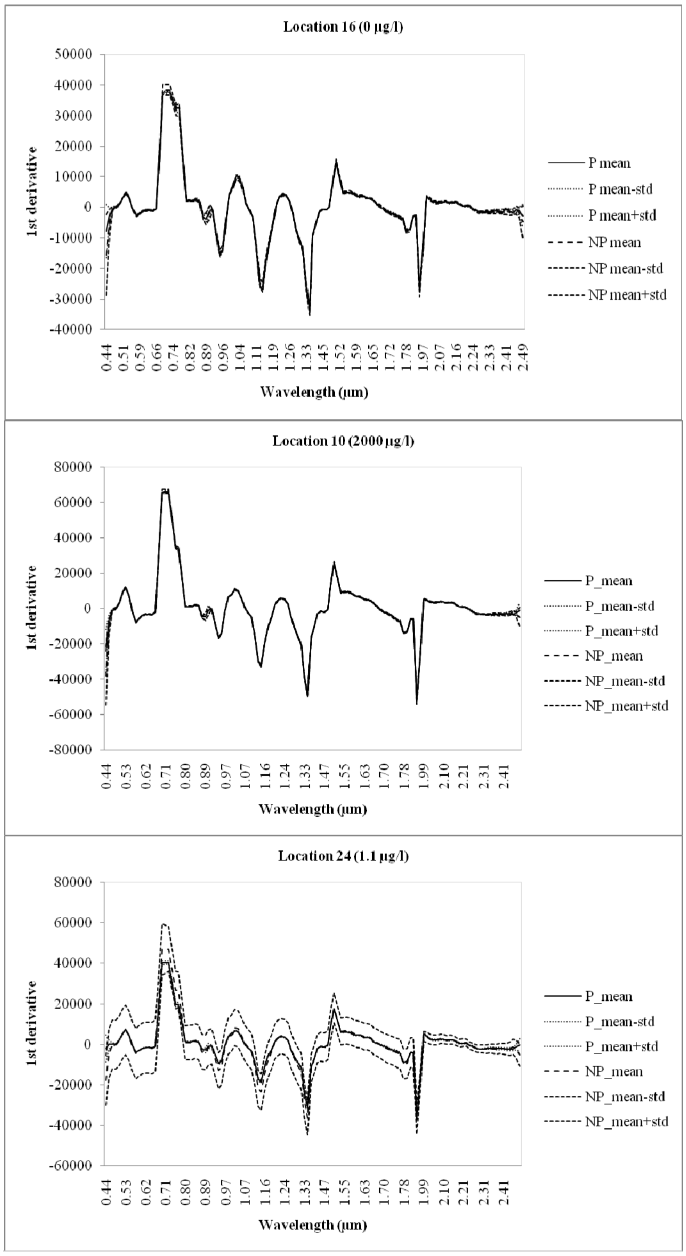


Figure 5.3 *Examples of the differences identified in the first derivative spectra between polluted (P) and non-polluted (NP) sites at the vegetated locations in the 2005 image analysis. No difference was identified at most of the sampled locations except at location 24 with HC concentration of 1.1 µg/l (lowest).*

5.1.2. QAF-analysis

No differences were identified in the absorption feature parameters P, D, W, A, and AS of the main vegetation absorption features between polluted and non-polluted sites at most of the sampled locations (Table 5.1). At location 24 (HC conc. of 1.1μg/l) however, the depth (D) and area (A) at wavelengths 0.44 μm, 0.67 μm, 1.4 μm and 1.9 μm were smaller for polluted than for non-polluted spectra with statistically significant difference ($p \leq 0.05$) (Table 5.1). After normalizing these particular absorption feature parameters of location 24, only the D at wavelength 0.67 μm was shallower for the polluted spectra than for the non-polluted one with statistically significant result ($p \leq 0.05$).

The feature identified at wavelength 2.4 - 2.5 μm (see section 5.2.1) showed no difference in most of the absorption feature parameters, except the depth (D) was shallower for the highly contaminated locations than for the cleaner ones (Figure 5.4).

Table 5.1 The differences in the absorption feature parameters (P, D, W, A, AS) in the main vegetation absorption features between polluted and non-polluted spectra at the sampled vegetated locations on the 2005 image (x=no difference, ↑=absorption feature larger for polluted spectra, ↓=absorption feature smaller for polluted spectra, numbers in the cells indicates acquired p-values). Statistically significant differences were identified at location 24 in the depth (D) and the area (A) of the main absorption features ($p \leq 0.05$).

		2005						
		Location (HC concentration in 2003 μg/l)						
Wavelength (μm)	**Parameter**	**16 (0)**	**11 (0.2)**	**21 (0.89)**	**24 (1.1)**	**8 (320)**	**10 (2000)**	**1 (32000)**
0.44	**P**	x	x	x	x	x	x	x
	D	x	x	x	↓< 0.01	x	x	x
	W	x	x	x	x	x	↑ 0.07	x
	A	↑ 0.11	x	x	↓< 0.01	x	x	x
	AS	x	x	x	x	x	x	x
0.67	**P**	x	x	x	x	x	x	x
	D	↑ 0.10	x	x	↓< 0.01	x	x	x
	W	x	x	x	x	x	x	x
	A	x	x	x	↓< 0.01	x	x	x
	AS	↑ 0.07	x	x	↓0.08	x	x	x
1.4	**P**	x	x	x	x	x	x	x
	D	x	x	x	↓< 0.01	x	x	x
	W	x	x	x	x	x	x	x
	A	x	x	x	↓< 0.01	x	x	x
	AS	x	x	x	x	x	↑ 0.12	x
1.9	**P**	x	x	x	x	x	x	x
	D	x	x	x	↓< 0.01	x	x	x
	W	x	x	x	x	x	x	x
	A	x	x	x	↓< 0.01	x	x	x
	AS	x	x	x	x	↑ 0.08	x	x

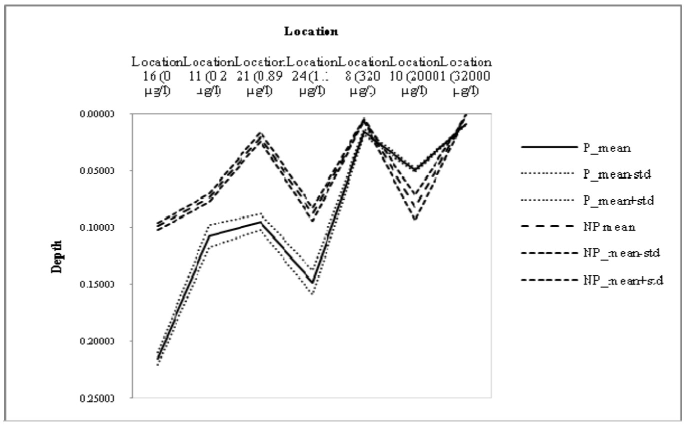


Figure 5.4. *The depth of the absorption feature at wavelength 2.4 – 2.5 µm of the polluted (P) and non-polluted (NP) spectra at all of the sampled vegetated locations along the pipeline in the 2005 image. The feature was shallower at the strongly polluted locations (8, 10 and 1) than at the less polluted areas (16, 11, 21 and 24).*

5.1.3. Vegetation health indices

Table 5.2 summarizes the results from the vegetation health index-analysis for all the sampled vegetated locations from the 2005 image and the results from the statistical analysis are summarized in Appendix 8. The index abbreviations can be found in Appendix 4.

Some of the indices such as BGI1, BGI2 and CTR1 showed no difference between polluted and non-polluted sites in any of the sampled locations (Table 5.2). A few indices, such as CRI2, GM and NDWI, gave contradicting results, while showing that the non-polluted sites around the pipeline were more stressed than the polluted sites in several of the sampled locations (Table 5.2). In contrast, the REP-index gave expected results for the two of the highly contaminated locations, showing vegetation stress at these sites on top of the pipeline (locations 8 and 10, see Table 5.2). The differences between polluted and non-polluted REP-index values were statistically significant at these two contaminated locations ($p \leq 0.05$) (Appendix 8).

At location 1, with the highest HC concentration (32000 µg/l), the REP-index showed no difference between polluted and non-polluted index values (Table 5.3). The three highly polluted locations (i.e. locations 8, 10 and 1) had different NDVI-values, which were associated with the amount of vegetation cover on the sites (see section 3.3.2. for further details). At locations 8 and 10 the NDVI-values were above 0.6, when at location 1 the index value was less than 0.6 (see Table 3.1). Hence, the REP-index could be able to differentiate vegetation stress related to HC pollution only in densely vegetated areas where the NDVI-values are higher than 0.6. In terms of testing this hypothesis, the index values of randomly collected non-polluted pixels (n=400) were plotted on the same graph (i.e. REP vs. NDVI). It was found that the REP index values were indeed more separable in NDVI > 0.6 and less distinguishable in NDVI < 0.6, despite of a few random outliers (Figure 5.5).

At location 24 (HC conc. of 1.1 µg/l) most of the tested vegetation indices indicated vegetation stress on top of the pipeline compared to the sites further away (Table 5.2). These results were similar to the findings from the normal-, CR-, 1st and 2nd derivative- and QAF-analysis for that particular location. At location 1, with the highest HC concentration (Table 3.1), several indices such as ARI2, EVI and GM showed vegetation stress on top of the pipeline (Table 5.2).

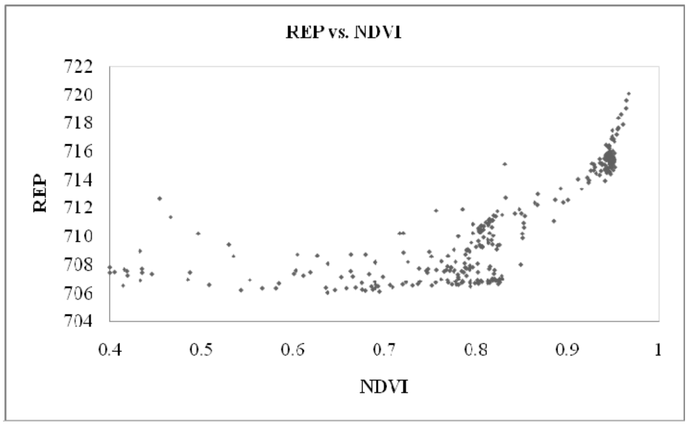
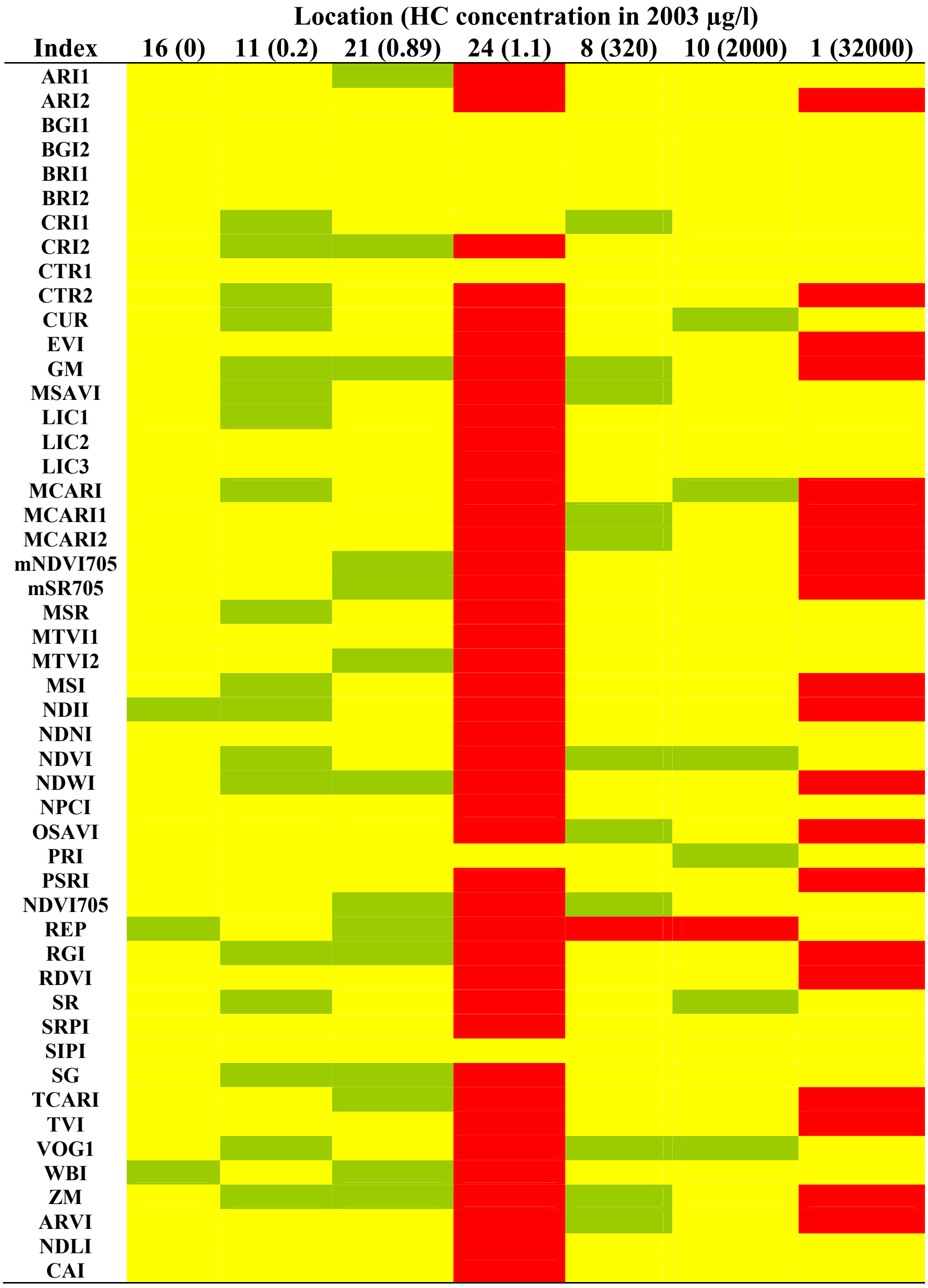


Figure 5.5 *The correlation between NDVI and REP for the randomly collected pixels from non-polluted areas (n=400). The REP-index became stronger above NDVI 0.6 indicating that the REP-index provides reliable results only in densely vegetated areas with NDVI-value > 0.6 and undistinguishable index values in NDVI < 0.6.*

Table 5.2 The performance of calculated vegetation indices in all of the sampled vegetated locations from the 2005 image. (Yellow=no difference, Green=unexpected difference (i.e. vegetation stress at non-polluted sites), Red=expected difference (i.e. vegetation stress at polluted sites)).



	Location (HC concentration in 2003 µg/l)						
Index	**16 (0)**	**11 (0.2)**	**21 (0.89)**	**24 (1.1)**	**8 (320)**	**10 (2000)**	**1 (32000)**
ARI1	🟨	🟨	🟩	🟥	🟨	🟨	🟨
ARI2	🟨	🟨	🟨	🟥	🟨	🟨	🟥
BGI1	🟨	🟨	🟨	🟨	🟨	🟨	🟨
BGI2	🟨	🟨	🟨	🟨	🟨	🟨	🟨
BRI1	🟨	🟨	🟨	🟨	🟨	🟨	🟨
BRI2	🟨	🟨	🟨	🟨	🟨	🟨	🟨
CRI1	🟨	🟩	🟨	🟨	🟩	🟨	🟨
CRI2	🟨	🟩	🟩	🟥	🟨	🟨	🟨
CTR1	🟨	🟨	🟨	🟨	🟨	🟨	🟨
CTR2	🟨	🟩	🟨	🟥	🟨	🟨	🟥
CUR	🟨	🟩	🟨	🟥	🟨	🟩	🟨
EVI	🟨	🟨	🟨	🟥	🟨	🟨	🟥
GM	🟨	🟩	🟩	🟥	🟩	🟨	🟥
MSAVI	🟨	🟩	🟨	🟥	🟩	🟨	🟨
LIC1	🟨	🟩	🟨	🟥	🟨	🟨	🟨
LIC2	🟨	🟨	🟨	🟥	🟨	🟨	🟨
LIC3	🟨	🟨	🟨	🟥	🟨	🟨	🟨
MCARI	🟨	🟩	🟨	🟥	🟨	🟩	🟥
MCARI1	🟨	🟨	🟨	🟥	🟩	🟨	🟥
MCARI2	🟨	🟨	🟨	🟥	🟩	🟨	🟥
mNDVI705	🟨	🟨	🟩	🟥	🟨	🟨	🟥
mSR705	🟨	🟨	🟩	🟥	🟨	🟨	🟥
MSR	🟨	🟩	🟨	🟥	🟨	🟨	🟨
MTVI1	🟨	🟨	🟨	🟥	🟨	🟨	🟨
MTVI2	🟨	🟨	🟩	🟥	🟨	🟨	🟨
MSI	🟨	🟩	🟨	🟥	🟨	🟨	🟥
NDII	🟩	🟩	🟨	🟥	🟨	🟨	🟥
NDNI	🟨	🟨	🟨	🟥	🟨	🟨	🟨
NDVI	🟨	🟩	🟨	🟥	🟩	🟩	🟨
NDWI	🟨	🟩	🟩	🟥	🟨	🟨	🟥
NPCI	🟨	🟨	🟨	🟥	🟨	🟨	🟨
OSAVI	🟨	🟨	🟨	🟥	🟩	🟨	🟥
PRI	🟨	🟨	🟨	🟨	🟨	🟩	🟨
PSRI	🟨	🟨	🟨	🟥	🟨	🟨	🟥
NDVI705	🟨	🟨	🟩	🟥	🟩	🟨	🟨
REP	🟩	🟨	🟩	🟥	🟥	🟥	🟨
RGI	🟨	🟩	🟩	🟥	🟨	🟨	🟥
RDVI	🟨	🟨	🟨	🟥	🟨	🟨	🟥
SR	🟨	🟩	🟨	🟥	🟨	🟩	🟨
SRPI	🟨	🟨	🟨	🟥	🟨	🟨	🟨
SIPI	🟨	🟨	🟨	🟨	🟨	🟨	🟨
SG	🟨	🟩	🟩	🟥	🟨	🟨	🟨
TCARI	🟨	🟨	🟩	🟥	🟨	🟨	🟥
TVI	🟨	🟨	🟨	🟥	🟨	🟨	🟥
VOG1	🟨	🟩	🟨	🟥	🟩	🟩	🟨
WBI	🟩	🟨	🟩	🟥	🟨	🟨	🟨
ZM	🟨	🟩	🟩	🟥	🟩	🟨	🟥
ARVI	🟨	🟨	🟨	🟥	🟩	🟨	🟥
NDLI	🟨	🟨	🟨	🟥	🟨	🟨	🟨
CAI	🟨	🟨	🟨	🟥	🟨	🟨	🟨

5.2. Discussion

No differences were identified between polluted and non-polluted sites in the normal, CR, 1st and 2nd derivative spectra, even at the highly contaminated locations where it would have been expected (see section 5.2.1). These results indicated that the HC pollution has not had an effect on the vegetation health and vigor over the sampled locations. Hussey and Simecek (1987) stated that the rooting system depths of different types of grasses vary from 0.6 – 1.8 m. All the most highly polluted locations (Locations 1, 8 and 10) were situated in meadows (Appendix 1) that can be assumed to consist of some variety of grass. It could therefore be that the pollution from 2 - 4 meter depth did not extend on the grass roots that would have caused vegetation stress on the sampled locations and further been detectable from the reflectance spectra of the aboveground grasses.

Another reason of not detecting differences between polluted and non-polluted sites at the highly contaminated locations in the CR, 1st and 2nd derivative spectra can be the yet too coarse spectral resolution of the airborne hyperspectral data. This is due to the fact that the spectral enhancement techniques, such as CR have been developed for laboratory spectra (Kokaly and Clark, 1999) with higher spectral resolution (~0.001 nm) than the HyMap data (10 – 15 nm) used in the study. It can therefore be that the spectral differences in the vegetation spectra caused by HC pollution were yet too subtle to be detected from the HyMap data with the aforementioned spectral techniques.

The results from the normal, CR, first and second derivative spectra-, QAF- and existing vegetation health index-analysis, indicated that the vegetation at location 24 was more stressed above the pipeline than further away (see Figures 5.1, 5.3, 5.4 and Table 5.1 and Table 5.2). Location 24 was the only site from the sampled locations that was situated on a potato field in 2005 (Appendix 1). It can therefore be that despite of the HC concentration level of 1.1 μg/l (i.e. less than the intervention value for benzene polluted groundwater, 30 μg/l), the effect of pollution is more severe for potatoes than for grasses. This is supported by Shirdam et al. (2008), who stated that HC pollutants have different effect on the growth of various species. According to Weaver (1926), the rooting system of potatoes may however only be as great as 0.9 – 1.2 meters. Hence, the vegetation stress observed at location 24 can be explained by other factors such as in field management practices, since it was unlikely that the rooting system of potatoes could extend up to the polluted groundwater (2 – 4 m). The results from the normalization of the depth (D) and area (A) in the main vegetation absorption features of location 24 (see section 5.2.2), supported this hypothesis. It was found that the differences between the polluted and non-polluted spectra in most of the main vegetation absorption features can be related to the changes in the soil background reflectance on top of the pipeline. As noted in chapter 4, the soils have been reworked and ploughed during and after the installation of the pipeline (see section 4.3.3). This can therefore have caused variation in the surface spectra on top of the pipeline, causing vegetation stress at location 24 which was unrelated to HC pollution.

In the QAF-analysis, the feature identified at wavelength 2.4 – 2.5 µm indicated shallower absorption feature depth especially at the highly polluted locations (Figure 5.4). This can be explained by water stress due to HC pollution, as the feature at this wavelength has been related to water (Jensen, 2000) and the less there is material present in a sample, the shallower the absorption feature is expected to become. The feature can however be related to noise as well, since SWIR region normally becomes noisier in hyperspectral data (Winkelmann, 2005). In this study, with small amount of drilling data one cannot conclude for certain if the feature was indirectly related to HC pollution or to that of noise. It is hence recommended that with larger drilling datasets the relationship between the absorption feature at wavelength 2.4 – 2.5 µm and HC pollution should be investigated further.

Noomen (2007) found that, HC gasses (especially ethane) caused increase in reflectance in the chlorophyll (at 0.4 µm and 0.6 µm) and water (at 1.4 µm and 1.9 µm) absorption features for wheat and maize, hampering the photosynthetic activity of the plants. In this study, there was no difference in the main vegetation absorption features between polluted and non-polluted spectra at most of the sampled locations (see section 5.2.2). These contrary findings can be related to the fact, that in the current study area the HC pollutant was not gas but liquid (i.e. mainly benzene). Eganhouse et al. (1996) stated that monoaromatic HCs, such as benzene, are the most soluble HCs and dissolve more easily into the groundwater than the HC gasses. The liquid benzene could have therefore migrated down to the groundwater having no effect on the shallow rooted aboveground vegetation. Another factor to further consider is that in this study most of the polluted locations were situated on meadows, whereas Noomen (2007) investigated the HC effect on wheat and maize. As mentioned above, HCs have different effect on various species (Shirdam et al., 2008). The contrary findings of the current study and Noomen (2007) can therefore be related to the differences in the sampled vegetation species.

In the vegetation index analysis, it was however found, that the REP-index indicated vegetation stress at the highly polluted locations that can be explained by HC pollution causing stress on the aboveground vegetation (Table 5.2). The ‘Red edge’ at 0.7 µm of the vegetation canopies has been found to be shifted to the shorter wavelengths due to natural (Reid et al., 1988 and Bammed and Birnie, 1994) or simulated gas leaks (e.g. Smith et al., 2004). In contrast, Yang et al. (1999;2000) found a shift of ‘Red edge’ towards longer wavelengths over an microseepage area. The authors hypothesized that their contrary findings to other studies could be due to HCs acting as fertlizers and enhancing the vegetation growth on the polluted sites (Yang et al. 1999;2000). In this study, the ‘Red edge’ was found to move to the shorter wavelengths that was indicated with lower REP-index values at the highly polluted areas. It can therefore be concluded that the HCs acted as stressors instead of fertilizers at the sampled polluted sites. In addition, this vegetation stress due to HC pollution was detectable from the HyMap data only with the REP-index and not with other spectral techniques such as CR, 1st and 2nd derivative analysis, despite of the sampled vegetation species and the length of their rooting system as was discussed above (see page 47).

In the current study, most of the tested vegetation indices (n=49) except the REP-index, gave no difference between polluted and non-polluted spectra at the sampled locations (see Table 5.2). This can be caused by the fact, that the other indices are calculated using only two or sometimes three narrow bands, whereas the REP-index, is computed with four narrow bands in the VIS (Red) and NIR region (see Appendix 4). Due to this, the REP-index has been chosen for example by van der Werff et al. (2008) to detect HC pollution indirectly through vegetation stress, since the REP-index allows the differentiation of subtle changes in the vegetation spectra to be more distinct compared with other indices. Based on this and the findings of the current study, the REP-index can therefore be recommended to be used in the detection of HC related vegetation stress. This is supported by the finding of van der Meijde et al. (2009), who indicated that the REP-index was showing the highest differentiation between polluted and clean sites, compared to the other vegetation indices they tested (e.g. CTR1 and CTR2, Carter, 1994). It should however be considered, that the REP-index is an indicator of decrease in chlorophyll content of plants (Guyot and Baret, 1988), which can be caused by other stressors such as water and nutrient deficiency as well. Noomen (2007), Smith et al. (2004), van der Werff (2006) and White (2007) indicated that contextual information such as pipeline maps should therefore be used. The findings of this study strongly support this, since it minimizes the false interpretations when detecting vegetation stress related to only HC pollution.

In the vegetation index analysis, it was further observed that the REP-index gave the strongest differentiation between polluted and non-polluted locations from the fields that had an NDVI > 0.6 (see Figure 5.5). Pu et al. (2003) and Guyot et al. (1992) noted that ‘Red edge’ parameters are relatively insensitive to the changes of soil cover percentage. According to Noomen (2007), NDVI > 0.6 corresponds to vegetation cover of approximately 40 – 80 %. It can therefore be that the REP-index is insensitive to vegetation stress when the vegetation cover is less than 40 % (i.e. NDVI < 0.6). In this case, the spectral techniques used for densely vegetated areas do not work as expected over partially vegetated sites when differentiating vegetation stress related to HCs. In chapter 4 it was further noticed that the remaining vegetation within the non-vegetated pixels hampers the detection of HC pollution (see section 4.2 for discussion). This resulted in a conclusion that the diagnostic HC features cannot be detected from partially vegetated sites with the spectral methods such as existing HC indices, since the vegetation cover over these sites is even higher than at the non-vegetated areas.

As the spectral techniques used either for non-vegetated or densely vegetated areas did not suit for partially vegetated sites, other methods should be found for such fields. In this study, only one known drilling location was situated in a field with NDVI < 0.6, at which remediation actions had been taken before the 2005 image acquisition (personal communication with Dr. Mark van der Meijde). Even though several indices were showing vegetation stress on that particular location (Table 5.2), one could not conclude for certain if the identified vegetation stress was related to HC pollution or man-induced clean up actions causing disturbances in the land cover. Since many fields (n ~ 15) along the pipeline were situated on fields with

NDVI < 0.6, it would however be crucial to find a technique that works in partially vegetation sites as well, in order not to miss any contaminated areas. In future studies it is therefore recommended that the relationship, between indices taking soil background reflectance into account (e.g. OSAVI, Rondeaux et al., 1996) and the vegetation stress caused by HC pollution on partially vegetated sites, should be investigated in detail.

6. HC pollution detection over time

HCs deplete over time due to natural attenuation (Osuji et al., 2006) that can be monitored with traditional investigation methods such as soil sampling and geochemical measurements. These methods, however, have been considered time-consuming and costly in nature (van der Meijde et al., 2009). The use of multi-temporal hyperspectral imagery in several applications has recently accelerated, due to increasing availability of two or more hyperspectral images from the same study area (Li et al., 2005). In man-induced HC contamination studies this has not yet been studied to large extent (see section 2.2.4).

In terms of investigating if two HyMap images can be used to detect the depletion of HC pollution over time, the same spectral techniques used in the analysis of the 2005 image were applied on the spectra sampled from the 2008 image in known locations, where HC concentration level had decreased drastically between years 2003 and 2008 (Table 3.1). An automatic change detection method was further proposed to detect the change over HC polluted areas over the whole study area (for further details, see section 3.3.8). In order to apply this methodology, an assumption was made that the vegetation growing in the relatively polluted soils above the pipeline, would stay more stressed over time than the vegetation growing in the clean sites, unless depletion of HCs had occurred. This was assumed to be applicable despite of the changing environmental conditions such as species type and distribution, soil quality and water supply in years 2005 (June) and 2008 (July). Consequently, this allowed the comparison of the two images taken in different time steps, since the interest of the current study were the relative rather than the absolute difference of reflectance values at the possibly polluted fields along the pipeline.

In this chapter, the acquired results from the spectral analysis of the 2008 image and the change detection analysis are depicted and discussed, in relation to the findings from the analysis of the 2005 image. The aim was to answer the research questions under objective 3 (see section 1.3).

6.1. Results

6.1.1. Visual interpretation

At seven locations out of ten (i.e. locations 8, 10, 11, 16, 28, 29 and 30), no distinct differences were found in the normal, CR and 1st and 2nd derivative spectra between polluted and non-polluted sites (see examples in Figure 6.1 and Appendix 7). These results were expected, since vegetation stress related to HC pollution was not found in the visual interpretation from the analysis of the 2005 image in any of these areas and HC concentration had decreased between the two years (for further details see Table 3.1 and section 5.1.1).

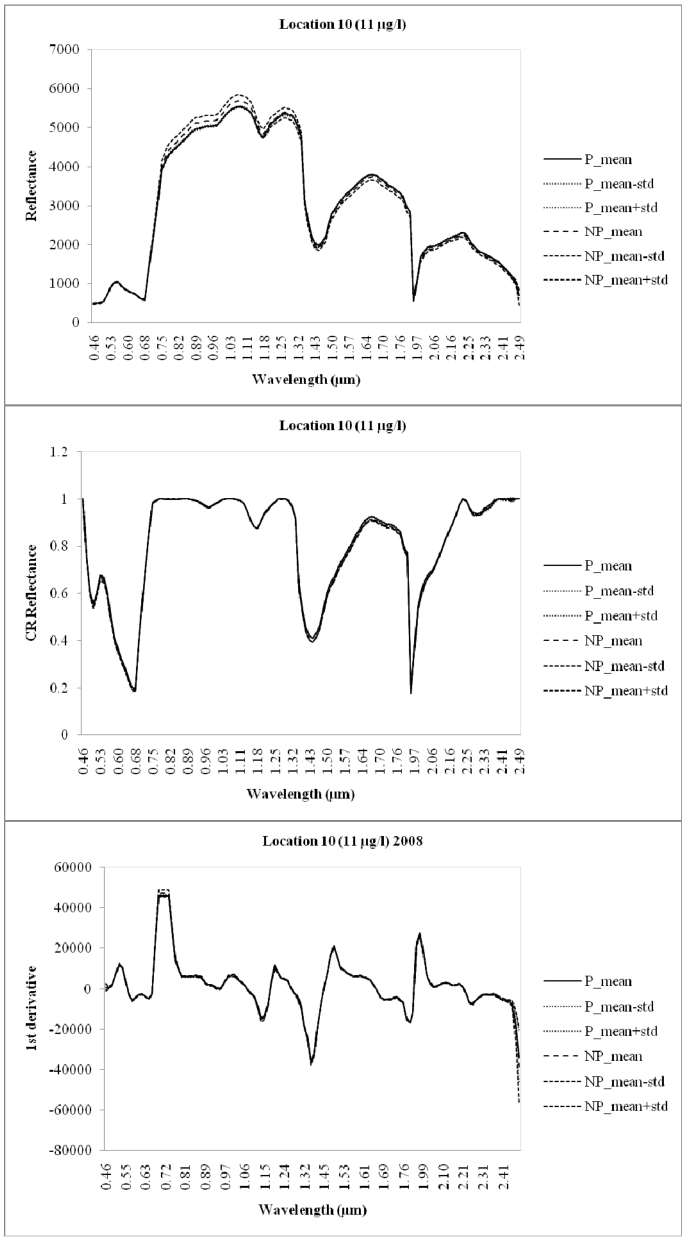


Figure 6.1 *Examples of the differences identified in the normal (top), the CR (middle) and the 1st derivative (bottom) spectra between the polluted (P) and non-polluted (N) sites at location 10 from the 2008 image analysis. No difference was identified between the P and NP spectra at most of the sampled locations where the HC concentration was low in both of the years or decreased between 2003 and 2008.*

At three locations out of ten (i.e. locations 1, 21 and 24), indicators of vegetation stress were however found from the visual interpretation of the collected spectra. The differences between polluted and non-polluted pixels were most visible in the main vegetation absorption features in the CR spectra (see example in Figure 6.2). This result was not as expected because HC pollution level was < 0.2 µg/l at all of these three locations, in year 2008 (Table 3.1).

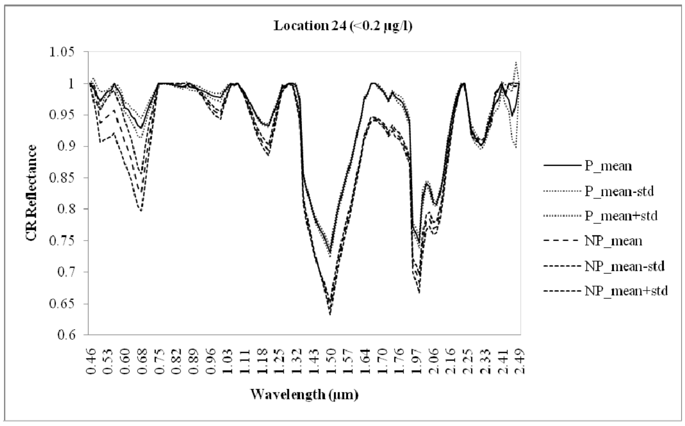


Figure 6.2. *Example of the differences identified in the continuum removed (CR) spectra between the polluted (P) and non-polluted (N) sites at location 24 from the 2008 image analysis. Differences were identified between the P and NP spectra at three of the sampled locations (Location 1, 21 and 24), where the HC concentration was low in both of the years (see Table 3.1. for the actual pollution level).*

6.1.2. QAF-analysis

At five locations out of ten (i.e. locations 11, 21, 24, 10 and 1, Table 3.1), the depth (D) and the area (A) of the absorption feature were significantly smaller between the polluted than the non-polluted spectra ($p \leq 0.05$) in at least one of the calculated wavelengths in the QAF-analysis of the 2008 image (Table 6.1). The normalization of these absorption feature parameters did not change the acquired results.

At the other five sampled locations (i.e. locations 8, 16, 28, 29 and 30, Table 3.1), no difference was identified in the main vegetation absorption features between the polluted and the non-polluted spectra from the analysis of the 2008 image (see for examples Table 6.2).

Table 6.1 The differences identified in the absorption feature parameters (P, D, W, A, AS) from the analysis of the 2008 image at several of the sampled locations (x=no difference, ↑=absorption feature larger for polluted spectra, ↓=absorption feature smaller for polluted spectra, numbers in the cells indicates acquired p-values). At most of the locations (e.g. at locations 11, 21, 24, 10 and 1), the depth (D) and the area (A) were significantly different between the collected spectra ($p \leq 0.05$) in at least one of the calculated wavelengths.

		2008						
		Location (HC concentration in 2008 µg/l)						
Wavelength (µm)	**Parameter**	**16 (0)**	**11 (<0.2)**	**21 (<0.2)**	**24 (<0.2)**	**8 (<0.2)**	**10 (11)**	**1 (<0.2)**
0.44	**P**	x	x	x	x	x	x	x
	D	x	x	↓< 0.01	x	x	x	x
	W	x	x	x	x	x	x	x
	A	x	x	↓< 0.01	x	x	x	x
	AS	x	x	x	x	x	x	x
0.67	**P**	x	x	x	x	x	x	x
	D	x	↓< 0.01	↓< 0.01	↓< 0.01	x	x	↓< 0.01
	W	x	x	x	x	x	x	x
	A	x	↓< 0.01	↓< 0.01	↓< 0.01	x	x	↓< 0.01
	AS	x	↓< 0.01	x	x	x	x	↓< 0.01
1.4	**P**	x	x	x	x	x	x	x
	D	x	↓< 0.01	↓< 0.01	↓< 0.01	x	↓< 0.01	↓< 0.01
	W	x	↓< 0.01	x	x	x	x	x
	A	x	↓< 0.01	↓< 0.01	↓< 0.01	x	x	↓< 0.01
	AS	x	x	x	x	x	x	x
1.9	**P**	x	x	x	x	x	x	x
	D	x	↓< 0.01	↓< 0.01	↓< 0.01	x	↓< 0.01	↓< 0.01
	W	x	x	x	x	x	x	x
	A	x	↓< 0.01	↓< 0.01	↓< 0.01	x	↓< 0.01	↓< 0.01
	AS	x	x	x	x	x	x	x

6.1.3. Vegetation health indices

In the vegetation index analysis, similar results were acquired for the sampled locations as was found from the QAF-analysis (see section 6.1.3).

At locations 8, 16, 28, 29 and 30 (Table 3.1), no difference was found in almost all of the vegetation health indices (see Table 6.2 and Appendix 9 for examples).

At locations 1, 10, 11, 21 and 24, several vegetation indices however showed vegetation stress on top of the pipeline (Table 6.2 and also Appendix 9), even though the HC concentration had decreased on these sites between years 2003 to 2008 (Table 3.1). The REP-index was not one of these indices.

Table 6.2 The performance of the calculated vegetation indices from the 2008 image analysis at several of the sampled locations. (Yellow=no difference, Green=unexpected difference (i.e. vegetation stress at non-polluted sites), Red=expected difference (i.e. vegetation stress at polluted sites)).

Index	**Location (HC concentration in 2008 µg/l)**						
	16 (0)	**11 (<0.2)**	**21 (<0.2)**	**24 (<0.2)**	**8 (<0.2)**	**10 (11)**	**1 (<0.2)**
ARI1	🟨	🟥	🟥	🟥	🟨	🟥	🟨
BGI2	🟨	🟨	🟥	🟨	🟨	🟨	🟨
BRI1	🟨	🟨	🟥	🟨	🟨	🟨	🟨
CRI2	🟩	🟨	🟥	🟥	🟨	🟨	🟨
CTR2	🟨	🟥	🟥	🟥	🟨	🟥	🟥
CUR	🟨	🟥	🟥	🟥	🟨	🟥	🟥
EVI	🟨	🟨	🟥	🟥	🟨	🟥	🟥
GM	🟨	🟥	🟥	🟥	🟨	🟥	🟥
MSAVI	🟨	🟨	🟥	🟥	🟨	🟥	🟥
LIC1	🟨	🟨	🟥	🟥	🟨	🟥	🟥
MCARI	🟨	🟨	🟥	🟥	🟨	🟥	🟥
MCARI1	🟨	🟨	🟥	🟥	🟨	🟥	🟥
MCARI2	🟨	🟨	🟥	🟥	🟨	🟥	🟥
mNDVI705	🟨	🟨	🟥	🟥	🟨	🟨	🟥
mSR705	🟨	🟨	🟥	🟥	🟨	🟨	🟥
MSR	🟨	🟥	🟥	🟥	🟨	🟥	🟥
MTVI1	🟨	🟨	🟥	🟥	🟨	🟥	🟥
MTVI2	🟨	🟨	🟥	🟥	🟨	🟨	🟥
MSI	🟨	🟥	🟥	🟥	🟨	🟥	🟥
NDII	🟨	🟥	🟥	🟥	🟨	🟥	🟥
NDNI	🟨	🟥	🟥	🟨	🟨	🟥	🟥
NDVI	🟨	🟥	🟥	🟥	🟨	🟥	🟥
NDWI	🟨	🟥	🟥	🟥	🟨	🟥	🟥
NPCI	🟨	🟨	🟥	🟨	🟨	🟨	🟨
OSAVI	🟨	🟥	🟥	🟥	🟨	🟥	🟨
PRI	🟨	🟨	🟥	🟥	🟨	🟨	🟨
PSRI	🟨	🟥	🟥	🟥	🟨	🟥	🟥
NDVI705	🟨	🟨	🟥	🟥	🟨	🟥	🟥
REP	🟨	🟨	🟨	🟩	🟨	🟨	🟨
RGI	🟨	🟨	🟥	🟥	🟨	🟥	🟨
RDVI	🟨	🟨	🟥	🟥	🟨	🟨	🟥
SR	🟨	🟥	🟥	🟥	🟨	🟥	🟥
SRPI	🟨	🟨	🟩	🟨	🟨	🟨	🟨
SIPI	🟨	🟨	🟥	🟩	🟨	🟨	🟨
SG	🟩	🟨	🟥	🟨	🟨	🟨	🟨
TCARI	🟨	🟨	🟥	🟥	🟨	🟨	🟥
TVI	🟨	🟨	🟥	🟥	🟨	🟨	🟥
VOG1	🟨	🟨	🟥	🟥	🟨	🟥	🟥
WBI	🟨	🟨	🟥	🟥	🟨	🟨	🟥
ZM	🟨	🟨	🟥	🟥	🟨	🟨	🟥
ARVI	🟨	🟨	🟥	🟥	🟨	🟨	🟥
NDLI	🟨	🟨	🟥	🟨	🟨	🟨	🟥
CAI	🟩	🟨	🟥	🟥	🟨	🟥	🟨

6.1.4. HC detection with automated image processing technique

The automated HC detection over the vegetated areas was performed with the REP-index, since it showed the expected result for the highly polluted locations in the 2005 image analysis (see section 5.1.3). The non-vegetated areas were masked out with NDVI-threshold < 0.6, since in chapter 5 it was found that this index did not perform as expected when NDVI was lower than 0.6 (see section 5.1.3). The reference REP-index value set for the growing region was 1 nm that was used to calculate the average REP-index values for the fields (for further details, see section 3.3.8).

In the change detection analysis, at two locations detected by van der Werff et al. (2008) as 'most likely' or 'likely' HC polluted sites in the analysis of the 2005 image, the vegetation had become healthier in the 2008 image (Figures 6.3 and 6.4, respectively). Depending on what threshold-value was used (e.g. < -0.4 or > 0.4 or < -0.6 or > 0.6), the size of the identified anomaly varied, especially at the stressed sites in the 2005 image (see Figures 6.3 and 6.4). At three other locations, detected as 'likely' HC polluted areas from the 2005 image by van der Werff et al. (2008), the vegetation was stressed in both of the years (see Appendix 10 for the images). The area of stressed vegetation had, however, become smaller in the 2008 image. At the remaining 23 locations detected by van der Werff et al. (2008), no change (i.e. more stressed or healthier vegetation) was identified in the change detection analysis.

In the accuracy assessment of the 2005 normalized REP-image with 22 additional drilling locations, the overall accuracy and the Kappa coefficient were 63.6 % and 0.27, respectively (see Appendix 11 and definition for the accuracy parameters in Table 3.4). The user accuracy for the non-polluted class 0 was 72.7 % and 54.5 % for the polluted class 1. The producer accuracies were 61.5 % for class 0 and 66.6 % for class 1.

The results from the accuracy assessments of the change detection map with different threshold-values are summarized in Table 6.3 (see also Appendix 12).

Table 6.3 Summary of the accuracy assessment for the change detection map using different threshold-values to define the change in the normalized images (0=change and 1=no change). The definition for the accuracy parameters can be found in Table 3.4.

Class	**Accuracy Measure**	**Index threshold-value for the normalized images**					
		< -0.1 or > 0.1	**< -0.2 or > 0.2**	**< -0.3 or > 0.3**	**< -0.4 or > 0.4**	**< -0.5 or > 0.5**	**< -0.6 or > 0.6**
	Overall (%)	47.0	47.0	52.9	58.8	52.9	52.9
	Kappa	-0.03	-0.03	0.07	0.13	-0.06	-0.06
0	**User (%)**	57.1	57.1	62.5	63.6	57.1	57.1
	Producer (%)	40.0	40.0	50.0	50.0	80.0	80.0
1	**User (%)**	40.0	40.0	44.4	70.0	33.3	33.3
	Producer (%)	57.1	57.1	57.1	42.9	14.3	14.3

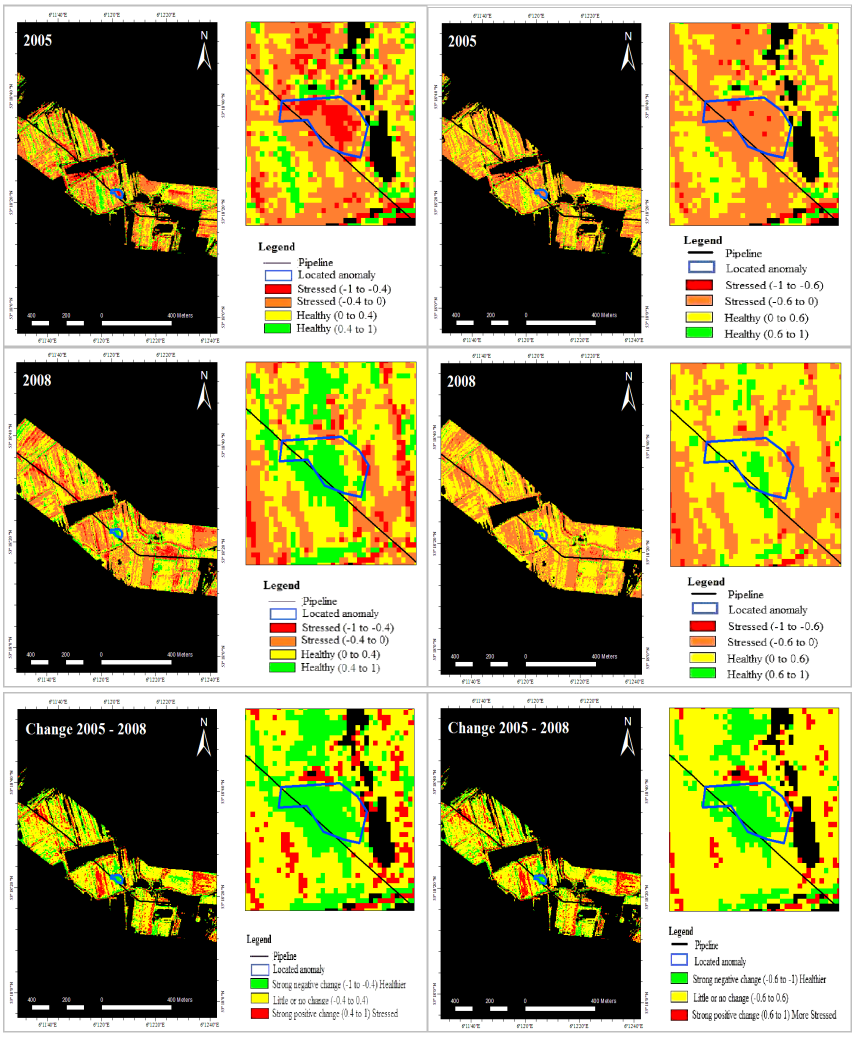


Figure 6.3 *The change occurred between years 2005 and 2008 at one 'most likely'-classified location detected by van der Werff et al. (2008) from the study area with the 2005 image previously. The vegetation has become healthier at the site as can be clearly seen between the 2005 (upper) and 2008 (middle) normalized REP-images and further in the derived change detection map (lowest). When using threshold-value 0.4 (left) the detected stressed locations especially in the 2005 image were larger than when using threshold-value 0.6 (right).*

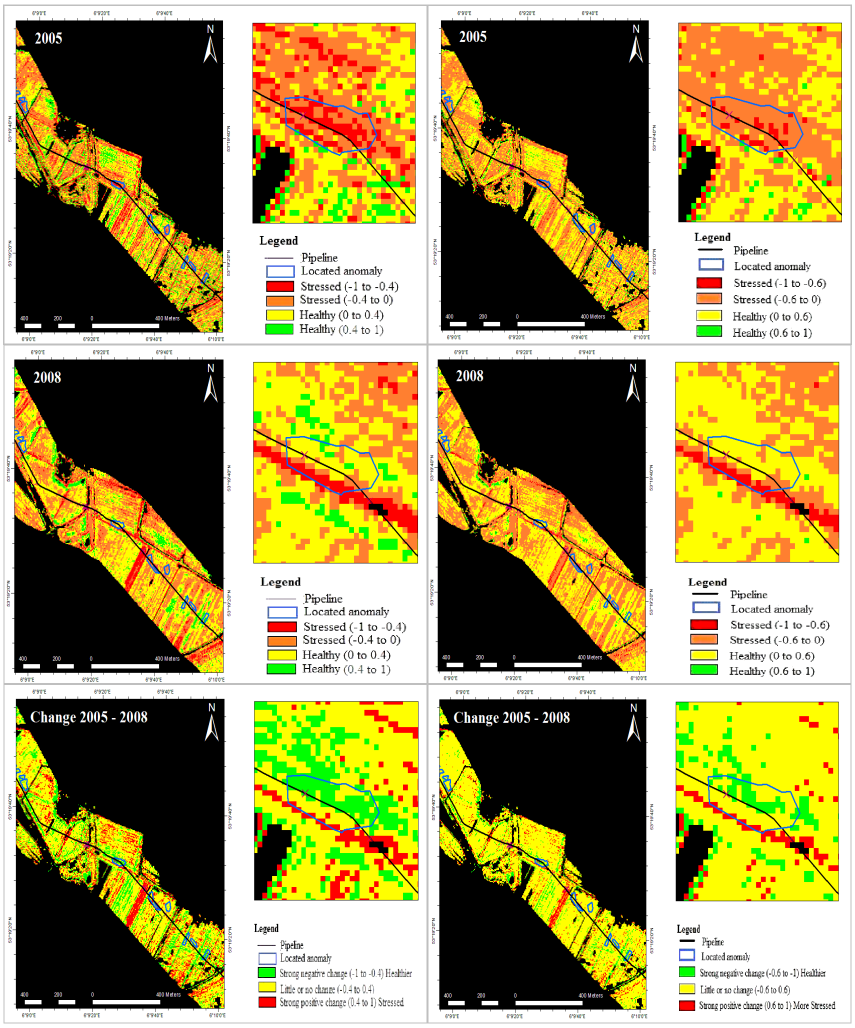


Figure 6.4 *The change occurred between years 2005 and 2008 at one 'likely'-classified location detected by van der Werff et al. (2008) from the study area with the 2005 image previously. The vegetation has become healthier at the site as can be clearly seen between the 2005 (upper) and 2008 (middle) normalized REP-images and further in the derived change detection map (lowest). When using threshold-value 0.4 (left) especially the detected stressed locations in the 2005 image were larger than when using threshold-value 0.6 (right).*

6.2. Discussion

6.2.1. Spectral analysis of the 2008 image

This section discusses the results acquired from the analysis of the 2008 image for the sampled drilling locations (Table 3.1), at which the same spectral techniques that were used in the 2005 image analysis, were applied. The results discussed here can be found in sections 6.1.1 - 6.1.3.

At locations 11 and 10, indicators of vegetation stress were identified in the QAF-analysis and in the vegetation index calculations from the 2008 image analysis (Tables 6.1 and 6.2, respectively). This was not expected since HC pollution was either low or decreased at both of these sites (Table 3.1). At location 11, the identified differences could be explained by other factors such as in-field management practices or different stressors (e.g. nutrient deficiency), since the pollution level was <0.2 μg/l in both of the years based on the geochemical data (Table 3.1). At location 10, differences between polluted and non-polluted spectra were identified especially in the main water absorption features in the SWIR region (Table 6.1). The vegetation can therefore have exhibited water stress due to long-term exposure of HC pollution as the HC concentration was yet elevated in year 2008 (10 μg/l). On the other hand, clear traces of tractor were found at the site in the detailed investigation of the drilling data in the change detection analysis (see Appendix 13). It was therefore not clear if the vegetation at location 10 was stressed in the 2008 image analysis due to HC pollution or because of the in-field management practices occurred at the site.

Based on the information given by the environmental consultancy DHV (2008), remediation actions had been taken at locations 1, 21 and 24 before year 2008. From the spectral analysis, indicators of vegetation stress were identified between polluted and non-polluted spectra at all of these sites (see sections 6.1.1 – 6.1.3). Since these locations were found clean in year 2008 based on the geochemical data (Table 3.1), the differences identified in the spectral analysis could be explained by the man-induced mixing of the ground due to the clean-up actions. The NDVI-values at these three remediated locations had decreased in three years from 0.5 – 0.9 to less than 0.2 (Table 3.1). This indicated that the fields above the pipeline were less covered with vegetation in the 2008 image. This strengthened the hypothesis that cleanup had taken place as one of the remediation methods used in the area was soil oxidation (personal communication with Dr. Mark van der Meijde) causing disturbances (i.e. stamping of vegetation down) on the land cover above the pipeline. It is therefore concluded that the detection of natural depletion of HC pollution from hyperspectral data is affected by the previous cleanup actions.

In the vegetation index analysis, the REP-index in contrast to almost all the other vegetation indices, did not indicate vegetation stress on top of the pipeline at locations 21 and 24 (Table 6.2). This can be related to the performance of the REP-index under NDVI < 0.6 (see section 5.1.3), since at location 24 the NDVI-value was less than 0.15 and at location 21 the NDVI-value was 0.2 (Table 3.1). This result strengthened the conclusion that indeed the REP-index does not perform properly under NDVI 0.6 (for further details, see discussion in 5.2). In fact, based on the analysis of the 2008 image, the results further referred that in between NDVI 0.2 – 0.6 the REP-index was not able to find any difference between vegetation stress and below NDVI of 0.15 it even gave the opposite results compared to other vegetation indices (see result for location 24 in Table 6.2). It can therefore be concluded that the REP-index should not be used in detection of vegetation stress related to HC pollution in areas with less vegetation cover as it can provide unreliable results in such areas.

6.2.2. Change detection analysis

This section provides a general discussion about the use of NDVI for quantifying the change in land cover. In addition, it discusses the results acquired from the change detection analysis produced with the automated image processing technique using both the 2005 and 2008 image as described in section 3.3.8. These results can be found in section 6.1.3.

In the current study, the NDVI-values were used as indicators of changes in land cover between years 2005 and 2008. It was chosen since it is one of the most commonly used vegetation indices and it can be used to delineate vegetated and non-vegetated areas from each other (i.e. the NDVI is an index indicating the greenness of land cover (Rouse et al., 1974). When using the NDVI to such purposes, several factors should be taken into account. For example the NDVI can be sensitive to soil background reflectance and it saturates in areas with high Leaf Area Index (Jensen, 2000). The change in land cover (or actually land use) could therefore be performed with other methods such as land use classifications. In this study, not enough data were available for proper validation for both of the years in case land use classifications would have been performed on the images (i.e. the land use map was available only for 2005), which would have allowed proper evaluation of the change in crop types between the two years. Since it is known that HCs have different effect on various species (Shirdam et al., 2008), the relationship between HCs and different crops could be investigated with land use classifications in future studies. In such cases, the hyperspectral airborne data would not just serve as information source of possible HC pollution, but also other type of relevant information such as the crop types could be acquired in the area that can be crucial in the current application.

In the detection of HC pollution with automated image processing technique, two locations were found in which the vegetation seemed to become clearly healthier between the 2005 and 2008 images (Figures 6.3 and 6.4). This can be explained by the natural attenuation of HCs at the sites (Osuji et al., 2006): when HC concentration decreases, the stress effect on plants decreases as the amount of a stressor is less (Shirdam et al., 2008). At few other locations, the vegetation was found to be stressed in both of the images but the stressed areas had become smaller in the 2008 scene (Appendix 10). This indicated that HC pollution existed at the locations but it had decreased to some extent, allowing vegetation to grow stronger. For validation of these sites, no drilling data for year 2008was currently available. Hence, one cannot conclude for certain if the detected changes were related to the changes in the vegetation health status due to HC pollution or because of some other factors (e.g. increase of water or nutrients at the sites allowing the vegetation to grow better in year 2008). The changes at these possibly polluted sites should therefore be further confirmed, in order to verify if hyperspectral imagery was able to correctly detect the change in the spatial extent of these possibly HC polluted areas.

In the accuracy assessment of the change detection map it was found that when using threshold-value < -0.4 or > 0.4, highest accuracy was acquired for most of the accuracy measures, if compared to the accuracies acquired with other threshold-values (Table 6.3). For example, the user accuracy indicating that a randomly chosen point on a map is actually that class in nature (Congalton, 1991), was higher for both of the classes when using < -0.4 or > 0.4 threshold-value. This result was important because if the user accuracy is low, both the financial and environmental costs of misclassifying the change in locations with strongly polluted material are high. In other words, if the produced map indicates that polluted location is clean even though this is not the case in reality, the risk of missing polluted sites is high. In contrast, if the clean locations are detected as polluted, further investigations are steered on already non-polluted sites and unnecessary expenses will be created. High user accuracy for both of the classes is therefore required.

Interesting findings were further made from the producer accuracies in the validation of the change detection map. They became larger for class 0 (no change) when higher threshold-values were used (i.e. above < -0.5 or > 0.5 and < -0.6 or > 0.6, see Table 6.4). The producer accuracy indicates how large percentage of a particular ground class is correctly classified (Congalton, 1991). In respect to the findings of this study, the higher user accuracies indicated that in most of the cases no changes had occurred when higher threshold-values were used. In other words, clean sites were clean in both of the years and polluted ones stayed polluted. Since it is known that HCs deplete over time naturally (Osuji et al., 2006) and the geochemical record of sample locations further confirmed this, it is unlikely that no changes has occurred over the HC polluted areas in the study area. It can hence be concluded that if larger threshold-values are used, the occurred changes can be missed further having serious environmental and/or economical consequences (i.e. either the polluted locations are not cleaned or clean locations are investigated without any need).

The selection of most suitable threshold-value for the change detection map is further discussed based on Figures 6.5 and 6.6. In these images it was clearly visible that the stressed anomalies in the 2005 image were smaller when using threshold-value < -0.6 or > 0.6 instead of < -0.4 or > 0.4 in both of the detected locations. Van der Werff et al. (2008) stated that the overestimation of polluted locations would be better than underestimation. This is clear since all of the polluted areas should be identified and further cleaned. Therefore, < -0.4 or > 0.4 could be used as threshold-value for the change detection map, since it did not under-predict the size of the possibly polluted locations. This was supported by the accuracy assessment of the change detection image, in which threshold-value 0.4 produced the best overall accuracy (Table 6.4). It should however be noted, that the overestimation of the size of the polluted locations also results in increase of clean up costs. Different threshold-values should therefore be tested when applying the change detection technique as preliminary detection and monitoring tool of possibly polluted areas. The spatial extent of these sites can then be defined in detail with ground truth measurements in order to optimize the clean up actions (i.e. clean only the polluted sites and leave the non-polluted untouched).

The Kappa coefficient, indicating how good the acquired map was compared to a map produced by chance (Congalton, 1991), had values varying from -0.06 to 0.13 in the accuracy assessment of the change detection image with different threshold-values (Table 6.3). In the accuracy assessment of the normalized 2005 REP-image, the Kappa coefficient was 50 % higher (0.27) than the best Kappa (0.13) acquired for the change detection image (Appendix 11). These results indicated that in the worst case the change detection map was 6 % worse and in the best case 13 % better than a map produced by chance. The normalized 2005 REP-image on the other hand, was 27 % better than a map produced by chance.

The acquired classification accuracy for the change detection image is affected by the nature of the groundwater validation data. Most of the drilling locations used for the validation were either situated on the edges, the whole field was stressed or clear traces of infield-management practices were found (see examples in Appendix 13). This gave mixed interpretation results unrelated to HC pollution. In addition, the amount of data was small and it was not evenly distributed along the pipeline. The data was collected from 2 – 4 m depth, making it difficult to establish the relationship between the ground truth data and the imagery that records the spectra on the surface. The discrepancy in the data acquisition time of the drilling data and the images may have further affected the results.

In the accuracy assessment of the 2005 REP-image, most of these aforementioned data discrepancies were not that severe. For example, the 22 drilling locations used in the accuracy assessment of the 2005 REP-image, were more equally distributed along the pipeline and some of the data points were collected almost on the surface (0 – 0.3 m) (Appendix 2). Even though the data acquisition time was different between the imagery (year 2005) and the drilling data (years 2002/2003), the reason why the acquired Kappa coefficient was higher for the 2005 REP-image than for the change detection image, is because the other factors (e.g. the distribution of sampling points and the sampling depth) were less discrepant. More drilling data that is corrected from the abovementioned problems is therefore needed in order to further validate the performance of the change detection method. For example, the locations at which possibly HC related changes were identified (Figures 6.3 and 6.4 and Appendix 10) should be investigated with ground truth data. This information would not just be important in terms of verifying the performance of the change detection method but also crucial for the economic side of the remediation actions if less clean up would be needed at these sites.

It should be acknowledged that the traditional methods such as soil sampling and drilling are already seen as time-consuming and costly in nature (van der Meijde et al., 2009). Hence, the use of non-destructive and possibly more cost-effective ways such as field spectrometer and ground penetrating radar (GPR) should be tested together with the images that can decrease the data discrepancy between the imagery and the drilling data. The abovementioned methods have been used in the investigation of HC pollution by many researchers with promising results (e.g. Atekwana et al., 2000 and Atekwana et al., 2004, Noomen 2007, van der Meijde et al., 2009). It is therefore recommended that the joint use of these different hyperspectral and geoelectrical techniques accompanied with testing of them in various study areas with different conditions (e.g. climatic settings, vegetation and soil types) should be conducted in the future studies. In this way, the possibility of using hyperspectral techniques in investigation of HC contaminated sites in different places is further evaluated. This can result in universal or adaptable ways that are cost-effective in operational use of detecting and monitoring HC pollution over time for remediation purposes.

7. Conclusions and recommendations

Several absorption features that can be associated with HC pollution were identified in the polluted spectra over the non-vegetated areas in the 2005 image analysis. These included wavelengths such as 2.1 µm, 2.2 µm, 2.3 µm and 2.4 µm that based on previous studies have been directly or indirectly related to HCs. When applying band ratios built on these particular absorption features or the existing HC indices from the literature, in most of the cases clear distinction between HC polluted and clean sites was not acquired over the non-vegetated areas in the 2005 image. In addition, some of the existing HC indices detected false anomalies such as roofs and roads. With band ratio 2.1/2.2 µm however, possible HC pollution was found at one site close to the pipeline, which should be further verified with ground truth data that was not available for the current study.

Various reasons were suggested why most of the spectral techniques applied over non-vegetated areas were not able to detect pollution. Firstly, the HCs from 2 – 4 m depth did not extend towards the surface in such large quantities, in order to be detected from the hyperspectral data. Secondly, the mixing of soil due to the installation of the pipeline and remaining vegetation within the pixels hampered the detection of unique HC absorption features. Finally, the spatial resolution of hyperspectral imagery (4 x 4 m) was too coarse for the detection of distinct and subtle HC absorption features over the non-vegetated areas due to the effect of 'mixed pixels'. This led into further conclusions that the techniques used over non-vegetated areas, such as existing HC indices and band ratios built on the diagnostic HC absorption features, do not work in the detection of HC pollution over partially vegetated sites, as one of the drawbacks of these techniques was the remaining vegetation signal within the pixels.

Over the vegetated locations sampled from the 2005 image, no difference was identified between polluted and non-polluted spectra with methods such as the normal -, CR -, 1st and 2nd derivative - and QAF - analysis. In the vegetation index analysis however, the REP – index was found to be the only index that was able to differentiate vegetation stress at the highly HC contaminated sites. Based on this, it was concluded that from all the tested spectral techniques and vegetation indices, the REP-index should be favored in the detection of vegetation stress caused by HC pollution. The REP – index should however not be used in areas with NDVI < 0.6, since it was found that the REP-index can provide unreliable results in NDVI-values < 0.6 and even opposite results under NDVI 0.2. The same spectral techniques used successfully in detection of HC pollution over the densely areas, cannot therefore be used in detection of HC pollution at sites with less vegetation cover (i.e. NDVI < 0.6). In addition, the pipeline maps should be used as an aid in the detection of HC related vegetation stress when using the REP-index. This is due to the fact that the REP-index is identifying all the stressed locations in the densely vegetated area that can be caused by other stressors such as in-field management practices as well (i.e. traces of tractors that have stamped the vegetation down). In this case, contextual information such as the pipeline maps should be used, in order to decrease the amount of misclassifications in the final classification products.

In this research, only one partially vegetated location (NDVI < 0.6) was available for the analysis, which was remediated before the acquisition of the image took place (year 2005). Even though many vegetation indices showed vegetation stress at that particular location, it could not be concluded for certain if these indices detected HC pollution or only the remediation actions taken at the site. Similarly, in the analysis for the 2008 image, indicators of vegetation stress were acquired at locations in which remediation actions had taken place. This hampered the result interpretation related to the change of natural attenuation of HCs. It can therefore be concluded that the change in land cover especially due to remediation actions taken at the sites are complicating the identification of HC pollution and its natural depletion from hyperspectral imagery.

The automated change detection analysis indicated that at several locations the vegetation had become clearly healthier between the two years. This indicated that HC pollution had decreased at these sites allowing improved growth of vegetation when the amount of stressor was less. In the accuracy assessment of the change detection map, it was found that threshold-value < -0.4 or > 0.4 gave the best overall accuracy compared to other threshold-values. The user accuracies for classes 1 and 2 (no change and change, respectively) were higher than 60 %. This result is important from economical and environmental point of view, as low user accuracies (i.e. close to 0 %) indicate misclassifications in the changes occurring for example over the strongly polluted areas. This is not desired as some of the yet polluted areas are missed or remediation actions are steered on locations which have actually become clean naturally. The threshold-value < -0.4 or > 0.4 should therefore be used for the change detection map, as it did not underestimate the polluted sites providing the highest accuracy for the important parameters from the application point of view. This should however be done cautiously as the possible overestimation of polluted sites results in increase of cleanup costs. In order to further verify these results, the sites at which vegetation was found to become healthier possibly due to HC pollution, should therefore be validated against geochemical measurements. This would not just provide information on the accuracy of the change detection method and the suitability of different thresholds, but also allow the site owners to find out if the extent of HC pollution level had decreased on these sites.

In terms of finally evaluating the applicability of hyperspectral RS techniques to investigate HC polluted areas over time for remediation purposes, few remarks are given. Hyperspectral imagery can be considered as a non-destructive and cost-effective tool to investigate HC polluted areas without fully replacing the traditional methods. These become a necessity especially in the validation part of the acquired results. It can therefore be concluded that together with other techniques (e.g. soil sampling, field spectrometer and GPR), hyperspectral imagery is a potential tool for the investigation of HC polluted areas. The automated image processing technique can be used as preliminary detection method in different time steps for possibly polluted locations, which should then be investigated in detail with other methods. This can decrease the remediation costs as more detailed inspection is steered only in the possibly polluted locations.

Based on the results acquired from this study and the given conclusions, further studies are recommended to focus on the following issues:

- Larger drilling dataset should be available for the analysis and validation, since at least 30 samples is recommended in each group to make statistically sound analyses (Crawley, 2002)
- The data discrepancy considering the data acquisition time between the imagery and the ground truth data should be minimized
- Field spectrometer measurements and Ground Penetrating Radar (GPR) could be applied in terms of minimizing the gap between the groundwater data collected from 2 – 4 m depth and the imagery acquired on the surface
- Sufficient amount of drilling data should be collected over different land covers, in order to improve the conclusions on the applicability of different spectral techniques in various land covers. This should focus especially on the non-vegetated and partially vegetated sites where indices taking into account the soil background reflectance (e.g. OSAVI) and their relationship to vegetation stress caused by HC pollution could be investigated in more detail.
- Land use classifications for both years could be performed in order to acquire information on the relationship between HC pollution and different crop types
- The identified possibly HC related anomalous regions in the change detection analysis should be investigated with ground truth data in terms of further evaluating the performance of the novel change detection method
- The novel multi-temporal hyperspectral RS technique and other spectral techniques should be tested together in a number of study sites and years in order to acquire universal and/or adaptable methods for areas with different land covers, climatic settings and other properties that could be taken into operational use when investigating HC contaminated sites over time for remediation purposes

8. References

Adams, R. S. and Ellis, R. 1960, 'Some physical and chemical changes in the soil brought about by saturation with natural gas', *Proceedings in Soil Science Society of America,* vol. 24, no., pp. 41 – 44.

Agostini, P., Carlon, C., Critto, A. and Marcomini, A. 2007, 'A step toward contaminated megasite management: six European experiences at comparison', in Velini, A. A. (ed) *Landfill research trends*, Nova Science Publishers Inc, New York, pp. 47 – 73.

Atekwana, E. A., Sauck, W. A. and Werkema, D. D. 2000,'Investigations of geoelectrical signatures at a hydrocarbon contaminated site', *Journal of applied Geophysics*, vol. 44, no. 2 – 3, pp. 167 – 180.

Atekwana, E. A., Werkema, D. D., Duris, J. W., Rossbach, S., Sauck, W. A., Cassidy, D. P., Means, J. and Legall, F. D. 2004,'In-situ apparent conductivity measurements and microbial distribution at a hydrocarbon-contaminated site', *Geophysics*, vol. 69, no. 1, pp. 56 – 63.

Bammel, B. H. and Birnie, R. W. 1994,'Spectral response of big sagebrush to hydrocarbon-induced stress in the Bighorn Basin, Wyoming', *Photogrammetric Engineering and Remote Sensing*, vol. 60, no. 1, pp. 87 – 96

Baumgardner, M. F., LeRoy, S. F., Biehl, L. L. and Stoner, E. R. 1985, 'Reflectance properties of soils', in Brady, N. C. (ed.) *Advances in Agronomy*, Academic Press Inc, London.

Ben-Dor, E., Inbar, Y. and Chen, Y. 1997,'The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400 – 2500 nm) during a controlled decomposition process, *Remote Sensing of the Environment*, vol. 61, no. 1, pp. 1 – 15.

Ben-Dor, E., Chabrillat, S., Demattê, J. A. M., Taylor, G. R., Hill, J., Whiting, M. L. and Sommer, S. 2009, 'Using Imaging Spectroscopy to study soil properties', *Remote Sensing of Environment*, vol. 113, Supplement 1, pp. S38-S55.

Bihong, F., Zheng, G., Nonomiya, Y., Wang, C., and Sun, G. 2007,'Mapping hydrocarbon induced mineralogical alteration in the northen Tian Shan using ASTER multispectral data', *Terra Nova*, 19, no. 4, pp. 225−231.

Borengasser, M., Hungate, W. S. and Watkins, R. 2008, *Hyperspectral remote sensing: principles and applications*, CRC Press, Boca Rato.

Broge, N. H. and Leblanc, E. 2000,'Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density', *Remote Sensing of Environment*, vol. 76, no. 2, pp. 156 – 172.

Broos, M. J., Aarts, L., van Tooren, C. F. and Stein, A. 1999, 'Quantification of the effects of spatially varying environmental contaminants into a cost model for soil remediation', *Journal of Environmental Management,* vol. 56, no. 2, pp. 133-145.

Calabrese, E. J. and Kostecki, P. T. 1993, *Hydrocarbon Contaminated soils, 3rd edn,* Lewish Publishers, Chelsea

Campbell, J. B. 2006, *Introduction to Remote Sensing*, 4th edn, The Guilford Press, New York.

Carlon, C., Pizzol, L., Critto, A. and Marcomini, A. 2008, 'A spatial risk assessment methodology to support the remediation of contaminated land', *Environment International*, vol. 34, no. 3, pp. 397 – 411.

Carter, G. A. 1993,'Responses of leaf spectral reflectance to plant stress', *American Journal of Botany*, vol. 80, no. 3, pp. 239 – 243.

Carter, G. A. 1994,'Ratios of leaf reflectances in narrow wavebands as indicators of plant stress', *International Journal of Remote Sensing*, vol. 15, no. 3, pp. 697 – 703.

Chemistry Glossary: Hydrocarbon definition, [online], 2009, available: http://chemistry.about.com/od/chemistryglossary/a/hydrocarbondef.htm [accessed 2009, August 22].

Chen, J. 1996,'Evaluation of vegetation indices and modified simple ratio for boreal applications', *Canadian Journal of Remote Sensing*, vol. 22, no. 3, pp. 229 – 242.

Choe, E., van der Meer, F., van Ruitenbeek, F., van der Werff, H., de Smeth, B. and Kim, Y. W. 2008, 'Mapping of heavy metal pollution in stream sediments using combined geochemistry, field spectroscopy, and hyperspectral remote sensing: A case study of the Rodalquilar mining area, SE Spain', *Remote Sensing of Environment,* vol. 112, no. 7, pp. 3222-3233.

Clark, R. N. and Roush, T. L. 1984,'Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications', *Journal of Geophysical Research*, vol. 89, no. B7, pp. 6329 – 6340.

Clevers, J. G. P. W., Kooistra, L. and Salas, E. A. L. 2004,'Study of heavy metal contamination in river floodplains using the red-edge position in spectroscopic data', *International Journal of Remote Sensing*, vol. 25, no. 9, pp. 3883 – 3895.

Cloutis, E. A. 1989, 'Spectral reflectance properties of hydrocarbons - remote-sensing implications', *Science,* vol. 245, no. 4914, pp. 165-168

Cocks, T., Jenssen, R., Stewart, A., Wilson, I., Shields, T. ,'The HyMap airborne hyperspectral sensor: the system, calibration and performance, in Schaepman, M., Schläpfer, D., Itten, K. (eds) *Proceedings of the 1st EARSeL workshop on Imaging Spectroscopy,6-8 October 1998, Zürich, Switzerland,* pp. 37 – 42, Remote sensing laboratories, University of Zürich, Switzerland.

Congalton, R. G. 1991,'A review of assessing the accuracy of classifications from remotely sensed data', *Remote Sensing of Environment,* vol. 37, no. 1, pp. 35 – 46.

Crawley, E. A. 2002, *Statistical computing*, John Wiley and Sons, Chichester.

Cundill, S. L. 2008: Investigation of soil reflectances for detecting hydrocarbon pipeline leakages, Msc Thesis, submitted to University of Southampton, University of Lund, University of Warsaw and ITC (International Institute for Geo-Information Science and Earth Observation, The Netherlands, p. 67.

Curran, P. J. 1994,'Imaging spectrometry', *Progress in Physical Geography*, vol. 18, no. 2, pp. 247 – 266.

Dai, X. and Khorram, S. 1999,'A feature-based image registration algorithm using improved chain-code representation combined with invariant moments', *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5-part-2, pp. 2351 – 2362.

Darvishzadeh, R., Atzberger, C. and Skidmore, A. K. 2006,'Hyperspectral vegetation indices for estimation of leaf area index', *In ISPRS mid-term symposium 2006 Remote Sensing: from pixels to processes*, 8 – 11 May, Enschede, The Netherlands, pp. 6.

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., Brown de Colstoun, E. and McMurtrey, J. E. 2000,'Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance', *Remote Sensing of Environment*, vol. 74, no. 2, pp. 229 – 239.

Daughtry, C. S. T., Hunt, E. R. and McMurtrey, J. E. 2004,'Assessing crop residue cover using shortwave infrared reflectance', *Remote Sensing of Environment*, vol. 90, no. 1, pp. 126 – 134.

Davidson, J. M. 2000 'Subsurface petroleum spills', in Lehr, J. H (ed) *Standard handbook of environmental science, health, and technology*, The McGraw-Hill Companies Inc, New York.

De Domenico, L., Crisafi, E., Magazzu, G., Puglisi, A. and La Rosa, A. 1994,'Monitoring of petroleum hydrocarbon pollution in surface waters by a direct comparison of fluorescence spectroscopy and remote sensing techniques', *Marine Pollution Bulletin*, vol. 28, no. 10, pp. 587 – 591.

Degermenci, O. 2001, 'From the Caspian Sea region to the European market - Feasibility of crude oil and natural gas transport via long-distance pipelines or routes through the Caspian Sea and their further continuation links westward', *Oil Gas-European Magazine,* vol. 27, no. 3, pp. 7-+.

Demattê, J. A. M., Fiorio, P. R. and Ben-Dor, E. 2009,'Estimation of soil properties by orbital and laboratory reflectance means and its relation with soil classification', *The open Remote Sensing Journal*, vol. 2, no.12, pp. 21 – 23.

DHV, 2008: GRE-leiding Anjum – Grijpskerk Rapportage monitoringsonderzoek 2008.

EEA (European Environment Agency): Expenditure on remediation of contaminated sites, [online], 2003,available:http://themes.eea.europa.eu/Specific_media/soil/indicators/clean_up%2C2003.11/expenditures_final_17_09_03.pdf, [accessed 2009, August 24].

EEA (European Environment Agency): Progress in Management of contaminated sites, [online], 2007,available:http://themes.eea.europa.eu/IMS/IMS/ISpecs/ISpecification20041007131746/IAssessment115261898983/view_content´ [accessed 2009, July 27].

Eganhouse, R. P., Dorsey, T. F., Phinney, C. S. and Westcott, A. M. 1996,'Processes affecting the fate of monoaromatic hydrocarbons in an aquifer contaminated by crude oil', *Environmental Science and Technology*, vol. 30, no. 11, pp. 3304 – 3312.

EIA (Energy Information Administration): International Energy statistics, [online], 2008, available:http://tonto.eia.doe.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=5&pid=53&aid, [accessed 2009, September 17].

Eismann, M. T. and Meola, J. 2008,'Hyperspectral change detection: methodology and challenges', *Geoscience and Remote Sensing Symposium IGARSS 2008*, vol. 2, pp. II-605-II-608.

Ellis, J. M., Davis, H. H and Zamudio, J. A. 2001, ´Exploring for onshore oil seeps with hyperspectral imaging´, *Oil and Gas Journal*, vol. 99, no. 37, pp. 49 – 58.

Elvidge, C. D. 1990,'Visible and near infrared reflectance of dry plant materials', *International Journal of Remote Sensing*, vol. 11, no. 10, pp. 1775 – 1795.

Encyclopaedia of the Nations: The Netherlands - Agriculture, [online], available: http://www.nationsencyclopedia.com/economies/Europe/The-Netherlands-AGRICULTURE.html, [accessed 2009, December 28].

ENSR International: An estimate of the national cost for remediation of MTBE releases from existing leaking underground storage tank sites, [online], 2005, available: http://www.aehs.com/reports/MTBEReport.pdf [accessed 2009, September 29].

Environmental Simulations International (ESI): Contaminated land – dealing with hydrocarbon contamination, 2008, [online], available: http://www.google.fi/#hl=fi&q=contaminated+land++dealing+with+hydrocarbon+contamination&meta=&fp=d651c94e2d327871, [accessed 2009, September 9].

Ferraro, G., Meyer-Roux, S., Muellenhoff, O., Pavliha, M., Svetak, J., Tarchi, D. and Topouzelis, K. 2009,'Long term monitoring of oil spills in European seas', *International Journal of Remote Sensing*, vol. 30, no. 3, pp. 627 – 645.

Filella, I. and Penuelas, J. 1998,'Visible and near-infrared reflectance techniques for diagnosing plant physiological status', *Trends in Plant Science*, vol. 3, no. 4, pp. 151 – 156.

Friend, D. J. 1996, *Remediation of Petroleum contaminated soils (Synthesis of highway practice / National Cooperative Highway program*, National Academy Press, Washington

Furby, S. L. and Campbell, N. A. 2001,'Calibration images from different dates to like-value digital counts', *Remote Sensing of Environment*, vol. 77, no. 2, pp. 186 – 196.

Gamon, J. A., Penuelas, J., and Field C. B. 1992,'A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency', *Remote Sensing of Environment*, vol. 41, no. 1, pp. 35 – 44.

Gamon, J. A. and Surfus, J. 1999,'Assessing leaf pigment content and activity with a reflectometer', *New Phytologist*, vol. 143, no. 1, pp. 105 – 117

Gao, B. C. 1995,'Normalized Difference Water Index for remote sensing of vegetation liquid water from space', *Proceedings in SPIE 2480,* pp. 225 – 236.

Gay, J. R. and Korre, A. 2006,'A spatially-evaluated methodology for assessing risk to a population from contaminated land', *Environmental Pollution*, vol. 142, no. 2, pp. 227 – 234.

Gingerich, J. C., Peshko, M. and Matthews, L. W., 2002, 'The development of new exploration technologies at Noranda: Seeing more with hyperspectral and deeper with 3-D seismic', *CIM Bulletin*, vol. 95, no. 1058, pp. 56 – 61.

Gitelson, A. A. and Merzlyak, M. N. 1997,'Remote estimation of chlorophyll content in higher plant leaves', *International Journal of Remote Sensing*, vol. 18, no. 12, pp. 2691 – 2697.

Gitelson, A. A., Merzlyak, M. N. and Chivkunova, O. B. 2001,'Optical properties and nondestructive estimation of Anthocyanin content in plant leaves', *Photochemistry and Photobiology*, vol. 71, no. 1, pp. 38 – 45.

Gitelson, A. A., Zur, Y., Chivkunova, O. B. and Merzlyak, M. N. 2002,'Assessing Carotenoid content in plant leaves with reflectance spectroscopy', *Photochemistry and Photobiology*, vol. 75, no. 3, pp. 272 – 281.

Guyot, G. and Baret, F. 1988,'Utlisation de haute resolution spectrale pour suivre l'etat des couverts vegetaux, *Proceedings in 4th international colloquium on spectral signatures of objects in remote sensing*, Aussois, France, 18-22 January 1988. ESA sp-287, pp. 279 – 286.

Guyot, G., Baret, F. and Jacquemoud, S. 1992, Imaging spectrometry for vegetation studies, in Toselli, F. and Bodechtel, J. (eds) *Imaging spectroscopy: Fundamentals and Prospective Applications*, Kluwer, Dordrecht, pp. 145 – 165.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J. and Dextraze, L. 2002,’Integrated narrow-band vegetation indices for prediction of chlorophyll content for application to precision agriculture’, *Remote Sensing of Environment*, vol. 81, no. 2 – 3, pp. 416 – 426.

Hall, F. G., Strebel, D. E., Nickerson, J. E. and Goetz, S. J. 1991,’Radiometric rectification: Toward a common radiometric response among multi-date, multi-sensor images, *Remote Sensing of Environment*, vol. 35, no. 1, pp. 11 – 27.

Hanson, R. S. and Hanson, T. E. 1996,’Methanotrophic bacteria’, *Microbiological review*, vol. 60, no. 2, pp. 439 – 471.

Hardisky, M. A., Klemas, V. and Smart, R. M. 1983,’The influence of soil salinity, growth form, and leaf moisture on the reflectance spectra of *Spartina Alterniflora* canopies’, *Photogrammetric Engineering and Remote Sensing*, vol. 49, pp. 77 – 83.

Hoeks, J. 1972, ´Changes in composition of soil air near leaks in natural gas mains´, *Soil Science*, vol. 113, no. 1, pp. 46 – 54.

Holzwarth, S. and Bachmann, M. 2009: Airborne Remote Sensing, HyEurope 2008, Summary report.

Hornback, J. M. 2005, *Organic Chemistry, 2nd edn,* Thomson Learning Inc, Belmont

Howari, F. M. 2004,’Investigation of hydrocarbon pollution in the vicinity of United Arab Emirates coasts using visible and near infrared remote sensing data’, *Journal of Coastal Research*, vol. 20, no. 4, pp. 1089 – 1095.

Hörig, B., Kuhn, F., Oschutz, F. and Lehmann, F. 2001, 'HyMap hyperspectral remote sensing to detect hydrocarbons', *International Journal of Remote Sensing,* vol. 22, no. 8, pp. 1413-1422.

Huete, A. R., Liu, H., Batchily, K. and van Leeuwen, W. 1997,’A comparison of vegetation indices over a global set of TM images for EOS-MODIS’, *Remote Sensing of Environment*, vol. 59, no. 3, pp. 440 – 451.

Hunt, G. R. 1980,’Electromagnetic radiation: the communication link in remote sensing’, in Siegal, B. S. and Gillespie, A. R. *Remote Sensing in Geology*, John Wiley & Sons Inc., New York, pp. 91 – 115.

Hunt, E. R. and Rock, B. N. 1989, ´Detection of changes in leaf water content using near – and middle – infrared reflectances, *Remote Sensing of the Environment*, vol. 30, pp. 43 – 45.

Hussey, M. A. and Simecek, S. 1987: Characterization of grass root systems, [online], available:http://forageresearch.tamu.edu/1987/Characterization%20of%20Grass%20Root%20Systems.pdf, [accessed 2009, December 20].

Jensen, J. R. 1996, *Introductory digital image processing: a remote sensing perspective*, 2nd edn, Prentice-Hall Inc, New Jersey.

Jensen, J. R. 2000, *Remote sensing of the environment: an Earth Resource Perspective*, Prentice Hall Inc, New Jersey.

Kaufman, Y. J. and Tanre, D. 1996,’Strategy of direct and indirect methods for correcting the aerosol effect on remote sensing: from AVHRR to EOS-MODIS’, *Remote Sensing of Environment*, vol. 55, no. 1, pp. 65 – 79.

Kemper, T. and Sommer, S. 2002, ‘Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy', *Environmental Science & Technology,* vol. 36, no. 12, pp. 2742-2747.

Keys, R. G. 1981,’Cubic convolution interpolation for digital image processing’, *IEEE Transactions on Acoustic Speech and Signal Processing*, vol. ASSP. 29, no. 6, pp. 1153 – 1160.

Khan, S. D. and Jacobson, S. 2008, 'Remote sensing and geochemistry for detecting hydrocarbon micro seepages', *Geological Society of America Bulletin*, vol. 120, no. 1-2, pp. 96-105.

Kim, J. and Corapcioglu, M. Y. 2003, 'Modeling dissolution and volatilization of LNAPL sources migrating on the groundwater table', *Journal of Contaminant Hydrology,* vol. 65, no. 1-2, pp. 137-158.

Kokaly, R. F. and Clark, R. N. 1999,’Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression, *Remote Sensing of Environment*, vol. 67, no. 3, pp. 267 – 287.

Kooistra, L. 2004: Incorporating spatial variability in ecological risk assessment of contaminated river floodplains Ph.D dissertation submitted to Utrecht University. 172pp.

Kooistra, L., Wanders, J., Epema, G. F., Leuven, R., Wehrens, R. and Buydens, L. M. C. 2003, 'The potential of field spectroscopy for the assessment of sediment properties in river floodplains', *Analytica Chimica Acta,* vol. 484, no. 2, pp. 189-200.

Kostecki, P. T. and Calabrese, E. J. 1991, *Hydrocarbon contaminated soils and groundwater: Analysis, Fate, Environmental and public health affects, Remediation, 1st edn*, Lewis Publishers, Inc. Michigan.

Kühn, F., Oppermann, K. and Hörig, B. 2004, 'Hydrocarbon Index - an algorithm for hyperspectral detection of hydrocarbons', *International Journal of Remote Sensing,* vol. 25, no. 12, pp. 2467-2473.

Laherrere, J.: Hydrocarbon resources, Forecast of oil and gas supply to 2050, [online], 2003, available:http://www.oilcrisis.com/laherrere/Petrotech090103.pdf, [accessed 2009, September 16].

Lehmann, T. M., Gönner, C. and Spitzer, K. 1999,'Survey: Interpolation methods in medical image processing', *IEEE Transactions on Medical Imaging*, vol.18, no. 11, pp. 1049 – 1075.

Li, H., Manjunath, B. S., and Mitra, S. K., 1995,'A contour based approach to multisensor image registration', *IEEE Transaction on Image Processing*, vol. 4, no. 3, pp. 320 – 334.

Li, L, Ustin, S. L. and Lay, M. 2005,´Application of AVIRIS data in detection of oil-induced vegetation stress and cover change at Jornada, New Mexico´, *Remote Sensing of Environment*, vol. 94, no. 1, pp. 1 – 16.

Lichtenthaler, H. K., Gitelson, A., Lang, M. 1996,'Non-destructive determination of chlorophyll content of leaves of a green and an aurea mutant of tobacco by reflectance measurements', *Journal of Plant Physiology*, vol. 148, no. 3 – 4, pp. 483 – 493.

Lin, H., Tao, S., Zuo, Q. and Coveney, R. M. 2007,'Uptake of polycyclic aromatic hydrocarbons by maize plants', *Environmental Pollution*, vol. 148, no. 2, pp. 614 – 619.

Luis., S. J. and Zemba, S. G. 'Estimating the impact of Source Depletion on Long-Term Risk Assessments', in Calabrese, E. J. and Kostecki, P. T. (eds) *Hydrocarbon contaminated soils vol. 3*, CRC Press, Massachusetts, pp. 243 – 266.

MacDonald, J. S., Ustin, S. J., and Schaepman, M. E., 2009,'The contributions of Dr. Alexander F. H. Goetz to imaging spectrometry', *Remote Sensing of Environment*, vol. 113, Supplement 1, pp. S2 – S4

Malley, D. F., Hunter, K. N. and Webster, G. R. B. 1999, 'Analysis of diesel fuel contamination in soils by near-infrared reflectance spectrometry and solid phase microextraction-gas chromatography', *Journal of Soil Contamination*, vol. 8, no. 4, pp. 481-489.

McCoy, R. M., Blake, J. G. and Andrews, K. L. 2001,'Detecting hydrocarbon microseepage using hydrocarbon absorption bands of reflectance spectra of surface soils', *Oil and Gas Journal*, vol. 99, no. 22, pp. 40 – 44.

Merton, R. 1998,'Monitring community hysteresis using spectral shift analysis and the red edge vegetation stress index', *In Proceedings of the 7th annual JPL airborne Earth Science Workshop*, NASA, JPL, pp. 1 – 14.

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. YU., 1999,'Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening', *Physiologia Plantarum*, vol. 106, no. 1, pp. 135 – 141.

Miller, J. R., White, H. P., Chen, J. M., Peddle, D. R., McDermid, G., Fournier, R. A., Shepherd, P., Rubinstein, I., Freemantle, J., Soffer, R. and LeDrew, E. 1997,'Seasonal change in understory reflectance of boreal forests and influence on canopy vegetation indices', *Journal of Geophysical Research*, vol. 102, no. D24, pp. 29,475 – 29,482.

Ministerie van Volkshuisvesting, Ruimtelijke ordening en milieubeheer: ANNEXES on circular on target values and intervention values for soil remediation, [online], 2000, available: http://www2.minvrom.nl/Docs/internationaal/annexS_I2000.pdf [accessed 2010, January 13].

Momeni, M. and Saradjian, M. R. 2007,' Evaluating NDVI-based emissivities of MODIS bands 31 and 32 using emissivities derived by Day/Night LST algorithm', *Remote Sensing of Environment*, vol. 106, no. 2, pp. 190 – 198.

Moore, D. S. 2000, *The Basic Practice of Statistics*, 2nd edn, W.H. Freeman Company, New York

Mutanga, O., Skidmore, A. K. and 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.

Mutanga, O., Skidmore, A. K., Kumar, L. and Ferwerda, J. 2005,'Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain', *International Journal of Remote Sensing*, vol. 26, no. 6, pp. 1093 – 1108.

Neue, H. U. 1993,'Methane emission from rice fields', *Bioscience*, vol. 43, no. 7, pp. 466 – 474.

Nielsen, E., Finco, M. and Hinkley, E. 2008, 'Change detection in image time-series affected by directional reflectance and phenological variability: Application to forest disturbance monitoring, *IGARSS International Geoscience and Remote Sensing Symposium*, vol. 4, pp. IV – 41 – IV – 44.

Noomen, M. F. 2007: Hyperspectral reflectance of vegetation affected by underground hydrocarbon gas seepage, PhD dissertation submitted to Wageningen University, pp. 151.

Noomen, M., van der Meer, F. and Skidmore, A. 2005,'Hyperspectral remote sensing for detecting the effects of three hydrocarbon gases on maize reflectance', in *Proceedings of the 31st international symposium on remote sensing of environment: global monitoring for sustainability and security*, St. Petersburg, Russia.

Noomen, M. F., Skidmore, A. K., van der Meer, F. D. and Prins, H. H. T. 2006, 'Continuum removed band depth analysis for detecting the effects of natural gas, methane and ethane on maize reflectance', *Remote Sensing of Environment,* vol. 105, no. 3, pp. 262-270.

Noomen, M. F., Smith, K. L., Colls, J. J., Steven, M. D., Skidmore, A. K. and Van der Meer, F. D. 2008, 'Hyperspectral indices for detecting changes in canopy reflectance as a result of underground natural gas leakage', *International Journal of Remote Sensing,* vol. 29, no. 20, pp. 5987-6008.

NRCAN (Natural Resources Canada): Canada Centre for Remote Sensing: Glossary of remote sensing terms, [online], 2005, available: http://www.ccrs.nrcan.gc.ca/glossary/index_e.php?id=68, [accessed 2009, November, 24].

Obire, O. and Nwaubeta, O. 2002,'Effects of refined petroleum hydrocarbon on soil physicochemical and bacteriological characteristics', *Journal of Applied Sciences and Environmental Management*, vol. 6, no. 1, pp. 39 – 44.

Olah, G. A. and Molnár, Á. 2003, *Hydrocarbon Chemistry, 2nd edn*, John Wiley and Sons, New Jersey

Organization of the Petroleum Exporting Countries (OPEC): World Oil Outlook 2009, [online], 2009, available: http://www.opec.org/library/world%20oil%20outlook/WorldOilOutlook09.htm [accessed 2009, September 9]

Orlov, D. S., Ammosova, M., Bocharnikova, A. and Lopukhina, O. V. 1999,'The use of reflectance of oil-polluted soils in monitoring from remote sensing imagery', *Mapping Sciences and Remote Sensing*, vol. 28, no. 5, pp. 270 – 275

Osuji, L. C., Udoetok, I. A. and Ogali, R. E. 2006, 'Attenuation of petroleum hydrocarbons by weathering: A case study', *Chemistry & Biodiversity,* vol. 3, no. 4, pp. 422-433.

Palmer, P. E., Hall, S. and Darby, J.: The future of petroleum hydrocarbon remediation: site closure through enhancement of in situ biological degradation, [online], 2002, available: http://ipec.utulsa.edu/Conf2002/palmer_hall_darby_13.pdf [accessed 2009, September 29].

Penn, B. S. 2002,'Band ratios as reconnaissance tool for hyperspectral data', *In Proceedings of JPL AVIRIS Workshop*, Pasadena, Jet Propulsion Laboratory

Penuelas, J. Gamon, J. A., Fredeen, A. L., Merino, J. and Field, C. B. 1994,'Reflectance indices associated with physiological changes in nitrogen and water-limited sunflower leaves', *Remote Sensing of Environment*, vol. 48, no. 2, pp. 135 – 146.

Penuelas, J., Filella, I., Lloret, P., Munoz, F and Vilajeliu, M. 1995,'Reflectance assessment of mite effects on apple trees', *International Journal of Remote Sensing*, vol. 16, no. 14, pp. 2727 – 2733.

Penuelas, J., Pinol, J., Ogaya, R. and Filella, I. 1997,'Estimation of plant water concentration by the reflectance Water Index (R900/R970)', *International Journal of Remote Sensing*, vol. 18, no. 13, pp. 2869 – 2875.

Poelman, J. N. B. 1966, *De bodem van Utrecht: toelichting bij blad 6 van de bodemkaart van Nederland: schaal 1:50 000*, Stichting voor Bodemkartering, Wageningen.

Pu, R., Gong, P., Biging, G. S. and Larrieu, M. R. 2003, 'Extraction of red edge optical parameters from Hyperion data for estimation of forest leaf area index', *IEEE Transactions on Geoscience and Remote Sensing,* vol. 41, no. 4, pp. 916 – 921.

Qi, J., Chehbouni, A., Huete, A. R., Keer, Y. H. and Sorooshian, S. 1994,'A modified soil vegetation adjusted index', *Remote Sensing of Environment*, vol. 48, no. 2, pp. 119 – 126.

Rama, R. N. 2008,'Development of a crop-specific spectral library and discrimination of various agricultural crop varieties using hyperspectral imagery', *International Journal of Remote Sensing*, vol. 29, no. 1, pp. 131 – 144.

Randolph, K., Wilson, J., Tedesco, L., Li, L. Pascual, D. D. and Soyeux, E. 2008,’Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll a, and phycocyanin, *Remote Sensing of Environment*, vol. 112, no. 11, pp. 4009 – 4019.

Reid, N., Iwashita, A, Yamashita, Y and Thompson, K. 1988,´High resolution imaging of geobotanical anomalies associated with subsurface hydrocarbons, *Proceedings in the Sixth Thematic Conference on Remote Sensing for Exploration Geology,* Houston, Texas, pp. 213 – 223.

Richter, R. and Schlaepfer, D. 2002,’Geo-atmospheric processing of airborne imaging spectrometry data, part 1: orthorectification’, *International Journal of Remote Sensing*, vol. 23, no. 23, pp. 2609 – 2630.

Rondeaux, G., Steven, M., Baret, F. 1996,’Optimization of soil-adjusted vegetation indices’, *Remote Sensing of Environment*, vol. 55, no. 2, pp. 95 – 107.

Rougean, J. L. and Breon, F. M. 1995,’Estimating PAR absorbed by vegetation from bidirectional reflectance measurements’, *Remote Sensing of Environment*, vol. 51, no. 3, pp. 375 – 384.

Rouse, J. W., Haas, R. H. and Schell, J. A. 1974. Monitoring the vernal advancement of retrogradiation of natural vegetation, Greenbelt, MD, USA.

Salem, F. and Kafatos, M. 2001,’Hyperspectral image analysis for oil spill mitigation’, *22nd Asian conference in Remote Sensing, 5–9 November, Singapore*, [online], available: http://www.crisp.nus.edu.sg/~acrs2001/pdf/108SALEM.PDF, [accessed 2009, December 18]

Schlaepfer, D. and Richter, R. 2002,’Geo-atmospheric processing of airborne imaging spectrometry data, part 2: atmospheric/topographic correction’, *International Journal of Remote Sensing*, vol. 23, no. 13, pp. 2631 – 2649.

Schmidt, K. S. and Skidmore, A. K. 2003,’Spectral discrimination of vegetation types in coastal wetland’, *Remote Sensing of Environment*, vol. 85, no. 1, pp. 92 – 108.

Schumacher, D. 1996, Hydrocarbon –induced alteration of soils and sediments, In Schumacher, D and Abrams, M. A. (eds), *Hydrocarbon Migration and its Near-Surface Expression*, AAPG Memoir 66, pp. 71 – 89.

Serrano, L., Penuelas, J. and Ustin, S. L. 2002,’Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data – decomposing biochemical from structural signals’, *Remote Sensing of Environment*, vol. 81, no. 2 – 3, pp. 355 – 364.

Shi, J. and Reichenbach, S. E. 2006,’Image interpolation by two-dimensional parametric cubic convolution’, *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1857 – 1870.

Shirdam, R., Daryabeigi Zand, A., Nabi Bidhendi, G. and Mehrdadi, N. 2008,’Phytoremediation of hydrocarbon-contaminated soils with emphasis on the effect of petroleum hydrocarbons on the growth of plant species’, *Phytoprotection*, vol. 89, no. 1, pp. 21 – 29.

Simonich S. and Hites, R. A. 1995,’Organic pollutant accumulation in vegetation’, *Environmental Science and Technology*, vol. 29, no. 12, pp. 2905 – 2914.

Sims, D. A. and Gamon, J. A. 2002,’Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages’, *Remote Sensing of Environment*, vol. 81, no. 3, pp. 337 – 354.

Smailbegovic, A., Michalski, J. and Rael, A. 2004,’Classification of geologic features in the open pit mines using high resolution HyperSpecTir imagery’, *Proceedings in IEEE Aerospace Conference*, vol. 3, pp. 1806 – 1811.

Smith, K.. L., Steven, M. and Colls, J. 2004,’Use of hyperspectral derivative ratios in the red edge region to identify plant stress responses to gas leaks’, *Remote Sensing of Environment*, vol. 92, no. 2, pp. 207 – 217.

Smith, K. L. Steven, M. D. and Colls, J. J. 2005,´Plant spectral responses to gas leaks and other stresses, *International Journal of Remote Sensing*, vol. 26, no. 18, pp. 4067 – 4081.

Sobrino, J. A. and Raissouni, N. 2000, ‘Toward remote sensing methods for land cover dynamic monitoring: application to Morocco', *International Journal of Remote Sensing*, vol. 21, no. 2, pp. 353-366.

Sorvari, J., Antikainen, R., Kosola, R., Hokkanen, M. L. and Haavisto, T. 2009,’eco-efficiency in contaminated land management in Finland – barriers and development needs’, *Journal of Environmental Management*, vol. 90, no. 5, pp. 1715 – 1727.

Taylor, A. W. 2000,'Hyperspectral remote sensing promotes: Early problem detection in facility and pipeline monitoring', *Pipeline and Gas Journal*, [online], available: http://www.oildompublishing.com/PGJ/pgjarchive/Oct2000/oct-2.pdf, [accessed 2009, September 29.

Ud Din, S., Al Dousari, A., Literathy, P. 2008,'Evidence of hydrocarbon contamination from the Burgan oil field, Kuwait – Interpretations from thermal remote sensing data', *Journal of Environmental Management*, vol. 86, no. 4, pp. 605 – 615.

U.S. Environmental Protection Agency (EPA): Hydrocarbons, [online], 2007, available:http://www.epa.gov/oms/invntory/overview/pollutants/hydrocarbons.htm [accessed 2009, August 22].

van den Berg, M., van de Meent, D., Peijnenburg, W. J. G. M., Sijm, D. T. H. M., Struijs, J. and Tas, J. W. 1995,Transport, accumulation and transformation processes, in van Leeuwen, C. J. and Hermens, J. L. M. *Risk assessment of chemicals: An introduction*, Kluwer Academic Publishers, Dordrecht.

van der Meer, F. D. 2001: Basic physics of spectrometry, in van der Meer, F. D. and de Jong, S. M. (eds), *Imaging spectrometry, Basic Principles and Prospective applications* Kluwer Academic Publisher, Netherlands, pp. 3 – 16.

van der Meer, F. D., van Dijk, P., van der Werff, H. and Yang, H. 2002,'Remote sensing and petroleum seepage: a review and case study', *Terra Nova*, vol. 14, no. 1, pp. 1 – 17.

van der Meer, F. 2004,'Analysis of spectral absorption features in hyperspectral imagery', *International Journal of Applied Earth Observation and Geoinformation*, vol. 5, no. 1, pp. 55 – 68.

van der Meer, F. D., Yang, H. and Kroonenberg, S. B. 2000a,'Imaging spectrometry and petroleum geology: hydrocarbon microseepage as a source of global methane production', *In Proceedings of the 28th International Symposium on Remote Sensing of the Environment: Information for sustainable development,* March 27 – 31, 2000, Cape Town, South Africa, Cat. 5, pp. 36 – 39.

van der Meer, F. D., van Dijk, P. M., Kroonenberg, S. B., Yang, H. and Lang, H. 2000b,'Hyperspectral hydrocarbon microseepage detection and monitoring: potentials and limitations', *In proceedings of the 2nd EARSeL workshop on imaging spectroscopy*, Enschede, ITC, July, pp. 1 – 9.

van der Meer, F. D., de Jong, S. and van der Werff, H. M. A. 2009: Continuum removal of spectral feature finding, Distance Education Course, Hyperspectral Remote Sensing CD.

van der Meijde, M., van der Werff, H. M. A., Jansma, P. F., van der Meer, F. D. and Groothuis, G. J. 2009,'A spectral-geophysical approach for detecting pipeline leakage', *International Journal of Applied Earth Observation and* Geoinformation, vol. 11., no. 1, pp. 77 – 82.

van der Werff, H. M. A. 2006: Knowledge based remote sensing of complex objects: recognition of spectral and spatial patterns resulting from natural hydrocarbon seepages. Ph.D dissertation submitted to Utrecht University. 138pp.

van der Werff, H. M. A., Bakker, W. H., Van der Meer, F. D. and Siderius, W. 2006, 'Combining spectral signals and spatial patterns using multiple Hough transforms: An application for detection of natural gas seepages', *Computers & Geosciences,* vol. 32, no. 9, pp. 1334-1343.

van der Werff, H. M. A., Noomen, M. F., van der Meijde, M. and van der Meer, F. D. 2007. Remote sensing of onshore hydrocarbon seepage: problems and solutions, in Teeuw, R. M. (ed) *Mapping hazardous terrain using remote sensing*, Geological Society, London, Special publications, 283, pp. 125 – 134.

van der Werff, H., van der Meijde, M., Jansma, F., van der Meer, F. and Groothuis, G. J. 2008,'A spatial-approach for visualization of vegetation stress resulting from pipeline leakage', *Sensors*, vol. 8, 1-x manuscripts, pp 1 – 17.

Weaver, J. E. 1926, *Root developments of field crops,* McGraw-Hill, New York, pp. 291.

White, D. C. 2007. Hyperspectral remote sensing of canopy scale vegetation stress associated with buried gas pipelines, Ph.D dissertation submitted to Newcastle University, 291pp.

WHO (World Health Organization): Environmental Health Criteria 150, Benzene, [online], 1993, available: http://www.inchem.org/documents/ehc/ehc/ehc150.htm [accessed 2009, September 17].

Winkelmann, K. H. 2005. On the applicability of Imaging Spectroscopy for the detection and investigation of contaminated sites with particular consideration given to the detection of fuel hydrocarbon contamination in soil, Ph.D dissertation submitted to Brandenburischen Technischen Universitat Cottbus. 245pp.

Vogelmann, J. E., Rock, B. N., Moss, D. M. 1993,'Red edge spectral measurements from sugar maple leaves', *International Journal of Remote Sensing*, vol. 14, no. 8, pp. 1563 – 1575.

Xu, D. Q., Ni, G. Q., Jiang, L. L., Shen, Y. T., Li, T., Ge, S. L., Shu, X. B. 2008, 'Exploring for natural gas using reflectance spectra of surface soils', *Advances in Space Research*, vol. 41, no. 11, pp. 1800 – 1817.

Yang, H., Zhang, J., Van der Meer, F. and Kroonenberg, S. B. 1999, 'Spectral characteristics of wheat associated with hydrocarbon micro seepages', *International Journal of Remote Sensing,* vol. 20, no. 4, pp. 807 – 813.

Yang, H., Zhang, J., Van der Meer, F. and Kroonenberg, S. B. 2000, 'Imaging spectrometry data correlated to hydrocarbon microseepage', *International Journal of Remote Sensing*, vol. 21, no. 1, pp. 197-202.

Yang, Q., Parvin, B., Mariano, A. J., Ryan, E. H., Evans, R. and Brown, O. B. 2004,´Seasonal and interannual studies of vortices in sea surface temperature data´, *International Journal of Remote Sensing*, vol. 25, no. 7 – 8, pp. 1371 – 1376.

Yaw Kwarteng, A. 1998, 'Multitemporal remote sensing data analysis of Kuwait's oil lakes', *Environmental International*, vol. 24, no. ½, pp. 121 – 137.

Zarco-Tejada, P. J., Berjon, A., Lopez-Lozano, R., Miller, J. R., Martin, P., Cachorro, V., Gonzales, M. R. and de Frutos, A. 2005,'Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy', *Remote Sensing of Environment*, vol. 99, no. 3, pp. 271 – 287.

Zarco-Tejada, P. J., Miller, J. R., Mohammed, G. H. and Noland, T. L. 2000,'Chlorophyll fluorescence effects on vegetation apparent reflectance: I. Leaf level measurements and simulation of reflectance and transmittance spectra', *Remote Sensing of Environment*, vol. 74, no. 3, pp. 582 – 595.

Zarco-Tejada, P. J., Miller, J. R., Mohammed, G. H., Noland, T. L. and Sampson, P. H. 2001,'Scaling-up and model inversion methods with narrow-band optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data', *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1491 – 1507.

9. Appendices

Appendix 1

Drilling data (groundwater samples, n=27) used in the analysis and validation

Point ID	**X**	**Y**	**HC conc. 2003 (µg/l)**	**HC conc. 2008 (µg/l)**	**Depth (m)**	**Land use in 2005 based on the land use map**	**ND VI 2005**	**ND VI 2008**	**Soil type***
0	205842	593688	< 0.2	< 0.2	1.9 - 2.9	Meadow	0.805	0.721	CHC
1	206093	593589	32 000	< 0.2	2.0 - 3.0	Meadow	0.498	0.682	CHC
13	206129	593584	0	3.8	2.0 - 3.0	Meadow	0.807	0.694	CHC
2	206128	593581	24 000	< 0.2	1.8 - 2.8	Meadow	0.830	0.676	CHC
12	206127	593577	0	< 0.2	2.0 - 3.0	Meadow	0.808	0.684	CHC
23	206131	593577	1900	480	4.0 - 5.0	Meadow	0.763	0.711	CHC
3	206139	593577	1.9	1000	4.0 - 5.0	Meadow	0.627	0.363	CHC
4	206192	593561	< 0.2	< 0.2	3.8 - 4.8	Meadow	0.769	0.730	CHC
5	206301	593510	4 700	< 0.2	1.5 - 2.5	Meadow	0.822	0.764	CHC
6	206356	593481	2 500	< 0.2	2.0 - 3.0	Meadow	0.770	0.760	CHC
7	206394	593448	2 500	< 0.2	1.8 - 2.8	Meadow	0.640	0.749	CHC
8	206440	593409	320	< 0.2	2.0 - 3.0	Meadow	0.802	0.755	CLC
15	206483	593379	0	0.55	2.0 - 3.0	Meadow	0.815	0.785	CHC
9	206481	593376	15 000	< 0.2	2.0 - 3.0	Meadow	0.822	0.795	CHC
14	206478	593374	0	< 0.2	2.0 - 3.0	Meadow	0.794	0.792	CHC
10	206563	593305	2000	11	2.0 - 3.0	Meadow	0.952	0.756	CLC
11	206615	593256	< 0.2	< 0.2	2.0 - 3.0	Meadow	0.952	0.763	CLC
28	208602	591805	0	< 0.2	2.0 - 3.0	Arable land	0.122	0.761	CLC
31	210120	591250	0	< 0.2	2.0 - 3.0	Meadow	0.791	0.919	CLC
30	211640	590680	94 000	< 0.2	2.5 - 3.5	Arable land	0.111	0.788	CLC
29	211820	590680	150	< 0.2	2.0 - 3.0	Arable land	0.120	0.888	CLC
21	211968	590427	0.89	< 0.2	3.0 - 4.0	Potato field	0.931	0.200	CHC
22	212754	589502	< 0.2	< 0.2	3.0 - 4.0	Arable land	0.649	0.264	CLC
24	213900	588895	1.1	< 0.2	2.8 - 3.8	Potato field	0.903	0.146	CLC
17	215235	588078	4 400	2	3.0 - 4.0	Arable land	0.736	0.415	CLC
16	216055	587935	0	< 0.2	3.2 - 4.2	Potato field	0.951	0.461	CHC
26	216204	587950	0	< 0.2	2.0 - 3.0	Meadow	0.673	0.800	CLC

(*CLC = Calcareous light clay and CHC = Calcareous heavy clay, based on digital soil map produced from Poelman, 1966)

Appendix 2

The drilling data on non-vegetated areas in the 2005 image used for validating the band ratio images

Non-vegetated validation points

Point ID	X	Y	HC concentration 2003 (µg/l)	Depth (m)	Non-polluted/Polluted
NP_1	211908	590531	0	1.6 - 1.9	Non-polluted
NP_2	211913	590523	0	3.0 - 3.5	Non-polluted
NP_3	211916	590520	0	2.5 - 3.0	Non-polluted
NP_4	208594	591811	0.05	2.0 - 2.5	Non-polluted
NP_5	212903	589391	0.05	3.0 - 4.0	Non-polluted
NP_6	212868	589417	0.05	2.25 - 2.75	Non-polluted
P_7	211887	590569	0.2 - 0.46	1.5 - 3.0	Polluted
P_8	211642	590673	1.2 - 6000	1.6 - 3.2	Polluted
P_9	213437	589124	0.39	2.5 - 3.0	Polluted

Details of the drilling data (n=22) on vegetated areas in the 2005 image used for validating the 2005 normalized index-image.

ID	x	y	Threshold *	Depth	2005** Normalized REP-index value
1_NP	204736	595467	<S	3.0 - 4.0	0.522
2_NP	204816	595315	<S	2.6 - 3.1	-0.069
3_NP	206264	593525	<S	1.9 - 2.4	-0.143
4_NP	207515	592475	<S	2.1 - 2.6	-0.229
5_NP	209558	591277	<S	2.0 - 2.5	0.36
6_NP	212286	589992	<S	2.5 - 3.0	0.889
7_NP	212304	589961	<S	3.5 - 4.0	0.12
8_NP	212731	589518	<S	2.1 - 2.6	-0.138
9_NP	214093	588798	<S	2.5 - 3.0	0.675
10_NP	214891	588293	<S	0.3 - 0.8	-0.185
11_NP	216069	587904	T-I	2.9 - 3.0	0.25
12_NP	216079	587945	<S	2.9 - 3.0	0.562
13_NP	216202	587951	<S	2.7 - 3.7	0.264
1_P	204816	595311	I-10xl	3.0 - 4.0	-0.043
2_P	212308	589961	I - 10xl	3.0 - 3.5	0.125
3_P	214891	588297	> 100xl	0 - 0.1	-0.375
4_P	214900	588290	I - 10xl	0.5 - 0.6	-0.328
5_P	214908	588286	I - 10xl	0.5 - 0.6	-0.388
6_P	214916	588283	10xl-100xl	0 - 0.1	-0.424
7_P	215264	588064	10xl-100xl	2.7 - 3.0	0.022
8_P	216069	587904	I - 10xl	2.9 - 3.0	0.25
9_P	216202	587955	10xl-100xl	3.1 - 4.1	-0.333

*(<S=under this level the soil is clean, T=Intermediate level, only needs to be cleaned in certain conditions and I=Intervention level[6])

**(Negative values indicated that vegetation was more stressed and positive values the opposite, see section 3.3.10. for further details.)

---

[6] http://www2.minvrom.nl/Docs/internationaal/annexS_I2000.pdf

Appendix 3

The existing HC indices and the band ratios built on this study and applied for the 2005 HyMap image over the non-vegetated areas.

Existing HC indices

Index name	Formula	Reference
HC Index (HI)	(λb – λa) * ((Rc-Ra)/(λc-λa)) + Ra - Rb	(Kühn et al., 2004)
Aromatics 1	Re (1.554 nm) – Re (1.694 nm)	(Winkelmann, 2005)
1200 nm Aromatics 1	Re (1.235 nm) – Re (1.148 nm)	(Winkelmann, 2005)
1200 nm Aromatics 2	Re (1.235 nm) – Re (1.133 nm)	(Winkelmann, 2005)
2150 nm Aromatics Wood	Re (2.004) – Re (2.150 nm)	(Winkelmann, 2005)
2300 nm HC	Re (2.219 nm) – Re (2.304)	(Winkelmann, 2005)
HC Band ratio 1	R (1650 nm) / R (1180 nm)	(McCoy et al., 2001)
HC Band ratio 2	R (1650 nm) / R (1380 nm)	(McCoy et al., 2001)
HC Band ratio 3	R (1650 nm) / R (2305 nm)	(McCoy et al., 2001)
HC Band ratio 4	R (1650 nm) / R (2245 nm)	(McCoy et al., 2001)

Re (λ) – relative reflectance at wavelength λ and R (λ) – reflectance at wavelength λ

Band ratio images built in the current study

Band ratio (identified wavelength µm)	Equation (HyMap band, µm)	High values or low values expected for polluted locations
1 (0.67 µm)	R (0.7852) / R (0.6790)	LOW
2 (2.2 µm)	R (2.1557) / R (2.2254)	HIGH
3 (2.35 – 2.37 µm)	R (2.3914) / R (2.3756)	HIGH
4 (2.4 µm)	R (2.3914) / R (2.4072)	LOW
5 (2.45 µm)	R (2.4546) / R (2.4387)	LOW

Appendix 4

Vegetation indices tested in the study

Index Name	Equation	Reference
Anthocyanin Reflectance Index (ARI1)	(1/R550)-(/R700)	(Gitelson et al., 2001)
Anthocyanin Reflectance Index 2 (ARI2)	R800*[(1/R550)-(1/R700)	(Gitelson et al., 2001)
Atmospherically Resistant Vegetation Index (ARVI)	(RNIR-(^{2}RRed-Rblue)/(RNIR+(^{2}RRed-Rblue)	(Kaufman and Tanre, 1996)
Blue/Green Index (BGI1)	(R400/R500)	(Zarco-Tejada et al., 2005)
Blue/Green Index 2 (BGI2)	(R450/R550)	(Zarco-Tejada et al., 2005)
Blue/Red Index (BRI1)	(R400/R690)	(Zarco-Tejada et al., 2005)
Blue/Red Index 2 (BRI2)	(R450/R690)	(Zarco-Tejada et al., 2005)
Cellulose Absorption Index (CAI)	0.5*(R2000+R2200) – R2100	(Daughtry et al., 2004)
Carotenoid reflectance Index 2 (CRI1)	(1/R510)-(1/550)	(Gitelson et al., 2002)
Carotenoid reflectance Index 2 (CRI2)	(1/R510)-(1/700)	(Gitelson et al., 2002)
Carter Index 1 (CTR1)	(R695/R420)	(Carter, 1994)
Carter Index 2 (CTR2)	(R695/R760)	(Carter, 1994)
Curvature Index (CUR)	R^2 683/(R675*R690)	(Zarco-Tejada et al., 2000)
Enhanced Vegetation Index (EVI)	2.5*(RNIR-Rred)/(RNIR+6*Rred-7.5*Rblue+1)	(Huete et al., 1997)
Gitelson and Merzylak 1 (GM1)	(R750/R550)	(Gitelson and Merzlyak, 1997)
Improved SAVI with self-adjusting factor L (MSAVI)	0.5*[2*R800+-((2*R800+1)^2-8*(R800-R670))$^{0.5}$]	(Qi et al., 1994)
Lichtenthaler Index 1 (LIC1)	(R800-R680)/(R800+R680)	(Lichtenthaler et al., 1996)
Lichtenthaler Index 2 (LIC2)	(R440/R690)	(Lichtenthaler et al., 1996)
Lichtenthaler Index 3 (LIC3)	(R440/R740)	(Lichtenthaler et al., 1996)
Modified C$_{ab}$ Absorption in Reflectance Index (MCARI)	[(R700-R670)-0.2*(R700-R550)]*(R700/R670)]	(Daughtry et al., 2000)
Modified chlorophyll absorption in reflectance Index (MCARI1)	1.2*[2.5*(R800-R670)-1.3*(R800-R550)]	(Haboudane et al., 2002)
Modified chlorophyll absorption in reflectance Index (MCARI2)	1.5*[2.5*(R800-R670)-1.3*(R800-R550)/√((2*R800+1)^2 – (6*R800-5*√R670))-0.5]	(Haboudane et al., 2002)
Modified Red Edge Normalized Difference Vegetation Index (mNDVI705)	(R750-R705)/(R750+R705-2*R445)	(Sims and Gamon, 2002)
Modified Simple Red Edge Ratio Index (mSR705)	(R750-R445)/(R705-R445)	(Sims and Gamon, 2002)
Modified Simple Ratio (MSR)	(RNIR/Rred-1)/((RNIR(Rred)$^{0.5}$+1)	(Chen, 1996)
Modified Triangular Vegetation Index (MTVI1)	1.2*[1.2*(R800-R550)-2.5*(R670-R550)]	(Haboudane et al., 2002)
Modified Triangular Vegetation Index (MTVI1)	1.5*[1.2*(R800-R550)-2.5*(R670-R550)/√((2*R800+1)^2 – (6*R800-5*√R670))-0.5]	(Haboudane et al., 2002)
Moisture Stress Index (MSI)	(R1599/R819)	(Hunt and Rock, 1989)
Normalized Difference Infrared Index (NDII)	(R819-R1649)/(R819+R1649)	(Hardisky et al., 1983)
Normalized Difference Lignin Index (NDLI)	(log(1/R1754)-log(1/R1680)) /(log(1/R1754)+log(1/R1680))	(Serrano et al., 2002)
Normalized Difference Nitrogen Index (NDNI)	(log(1/R510)-log(1/R1680)) /(log(1/R1510)+log(1/R1680))	(Serrano et al., 2002)
Normalized Difference Vegetation Index (NDVI)	(RNIR – Rred)/(RNIR + Rred)	(Rouse et al., 1974)
Normalized Difference Water Index (NDWI)	(R857-R1241)/(R857+R1241)	(Gao, 1995)
Normalized Pigment Chlorophyll Index (NPCI)	(R680 – R430) / R680 + R430)	(Penuelas et al., 1994)
Optimized Soil-Adjusted Vegetation Index (OSAVI)	(1+0.16)*(R800-R670)/(R800+R670+0.16)	(Rondeaux et al., 1996)
Photochemical reflectance Index (PRI)	(R531-R570)/(R531+R570)	(Gamon et al., 1992)
Plant senescence Reflectance Index (PSRI)	(R680-R500)/R750	(Merzlyak et al., 1999)
Red Edge Normalized Difference Vegetation Index (NDVI705)	(R750-R705)/(R750+R705)	(Sims and Gamon, 2002)

Red Edge Position Index (REP)	700+40*[((R670+R780/2)-R700)/ (R740-R700)]	(Guyot and Baret, 1988)
Red/Green Index (RGI)	(R690/R550)	(Zarco-Tejada et al., 2005)
Renormalized Difference Vegetation Index (RDVI)	(R800 – R670)/√(R800 – R670)	(Rougean and Breon, 1995)
Simple Ratio Index (SR)	(RNIR/Rred)	(Rouse et al., 1974)
Simple Ratio Pigment Index (SRPI)	R430/R680	(Penuelas et al., 1995)
Sum Green Index (SG)	Average of R500 to R600	(Gamon and Surfus, 1999)
Transformed CARI (TCARI)	3*[(R700-R670)-0.2*(R700-R550)*(R700/R670)]	(Haboudane et al., 2002)
Triangular Vegetation Index (TVI)	0.5*[120*(R750-R550)-200*(R670-R550)]	(Broge and Leblanc, 2000)
Vogelmann Index (VOG1)	(R740/R720)	(Vogelmann et al., 1993)
Water Band Index (WBI)	(R900/R970)	(Penuelas et al., 1997)
Zarco and Miller (ZM)	(R750/R710)	(Zarco-Tejada et al., 2001)

(R – Reflectance at specified wavelength

Appendix 5

HyMap bands (2005 and 2008) used for the calculation of vegetation indices (x=not needed for the calculation)

Index	Eq.1	HyMap Band1 (µm)		Eq. 2	HyMap Band2 (µm)		Eq.3	HyMap Band3 (µm)		Eq. 4	HyMap Band4 (µm)	
		2005	2008		2005	2008		2005	2008		2005	2008
ARI1	0.55	0.5568	0.5602	0.7	0.6944	0.6919	x	x	x	x	x	x
ARI2	0.8	0.8003	0,792	0.55	0.5568	0.5602	0.7	0.6944	0.6919	x	x	x
BGI1	0.4	0.4365	0.4572	0.55	0.5568	0.5602	x	x	x	x	x	x
BGI2	0.45	0.452	0.4572	0.55	0.5568	0.5602	x	x	x	x	x	x
BRI1	0.4	0.4365	0.4572	0.69	0.6944	0.6919	x	x	x	x	x	x
BRI2	0.45	0.452	0.4572	0.69	0.6944	0.6919	x	x	x	x	x	x
CRI1	0.51	0.5101	0.5157	0.55	0.5568	0.5602	x	x	x	x	x	x
CRI2	0.51	0.5101	0.5157	0.7	0.6944	0.6919	x	x	x	x	x	x
CTR1	0.695	0.6944	0.6919	0.42	0.4365	0.4572	x	x	x	x	x	x
CTR2	0.695	0.6944	0.6919	0.76	0.7552	0.7494	x	x	x	x	x	x
CUR	0.675	0.679	0.6773	0.69	0.6944	0.6919	0.683	0.679	0.6773	x	x	x
EVI	RNIR	0.8003	0.792	Rred	0.6487	0.6482	Rblue	0.452	0.4572	x	x	x
GM	0.75	0.7552	0.7494	0.55	0.5568	0.5602	x	x	x	x	x	x
MSAVI	0.8	0.8003	0.792	0.67	0.679	0.6773	x	x	x	x	x	x
LIC1	0.8	0.8003	0.792	0.68	0.679	0.6773	x	x	x	x	x	x
LIC2	0.44	0.4365	0.4572	0.69	0.6944	0.6919	x	x	x	x	x	x
LIC3	0.44	0.4365	0.452	0.74	0.74	0.7351	x	x	x	x	x	x
MCARI	0.7	0.6944	0.699	0.67	0.679	0.6773	0.55	0.5568	0.5602	x	x	x
MCARI1	0.8	0.8003	0.792	0.67	0.679	0.6773	0.55	0.5568	0.5602	x	x	x
MCARI2	0.8	0.8003	0.792	0.67	0.679	0.6773	0.55	0.5568	0.5602	x	x	x
mNDVI705	0.75	0.7552	0.7494	0.705	0.7098	0.7063	0.445	0.452	0.4572	x	x	x
mSR705	0.75	0.7552	0.7494	0.445	0.452	0.4572	0.705	0.7098	0.7063	x	x	x
MSR	RNIR	0.8003	0.792	Rred	0.679	0.6773	x	x	x	x	x	x
MTVI1	0.8	0.8003	0.792	0.55	0.5568	0.5602	0.67	0.679	0.6773	x	x	x
MTVI2	0.8	0.8003	0.792	0.55	0.5568	0.5602	0.67	0.679	0.6773	x	x	x
MSI	1.599	1.5879	1.585	0.819	0.8156	0.8206	x	x	x	x	x	x
NDII	0.819	0.8156	0.8206	1.649	1.6397	1.6372	x	x	x	x	x	x
NDNI	1.51	1.1568	1.1516	1.68	1.6776	1.6755	x	x	x	x	x	x
NDVI	RNIR	0.8003	0.792	Rred	0.679	0.6773	x	x	x	x	x	x
NDWI	0.857	0.8613	0.8631	1.241	1.2432	1.2348	x	x	x	x	x	x
NPCI	0.68	0.679	0.6773	0.43	0.4365	0.4572	x	x	x	x	x	x
OSAVI	0.8	0.8003	0.792	0.67	0.679	0.6773	x	x	x	x	x	x
PRI	0.531	0.5259	0.5308	0.57	0.5719	0.5751	x	x	x	x	x	x
PSRI	0.68	0.679	0.6773	0.5	0.4952	0.5009	0.75	0.7552	0.7494	x	x	x
NDVI705	0.75	0.7552	0.7494	0.70	0.7098	0.7063	x	x	x	x	x	x
REP	0.67	0.679	0.6773	0.78	0.7852	0.792	0.7	0.7098	0.7063	0.74	0.74	0.735
RGI	0.69	0.6944	0.6919	0.55	0.5568	0.5602	x	x	x	x	x	x
RDVI	0.8	0.8003	0.792	0.67	0.679	0.6773	x	x	x	x	x	x
SR	RNIR	0.8003	0.792	Rred	0.679	0.6773	x	x	x	x	x	x
SRPI	0.43	0.4365	0.4572	0.68	0.679	0.6773	x	x	x	x	x	x
SIPI	0.8	0.8003	0.792	0.45	0.452	0.4572	0.65	0.6487	0.6482	x	x	x
SG	0.5	0.4952	0.5009	0.6	0.6025	0.6046	x	x	x	x	x	x
TCARI	0.7	0.7098	0.7063	0.67	0.679	0.6773	0.55	0.5568	0.5602	x	x	x
TVI	0.75	0.7552	0.7494	0.55	0.5568	0.5602	0.67	0.679	0.6773	x	x	x
VOG1	0.76	0.74	0.7351	0.72	0.7249	0.7207	x	x	x	x	x	x
WBI	0.9	0.8948	0.8953	0.97	0.9746	0.9722	x	x	x	x	x	x
ZM	0.75	0.7552	0.7494	0.71	0.7098	0.7063	x	x	x	x	x	x
ARVI	RNIR	0.8003	0.792	Rred	0.679	0.6773	Rblue	0.452	0.4572	x	x	x
NDLI	1.754	1.7519	1.7497	1.68	1.6776	1.6755	x	x	x	x	x	x
CAI	2	1.9906	1.9888	2.2	2.2074	2.2094	2.1	2.1017	2.102	x	x	x

Appendix 6

The HyMap bands (2005) that were used to quantitatively analyze the possibly HC related absorption features identified from the CR spectra (see section 4.1.2).

No	VIS/NIR /SWIR	Wavelength (µm)	HyMap wavelength (µm)
1	VIS	0.45 - 0.5	0.43 – 0.55
2	VIS	0.67	0.55 – 0.70
3	NIR	0.9	0.87 – 1.07
4	NIR	1.15	1.03 – 1.27
7	SWIR	2.1	2.10 – 2.13
8	SWIR	2.2	2.17 – 2.29
9	SWIR	2.3	2.29 – 2.32
10	SWIR	2.35 – 2.37	2.35 – 2.37
11	SWIR	2.4	2.39 – 2.42
12	SWIR	2.45	2.43 – 2.47

The results from the absorption feature parameter calculations and statistical testing in wavelengths 0.45 µm, 0.67 µm, 0.89 µm, 1.15 µm, 2.1 µm, 2.2 µm, 2.3 µm, 2.35 – 2.37 µm, 2.4 µm and 2.45 µm for the average polluted and non-polluted spectra collected from three non-vegetated locations along the pipeline (table continues on the next page, see section 4.1.2 for summary of the results).

Wavelength (µm)	Parameter	Non-polluted /Polluted	28 (µg/l)	p-value	29 (150 µg/l)	p-value	30 (94000 µg/l)	p-value
0.45	**P**	**NP**	0.4752	0.07	0.4725	0.68	0.4776	0.48
		P	0.4627		0.4760		0.4904	
	BD	**NP**	0.0608	0.96	0.0833	0.76	0.0740	0.57
		P	0.0660		0.0643		0.0771	
	W	**NP**	0.0625	0.73	0.0648	0.80	0.0670	0.26
		P	0.0618		0.0634		0.0585	
	A	**NP**	0.0032	x	0.0046	0.56	0.0040	0.59
		P	0.0032		0.0034		0.0042	
	AS	**NP**	-0.7202	0.72	-0.6558	0.67	-0.5784	0.10
		P	-0.7262		-0.6307		0.1128	
0.67	**P**	**NP**	0.6790	x	0.6790	X	0.6790	x
		P	0.6790		0.6790		0.6790	
	BD	**NP**	0.0662	0.98	0.0853	0.01	0.0717	0.06
		P	0.0633		0.1053		0.0812	
	W	**NP**	0.1014	0.88	0.1071	x	0.1071	x
		P	0.1012		0.1071		0.1071	
	A	**NP**	0.0088	0.35	0.0077	0.01	0.0064	0.11
		P	0.0053		0.0094		0.0071	
	AS	**NP**	0.3005	0.48	0.4027	0.18	0.3711	0.74
		P	0.2590		0.4093		0.3649	
0.89	**P**	**NP**	0.9018	0.85	0.9047	0.70	0.9108	0.16
		P	0.9110		0.9024		0.9050	
	BD	**NP**	0.0158	0.06	0.0144	0.34	0.0127	0.30
		P	0.0116		0.0125		0.0144	
	W	**NP**	0.0786	0.39	0.0790	0.87	0.0774	0.27
		P	0.0774		0.0784		0.0805	
	A	**NP**	0.0009	0.21	0.0009	0.76	0.0008	0.33
		P	0.0007		0.0008		0.0009	
	AS	**NP**	-0.7896	0.13	-0.6590	0.80	-0.4957	0.19
		P	-0.5430		-0.6800		-0.6552	
1.15	**P**	**NP**	1.0597	0.28	1.1568	x	1.1568	0.34
		P	1.1568		1.1568		1.1532	
	BD	**NP**	0.0229	0.07	0.0290	0.92	0.0264	0.07
		P	0.0253		0.0292		0.0254	
	W	**NP**	0.1430	0.32	0.1442	0.22	0.1461	0.47
		P	0.1423		0.1476		0.1481	
	A	**NP**	0.0027	0.94	0.0036	x	0.0033	0.16
		P	0.0028		0.0036		0.0032	
	AS	**NP**	-0.0653	0.12	-0.0115	0.14	-0.0009	0.19
		P	-0.1378		-0.0529		-0.0522	

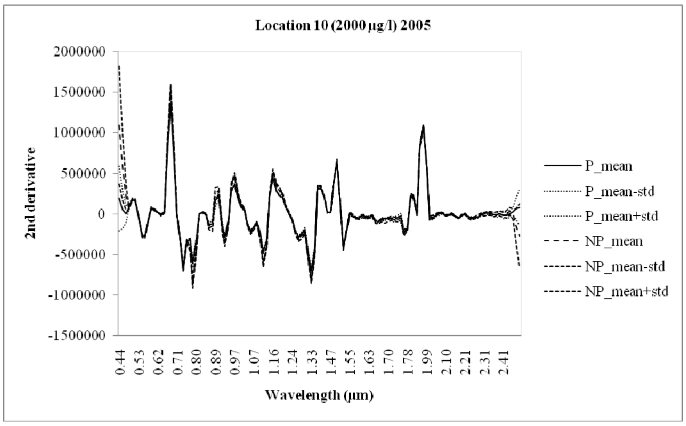
Appendix 6 (continue...)

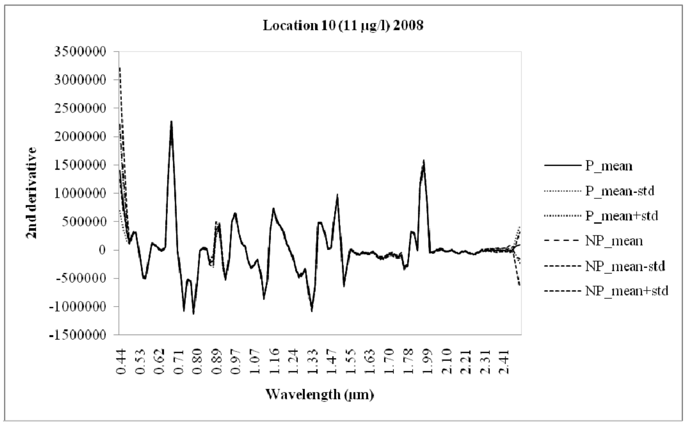
Wavelength (µm)	Parameter	Non-polluted/ Polluted	28 (µg/l)	p-value	29 (150 µg/l)	p-value	30 (94000 µg/l)	p-value
2.1	**P**	**NP**	2.1198	x	2.1198	x	2.1198	x
		P	2.1198		2.1198		2.1198	
	BD	**NP**	0.0007	0.13	0.0014	0.87	0.0011	0.15
		P	0.0001		0.0015		0.0000	
	W	**NP**	0.0540	x	0.0540	x	0.0540	x
		P	0.0540		0.0540		0.0540	
	A	**NP**	0.0000	x	0.0000	0.67	0.0000	0.78
		P	0.0000		0.0002		0.0001	
	AS	**NP**	0.0000	x	0.0000	x	0.0000	x
		P	0.0000		0.0000		0.0000	
2.2	**P**	**NP**	2.2224	x	2.2254	x	2.2254	x
		P	2.2254		2.2254		2.2254	
	BD	**NP**	0.0406	0.06	0.0704	0.055	0.0662	0.13
		P	0.0468		0.0659		0.0596	
	W	**NP**	0.0693	0.07	0.0857	0.97	0.0874	0.34
		P	0.0765		0.0855		0.0936	
	A	**NP**	0.0020	0.06	0.0044	**0.01**	0.0043	**0.03**
		P	0.0026		0.0038		0.0040	
	AS	**NP**	0.1626	0.87	-0.0237	0.06	-0.1012	0.40
		P	0.1659		-0.0277		-0.1801	
2.3	**P**	**NP**	2.3190	0.49	2.3098	0.87	2.3167	0.88
		P	2.3098		2.3122		2.3181	
	BD	**NP**	0.0042	0.38	0.0056	0.43	0.0026	0.44
		P	0.0018		0.0090		0.0016	
	W	**NP**	0.0314	0.27	0.0235	0.14	0.0207	0.29
		P	0.0185		0.0484		0.0126	
	A	**NP**	0.0001	0.36	0.0001	0.87	0.0001	0.63
		P	0.0000		0.0002		0.0000	
	AS	**NP**	-0.1752	0.17	-0.4286	0.78	-0.1361	0.61
		P	0.0040		-0.4131		-0.2500	
2.35 - 2.37	**P**	**NP**	2.3660	0.92	2.3650	0.98	2.3675	0.99
		P	2.3654		2.3675		2.3648	
	BD	**NP**	0.0087	0.12	0.0082	0.58	0.0047	0.06
		P	0.0126		0.0046		0.0011	
	W	**NP**	0.0398	0.36	0.0398	0.76	0.0338	0.23
		P	0.0530		0.0406		0.0188	
	A	**NP**	0.0002	0.17	0.0003	0.95	0.0001	0.07
		P	0.0002		0.0001		0.0000	
	AS	**NP**	0.2351	0.87	-0.0452	0.11	0.0614	0.69
		P	0.0000		-0.1667		0.0000	
2.4	**P**	**NP**	2.4139	0.58	2.4027	0.87	2.3192	0.34
		P	2.4135		2.4033		2.4052	
	BD	**NP**	0.0073	0.74	0.0079	0.85	0.0144	0.52
		P	0.0088		0.0105		0.0181	
	W	**NP**	0.0338	0.62	0.0348	0.27	0.0482	0.24
		P	0.0300		0.0453		0.0548	
	A	**NP**	0.0001	0.85	0.0002	x	0.0003	x
		P	0.0002		0.0002		0.0003	
	AS	**NP**	0.1030	0.58	-0.1878	0.67	0.5009	0.52
		P	0.4509		-0.1667		0.7143	
2.45	**P**	**NP**	2.4387	0.74	2.4346	0.78	2.4387	0.76
		P	2.4387		2.4387		2.4546	
	BD	**NP**	0.0083	0.62	0.0000	0.08	0.0032	0.11
		P	0.0059		0.0008		0.0046	
	W	**NP**	0.0474	x	0.0000	0.07	0.0474	0.85
		P	0.0474		0.0474		0.0472	
	A	**NP**	0.0002	0.81	0.0001	x	0.0001	x
		P	0.0001		0.0001		0.0001	
	AS	**NP**	-1.0000	0.12	1.0000	x	-1.0000	x
		P	-0.6789		1.0000		-1.0000	

Appendix 7

Examples of the differences identified in the 2nd derivative spectra between polluted and non-polluted sites at the locations 10 and 24 in the analysis of the 2005 and 2008 images (see sections 5.1.1 and 6.1.1. for summary of the results).

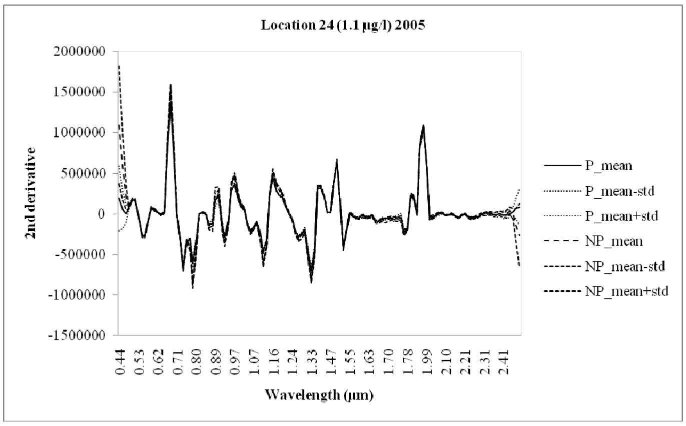
2nd derivative at location 10 from the analysis of the 2005 (right) and 2008 (left) images

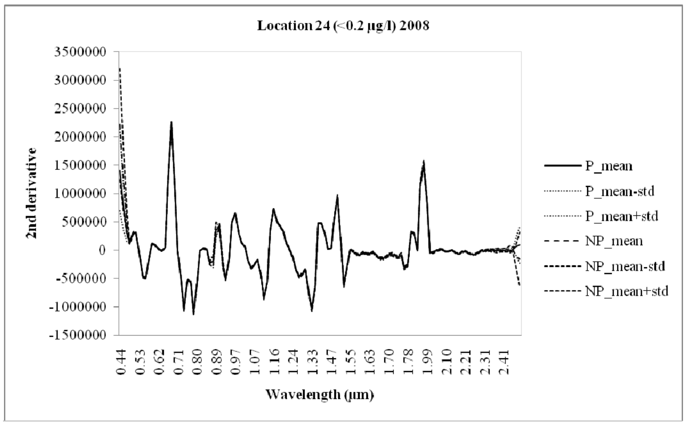




Appendix 7 (continues…)

2nd derivative at location 24 from the analysis of 2005 (right) and 2008 (left) images





Appendix 8

The results from the statistical analysis of the existing vegetation health indices which were calculated for the spectra collected from vegetated areas in the analysis of the 2005 image (see also section 5.1.3.).

	Location (HC concentration in 2003 µg/l)						
Index	**16 (0)**	**11 (0.2)**	**21 (0.89)**	**24 (1.1)**	**8 (320)**	**10 (2000)**	**1 (32000)**
ARI1	x	x	x	< 0.01	x	x	x
ARI2	x	x	x	< 0.01	x	x	< 0.01
BGI1	x	x	x	x	x	x	x
BGI2	x	x	x	x	x	x	x
BRI1	x	x	x	x	x	x	x
BRI2	x	x	x	x	x	x	x
CRI1	x	x	x	x	x	x	x
CRI2	x	x	x	< 0.01	x	x	x
CTR1	x	x	x	x	x	x	x
CTR2	x	x	x	< 0.01	x	x	0.08
CUR	x	x	x	< 0.01	x	x	x
EVI	x	x	x	< 0.01	x	x	0.01
GM	x	x	x	< 0.01	x	x	0.02
MSAVI	x	x	x	< 0.01	x	x	x
LIC1	x	x	x	< 0.01	x	x	x
LIC2	x	x	x	< 0.01	x	x	x
LIC3	x	x	x	< 0.01	x	x	x
MCARI	x	x	x	< 0.01	x	x	0.03
MCARI1	x	x	x	< 0.01	x	x	0.03
MCARI2	x	x	x	< 0.01	x	x	0.09
mNDVI705	x	x	x	< 0.01	x	x	0.02
mSR705	x	x	x	< 0.01	x	x	0.02
MSR	x	x	x	< 0.01	x	x	x
MTVI1	x	x	x	< 0.01	x	x	x
MTVI2	x	x	x	< 0.01	x	x	x
MSI	x	x	x	< 0.01	x	x	< 0.01
NDII	x	x	x	< 0.01	x	x	< 0.01
NDNI	x	x	x	< 0.01	x	x	x
NDVI	x	x	x	< 0.01	x	x	x
NDWI	x	x	x	< 0.01	x	x	< 0.01
NPCI	x	x	x	< 0.01	x	x	x
OSAVI	x	x	x	< 0.01	x	x	0.04
PRI	x	x	x	x	x	x	x
PSRI	x	x	x	< 0.01	x	x	0.07
NDVI705	x	x	x	< 0.01	x	x	x
REP	x	x	x	< 0.01	0.03	< 0.01	x
RGI	x	x	x	< 0.01	x	x	< 0.01
RDVI	x	x	x	< 0.01	x	x	0.01
SR	x	x	x	< 0.01	x	x	x
SRPI	x	x	x	< 0.01	x	x	x
SIPI	x	x	x	x	x	x	x
SG	x	x	x	< 0.01	x	x	x
TCARI	x	x	x	< 0.01	x	x	0.03
TVI	x	x	x	< 0.01	x	x	0.03
VOG1	x	x	x	< 0.01	x	x	x
WBI	x	x	x	< 0.01	x	x	x
ZM	x	x	x	< 0.01	x	x	0.04
ARVI	x	x	x	< 0.01	x	x	0.07
NDLI	x	x	x	< 0.01	x	x	x
CAI	x	x	x	< 0.01	x	x	x

Appendix 9

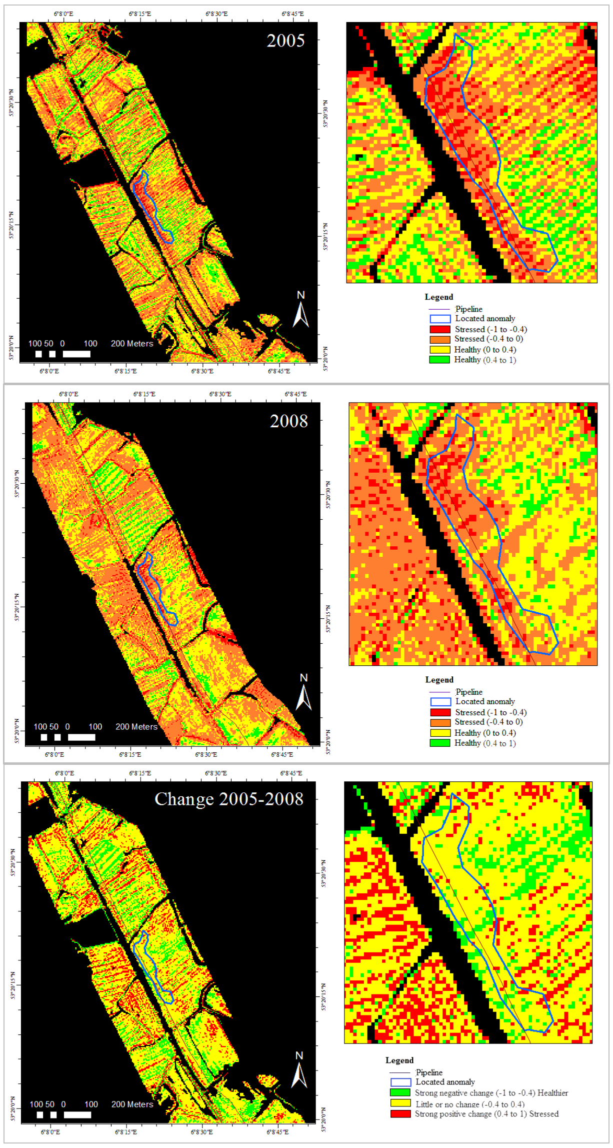
The results from the statistical analysis of the existing vegetation health indices which were calculated for the spectra collected from the 2008 image at some of the sampled locations (see also section 6.1.3).

	Location (HC concentration in 2008 µg/l)						
Index	**16 (0)**	**11 (<0.2)**	**21 (<0.2)**	**24 (<0.2)**	**8 (<0.2)**	**10 (11)**	**1 (<0.2)**
ARI1	x	< 0.01	< 0.01	< 0.01	x	< 0.01	x
BGI2	x	x	< 0.01	x	x	x	x
BRI1	x	x	< 0.01	x	x	x	x
CRI2	x	x	< 0.01	< 0.01	x	x	x
CTR2	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
CUR	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
EVI	x	x	< 0.01	< 0.01	x	< 0.01	< 0.01
GM	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
MSAVI	x	x	< 0.01	< 0.01	x	< 0.01	< 0.01
LIC1	x	x	< 0.01	< 0.01	x	< 0.01	< 0.01
MCARI	x	x	< 0.01	< 0.01	x	< 0.01	< 0.01
MCARI1	x	x	< 0.01	< 0.01	x	< 0.01	< 0.01
MCARI2	x	x	< 0.01	< 0.01	x	< 0.01	< 0.01
mNDVI705	x	x	< 0.01	< 0.01	x	x	< 0.01
mSR705	x	x	< 0.01	< 0.01	x	x	< 0.01
MSR	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
MTVI1	x	x	< 0.01	< 0.01	x	0.01	< 0.01
MTVI2	x	x	< 0.01	< 0.01	x	x	< 0.01
MSI	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
NDII	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
NDNI	x	< 0.01	< 0.01	x	x	< 0.01	< 0.01
NDVI	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
NDWI	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
NPCI	x	x	< 0.01	x	x	x	x
OSAVI	x	< 0.01	< 0.01	< 0.01	x	< 0.01	x
PRI	x	x	< 0.01	< 0.01	x	x	x
PSRI	x	< 0.01	< 0.01	< 0.01	x	0.01	< 0.01
NDVI705	x	x	< 0.01	< 0.01	x	< 0.01	< 0.01
REP	x	x	x	x	x	x	x
RGI	x	x	< 0.01	< 0.01	x	< 0.01	x
RDVI	x	x	< 0.01	< 0.01	x	x	< 0.01
SR	x	< 0.01	< 0.01	< 0.01	x	< 0.01	< 0.01
SRPI	x	x	x	x	x	x	x
SIPI	x	x	< 0.01	x	x	x	x
SG	x	x	< 0.01	x	x	x	x
TCARI	x	x	< 0.01	< 0.01	x	x	< 0.01
TVI	x	x	< 0.01	< 0.01	x	x	< 0.01
VOG1	x	x	< 0.01	< 0.01	x	< 0.01	< 0.01
WBI	x	x	< 0.01	< 0.01	x	x	< 0.01
ZM	x	x	< 0.01	< 0.01	x	x	< 0.01
ARVI	x	x	< 0.01	< 0.01	x	x	< 0.01
NDLI	x	x	< 0.01	x	x	x	< 0.01
CAI	x	x	< 0.01	< 0.01	x	< 0.01	x

Appendix 10

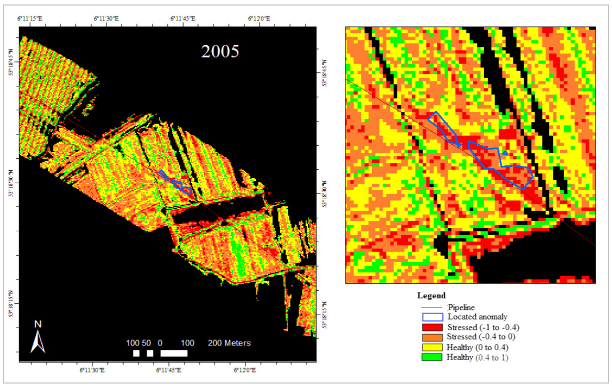
Other three anomalies located by van der Werff et al. (2008) with the 2005 image previously, in which the vegetation was found to be stressed in both of the years in the current study. The stress seemed to become less strong in the 2008 image, indicating decrease of HC pollution on the site. The maps for the other two anomalies are shown in the next two pages. For summary of the results see section 6.1.4 and discussion in section 6.2.2)

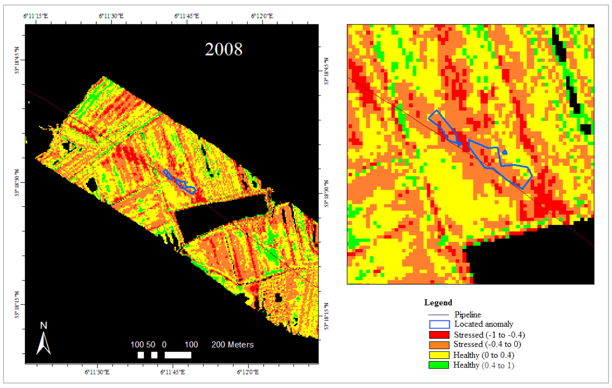
Anomaly 1

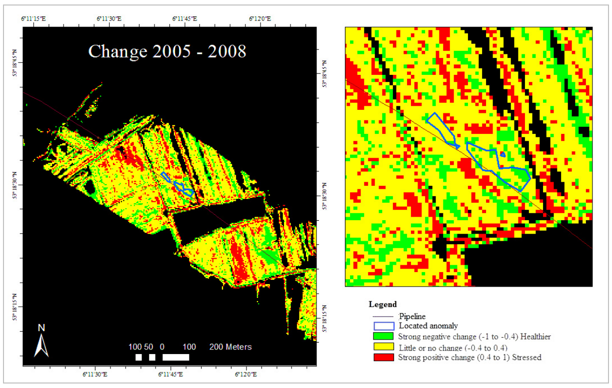


Appendix 10 (continues…)

Anomaly 2

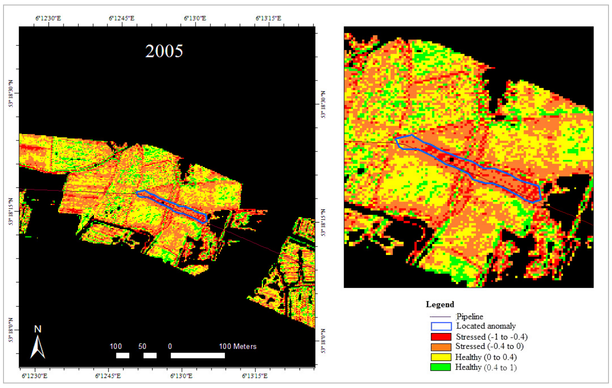


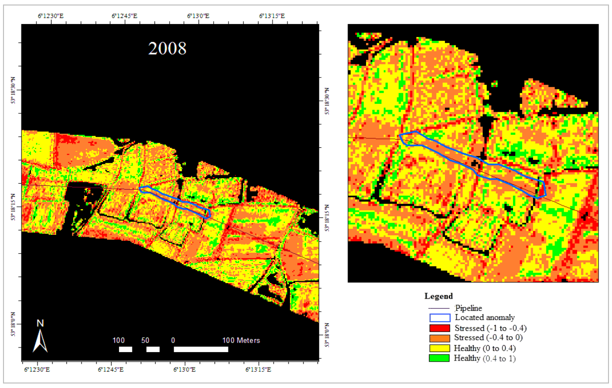


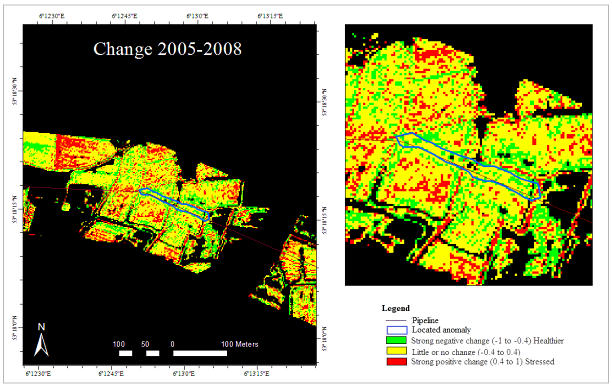


Appendix 10 (continues…)

Anomaly 3







Appendix 11

Confusion matrix calculated for the normalized 2005 REP- image. Negative values in the image indicated vegetation stress (class 1=polluted) and positive values healthy (class 2=non-polluted) vegetation in relation to the surrounding field (see section 3.3.8.) This matrix was calculated with the additional drilling data over vegetated areas from 2002/2003 (see details in Appendix 2). See discussion in section 6.2.2.

2005		**Ground**		**Total**
		1	**0**	
Map	**1**	6	5	11
	0	3	8	11
Total		9	13	22
Overall				63.6
Kappa				0.27
Class	User	Producer		
1	54.5	66.6		
0	72.7	61.5		

Appendix 12

Confusion matrices calculated for the change detection map in section 3.3.8 when investigating HC related vegetation stress over the whole study area with two normalized REP-images. The index-threshold-values for defining the change were varied (for further details, see section 3.3.8). The summary of the results can be found in section 6.1.4 and discussion in section 6.2.2.

Confusion matrix A		**Index Threshold-values**	**Definition**	
		-0.1 to 0.1 -1 < -0.1 or 0.1 < 1	0= little or no change 1=change	
		Ground truth		**Total**
		0	**1**	
Map	**0**	4	3	7
	1	6	4	10
Total		10	7	17
Overall (%)				47.0
Overall Kappa				- 0.03
Class	**User (%)**	**Producer (%)**		
1	57.1	40.0		
2	40.0	57.1		

Confusion matrix B		**Index Threshold-values**	**Definition**	
		-0.2 to 0.2 -1 < -0.2 or 0.2 < 1	0= little or no change 1=change	
		Ground truth		**Total**
		0	**1**	
Map	**0**	4	3	7
	1	6	4	10
Total		10	7	17
Overall (%)				47.0
Overall Kappa				-0.03
Class	**User (%)**	**Producer (%)**		
1	57.1	40.0		
2	40.0	57.1		

Appendix 12 (continues…)

Confusion matrix C		**Threshold-values**	**Definition**	
		0.3 to 0.3 1 < -0.3 or 0.3 < 1	0= little or no change 1=change	
		Ground truth		**Total**
		0	**1**	
Map	**0**	5	3	8
	1	5	4	9
Total		10	7	17
Overall (%)				52.9
Overall Kappa				0.07
Class	**User (%)**	**Producer (%)**		
0	62.5	50		
1	44.4	57.1		

Confusion matrix D		**Index Threshold-values**	**Definition**	
		-0.4 to 0.4 -1 < -0.4 or 0.4 < 1	0= little or no change 1=change	
		Ground truth		**Total**
		0	**1**	
Map	**0**	7	4	11
	1	3	3	6
Total		10	7	17
Overall (%)				58.8
Overall Kappa				0.13
Class	**User (%)**	**Producer (%)**		
1	63.6	50.0		
2	70.0	42.9		

Confusion matrix E		**Threshold-values**	**Definition**	
		0.5 to 0.5 1 < -0.5 or 0.5 < 1	0= little or no change 1=change	
		Ground truth		**Total**
		0	**1**	
Map	**0**	7	5	12
	1	3	2	5
Total		10	7	17
Overall (%)				52.9
Overall Kappa				-0.01
Class	**User (%)**	**Producer (%)**		
0	58.3	70.0		
1	40.0	28.6		

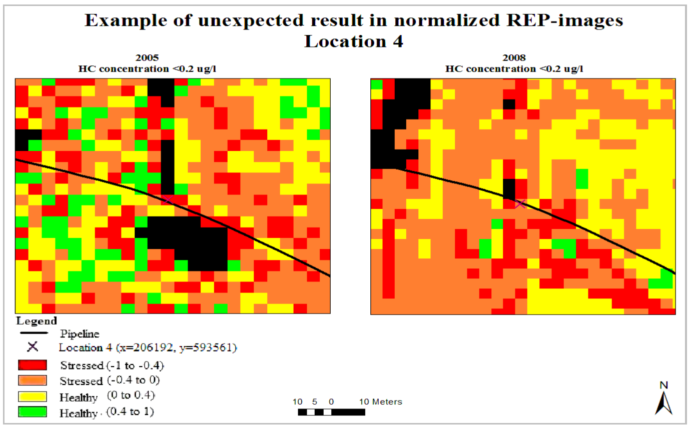
Confusion matrix F		**Index Threshold-values**	**Definition**	
		-0.6 to 0.6 -1 < -0.6 or 0.6 < 1	0= little or no change 1=change	
		Ground truth		**Total**
		0	**1**	
Map	**0**	8	6	14
	1	2	1	3
Total		10	7	17
Overall (%)				52.9
Overall Kappa				-0.06
Class	**User (%)**	**Producer (%)**		
0	57.1	80.0		
1	33.3	14.3		

Appendix 13

Interpretation of the normalized REP-images and the change detection map for the known locations in which remarks has been given for the result, see discussion in section 6.2.2.

Point ID	HC 2003 (µg/l)	HC 2008 (µg/l)	2005 pixel value	2008 pixel value	2005-2008 pixel value	Expected result for 2005/2008	Map result for 2005/2008	Expected YES/NO	Remarks / Explanation
0	< 0.2	< 0.2	– 0.36	0.02	– 0.38	+ +	- +	NO	On the edge
1	32 000	< 0.2	NaN	0.10	NaN	- +	x +	YES	Remediated before 2008
13	0	3.8	0.08	0.26	– 0.18	+ +	+ +	YES	Close to edge, remediation actions taken on the sites before 2008 (these three locations were close to each other)
2	24 000	< 0.2	0.02	0.37	– 0.35	- +	+ +	NO	
12	0	< 0.2	0.29	0.12	0.17	+ +	+ +	YES	
23	1900	480	– 0.05	– 0.10	0.05	- -	- -	YES	HC pollution caused stress but location also on the edge
3	1.9	1000	– 0.53	NaN	NaN	+ -	- x	NO	On the edge
4	< 0.2	< 0.2	– 0.55	– 0.25	– 0.30	+ +	- -	NO	On the edge
5	4 700	< 0.2	– 0.31	0.16	– 0.47	- +	- +	YES	vegetation healthier due to HC conc.decrease but also on the edge
6	2 500	< 0.2	0.44	– 0.10	0.54	- +	+ -	NO	On the edge
7	2 500	< 0.2	0.35	0.20	0.15	- +	+ +	NO	On the edge
8	320	< 0.2	0.06	0.04	0.02	- +	+ +	NO	Tracks on the field in 2008
15	0	0.55	0.02	– 0.66	0.68	+ +	+ -	NO	Whole field was stressed in 2008 (these three locations were close to each other)
9	15 000	< 0.2	0.20	– 0.44	0.64	- +	+ -	NO	
14	0	< 0.2	– 0.05	– 0.55	0.50	+ +	- -	NO	
10	2000	11	0.03	– 0.13	0.16	- +	+ -	NO	Tracks on the field in 2008
11	< 0.2	< 0.2	– 0.00	– 0.21	0.21	+ +	- -	NO	Tracks on the field in 2008
28	0	< 0.2	NaN	0.00	NaN	+ +	x +	YES	Location was clean and NOT on the edge
31	0	< 0.2	– 0.13	0.59	– 0.72	+ +	- +	NO	On the edge
30	94 000	< 0.2	NaN	– 0.33	NaN	- +	x -	NO	Whole field was stressed in 2008
29	150	< 0.2	NaN	– 0.32	NaN	- +	x -	NO	Whole field was stressed in 2008
21	0.89	< 0.2	0.61	NaN	NaN	+ +	+ x	YES	Location clean and NOT on the edge
22	< 0.2	< 0.2	– 0.51	NaN	NaN	+ +	- x	NO	On the edge
24	1.1	< 0.2	– 0.29	NaN	NaN	+ +	- x	NO	Tracks on the field in 2005
17	4 400	2	– 0.46	NaN	NaN	- +	- x	YES	HC pollution caused stress but location also on the edge
16	0	< 0.2	0.10	NaN	NaN	+ +	+ x	YES	Location clean and NOT on the edge
26	0	< 0.2	0.26	0.29	– 0.03	+ +	+ +	YES	Location clean but on the edge

+ = Healthy, - = Stressed and x = No data for 2005 and 2008 images, NOTE! The opposite interpretation for the change detection map (i.e. - = Healthy and + = Stressed) see section 3.3.10.



Example of one location (4): *Vegetation stress was found at a clean site from the normalized REP-images in the analysis of the 2005 (left) and 2008 (right) images. This was unexpected as the location was clean in both of the years based on the geochemical data but the location was situated on the edge, which could have affected the accuracy assessment of the change detection image (see discussion in section 6.2.2).*