D6.1 – Algorithm Theoretical Basis Document

Period Number: 2

Due date: 31 January 2009


Date of preparation: December 2008

Date of submission: (SESAM)

Start date of project: 01/02/2007

Duration: 48 months

Antonio J. Plaza

University of Extremadura

Revision [4.4]

Wageningen University (WUR)

<table>
<thead>
<tr>
<th>Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissemination in level</td>
</tr>
<tr>
<td>PU    Public</td>
</tr>
<tr>
<td>PP    Restricted to other programme participants (including the Commission Services)</td>
</tr>
<tr>
<td>RE    Restricted to a group specified by the consortium (including the Commission Services)</td>
</tr>
<tr>
<td>CO    Confidential, only for members of the consortium (including the Commission Services)</td>
</tr>
</tbody>
</table>
## Contributors List

<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silvia Huber, Michael Schaeppman, Jan Clevers, Zbynek Malenovsky Véronique Carrère Paolo Gamba Bogdan Zagajewski</td>
<td></td>
<td>Wageningen University</td>
</tr>
<tr>
<td></td>
<td></td>
<td>University of Nantes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>University of Pavia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>University of Warsaw</td>
</tr>
</tbody>
</table>

## Document Signature Table

### APPROVAL TABLE

<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paolo Gamba</td>
<td>Quality Manager</td>
<td></td>
</tr>
<tr>
<td>Antonio J. Plaza</td>
<td>Project Coordinator</td>
<td></td>
</tr>
</tbody>
</table>

## Distribution List

### Internal Distribution

<table>
<thead>
<tr>
<th>Name</th>
<th>No. Copies</th>
</tr>
</thead>
<tbody>
<tr>
<td>All partners</td>
<td>1</td>
</tr>
</tbody>
</table>

### External Distribution

<table>
<thead>
<tr>
<th>Company</th>
<th>Name</th>
<th>No. Copies</th>
</tr>
</thead>
</table>

## Document Record

<table>
<thead>
<tr>
<th>Issue / Rev.</th>
<th>Class (R = Review / A = Approval)</th>
<th>Date</th>
<th>Reason for Change</th>
<th>Changed Pages/Paragraphs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.9</td>
<td></td>
<td>22/02/2009</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

1 INTRODUCTION .......................................................................................................................... 6
   1.1 SCOPE OF HYPER-I-NET ........................................................................................................... 6
   1.2 SCOPE OF THIS DOCUMENT .................................................................................................... 6
   1.3 DOCUMENT STRUCTURE ........................................................................................................... 7

2 ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD) ........................................................... 9
   2.1 ATMOSPHERE ......................................................................................................................... 9
      2.1.1 Aerosols and air quality ....................................................................................................... 9
      2.1.2 Trace gases ........................................................................................................................ 9
      2.1.3 Cloud properties ............................................................................................................... 10
   2.2 BIOSPHERE ............................................................................................................................. 10
      2.2.1 Pigments ........................................................................................................................... 10
      2.2.2 Non-pigments ................................................................................................................... 26
      2.2.3 Structural characteristics .................................................................................................. 31
      2.2.4 Non-photosynthetic vegetation (NPV) .............................................................................. 36
      2.2.5 Radiation interception (Albedo, fPAR, fAPAR, Light use efficiency, Photosynthesis,
          Fluorescence) ....................................................................................................................... 38
      2.2.6 Biological diversity and invasive species ........................................................................... 40
      2.2.7 Environmental stress ......................................................................................................... 44
   2.3 CRYOSPHERE .......................................................................................................................... 46
      2.3.1 Snow dust content ............................................................................................................. 47
      2.3.2 Snow algae ......................................................................................................................... 47
      2.3.3 Snow albedo ....................................................................................................................... 47
      2.3.4 Snow grain size .................................................................................................................. 48
   2.4 GEOSPHERE ............................................................................................................................. 48
      2.4.1 Mineral exploration – geological mapping ............................................................................ 49
      2.4.2 Environmental geology ...................................................................................................... 59
         2.4.2.1 Acid Mine drainage – Soil contamination ..................................................................... 59
         2.4.2.2 Production of pH maps .................................................................................................. 63
      2.4.3 Geologic hazards ............................................................................................................... 63
         2.4.3.1 Expansive clay soils ....................................................................................................... 64
         2.4.3.2 Oil seeps ......................................................................................................................... 64
      2.4.4 Regolith mapping ............................................................................................................... 67
2.4.4.1 Kaolin disorder ........................................................................................................................................................................68
2.4.4.2 Micas composition and metasomatic hydrothermal alteration .....................................................................................................................68
2.4.4.3 Mapping non magnetic, talc-bearing ultramafic rocks ......................................................................................................................70

2.5 HYDROSPHERE ...............................................................................................................................................................................................71

2.5.1 Bathymetry, bottom types (albedo, sand, silt, macrophytes) and water column optical properties .................................................................................................................................................................................................71
2.5.2 Phytoplankton, suspended sediments, dissolved organic material (DOM), fCover of bottom green macrophytes ..................................................................................................................................................................................72

2.6 PEDOSPHERE .................................................................................................................................................................................................73

2.6.1 Soil degradation (salinity, erosion, deposition) ......................................................................................................................................................73
2.6.1.1 Soil characterization ..............................................................................................................................................................................74
2.6.1.2 Soil salinity ..........................................................................................................................................................................................78
2.6.1.3 Soil erosion and deposition ..............................................................................................................................................................81

2.7 URBAN SPHERE .........................................................................................................................................................................................94

2.7.1 Urban materials and biotopes ..................................................................................................................................................................95
2.7.2 Mapping and cartographic feature extraction .........................................................................................................................................96

REFERENCES .................................................................................................................................................................................................................99
LIST OF TABLES

Table 2-1 Spectral indices developed as chlorophyll indicators................................................................. 11
Table 2-2 Spectral indices developed as carotenoid indicators. ................................................................. 13
Table 2-3 Spectral indices developed as anthocyanin indicators................................................................. 14
Table 2-4 Different methods applied to determine the REIP. ................................................................. 21
Table 2-5 Spectral indices developed as water indicators. ................................................................. 27
Table 2-6 Spectral indices developed as carbon indicators. ................................................................. 27
Table 2-7 Spectral indices developed as nitrogen indicators. ................................................................. 28
Table 2-8 Methods to calculate water thickness .................................................................................. 29
Table 2-9 Spectral indices developed as structural canopy indicators. ............................................... 32
Table 2-10 Indices developed as NPV indicators. ............................................................................... 36
Table 2-11 Indices developed as indicators of radiation interception .................................................. 39
Table 2-12 Indices used in biological diversity and invasive species studies ........................................ 41
Table 2-13 Indices used as stress indicators ................................................................................... 45
1 INTRODUCTION

1.1 SCOPE OF HYPER-I-NET

The Hyperspectral Imaging Network (HYPER-I-NET) is a four-year (2007-2010) FP6 Marie Curie Research Training Network designed to build an interdisciplinary European research community focusing on hyperspectral imaging activities. Imaging spectroscopy, also known as hyperspectral imaging, deals with the measurement, analysis and interpretation of spectra acquired from a given scene (or specific object) at a short, medium or long distance by an airborne or satellite imaging spectroradiometer. In contrast to multispectral sensors, imaging spectroradiometers sample continguously in the visible and infrared part of the electromagnetic spectrum using dozens to hundreds of narrow spectral bands. Using the reflectance information from these narrow spectral bands (usually only few nanometers wide) has opened ground-breaking perspectives in all fields of Earth Science. The core strategy of HYPER-I-NET is to create a powerful interdisciplinary synergy between different domains of expertise within Europe and use it to break new grounds in areas related with hyperspectral imaging. This is achieved by providing a well structured multidisciplinary training programme encompassing all aspects of the hyperspectral data processing chain: sensor design, calibration and flight operation, data collection, processing, interpretation and dissemination.

1.2 SCOPE OF THIS DOCUMENT

Myriads of methods and algorithms are documented in literature to explore imaging spectroscopy data for various applications, ranging from simple two-band ratios to complex numerical inversions of physically-based radiative transfer models. This Algorithm Theoretical Basis Document (ATBD) serves as a guideline for application-oriented users working with hyperspectral data to get an overview on often used methods within thematic domains and to facilitate the selection of a methodological tool according to their needs. The document lists most commonly used methods for hyperspectral data analysis covering applications within the classical spheres of Earth Science: i) Atmosphere, ii) Biosphere, iii) Cryosphere, iv) Geosphere, v) Hydrosphere, vi) Pedosphere and vii) Urban Sphere. This thematic structure was chosen to emphasise applications rather than methodologies. It should encourage new users of hyperspectral data to get a fast overview of commonly used methods in their field of interest.

The ATBD skeleton is based on the classical domains and each domain is further subdivided into main areas of application relevant to imaging spectroscopy (e.g., biosphere is subdivided into pigments, non-
pigments etc.). Finally, most commonly applied methods are listed under each area of application, complemented with a brief introduction, references for further reading and application-oriented examples. Most complete information is provided for spheres covered by expert knowledge of Hyper-I-NET consortium partners (= Biosphere, Geo/Pedosphere and Urban Sphere). Still, for the sake of completeness, all major domains are included into the ATBD and basic information on relevant areas of application for imaging spectroscopy and references for further reading are given.

On purpose, the itemised methods are only briefly discussed in the ATBD, instead the reader is referred to publications and websites for detailed and complete methodological descriptions. Further, many methods and algorithms listed in this document are used in different domains and for various applications. To avoid recurrences, the methods are only introduced once - at their first appearance. Otherwise a cross-reference is used to point to the section in the ATBD where the methodology has been introduced.

All referenced publications in the ATBD are linked to an appended EndNote database, termed Application Library (AL – D6.2). For each Earth Science domain a separate AL was created. The ALs contain regular reference information, an abstract, a link to a PDF of the reference and further particulars on application category, method, object and scale of investigation. This allows the ALs to be browsed according to requirements, for instance by method or application. As the ALs are seen as open schemes, any information can be added or modified at a later stage.

We emphasise that the ATBD provides a starting point in quest of useful methods for hyperspectral data analysis but it is not intended to be exhaustive. The choice of methods for each application basically reflects the publications record within the domain. Via the literature search platform SCOPUS (http://www.scopus.com) publications were sought for all application areas. Papers were selected based on relevance in terms of authorship, number of citations and actuality.

### 1.3 DOCUMENT STRUCTURE
The document includes the following sections:

- **Section 1 - Introduction**
- **Section 2 – Algorithm Theoretical Basis Document (ATBD)**

The present section.
given and some papers listed in which the described method has been applied.
2 ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD)

2.1 ATMOSPHERE

Hyperspectral data provide a way to monitor a complex range of atmospheric phenomena including water vapour throughout most of the atmosphere, trace gases, oxygen, cloud properties, aerosols and ozone. Water vapour is of interest as a critical component of the atmosphere and the primary absorber across the spectral range. Hyperspectral sensors provide core measurements for climate monitoring and prediction needs and are necessary to understand the processes and feedbacks that cannot be observed with long records of broadband multi-spectral data (Green et al., 1998). For more information we refer to Fishman et al. (2008) and Smith et al. (2005).

2.1.1 Aerosols and air quality

Aerosols are solid or liquid particles dispersed in the air, and include dust, soot, sea salt crystals, spores, bacteria, viruses and a vast number of other microscopic particles. Atmospheric aerosols can scatter or absorb solar radiation, which modifies the radiative balance of the atmosphere and consequently influences the climate. Hence, remote sensing of aerosol over the land and ocean is essential to obtain the global aerosol budget and to estimate the contribution of anthropogenic emission and climate forcing (http://www.nerc.ac.uk/research/programmes/appraise/). For further reading we refer to Kaufman et al. (1997; 2002), King et al. (1999), Tanre et al. (1997) and Weinhold (2008).

2.1.2 Trace gases

Most of the earth’s atmosphere is made up of nitrogen (78.1 % by volume) and oxygen (20.9 %). The remaining 1% of the atmospheric gases are known as trace gases because they are present in such small concentrations. Despite their relative scarcity, the most important trace gases are the greenhouse gases, including carbon dioxide, methane, nitrous oxide, ozone and water vapour. Water vapour is one of the most abundant atmospheric constituents and in fact the most important greenhouse gas. More than 99 % of water vapour is located in the troposphere where it significantly contributes to atmospheric chemistry, weather, energy exchanges within the atmosphere and climate. Its large spatial and temporal variability makes water vapour a tracer for tropospheric changes and especially important for global models which aim to predict climate (Noel et al., 2004). Some applications where
Hyperspectral data was used to gain information on trace gases can be found in (Burke et al., 2005), (Spinetti et al., 2008) and (Marion et al., 2004).

### 2.1.3 Cloud properties

Clouds are a major driving force of the climate system, controlling the planetary albedo. They strongly modulate the energy balance of the Earth through absorption and scattering of solar radiation and absorption and emission of terrestrial radiation and they regulate, together with precipitation, the hydrological cycle. Although the importance of clouds is widely recognised, their impact is associated with great uncertainties due to the complexity and space-time variation of cloud phenomena. Inhomogeneous clouds, for instance, have complex effects on fluxes and can be challenging to parameterize hence emphasising the need for a thorough understanding of the radiative effects of inhomogeneous clouds. While early investigations have focused on the total cloud cover by neglecting the effect of cloud type variation, it is now clear that both can equally contribute to determine the cloud effects on climate. Remote sensing can yield cloud properties such as cloud top pressure, effective cloud amount, cloud phase, cloud optical thickness and effective particle size. Some examples of remote sensing of cloud properties can be found in (Bo-Cai et al., 2002; Li et al., 2004).

### 2.2 BIOSPHERE

Key properties of the biosphere, in particular of vegetation canopies, are required as inputs to ecological, hydrological, agro-meteorological and other process models for a better understanding of biogeochemical cycles and the role of the biosphere in climate change and for sustainable ecosystem management. Changes in the structure and function of the Earth’s biosphere can be monitored with satellite and airborne remote sensing systems, mostly using empirical methods such as vegetation indices or by the inversion of physically based canopy reflectance models. Most important applications within the biosphere domain are related to pigments, non-pigments, vegetation structure, radiation interception and biological diversity.

#### 2.2.1 Pigments

Pigments are fundamental determinants of light capture and utilization and they provide protection against the harmful effects of high radiation. Among the most important plant pigments are
chlorophylls, carotenoids and anthocyanins. Chlorophyll is the most abundant pigment present in every plant that performs photosynthesis. Chlorophyll absorbs light most strongly in the blue (450-520 nm) and red (630-690 nm) but a bit less in the green portions of the electromagnetic spectrum. Carotenoids are usually red, orange, or yellow pigments that are naturally occurring in chromoplasts of plants. Carotenoids serve two key roles in plants: they absorb light energy for use in photosynthesis and they protect chlorophyll from photo damage. Anthocyanins are water-soluble flavonoid pigments that may appear from red to blue. They occur in all tissues of higher plants, providing colour in leaves, stems, roots, flowers and fruits, though not always in sufficient quantities to be noticeable. Anthocyanins are most visible in the petals of flowers, where they may make up as much as 30 per cent of the dry weight of the tissue. Different pigment levels can reveal conditions of stress, photosynthetic activity or biological diversity (Ustin et al., 2004).

A. Vegetation Indices (VIs)

Most commonly used vegetation indices are based on discrete red and NIR wavebands, because vegetation exhibits unique reflectance properties in these spectral regions. Generally, indices are divided into ratio and orthogonal indices. Whereas ratio indices are calculated independently of soil reflectance properties, the orthogonal indices refer to a base line specific to the soil background. More recently, indices have emerged that can be considered hybrid versions of the classic ratio and orthogonal indices.

<table>
<thead>
<tr>
<th>Spectral index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modified Simple Ratio</td>
<td>BmSR</td>
<td>( \frac{(BR_{\lambda_2} - BR_{\lambda_4})}{(BR_{\lambda_2} - BR_{\lambda_5})} )</td>
<td>(Le Maire et al., 2004)</td>
</tr>
<tr>
<td>Derivatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorophyll Index</td>
<td>Chgreen</td>
<td>( \frac{R_{\lambda_{541-560}}^{-1} - R_{\lambda_{NNR}}^{-1}}{R_{\lambda_{NNR}}} )</td>
<td>(Gitelson et al., 2006)</td>
</tr>
<tr>
<td>using Green Reflectance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorophyll Index</td>
<td>Chred</td>
<td>( \frac{R_{\lambda_{690-725}}^{-1} - R_{\lambda_{NNR}}^{-1}}{R_{\lambda_{NNR}}} )</td>
<td>(Gitelson et al., 2006)</td>
</tr>
<tr>
<td>using Red Edge Reflectance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double Difference</td>
<td>DD</td>
<td>( (R_{\lambda_{i+5}} - R_{\lambda_{i}}) - (R_{\lambda_{i+5}} - R_{\lambda_{j}}) )</td>
<td>(Le Maire et al., 2004)</td>
</tr>
<tr>
<td>Eucalyptus Pigment Index</td>
<td>EPI</td>
<td>( \alpha \left( \frac{R_{672}}{(R_{550} \times R_{708})^{\beta}} \right) )</td>
<td>(Datt, 1998)</td>
</tr>
<tr>
<td>HYPER-I-NET</td>
<td>Algorithm Theoretical Basis Document</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doc. N.: D6.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issue: 1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date: FEB 2009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page: 12 of 117</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Green NDVI  
\[
g_{\text{NDVI}} = \frac{(R_{750} - R_{\text{GREEN}})}{(R_{750} + R_{\text{GREEN}})}
\]  
(Datt, 1998, 1999; Gitelson et al., 1996)

Red Edge Inflexion Point Position  
\[
R^*(\lambda) = 0
\]  
(Curran et al., 1995; Horler et al., 1983; Kochubey and Kazantsev, 2007; Lichtenthaler, 1987)

Logarithm of Reciprocal Reflectance  
\[
\text{LRR} = \log R_{737}^{-1}
\]  
(Yoder and Pettigrew-Crosby, 1995)

Modified Normalised Difference  
\[
m_{\text{ND705}} = \frac{(R_{750} - R_{705})}{(R_{750} + R_{705} - 2 \times R_{445})}
\]  
(Sims and Gamon, 2002)

Modified Red-Edge NDVI  
\[
m_{\text{NDI}} = \frac{(R_{750} - R_{705})}{(R_{750} + R_{705} - 2 \times R_{445})}
\]  
(Sims and Gamon, 2002)

Red-Edge NDVI  
\[
m_{\text{NDVI}} = \frac{(R_{750} - R_{705})}{(R_{750} + R_{705})}
\]  
(Datt, 1999; Gamon and Surfus, 1999; Gitelson and Merzlyak, 1994; Sims and Gamon, 2002)

Modified Red-Edge Ratio  
\[
m_{\text{SR}} = \frac{(R_{750} - R_{445})}{(R_{750} - R_{445})}
\]  
(Sims and Gamon, 2002)

Modified Simple Ratio  
\[
m_{\text{SR705}} = \frac{(R_{750} - R_{445})}{(R_{750} - R_{445})}
\]  
(Sims and Gamon, 2002)

Pigment Specific Normalized Difference  
\[a) (R_{800} - R_{675})/(R_{800} + R_{675})
\]  
(Blackburn, 1998)

\[b) (R_{800} - R_{650})/(R_{800} + R_{650})
\]  
(Blackburn, 1998, 1999; Sims and Gamon, 2002)

Pigment Specific Simple Ratio  
\[a) R_{800}/R_{675}
\]  
(Blackburn, 1999; Chappelle et al., 1992)

\[b) R_{800}/R_{650}
\]  
(Blackburn, 1999; Chappelle et al., 1992)

Ratio Analysis of Reflectance Spectra  
\[a) R_{675}/R_{700}
\]  
(Blackburn, 1999; Chappelle et al., 1992)

\[b) R_{675}/(R_{650} \times R_{700})
\]  
(Blackburn, 1999; Chappelle et al., 1992)

Reciprocal Reflectance  
\[
R_{700}^{-1}
\]  
(Gitelson et al., 1999)

Summed Reflectance Index  
\[
s_{700} = \int_{700}^{750} (R\lambda/R_{555} - 1)d\lambda
\]  
(Gitelson and Merzlyak, 1994)
\[ s_2 = \int_{700}^{750} (R\lambda/R705 - 1) d\lambda \]

In this table \( R\lambda \) = reflectance at wavelength \( \lambda \); \( B\lambda \) = first derivative of reflectance at wavelength \( \lambda \); \( R_{\text{GREEN/NIR}} \) = reflectance at a green and NIR wavelength, resp., and \( \alpha, \beta \) = coefficients of power curve.

**Examples** and additional indices can be found in the following literature:

- (Blackburn, 2007b) Hyperspectral remote sensing of plant pigments.
- (Gitelson et al., 2006) Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves.
- (Le Maire et al., 2004) Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements.
- (Gitelson et al., 2003) Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves.
- (Haboudane et al., 2002) Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture.
- (Zarco-Tejada et al., 2001) Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data.
- (Gitelson and Merzlyak, 1994) Quantitative estimation of chlorophyll-a using reflectance spectra - experiments with autumn chestnut and maple leaves.

**Table 2-2** Spectral indices developed as carotenoid indicators.

<table>
<thead>
<tr>
<th>Spectral index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carotenoid Reflectance Index</td>
<td>CRI</td>
<td>a) ( R_{510}^{-1} - R_{550}^{-1} )</td>
<td>(Gitelson et al., 2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) ( R_{510}^{-1} - R_{700}^{-1} )</td>
<td></td>
</tr>
<tr>
<td>Eucalyptus Pigment Index</td>
<td>EPI</td>
<td>( \alpha(R672/(R550 \times R708))^{\beta} )</td>
<td>(Datt, 1998)</td>
</tr>
<tr>
<td>Modified Carotenoid Reflectance Index</td>
<td>mCRI</td>
<td>a) ( (R_{510-520}^{-1} - R_{560-570}^{-1}) \times R_{\text{NIR}} )</td>
<td>(Gitelson et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) ( (R_{510-520}^{-1} - R_{690-710}^{-1}) \times R_{\text{NIR}} )</td>
<td></td>
</tr>
<tr>
<td>Photosynthetic Reflectance Index</td>
<td>PRI</td>
<td>( (R531 - R570)/(R531 + R570) )</td>
<td>(Gamon et al., 1990)</td>
</tr>
<tr>
<td>Pigment Specific</td>
<td>PSSR</td>
<td>( R800 / R500 )</td>
<td>(Blackburn, 1998)</td>
</tr>
</tbody>
</table>

© HYPER-I-NET February 09
Simple Ratio

Ratio Analysis of Reflectance Spectra

Structure Insensitive Pigment Index

In this table \( R_\lambda \) = reflectance at wavelength \( \lambda \) and \( \alpha, \beta \) = coefficients of power curve.

<table>
<thead>
<tr>
<th>Spectral index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthocyanin Reflectance Index</td>
<td>ARI</td>
<td>( R_{510}^{-1} - R_{700}^{-1} )</td>
<td>(Gitelson et al., 2001)</td>
</tr>
<tr>
<td>Modified Anthocyanin Reflectance Index</td>
<td>mARI</td>
<td>((R_{530-570}^{-1} - R_{690-710}^{-1}) \times R_{NIR})</td>
<td>(Gitelson et al., 2006)</td>
</tr>
<tr>
<td>Red:Green Ratio</td>
<td>RGR</td>
<td>(R_{RED} / R_{GREEN})</td>
<td>(Gamon and Surfus, 1999; Sims and Gamon, 2002)</td>
</tr>
</tbody>
</table>

In this table \( R_\lambda \) = reflectance at wavelength \( \lambda \) and \( R_{\text{GREEN/RED/NIR}} \) = reflectance at a green, red and NIR wavelength, respectively.

### B. Radiative Transfer (RT) models

A number of physically-based models, which account for the interactions of incident radiation with the Earth surface, have been developed. Radiative transfer (RT) models have been used in both forward and inverse mode. In forward mode, RT model simulation allows for the validation and intercomparison of different RT model implementations (Myneni et al., 1995) and sensitivity studies of canopy parameters relative to diverse observation specifications, for an improved understanding of the remote sensing signals and an optimized instrument design of future Earth Observation systems (Bacour et al., 2002). For the retrieval of vegetation properties, RT models need to be inverted with remotely sensed data as input. For a successful inversion, one has choose an appropriate and properly validated RT model, which matches spatial scale and correctly represents the radiative transfer of the observed target (Pinty and Verstraete, 1992). For the inversion process, a wide range of minimization techniques have been used: classical iterative optimization, simulated annealing, genetic algorithms, look up tables, Monte-Carlo Markov Chains (MCMC) and Generalized Likelihood Uncertainty Estimation (GLUE). However, classical iterative optimization techniques (OPT) and look up tables (LUT) have been the most widely used (Liang, 2004).
Models are traditionally being developed at the leaf and canopy level for the RT within vegetation. Leaf RT models physically simulate directional-hemispherical reflectance and transmittance of plant leaves, which can be used by canopy level RT models to compute phase functions for multiple scattering. This model coupling represents a way how RT models can be inverted to retrieve information on the leaf anatomy or biochemical constituents (Malenovsky, 2006a).

Canopy RT models can be classified either based on their dimensionality or based on the RT solution. In terms of dimensionality, there are two types of models: i) one-dimensional (1D) models that require homogeneity in the vertical dimensions (e.g., grassland, agro-ecosystem), such as turbid medium models and ii) three-dimensional (3D) models that handle heterogeneity and discontinuity in both horizontal and vertical dimensions (e.g., forest ecosystems), such as geometric optical or hybrid models. Hybrid models have been most recently developed. They combine features from turbid medium and geometric models: a forest stand, for instance, is modeled as geometric objects (tree crowns) with a given tree distribution pattern and, differently to the geometric-optical models, an internal architectural structure in the tree crowns. The internal structure of crowns is believed to be a significant factor in determining the directional reflectance behavior of a canopy (Chen and Leblanc, 1997) and thus the internal structure of the tree crowns in these models can range from a turbid medium to some level of grouped architecture (Kuusk and Nilson, 2000; Li et al., 1995) with mathematical complexity and computation time increasing simultaneously with the degree of grouping.

Based on the RT solution, models can be divided into the following six categories: i) two/three/four-flux approximate solutions (Kubelka-Munk theory), ii) numeric solutions (Discrete ordinates, Gauss-Siedel algorithm), iii) ray tracing (Monte Carlo simulation), iv) method of radiosity (realistic scene rendering), v) stochastic solution (mean intensity - first moment of radiation field) and vi) recollision probability (spectral invariant).

To benchmark so many RT models designed by different authors for different purposes, the radiation transfer model intercomparison (RAMI) activity was initiated. As an open-access, on-going activity, RAMI operates in successive phases, each one aiming at re-assessing the capability, performance and agreement of the latest generation of RT models (http://rami-benchmark.jrc.ec.europa.eu/HTML/Home.php).

RT model inversion techniques stand for a more generic physical approach to quantify vegetation properties from hyperspectral data than statistical approaches, based on empirical calibrations. Generally, all leaf and canopy parameters used for the parameterisation of a RT model that have a significant impact on the plant’s spectral signature (e.g., specific absorption feature) are retrievable by inversion of a RT model. Because RT models are widely used for chlorophyll retrieval, most details are given in this section of the ATBD document.
An overview of the theory of radiative transfer is available in (Liang, 2004). Detailed information about mathematical canopy RT modelling can be found in (Myneni et al., 1995) and the textbooks of (Myneni and Ross, 1991) or (Asrar, 1989).

I. Leaf-level models

i. PROSPECT (Leaf Optical Properties Spectra)

PROSPECT is a radiative transfer model for individual leaves that simulates leaf reflectance and leaf transmittance as a function of leaf biochemical and structural properties. It is based on the generalized plate model (Allen et al., 1969), but considers a compact theoretical plant leaf as a pile of transparent plates with rough plane parallel surfaces separated by layers of air. PROSPECT allows computation of the 400-2500 nm reflectance and transmittance spectra of very different leaves using four or five input variables: leaf mesophyll structural parameter N, dry matter content, water and chlorophyll content, and in the latest version also content of carotenoids. The PROSPECT outputs can be used directly as inputs into canopy models (e.g., SAIL, DART) (Jacquemoud et al., 1995). For downloading the PROSPECT code, visit http://teledetection.ipgp.jussieu.fr/prosail/.

Examples

- (Moorthy et al., 2008) Estimating chlorophyll concentration in conifer needles with hyperspectral data: An assessment at the needle and canopy level.
- (Malenovsky et al., 2006) Applicability of the PROSPECT model for Norway spruce needles.
- (Le Maire et al., 2004) Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements.
- (Jacquemoud et al., 1996) Estimating leaf biochemistry using the PROSPECT leaf optical properties model.

ii. LIBERTY (Leaf Incorporating Biochemistry Exhibiting Reflectance and Transmittance Yields)

LIBERTY is a radiative transfer model developed to simulate coniferous needle optical properties. The cellular structure of needles are represented by powders (grains with inter granular air cavities) in LIBERTY following Melamed's theory of the interaction of light with powders (Melamed, 1963). LIBERTY is a general purpose model for predicting the reflectance and transmittance spectra of a coniferous needle, or stack of coniferous needles in the visible and near infrared wavelengths (400 - 2500 nm). As
needles are treated as an aggregation of cells, with multiple radiation scattering between cells, output spectra are a function of three structural parameters and the combined absorption coefficients of needle biochemical constituents. LIBERTY can be used together with the models of vegetation canopy or even ecosystem (Dawson et al., 1998). For more information and to download the source code, visit http://www.geos.ed.ac.uk/homes/s0455489/LIBERTY.html.

Examples

- (Moorthy et al., 2008) Estimating chlorophyll concentration in conifer needles with hyperspectral data: An assessment at the needle and canopy level.
- (Coops and Stone, 2005) A comparison of field-based and modelled reflectance spectra from damaged Pinus radiata foliage.

iii. RAYTRAN (Ray tracing model)

Among various approaches, only ray tracing techniques can account for the complexity of internal leaf structure as it appears in a photomicrograph. They require a detailed description of individual cells and their unique arrangement inside leaf tissues. The optical constants of leaf materials (cell walls, cytoplasm, pigments, air cavities, etc.) also have to be defined. Using the laws of reflection, refraction, and absorption, it is then possible to simulate the propagation of individual photons incident on the leaf surface. Once a sufficient number of rays have been simulated, statistically-valid estimates of the radiation transfer in a leaf may be deduced. This ray-tracing code allows the explicit representation of radiation transfer in arbitrarily complex scenes. It implements a Monte Carlo approach where the behaviour of millions of individual rays is followed as they travel through the computer simulated leaf, but also simulation of photon transfer through a vegetation canopy is possible. This model implements the most detailed and most faithful simulations of radiation transfer, but it is rather computationally expensive. It has the capability to run in parallel on a cluster of machines using the Message Passing Interface (MPI). Due to its modular nature, arbitrarily complex canopy scenes may be entered, and further measurement types may always be added. RAYTRAN is currently not distributed publicly (Govaerts and Verstraete, 1998). More information about RAYTRAN can be found at http://yves-govaerts.com/1/rtm_ray.php.

Examples

- (Ustin et al., 2001) Simulation of photon transport in a three-dimensional leaf: implications for photosynthesis.
iv. Other leaf RT models

- LEAFMOD – Leaf Experimental Absorptivity Feasibility Model; (Ganapol et al., 1998).
- SLOPE – Stochastic model for leaf optical properties; (Maier et al., 1999).

II. Canopy-level models

i. 4SAIL2 (SAIL, SAILH, 4SAIL - Scattering by Arbitrary Inclined Leaves)

4SAIL2 is a surface reflectance model that was evolved from the GeoSAIL model (Verhoef and Bach, 2003) through an extension with a non-lambertian soil BRDF model and the additional consideration of vegetation with ground or crown coverage lower than 1. These extensions allow the more realistic simulations of directional acquisitions especially in forest environment. 4SAIL2 follows a 4-stream concept, which divides the modelled fluxes in their direct and diffuse, upward and downward contributions. The input parameters to 4SAIL2 describe structural and physiological information on the vegetation, soil properties and the observation geometry. 4SAIL2 incorporates a sub-model for the soil reflectance, which varies with moisture.

The canopy model is a two-layer version of the SAIL model, which incorporates crown clumping effects and this way accommodates the modeling of forest stands (Verhoef and Bach, 2007). The canopy is modelled in 2 layers to mimic the vertical leaf colour gradient often seen in agricultural canopies. The structural properties in both layers (leaf angle distribution and leaf size) are assumed to be identical in both layers, but the LAI of green and brown leaves may differ. The spectral information on the leaf optical properties (i.e., spectral reflectance and transmittance of green and brown leaves) is often calculated using the PROSPECT model (Verhoef, 1984; Verhoef and Bach, 2003, 2007). The RAMI benchmark is available at [http://rami-benchmark.jrc.ec.europa.eu/HTML/RAMI3/MODELS/4SAIL2/4SAIL2.php](http://rami-benchmark.jrc.ec.europa.eu/HTML/RAMI3/MODELS/4SAIL2/4SAIL2.php). The 4SAIL code can be downloaded at [http://teledetection.ipgp.jussieu.fr/prosail/](http://teledetection.ipgp.jussieu.fr/prosail/).

Examples

- (Moorthy et al., 2008) Estimating chlorophyll concentration in conifer needles with hyperspectral data: An assessment at the needle and canopy level.
- (Zarco-Tejada et al., 2001) Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data.
ii. **DART (Discrete Anisotropic Radiative Transfer)**

The DART model simulates radiative transfer of complex 3-D scenes, such as urban or natural landscapes possibly with topography and atmosphere for any wavelength in the optical domain. It is equipped with innovative multispectral approaches as ray tracing, exact kernel and discrete ordinate techniques. The present DART version has two functioning modes: i) Mode (R) simulates the reflectance of direct sun and/or atmosphere radiation and ii) Mode (T) simulates scene thermal emission, possibly with sun emission (e.g., for simulating the 3-4 μm domain). In (R) mode, DART can use a Monte Carlo module to simulate very accurately multiple scattering for reference purposes. DART can simulates any landscape as a 3-D matrix of cells that contain turbid material and triangles or parallelograms. "Turbid" material is used for simulating vegetation canopies (e.g., tree crowns, grass or agricultural crops) and the atmosphere. Triangles and parallelograms are used for simulating translucent and opaque surfaces that make up topography, urban elements and 3-D vegetation structures (e.g., trunks and branches of trees). DART can use structural and spectral data bases (e.g., atmosphere, vegetation, soil). A Graphic User Interface (GUI) is used to input and display computer scene parameters and simulations. DART is freely available for scientific activities (Gastellu-Etchegorry et al., 1996; Gastellu-Etchegorry et al., 2004). To access the model code and for more information we refer to [http://www.cesbio.ups-tlse.fr/us/dart.htm](http://www.cesbio.ups-tlse.fr/us/dart.htm) and for the RAMI benchmark results to [http://rami-benchmark.jrc.ec.europa.eu/HTML/RAMI3/MODELS/DART/DART.php](http://rami-benchmark.jrc.ec.europa.eu/HTML/RAMI3/MODELS/DART/DART.php).

**Examples**

- (Malenovsky, 2006b) A new hyperspectral index for chlorophyll estimation of a forest canopy: area under curve normalized to maximal band depth between 650-725 nm.

iii. **FLIGHT (Forest Light Interaction Model)**

The FLIGHT model is based on Monte Carlo simulation of photon transport (Disney et al., 2000; North, 1996) and has been developed to simulate the observed reflectance and LiDAR response of three-dimensional vegetation canopies. Evaluation of bidirectional reflectance is achieved by simulation of the photon free-path within a canopy representation, and simulation of the chain of scattering events incurred by a photon in its path from the source to the receiver or to its absorption. Multiple scattering between different foliage elements and with the ground is thus modelled. LiDAR waveform response is simulated by including a time variable in the calculation. Foliage is approximated by structural
parameters of area density, angular distribution, and size, and optical properties of reflectance and transmittance. Leaves are approximated as bi-Lambertian scatterers, and the angular distribution of incident diffuse light may be modelled explicitly (North, 1996). For more information about the FLIGHT model, also visit http://rami-benchmark.jrc.ec.europa.eu/HTML/RAMI3/MODELS/FLIGHT/FLIGHT.php.

Examples

- (Prieto-Blanco et al., 2009) Satellite-driven modelling of Net Primary Productivity (NPP): Theoretical analysis.
- (Kempeneers et al., 2008) Model inversion for chlorophyll estimation in open canopies from hyperspectral imagery.

iv. **SPRINT3 (Spreading of Photons for Radiation Interception)**

The SPRINT model is universal in the sense that it can emulate turbid medium, geometrical, hybrid and computer simulation models. A part of the model is the generation of the kilometer-level scenes, which can contain various kinds of objects. It takes only a few minutes to calculate BRDF on a normal computer and this independently of the complexity of the scene. The basic model consists of several modules which are described separately in (Thompson and Goel, 1998) and (Goel and Thompson, 2000). The RAMI results of SPRINT3 are accessible at http://rami-benchmark.jrc.ec.europa.eu/HTML/RAMI3/MODELS/Sprint3/Sprint3.php.

Examples

- (Zarco-Tejada et al., 2004) Needle chlorophyll content estimation through model inversion using hyperspectral data from boreal conifer forest canopies.

v. **Other canopy RT models:**

- DISORD – DIScrete ORDinates method; (Myneni and Asrar, 1993; Myneni et al., 1992).
- RGM – Radiosity-Graphics combined Method; (Qin and Gerstl, 2000).
- SMRT – Stochastic Mixture Radiative Transfer model, (Shabanov et al., 2007; Shabanov et al., 2000).
- PARAS – Simple semi-physical PARAMeterization model; (Rautiainen and Stenberg, 2005).

C. **Red-edge inflection point (REIP)**
Clevers (1999) showed that imaging spectroscopy might provide additional information at the red-edge region, not covered by the information derived from a combination of a NIR and a VIS spectral band. Collins (1978) and Horler et al. (1983) were among the first researchers to point out the importance of the red-NIR wavelength transition for vegetation studies. At red wavelengths, reflectance is low due to absorbance by chlorophyll pigments while in NIR wavelengths reflectance is high due to scattering inside the leaf and multiple reflections inside the canopy, resulting in a steep rise in reflectance between 670 and 780 nm. Both the position and the slope of this red-edge change under stress conditions, resulting in a shift of the slope towards shorter wavelengths (e.g., Horler et al., 1983). The red-edge inflection point (REIP) is defined as the position of the inflection point of the red-NIR slope. This REIP shift due to stress conditions can be caused both by a decrease in leaf chlorophyll concentration and by a decrease in LAI (see Clevers and Jongschaap, 2001). These are the main variables determining the REIP. The REIP can be studied by plotting dR/dλ, the first derivative of reflectance (R) with respect to wavelength (λ), as a function of λ. Alternatively, in many studies simple functions have been fitted to the reflectance spectrum in the red-edge region, and subsequently the wavelength belonging to the maximum slope has been extracted from such an analytical expression. This can then be used for estimating leaf chlorophyll content (Baranoski and Rokne, 2005).

The inflection point is defined as the maximum slope (maximum first derivative) of the reflectance spectrum in the red-edge region. Derivative spectra may be calculated, but accurate determination of the REIP requires a large number of spectral measurements in narrow bands. However, the calculated REIP will depend on the spectral sampling interval applied and the REIP will always be one of the wavelengths or the mid-value of the sampling interval employed by the remote sensing sensor used. To overcome this dependency on the sampling interval high-order curve fitting techniques have been employed to fit a continuous function to the derivative spectrum ((Horler et al., 1983), (Demetriadessshah et al., 1990)). However, existing curve fitting techniques are very complex. (Dawson and Curran, 1998) presented a technique based upon a three-point Lagrangian interpolation technique for locating the REIP in spectra that have been sampled coarsely. It fits a second-order polynomial curve to three spectral bands, from which the REIP is determined.

Although an increasing number of airborne and spaceborne imaging spectrometers have become available, their spectral resolution is not fine enough for an accurate determination of the REIP using derivative spectra. Therefore, fitting a mathematical function to a few measurements in the red-edge region is often applied to estimate the REIP. The above-mentioned Lagrangian technique of Dawson and Curran (1998) is an example.

**Table 2-4** Different methods applied to determine the REIP.
### Example

- (Le Maire et al., 2004) Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements.

- (Sims and Gamon, 2002) Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages.

- (Richardson et al., 2002) An evaluation of noninvasive methods to estimate foliar chlorophyll content.

- (Zarco-Tejada et al., 2001) Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data.

- (Filella and Penuelas, 1994) The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status.

### D. Derivative analysis

Derivative analysis can be an effective tool to analyze hyperspectral data with a different emphasis than traditional remote sensing algorithms. Treating hyperspectral data as truly continuous allows access to information that is often suppressed by the standard analysis methods. Derivatives of second order or higher should be relatively insensitive to variations in illumination intensity whether caused by changes in sun angle, cloud cover, or topography. At the spectral sampling interval typical of hyperspectral systems, derivatives should also be relatively less sensitive to the spectral variations of sunlight and skylight (Tsai and Philpot, 1998). In most studies first- and second order derivatives have been applied.
A first derivative spectrum displays the variation with wavelength in the slope of the original reflectance spectrum and likewise, a second derivative calculates the slope of the first derivative (Blackburn, 2007b). However, derivatives are notoriously sensitive to noise and consequently smoothing or otherwise minimizing random noise is a major issue. Among the various methods for smoothing spectral data, algorithms based on least-square fits are most commonly selected, e.g., the Savitzky-Golay smoothing filter. For more details we refer to Tsai and Philpot (1998) and Dixit and Ram (1985).

**Examples**

- (Guan et al., 2005) Algorithms for the estimation of the concentrations of chlorophyll a and carotenoids in rice leaves from airborne hyperspectral data.
- (Curran et al., 2001) Estimating the foliar biochemical concentration of leaves with reflectance spectrometry testing the Kokaly and Clark methodologies.
- (Blackburn, 1999) Relationships between spectral reflectance and pigment concentrations in stacks of deciduous broadleaves.
- (Yoder and Pettigrew-Crosby, 1995) Predicting Nitrogen and Chlorophyll Content and Concentrations from Reflectance Spectra (400-2500 nm) at Leaf and Canopy Scales.
- (Curran et al., 1992) Reflectance Spectroscopy of Fresh Whole Leaves for the Estimation of Chemical Concentration.

**E. Continuum removal analysis**

Continuum removal is a normalization technique aimed at enhancing the spectral features of interest while minimizing extraneous factors, such as atmospheric absorptions, anisotropic effects or soil background effects. It allows comparison of individual absorption features from a common baseline (Kokaly and Clark, 1999). The observed spectral continuum is considered an estimate of the other absorptions present in the spectrum, not including the one of interest (Clark and Roush, 1984b). The continuum is a convex hull fitted over a spectrum. The hull can be fitted either over the whole spectrum or over specific absorption features of interest by using straight-line segments that connect local spectral maxima (endpoints of continuum lines). Originally, continuum removal was successfully used in the field of geology (Clark and Roush, 1984b) until Kokaly and Clark (1999) demonstrated its use for foliar biochemistry retrieval. The maximum band depth, the area under the continuum, and the band depth normalized to the area (Curran et al., 2001) have been used thus far for estimating foliar biochemicals like chlorophyll.
Examples

- (Mutanga and Skidmore, 2007) Red edge shift and biochemical content in grass canopies.
- (Noomen et al., 2006) Continuum removed band depth analysis for detecting the effects of natural gas, methane and ethane on maize reflectance.
- (Kokaly et al., 2003) Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data.

F. Artificial Neural Network (ANN)

Artificial neural network (ANN) is an effective information processing technique for solving nonlinear problems that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this method is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working coevally to solve specific problems. ANNs learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. A trained network (sets of patterns and trained neurons) could be successfully used for classification of trained objects located on unrecognised areas with very high accuracy. More detailed information is given in Fausett (1994), Grossberg (1988) and Atkinson and Tatnall (1997).

Examples

- (Shi and Sun, 2007) Leaf chlorophyll and leaf water estimation with artificial neural networks (ANN) from simulated PROSPECT reflectance data.
- (Huang et al., 2004) Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis.

G. Wavelet decomposition

Wavelets are mathematical functions that decompose a signal into component sub-signals and then study each component with a resolution matched to its scale. Wavelets have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. When applied to reflectance spectra, a model can be established between wavelet coefficients that are assigned to the component sub-signals and the pigment constituents of interest. In the context of the remote sensing of foliar chlorophyll, wavelet analysis has the potential to capture
much more of the information contained with reflectance spectra than previous analytical approaches which have tended to focus on using a small number of optimal wavebands while discarding the majority of the spectrum (Blackburn, 2007a). Wavelet analysis (WA) was developed independently in several scientific fields but interchanges between them during the last decade have led to a wide range of applications of this signal processing technique (Blackburn, 2007a; Graps, 1995). For an extended discussion on wavelet analysis we refer to Graps (1995) or http://www.amara.com/IEEEwave/IEEEwavelet.html, Meyer and Ryan (1993) and Walnut (2002).

**Examples**

- (Blackburn and Ferwerda, 2008) Retrieval of chlorophyll concentration from leaf reflectance spectra using wavelet analysis.
- (Blackburn, 2007a) Wavelet decomposition of hyperspectral data: a novel approach to quantifying pigment concentrations in vegetation.

### H. Regression analysis

Regression analysis is a statistical tool for the investigation of relationships between variables. In this sense it is not a stand-alone tool to directly analyse imaging spectrometer data but rather links measurements obtained from different sources. Multiple linear regression (more than one explanatory variable) is widely used to develop statistical models between plant biochemical and structural properties and hyperspectral measurements. The statistical approach assumes that a foliar spectrum is the sum of each chemical constituent, weighted by its concentration. A calibration equation between the chemical of interest (e.g., chlorophyll concentration) and the (transformed) spectral signature at different wavelengths (explanatory variables) is developed. Regression analysis is extensively explained in (Draper and Smith, 1966). Also in multiple linear regression the number of explanatory variables can not be arbitrarily high. Especially when using high-dimensional hyperspectral signatures to develop regression models, the number of bands used in the model must be reduced to a reasonable number and be in due proportion to the number of observations. This has to be done to avoid data overfitting, to optimize the accuracy and predictive power (generalization) of a model and to remove irrelevant and redundant data to reduce computational costs (Hruschka, 1987). The selection of predictors (feature subset selection or variable selection) can be made arbitrarily, automatically, by following physiological principles, or by following specific shrinkage rules (Dash and Liu, 1997).

**Examples**
2.2.2 Non-pigments

The most important non-pigment biochemical components of green foliage are nitrogen, carbon and water. Proteins are the major nitrogen bearing leaf constituents, typically holding 70–80 % of all nitrogen. An additional 5–10 % of nitrogen is allocated to chlorophyll and lipoproteins. Even though nitrogen contributes only very little to leaf dry weight (<4 %), numerous studies have described foliar nitrogen to be closely related to ecosystem functions such as photosynthesis, ecosystem productivity, soil respiration and hence the decay rate of leaf litter but also to soil carbon-to-nitrogen (C:N) ratios and nitrate production. Therefore, estimates of canopy nitrogen content are essential for a broad range of environmental applications including ecosystem modelling. Important nitrogen (protein) absorption features occur between 910 and 2350 nm, with two major features at 1510 and 2180 nm (Curran, 1989).

Carbon is allocated mainly to cellulose (up to 65 %) and lignin (up to 20 %). Cellulose is the structural component of the primary cell wall of green plants and lignin fills the spaces in the cell wall between cellulose and the other components. Foliar carbon affects leaf decomposition rates, soil C:N ratios and soil respiration. The three major absorption features for carbon occur at 1690, 1780 and 2100 nm (Curran, 1989).

Water content accounts for 40 to 80 % of the fresh weight of green leaves. Leaf water is an important factor regulating photosynthesis, canopy temperature and moisture stress. Lack of water restricts the transpiration, which induces closure of stomata and results in reduced evaporation from the leaf surface and thus decreases photosynthesis. Liquid water absorbs solar radiation strongly in a series of absorption features in the near and shortwave infrared spectral region, whereas the primary and secondary effects of water content on leaf reflectance are greatest in spectral bands centred at 1450, 1940 and 2500 nm (Carter, 1991). However, at 970 and 1200 nm less pronounced water absorption features occur, that are still clearly observable (Danson et al., 1992). The dominant features of water absorption, however, obscure biochemical features related to lignin and other carbon constituents.
Empirical methods and model inversions are widely used to obtain information on water content of vegetation at the leaf and canopy level. Often the same methods are used for both leaf and canopy level, deriving respectively water concentration (per unit leaf area) and water content (per unit ground surface). But rarely are indices developed as carbon or nitrogen indicators and model inversions are only partly capable to capture information on carbon and nitrogen content (Ustin et al., 1998).

### A. Vegetation Indices (VIs)

**Table 2-5** Spectral indices developed as water indicators.

<table>
<thead>
<tr>
<th>Spectral Index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moisture Stress Index</td>
<td>MSI</td>
<td>$\frac{R_{1599}}{R_{819}}$</td>
<td>(Hunt and Rock, 1989)</td>
</tr>
<tr>
<td>Normalized Difference Infrared Index</td>
<td>NDII</td>
<td>$(R_{819} - R_{1649})/(R_{819} + R_{1649})$</td>
<td>(Hardisky et al., 1983)</td>
</tr>
<tr>
<td>Normalized Difference Vegetation Index</td>
<td>NDVI</td>
<td>$(R_{895} - R_{675})/(R_{895} + R_{675})$</td>
<td>(Penuelas et al., 1993)</td>
</tr>
<tr>
<td>Normalized Difference Water Index</td>
<td>NDWI</td>
<td>$(R_{857} - R_{1241})/(R_{857} + R_{1241})$</td>
<td>(Gao, 1996)</td>
</tr>
<tr>
<td>Simple Ratio Water Index</td>
<td>SRWI</td>
<td>$R_{858,5}/R_{1240}$</td>
<td>(Zarco-Tejada and Ustin, 2001)</td>
</tr>
<tr>
<td>Water Absorption Index</td>
<td>WAI</td>
<td>$R_{895}/R_{972}$</td>
<td>(Penuelas et al., 1993)</td>
</tr>
<tr>
<td>Water Index</td>
<td>WI</td>
<td>$R_{900}/R_{970}$</td>
<td>(Penuelas et al., 1997)</td>
</tr>
</tbody>
</table>

In the stable $R_\lambda = \text{reflectance at wavelength } \lambda$.

**Examples**

- (Serrano et al., 2000) Deriving water content of chaparral vegetation from AVIRIS data.

**Table 2-6** Spectral indices developed as carbon indicators.

<table>
<thead>
<tr>
<th>Spectral Index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellulose Absorption Index</td>
<td>CAI</td>
<td>$0.5(R_{200} + R_{2200}) - R_{2100}$</td>
<td>(Daughtry, 2001; Daughtry et al., 1996)</td>
</tr>
</tbody>
</table>
Examples

- (Cheng et al., 2008) Water content estimation from hyperspectral images and MODIS indexes in Southeastern Arizona.
- (Chen et al., 2005) Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands.
- (McMorrow et al., 2004) Hyperspectral indices for characterizing upland peat composition.
- (Nagler et al., 2003) Cellulose absorption index (CAI) to quantify mixed soil-plant litter scenes.
- (Serrano et al., 2002) Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals.

<table>
<thead>
<tr>
<th>Spectral Index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized Ratio Index</td>
<td>NRI</td>
<td>( \frac{R_{ij} - R_{ij}}{R_{ij} + R_{ij}} )</td>
<td>(Sims and Gamon, 2002)</td>
</tr>
<tr>
<td>Normalised Difference Red Edge</td>
<td>NDRE</td>
<td>( \frac{(R_{790} - R_{720})}{(R_{790} + R_{720})} )</td>
<td>(Barnes et al., 2000)</td>
</tr>
</tbody>
</table>

### B. Radiative transfer (RT) models

For more details see

In this table \( R_{\lambda} = \) reflectance at wavelength \( \lambda \) and \( R_{\text{GREEN/RED/NIR}} = \) reflectance at a green, red and NIR wavelength, respectively.

**Radiative Transfer (RT) models** under the pigments section.

#### i. Prospect-SAIL(H)

Examples

- (Zarco-Tejada et al., 2003) Water content estimation in vegetation with MODIS reflectance data and model inversion methods.
- (Ceccato et al., 2001) Detecting vegetation leaf water content using reflectance in the optical domain.
(Gond et al., 1999) Seasonal variations in leaf area index, leaf chlorophyll, and water content; scaling-up to estimate fAPAR and carbon balance in a multilayer, multispecies temperate forest.

C. Equivalent water thickness (EWT)

Narrow-band spectroscopy measurements like the Equivalent Water Thickness (EWT) or the Water Thickness (WT) (Green et al., 1993; Roberts et al., 1998a) are obtained by modelling atmospheric water vapour and liquid water in imaging spectrometer data. Equivalent Water Thickness is a measure of the thickness of water that would be required to mimic water absorption features in vegetation spectra. Green and co-workers (1993) developed spectrum-matching techniques for retrieving apparent surface reflectance, mapping column water vapor and liquid water from hyperspectral data using a modified version of Modtran radiative transfer code. In order to separate water vapor in the atmosphere from liquid water in the landscape, a simple model was incorporated for the expression of liquid water in a reflectance spectrum. This model makes the assumption that the depth of the liquid water band across the 865 to 1035 nm region can be approximated using Beer-Lambert’s law for exponential extinction in an absorbing or scattering medium. Based on this model, the depth of the water band will vary as a function of the strength of the absorber (described by the absorption coefficient for liquid water) and the pathlength of light within an absorbing/scattering element (Roberts et al., 1998a).

Table 2-8 Methods to calculate water thickness.

<table>
<thead>
<tr>
<th>Method</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equivalent Water Thickness</td>
<td>EWT</td>
<td>$R_{867} \sim R_{1049}$</td>
<td>(Gao and Goetz, 1995; Green et al., 1993)</td>
</tr>
<tr>
<td>Water Thickness</td>
<td>WT</td>
<td>$R_{867} \sim R_{1088}$</td>
<td>(Roberts et al., 1998a)</td>
</tr>
</tbody>
</table>

D. Regression models

For methodological details see Regression under the pigments section.

Examples

- (Martin et al., 2008) A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems.
- (Ferwerda et al., 2006) A bootstrap procedure to select hyperspectral wavebands related to tannin content.
• (Coops et al., 2003) Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data.
• (Smith et al., 2002) Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen.
• (Serrano et al., 2002) Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals.
• (Fourty and Baret, 1998) On spectral estimates of fresh leaf biochemistry.
• (Martin and Aber, 1997) High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes.

E. Artificial neural networks (ANN)

For methodological details see Artificial Neural Network (ANN) under the pigments section.

Examples

• (Shi and Sun, 2007) Estimating leaf biochemical information from leaf reflectance spectrum using artificial neural network.
• (Mutanga and Skidmore, 2004) Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa.

F. Derivative analysis

Danson et al. (1992) showed that the first derivative of the reflectance spectrum corresponding to the slopes of the absorption features at 970 and 1200 nm provides better correlations with leaf water content than those obtained from the direct correlation with reflectance. (Rollin and Milton, 1998) found moderate correlations between the first derivative at the left slope of both absorption features and canopy water content (CWC) for a grassland site in the UK. Recently, (Clevers et al., 2008) applied derivatives in a preliminary study at the field and airborne level. These studies showed that spectral derivatives at the slopes of the 970 nm and 1200 nm absorption feature have good potential as predictors of CWC.

For methodological details, the reader is referred to Derivative analysis in the pigments section.

Examples
- (Clevers et al., 2008) Using spectral information from the NIR water absorption features for the retrieval of canopy water content.
- (Serrano et al., 2002) Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals.
- (Johnson and Billow, 1996) Spectrometric estimation of total nitrogen concentration in Douglas-fir foliage.

G. Continuum removal analysis

For methodological details see Continuum removal analysis in the pigments section.

Examples

- (Huber et al., 2008) Estimating foliar biochemistry from hyperspectral data in mixed forest canopy.
- (Mutanga et al., 2005) Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain.
- (Huang et al., 2004) Estimating foliage nitrogen concentration in Eucalypt trees from HYMAP data using continuum removal analysis.

2.2.3 Structural characteristics

Canopy structural characteristics, i.e., the vertical and spatial distribution, the orientation and density of foliage and its supporting structures, determines the pattern of light attenuation and consequently also the distribution of photosynthesis, respiration, transpiration and nutrient cycling in the canopy (Ross, 1981). Most commonly retrieved structural information from remote sensing observations is fractional cover (fcover) and leaf area index (LAI). Both fcover and LAI are key canopy biophysical variables required by many process models to describe energy and mass exchanges in the soil-plant-atmosphere.
system. Generally, structure enhances canopy reflectance in spectral regions where the scatterers are bright (e.g., NIR for green leaves), and enhances canopy absorption in dark regions (e.g., 680 nm for green leaves) (Asner, 1998). Highly sensitive to changes in canopy LAI are the red edge, the 1275–1350 nm and the 1500-1900 nm region, while the VIS and 1900-2450 nm region is insensitive if LAI > 1 (Asner, 1998).

The common and widely used approach has been to develop relationships between ground-measured structural properties and vegetation indices. Consequently, a large number of relationships have been established (Haboudane et al., 2004). Only in recent years studies have focused on inversion of sophisticated radiative transfer models to retrieve canopy structural attributes.

A. Spectral indices

<table>
<thead>
<tr>
<th>Spectral Index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrow Band Normalized Difference Vegetation Index</td>
<td>Narrow Band NDVI</td>
<td>$\frac{R_{\lambda j} - R_{\lambda i}}{R_{\lambda j} + R_{\lambda i}}$</td>
<td>(Thenkabail et al., 2000)</td>
</tr>
<tr>
<td>Narrow Band Transformed Soil Adjusted Vegetation Index</td>
<td>Narrow Band TSAVI</td>
<td>$\frac{a(R_{\text{NIR}} - aR_{\text{RED}} - b)}{(R_{\text{RED}} + aR_{\text{NIR}} - ab)}$</td>
<td>(Baret et al., 1989; Thenkabail et al., 2000)</td>
</tr>
<tr>
<td>Modified Simple Ratio</td>
<td>MSR</td>
<td>$\left( \frac{R_{800}}{R_{670}} - 1 \right) / \sqrt{\left( \frac{R_{800}}{R_{670}} + 1 \right)}$</td>
<td>(Chen, 1996)</td>
</tr>
<tr>
<td>Soil-Adjusted Vegetation Index</td>
<td>SAVI2</td>
<td>$\frac{R_{\text{NIR}}}{R_{\text{RED}}} + b/a$</td>
<td>(Major et al., 1990)</td>
</tr>
<tr>
<td>Modified Soil-Adjusted Vegetation Index</td>
<td>MSAVI</td>
<td>$\frac{1}{2} \left[ 2R_{800} + 1 - \sqrt{(2R_{800} + 1)^2 - 8(R_{800} - R_{670})} \right]$</td>
<td>(Qi et al., 1994)</td>
</tr>
<tr>
<td>Soil and Atmospherically Resistant VI</td>
<td>SARVI</td>
<td>$(1 + L)(R_{800} - R_{\text{vb}})/(R_{800} + R_{\text{vb}} + L)$</td>
<td>(Kaufman and Tanre, 1992)</td>
</tr>
<tr>
<td>Modified Chlorophyll Absorption Ratio Index</td>
<td>MCARI1</td>
<td>$2.5(R_{800} - R_{670}) - 1.3(R_{800} + R_{550})$</td>
<td>(Haboudane et al., 2004)</td>
</tr>
<tr>
<td>Triangular Vegetation Index</td>
<td>TVI</td>
<td>$0.5 \left[ 2(0.5 - 0.750 - 200(R_{670} + R_{550})) \right]$</td>
<td>(Brogue and Leblanc, 2000)</td>
</tr>
</tbody>
</table>
In this table $R_\lambda$ is reflectance at wavelength $\lambda$; $a$, $b$ = intercept and slope of soil lines, and $L$ = constant to minimize the soil effect. For the function $\gamma$ the authors recommend the value of 1.

### Examples

- (Le Maire et al., 2008) Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass.
- (Haboudane et al., 2004) Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture.
- (Broge and Leblanc, 2001) Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density.
- (Thenkabail et al., 2000) Hyperspectral vegetation indices and their relationships with agricultural crop characteristics.

### B. Radiative transfer (RT) models

#### i. PROSPECT-SAIL(H) (also termed PROSAIL)

For methodological details we refer to

In this table $R_\lambda$ = reflectance at wavelength $\lambda$ and $R_{\text{GREEN/RED/NIR}}$ = reflectance at a green, red and NIR wavelength, respectively.

### Radiative Transfer (RT) models in the pigments section.

#### Examples

- (Darvishzadeh et al., 2008) Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland.
- (Bacour et al., 2006) Neural network estimation of LAI, fAPAR, fCover and LAI×Cab, from top of canopy MERIS reflectance data: Principles and validation.
- (Schlerf and Atzberger, 2006) Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data.
- (Jacquemoud et al., 1995) Extraction of vegetation biophysical parameters by inversion of the PROSPECT + SAIL models on sugar beet canopy reflectance data.

### ii. DART

For methodological details we refer to

In this table $R_\lambda$ = reflectance at wavelength $\lambda$ and $R_{\text{GREEN/RED/NIR}}$ = reflectance at a green, red and NIR wavelength, respectively.

#### Radiative Transfer (RT) models

in the pigments section.

#### Examples

- (Duthoit et al., 2008) Assessing the effects of the clumping phenomenon on BRDF of a maize crop based on 3D numerical scenes using DART model.
- (Kimes et al., 2002) Recovery of forest canopy characteristics through inversion of a complex 3D model.

### iii. GeoSAIL

The GeoSail model combines a geometric model that calculates the amount of shadowed and illuminated components in a scene with a turbid media model that calculates the reflectance and transmittance of the tree crowns. It is designed to use canopy component optical properties, tree shape, solar zenith angle, and canopy cover to calculate scene reflectance and the fraction of absorbed or intercepted photosynthetically active radiation (fAPAR and fIPAR) for forest stands. It combines the Scattering from Arbitrarily Inclined Leaves (SAIL) model (Verhoef, 1984) with the Jasinski geometric model (Jasinski and Eagleson, 1989, 1990). The SAIL model provides the within-tree radiative transfer calculations and Jasinski's model combines the SAIL results into a scene reflectance. GeoSail makes several assumptions, which allows the model to be computationally simple, yet provides reasonable descriptions of forest canopy reflectance. In GeoSail, all trees have the same shape and size; trees do not shadow each other; tree crowns do not overlap each other; the size of the trees is small relative to the size of the pixel; the illuminated canopy, illuminated background, and shadows each has a single reflectance; and the model is limited to nadir views. These simplifications impose some limitations on the use of GeoSail; for example, the requirement that trees do not cast shadows on each other means
that GeoSail can underestimate shadow amounts in a scene when canopy cover is high or the solar zenith angle is large (Huemmrich, 2001).

**Examples**

- (Koetz et al., 2004) Radiative transfer modeling within a heterogeneous canopy for estimation of forest fire fuel properties.
- (Hall et al., 1997) Physically based classification and satellite mapping of biophysical characteristics in the southern boreal forest.

**iv. FLIGHT**

For methodological details we refer to

In this table $R_\lambda$ = reflectance at wavelength $\lambda$ and $R_{\text{GREEN/RED/NIR}}$ = reflectance at a green, red and NIR wavelength, respectively.

**Radiative Transfer (RT) models** in the pigments section.

**Examples**

- (Kobayashi et al., 2007) Reflectance seasonality and its relation to the canopy leaf area index in an eastern Siberian larch forest: Multi-satellite data and radiative transfer analyses.
- (Koetz et al., 2004) Radiative transfer modeling within a heterogeneous canopy for estimation of forest fire fuel properties.

**C. Support vector machines (SVM)**

Support vector machines (SVMs) are a set of non-parametric methods used for classification and regression. SVM algorithms were inspired by the statistical learning theory which began in the 1960s to solve pattern recognition problems (Cortes and Vapnik, 1995; Vapnik, 2000). The theoretical development of SVMs has been addresses by Vapnik (2000) and several other papers (Bazi and Melgani, 2006; Burges, 1989; Huang et al., 2002). An extensive description is also given in (Cristianini and Shawe-Taylor, 2000). The basic idea of SVM algorithms is to locate the optimal boundaries between linearly separable classes. If classes are separable in two dimensions they can be split by a line. For linearly non-separable cases, the original data is mapped into higher-dimensional feature space through nonlinear transformations (*kernel* functions). Then, in this new feature space, an optimal linearly separating *hyperplane* is constructed. There are many possible solutions for a separating
Hyperplane but the problem is to find the optimal one (= ill-posed problem). The optimal hyperplane is the one with maximal margin of separation between the two classes (Vapnik, 2000). For SVMs only difficult points close to the decision boundary influence optimality, whereas in linear regression all points are involved. These points that lie closest to the decision surface (hyperplane) are termed "support vectors". They carry all relevant information about the classification problem and they have direct bearing on the optimum location of the decision surface by changing the position of the dividing hyperplane if removed. The number of support vectors is thus small as the decision surface is fitted only to these vectors (Vapnik, 2000). Therefore, SVMs are independent of the dimensionality of feature space, a major advantage of this method (Pal and Mather, 2003).

Examples

- (Durbha et al., 2007) Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer.
- (Camps-Valls et al., 2006) Robust support vector regression for biophysical variable estimation from remotely sensed images.

2.2.4 Non-photosynthetic vegetation (NPV)

Non-photosynthetic vegetation (NPV) such as woody material, leaf and standing litter and dry grass comprise a substantial portion of the canopy in many ecosystems including grasslands, shrublands, savannas and dry woodlands. Monitoring the fractional cover of NPV is important for carbon and water modeling, grazing management, fire risk assessment and erosion control. Estimates of aboveground carbon storage in the form of NPV can improve predictions of ecosystem processes (Asner et al. 1998, 1999). In semiarid shrublands and savannas, the abundance of potential fuel, in the form of NPV, can be used to assess wildfire risk (Asner et al. 1998, Roberts et al. 1998). The spatial patterns of dry plant residues in shrub and grassland ecosystems help to provide indicators of climate change and desertification (Ustin and Costick 1999, Asner and Green 2001). NPV affects the radiation field through its reflectance and transmittance characteristics (Asner, 1998; Ross, 1981). It was shown that stem surfaces significantly contribute to both low and high LAI canopy reflectance spectra with highest variation in the near-infrared and lowest in the visible spectral range (Asner, 1998). Stem effects are enhanced in the NIR because photons penetrate and exit the canopy more effectively in that spectral region, allowing maximum interaction of stem material in the radiation field. In the SWIR, effects of
woody stems are dampened (yet remain significant) by the strong water absorption features present in green foliage (Asner, 1998).

A. Spectral indices

Table 2-10 Indices developed as NPV indicators.

<table>
<thead>
<tr>
<th>Index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized Difference NPV Index</td>
<td>NDNPVI</td>
<td>$\frac{NPV_{\text{back}} - NPV_{\text{nadir}}}{NPV_{\text{back}} + NPV_{\text{nadir}}}$, where 'back' corresponds to the fractions at the -55° viewing angle.</td>
<td>(Verrelst et al., 2007)</td>
</tr>
</tbody>
</table>

B. Spectral mixture analysis (SMA)

Spectral mixture analysis (SMA) exploits the theory that the reflectance spectrum of any pixel can be represented as a mixture of signals contributed by all spectrally pure features, called endmembers, within the instantaneous field of view (IFOV) of the sensor at a given time (Adams et al., 1986; Boardman, 1993). Assuming linear mixing, each endmember weight is directly proportional to the area the pixel containing that endmember. If the spectra of all endmembers in the scene are known, then their abundances within each pixel can be calculated from each pixel's spectrum. Unmixing simply solves a set of $n$ linear equations for each pixel, where $n$ is the number of bands in the image. The unknown variables in these equations are the fractions of each endmember in the pixel. To be able to solve the linear equations for the unknown pixel fractions it is necessary to have more equations than unknowns, which means that more bands than endmembers are needed. The results of spectral unmixing include one abundance image for each endmember. The pixel values in these images indicate the percentage of the pixel made up of that endmember. In addition, an error image is calculated to help evaluate the success of the unmixing analysis (Adams et al., 1986; Boardman, 1993). The simple mixing model concept is limited because it fails to account for the fact that the number of materials within the field of view and the spectral contrast between those materials is variable. For example, a forested ecosystem might be best described as a mixture of green leaves and shade while a shrubland may require four endmembers. Hence, the accuracy of fractions will be highest when the exact number that is required to account for the spectral variability are utilized in the model; too few endmembers will partition the unmodeled endmember into the fractions (creating a fraction error) and increase the RMS;
too many endmembers will make the model sensitive to instrumental noise, atmospheric contamination and natural variability in spectra, resulting in fraction errors. Fraction errors are commonly expressed as negative fractions in models that are unconstrained to force positive fractions. Another disadvantage to the simple mixing model concept is that it fails to account for subtle spectral differences (Roberts et al., 1998b).

As a further refinement of SMA, the multiple endmember spectral mixture analysis (MESMA) minimizes such fraction errors while permitting a large number of endmembers to be modeled across a scene. Further, the number and identity of endmembers can vary on a per pixel basis. Another use of SMA is for deriving the spectral signatures of all endmembers in a scene when the abundances within a pixel are known. An example is the fusion of an image with a high spatial resolution, but a low spectral resolution, with an image with medium spatial resolution, but high spectral resolution as described by (Zurita-Milla et al., 2008). This technique has subsequently been applied to time series of MERIS images for obtaining continuous fields of canopy chlorophyll and of the fAPAR.

Examples

- (Verrelst et al., 2007) Angular unmixing of photosynthetic and non-photosynthetic vegetation within a coniferous forest using CHRIS-PROBA.
- (Jia et al., 2006) Assessing spatial patterns of forest fuel using AVIRIS data.
- (Asner et al., 2003) Desertification in Central Argentina: changes in ecosystem carbon and nitrogen from imaging spectroscopy.
- (Roberts et al., 1998b) Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models.
- (Roberts et al., 1993) Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data.

2.2.5 Radiation interception (Albedo, fPAR, fAPAR, Light use efficiency, Photosynthesis, Fluorescence)

Human activities have changed Earth’s terrestrial surface for instance by forest clearance or farming activities and as a consequence thereof the reflection of sunlight from land surfaces (albedo) has been altered. The impact of albedo alterations on regional and local climate is an active area of research, especially changes in climate in response to changes in vegetation canopies and built structures. These
changes alter the surface heat balance not only by modified surface albedo, but also by altering evaporative heat transfer caused by evapotranspiration from vegetation and changed surface roughness (http://www.eoearth.org/article/Albedo).

Another critical component of climate change research is the estimation and monitoring of plant photosynthesis or gross primary productivity (GPP). The light use efficiency (LUE) approach which expresses GPP as product of the photosynthetically active radiation (PAR), the fraction of PAR being absorbed by the plant canopy (fPAR or fAPAR) and the efficiency \( e \) with which this absorbed PAR can be converted into biomass. fAPAR expresses the canopy’s energy absorption capacity and depends both on canopy structure, leaf and soil optical properties and irradiance conditions. Hence, fAPAR is a key variable in carbon cycle models implementing land surface process schemes.

Hyperspectral remote sensing offers the possibility of sensing the changes in the xanthophyll cycle, which is closely coupled to photosynthesis. Several studies have shown that the photochemical reflectance index (PRI) based on the reflectance at 531 nm is strongly correlated with the LUE over time. A second hyperspectral approach relies on the remote detection of fluorescence, which is directly related to the efficiency of photosynthesis (Grace et al., 2007).

A. Spectral indices

<table>
<thead>
<tr>
<th>Index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photochemical Reflectance Index</td>
<td>PRI</td>
<td>( \frac{R_{570} - R_{531}}{R_{570} + R_{531}} )</td>
<td>(Gamon et al., 1992)</td>
</tr>
</tbody>
</table>

Examples

- (Zagajewski and Jarocinska, 2008) Analysis of plant condition of the Bystrzanka catchment.
- (Rascher et al., 2007) Monitoring spatio-temporal dynamics of photosynthesis with a portable hyperspectral imaging system.
- (Grace et al., 2007) Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence.

B. Radiative transfer (RT) models
The reader is referred to

In this table $R_{\lambda}$ = reflectance at wavelength $\lambda$ and $R_{\text{GREEN/RED/NIR}}$ = reflectance at a green, red and NIR wavelength, respectively.

Radiative Transfer (RT) models in the pigments section for methodological details.

Examples
- (Prieto-Blanco et al., 2009) Satellite-driven modelling of Net Primary Productivity (NPP): Theoretical analysis.
- (Gond et al., 1999) Seasonal variations in leaf area index, leaf chlorophyll, and water content; scaling-up to estimate fAPAR and carbon balance in a multilayer, multispecies temperate forest.

C. FluorMOD Fluorescence leaf-canopy model
The FluorMOD integrated model simulates solar-induced chlorophyll fluorescence in vegetation at leaf and canopy levels using atmospheric inputs and inputs from two sub-component models – the FluorMODleaf for leaf-level fluorescence and the FluorSAIL for the canopy level. These models are linked by FluorMODgui, a graphical user interface which permits the leaf and canopy models to be run either individually or as an integrated unit. The integrated model is applicable to broad-leaf species. Outputs available in the 400-1000 nm spectral range are the apparent leaf spectral reflectance and transmittance, and the canopy apparent reflectance, with and without fluorescence effects (Miller et al., 2005). The models and documentation can be accessed from the web page at http://www.ias.csic.es/fluormod.

Examples
- (Middleton et al., 2008) Comparison of measurements and FluorMOD simulations for solar-induced chlorophyll fluorescence and reflectance of a corn crop under nitrogen treatments.
- (Zarco-Tejada et al., 2002) Vegetation stress detection through chlorophyll $a + b$ estimation and fluorescence effects on hyperspectral imagery.

D. Artificial Neural Network (ANN)
For more information the reader is referred to the Artificial Neural Network (ANN) section in the pigments paragraph.
Examples

- (Bacour et al., 2006) Neural network estimation of LAI, fAPAR, fCover and LAI×Cab, from top of canopy MERIS reflectance data: Principles and validation.

2.2.6 Biological diversity and invasive species

Mapping biological diversity is a high priority for conservation research, management and policy development. Diversity maps set a starting point from which temporal changes in the landscape can be tracked, giving communities feedback regarding how their management decisions are affecting local ecosystems. Imaging spectroscopy has the potential to resolve the spatial and spectral details of vegetation canopy structure and chemistry required to detect canopy-level diversity in complex, layered ecosystems and to provide a biophysically-based approach for mapping biodiversity (Carlson et al., 2007). In general there are two approaches to assess biological diversity: direct remote sensing of species assemblages and communities or an indirect estimation through the use of other environmental variables, including primary productivity, biochemistry, climate and habitat structure (Kalacska et al., 2007; Turner et al., 2003). A new spectranomics approach was recently presented that combines taxonomic, chemical and spectral remote sensing perspectives to map canopy diversity over large geographic regions (Asner and Martin, 2009). For more information on this project, visit http://spectranomics.stanford.edu/ However, successful canopy spectranomics requires the fusion of biogeochemistry, taxonomy, community ecology, and physically based remote sensing (Asner and Martin, 2009).

A second area of recent interest has been the mapping of invasive species from imaging spectroscopy data. Invasive species are a major component of anthropogenic global change and they have large associated environmental, economic, and cultural costs. Over time, the numbers of invasive species and of ecosystems impacted by them have increased. It is essential to know where infestations occur in order to manage them and to understand their causes and their consequences (Andrew and Ustin, 2008). Several studies showed that with spectral information provided by hyperspectral sensors the species-level detection necessary to map invasive weeds is feasible (Asner and Vitousek, 2005; Lawrence et al., 2006; Underwood et al., 2003).

A. Vegetation indices
Table 2-12 Indices used in biological diversity and invasive species studies.

<table>
<thead>
<tr>
<th>Index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canopy Structure Index</td>
<td>CSI</td>
<td>[2sSR - sSR^2 + SWT^2, \text{where:}]</td>
<td>(Sims and Gamon, 2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[sSR = \frac{R_{800}}{R_{680}} - \frac{1}{\left(\frac{R_{800}}{R_{630}} - 1\right)_{\text{max}}}]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[SWI = \frac{R_{900}}{R_{1130}} - \frac{1}{\left(\frac{R_{900}}{R_{1130}} - 1\right)_{\text{max}}}]</td>
<td></td>
</tr>
<tr>
<td>Modified Single Ratio</td>
<td>MSR</td>
<td>[\frac{R_{800}}{R_{680} - 1} \sqrt{\frac{R_{800}}{R_{680}} + 1}]</td>
<td>(Chen, 1996)</td>
</tr>
<tr>
<td>Normalized Difference</td>
<td>ND_{705}</td>
<td>[(R_{750} - R_{705})/(R_{750} + R_{705})]</td>
<td>(Sims and Gamon, 2002)</td>
</tr>
<tr>
<td>Vegetation Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canopy Normalized Difference</td>
<td>ND_{canopy}</td>
<td>[(R_{800} - R_{680})/(R_{800} + R_{680})]</td>
<td>(Sims and Gamon, 2003)</td>
</tr>
<tr>
<td>Vegetation Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canopy Simple Ratio</td>
<td>SR_{canopy}</td>
<td>[R_{800}/R_{680}]</td>
<td>(Sims and Gamon, 2002)</td>
</tr>
<tr>
<td>Simple Ratio</td>
<td>SR</td>
<td>[R_{750}/R_{705}]</td>
<td>(Sims and Gamon, 2002)</td>
</tr>
</tbody>
</table>

Examples

- (Kalacskai et al., 2007) Ecological fingerprinting of ecosystem succession: Estimating secondary tropical dry forest structure and diversity using imaging spectroscopy.
- (Carter et al., 2005) Indicators of plant species richness in AVIRIS spectra of a mesic grassland.
- (Andrew and Ustin, 2008) The role of environmental context in mapping invasive plants with hyperspectral image data.
- (Underwood et al., 2003) Mapping nonnative plants using hyperspectral imagery.

B. Derivative analysis
See **Derivative analysis** in the pigments section for more details.

**Examples**

- (Carlson et al., 2007) Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests.

**C. Principal component analysis (PCA)**

The aim of principal component analysis (PCA) is to condense and reorganise intercorrelated data into a few independent variables, called principal components (Pearson, 1901). It is a vector space transform often used to reduce multidimensional data sets to lower dimensions for analysis which may better capture the essential information. PCA transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. This transformation is a rotation of the original axes to new orientations that are orthogonal to each other and therefore there is no correlation between variables. For more methodological details we refer to Shlens (2005) and Jolliffe (2002).

**Examples**

- (Castro-Esau et al., 2004) Discrimination of lianas and trees with leaf-level hyperspectral data.
- (Underwood et al., 2003) Mapping nonnative plants using hyperspectral imagery.

**D. Spectral angle mapper (SAM)**

The Spectral Angle Mapper (SMA) algorithm determines the similarity between a reference spectrum and image spectrum by calculating the ‘spectral angle’ between them. The smaller the spectral angle, the higher the similarity between pixel and reference spectra. The spectral angle is relatively insensitive to changes in pixel illumination because increasing or decreasing illumination does not change the direction of the vector, only its magnitude (i.e., a darker pixel will plot along the same vector, but closer to the origin). The spectral angle calculation is based on all of the bands in the explored image. In the case of a hyperspectral image, a spectral ‘hyper-angle’ is calculated between each pixel and each
reference (Kruse et al., 1993; Yuhas et al., 1992). It is a very useful tool for the creation of patterns for classification (Zagajewski and Olesiuk 2008).

Examples

- (Hestir et al., 2008) Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem.
- (Lass et al., 2005) A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (Centaurea maculosa) and babysbreath (Gypsophila paniculata) with a hyperspectral sensor.

E. Mixture Tuned Matched Filtering (MTMF)

Mixture Tuned Matched Filtering (MTMF) performs a partial unmixing analysis that finds the sub-pixel abundances of reference spectra (endmembers). The MTMF outputs appear as gray-scale images with values ranging from zero to one (MF-score). These images provide a means of estimating the relative degree of match to the reference spectrum and approximate sub-pixel abundance, where 1.0 is a perfect match with the reference spectrum and 0.0 is no match situation. The higher the MF-score, the higher is the abundance of the material in the pixel. As it stands, the MTMF procedure does not require full knowledge of all endmembers within the image and achieves results according to the selected endmembers (Boardman, 1998; Boardman et al., 1995).

Examples

- (Andrew and Ustin, 2008) The role of environmental context in mapping invasive plants with hyperspectral image data.
- (Glenn et al., 2005) Hyperspectral data processing for repeat detection of small infestations of leafy spurge.

F. Spectral mixture analysis

For methodological details we refer to Spectral mixture analysis (SMA) in the canopy structure section.

Examples
2.2.7 Environmental stress

Biotic and abiotic constraints such as insufficient water or nutrients, adverse climatic conditions, plant diseases, insect damage and extended ozone exposure or higher radiation in upper levels of life habitats, all limit growth of plants at some stage and cause plants to be stressed. Physiological, morphological and anatomical changes take place within plants as a result of stress. In case of restricted transpiration owing to lack of water or a vascular disease, leaf temperatures will increase because of less cooling by transpired water as it evaporates from the leaf surfaces. Leaf color may change as a result of physiological changes or anthropogenic input parameters caused by a water deficiency, a change in nutrient status, biotic or chemical factors (e.g. herbicides, heavy metals, dioxides, hydrocarbons). Plant pathogens and insects may change leaf structure by causing chemical and physical changes within plant cells or by growing on plant surfaces. Morphological changes such as leaf curl or droop may result from the action of any of several stress factors (Jackson, 1986). Ozone can cause a range of adverse effects, including premature leaf aging, early onset of senescence, growth and yield reductions and altered sensitivity to biotic and abiotic stresses (Campbell et al., 2007). As changes in natural and cultivated ecosystems occur, alterations in foliar chemistry and membrane structure result in changes in vegetation spectral signatures so that spectral observations could offer a sensitive, physiologically based indicator of vegetation stress. For instance, an excessive amount of carotenoids is often typical for senescent or stressed plants (e.g., high solar radiation or chemical factors) and altered chlorophyll concentration is denoting for heavy metal contaminated plants: ions of metals replace Mg\(^{2+}\) or block other processes, which in turn decrease the chlorophyll content.

Generally, aging and stress increase reflectance over the visible and shortwave-infrared and decrease it in the near-infrared spectral region. Decline in chlorophyll (chlorosis) increases reflectance across the visible and near-infrared spectrum and shifts the red edge toward shorter wavelengths (blue shift). Measuring such shifts in wavelength position requires instruments with spectrally high-resolution capabilities, because such shifts in wavelengths are often less than 25 nm. Moreover, stress analysis is a multi-factor investigation of complex indicators and different processes involved (photosynthesis, water and temperature management, nutrient migration). Only an integrated approach of integrating remotely sensed information with plant physiological and biometrical parameters can successfully
monitor plant conditions. However, the main problem remains detecting stress as early as possible in order to initiate management practices to minimize stress effects, for instance on the harvestable crop yield (Ustin et al., 2004).

A. Vegetation indices

Table 2-13 Indices used as stress indicators.

<table>
<thead>
<tr>
<th>Index</th>
<th>Acronym</th>
<th>Formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrogen Stress Index</td>
<td>NSI</td>
<td>Based on the modified spectral ratio planar index (mSRPi) and the canopy chlorophyll content index (CCI)</td>
<td>(Barnes et al., 2000; Clarke et al., 2001; Rodriguez et al., 2006)</td>
</tr>
<tr>
<td>Photochemical Reflectance Index</td>
<td>PRI</td>
<td>$\frac{R_{570} - R_{531}}{R_{570} + R_{531}}$</td>
<td>(Gamon et al., 1992)</td>
</tr>
</tbody>
</table>

Examples

- (Suarez et al., 2008) Assessing canopy PRI for water stress detection with diurnal airborne imagery.
- (Tilling et al., 2007) Remote sensing of nitrogen and water stress in wheat.
- (Campbell et al., 2007) Assessment of vegetation stress using reflectance or fluorescence measurements.
- (Asner et al., 2004) Drought stress and carbon uptake in an Amazon forest measured with spaceborne imaging spectroscopy.
- (Campbell et al., 2004) Detection of initial damage in Norway spruce canopies using hyperspectral airborne data.

B. Red edge inflection point (REIP)

The measurement of the red-edge inflection point is one of the fastest ways to determine long-term stresses of plants. Generally, the spectral characteristics in the red-NIR region of stressed plants are significantly shifted to shorter wavelength regions making it a useful tool to observe plant condition. For methodological details see Red-edge inflection point (REIP) in the pigments section.

Examples
- (Clevers et al., 2004) Study of heavy metal contamination in river floodplains using the red-edge position in spectroscopic data.
- (Kooistra et al., 2004) Exploring field vegetation reflectance as an indicator of soil contamination in river floodplains.

2.3 CRYOSPHERE

Snow plays an important interactive role in the Earth's radiation balance, because snow has a higher albedo than any other natural surface. Over 30 per cent of the Earth's land surface is seasonally covered by snow. Snow cover represents a changing atmospheric output resulting from variability in the Earth's climate, and it is also a changing boundary condition in climate models. Thus understanding of global and regional climates and assessment of water resources require that we monitor the temporal and spatial variability of the snow cover over land areas, from the scale of small drainage basins to continents. For further reading we refer to Dozier and Painter (2004), Tedesco and Kokhanovsky (2007) and Aoki et al. (2007).

2.3.1 Snow dust content

Worldwide dust emission events frequently deposit in mountain snow cover. The inference of the radiative forcing of soot and dust in snow has taken on importance in recent years because of the realization of the climate effects that they have (Hansen and Nazarenko, 2004). Dust loading at the snow surface positively forces troposphere temperatures through direct and indirect effects. Absorption by dust in the snow increases near-surface snowpack temperatures, decreasing the column cold content of the snowpack and increasing the energy available for melt. To date, determining the presence of impurities in snow remains an unsolved issue. For further reading we refer to Painter et al. (2007).

2.3.2 Snow algae

Snow algae are cold-tolerant algae and cyanobacteria that grow on snow and ice during alpine and polar summers. Description of spatial distributions of snow algal blooms is important for understanding snow algae's unique life in an extremely cold environment and for determining the effect of algae through the reduction of surface albedo. Microbial processes, including heterotrophy, photosynthesis, and nutrient cycling, occur in melting snow and are important factors in estimating carbon budgets and
CO2 flux. Moreover, variations in snow algal biomass and species composition may reflect regional environmental or climate change (Painter et al., 2001).

### 2.3.3 Snow albedo

Snow-covered surfaces have a very high surface albedo, thereby allowing little energy to be absorbed by the snowpack. As the snowpack ages and/or begins to melt, the snow albedo decreases and more solar energy is absorbed by the snowpack. Therefore, accurate estimation of snow albedo is essential for monitoring the state of the cryosphere. Surface albedo is an important climate parameter, as it influences the amount of solar radiation absorbed by the surface. For snow-covered surfaces, the albedo may be greater than 0.80, thereby allowing very little solar energy to be absorbed by the snowpack. As the snow ages and/or begins to melt, the albedo is reduced considerably, leading to enhanced absorption of solar radiation which further reduces the surface albedo. Consequently, snow melt comprises an unstable, positive feedback component of the earth’s climate system, which amplifies small perturbations to that system (Dozier and Painter, 2004; Stroeve and Nolin, 2002).

### 2.3.4 Snow grain size

Snow grain size is the primary parameter controlling broadband albedo and hence crucial to know for calculating the snowpack’s absorption of solar radiation. Broadband albedo represents the fraction of total solar radiation that is reflected by the surface. Because the rate of grain growth is exponentially proportional to snow temperature, changes in grain size are useful indicators of thermodynamic processes in the snowpack. Changes in snow grain size can help identify ice sheet surface features, such as melt areas, snow dunes and blue ice regions, and often indicate changes in snowpack energy balance (Nolin and Dozier, 2000). Imaging spectrometers provide the spectral leverage necessary to infer measurements of effective grain size to estimate the surface permeability and to infer spectral albedo (Li et al., 2001; Painter et al., 2003). For further reading we refer to Dozier and Painter (2004).

### 2.4 GEOSPHERE

Geologists are generally interested in mapping minerals or rocks for various applications. This means identifying rock types and minerals of interest, mapping mineral concentrations and estimating their contribution at the sub-pixel scale (% coverage) or obtaining information on physical parameters such as grain size and moisture content or on soil quality. Some of these subjects are also relevant to Soil Sciences or Pedology, hence some applications and specific algorithms will be developed in the Pedosphere section of this document.
For most applications, spectral analysis will focus on four main objectives:

- Extraction of information on surface composition which translates into specific absorption features to be able to identify minerals;
- Estimation of concentration through the strength of the absorption features (one has to watch for possible saturation that might affect the relationship);
- Mapping grain size that can modify both the strength and shape of absorption in the case of coarse grains and the general shape of the spectrum because of scattering effects in the case of fine grains;
- Estimation of moisture content using the depth of the specific water absorption features or the evolution of the level of reflectance and general shape of the continuum.

Finally, due to improvements in instrument quality, there is some interest in getting insights on crystallinity of some minerals or on substitutions in the mineral structure which can provide information on regolith transport or source of mineralization.

### 2.4.1 Mineral exploration – geological mapping

Geologic formations are of course composed of minerals so mineral maps can be used to map the extent of geologic formations. They can also be used to subdivide geologic member units if there is a compositional change. Finally, some minerals can be indicative of specific geochemical conditions and therefore be used as markers or tracers for such conditions. Most metallic ore deposit models also include common elements that are relevant to remote sensing studies: tectonic setting, host rocks, source rocks, alteration and structural control.

Many ore deposits are first detected in the field by the recognition of hydrothermally altered host rocks. Alteration halos may be manifested as spectacular color anomalies such as are produced by interaction of acidic fluids with mafic wall rocks, or they might be very subtle and difficult to detect. Zonation in alteration halos, manifested by successions of mineral assemblages or spatial gradations in composition and crystallinity of alteration minerals, reflect chemical and temperature gradients in hydrothermal systems and can provide important information relating to the type and position of the ore body as well as the ore-forming process itself.

For example, kimberlite bodies present a great economic interest as kimberlite dykes can bear diamonds. When weathered, they can also show elevated levels of nickel (Ni), chromium (Cr) and niobium (Nb). Alteration of kimberlite immediately after emplacement is common, resulting in extensive serpentinisation of olivine and monticellite via reaction with meteoric and ground waters. The high CO₂ content of magmatic fluids can result in widespread calcite crystallization during early alteration. Clinopyroxene is altered to chlorite and continued supergene weathering of kimberlite generates a
series of alteration products that include talc, vermiculite, dolomite and magnesium-iron (Mg-Fe) smectite, which under intense or prolonged weathering are reduced to kaolin and Fe oxyhydroxides. Minerals found in altered kimberlite have in common a strong absorption feature in the 2.3-2.33 µm wavelength range (Keeling et al., 2005).

Additionally, altered mineralized systems such as those at Cuprite, Goldfield, Yerrington, and Virginia City, in Nevada, have become standard test sites to evaluate new airborne and satellite sensors and specific algorithms developed for lithologic / mineral mapping (Swayze et al., 1998).

Most approaches entail qualitative distinctions based on overall albedo and spectral characteristics unique to particular rock units. The VNIR bands are useful for mapping color variations in the visible portion of the spectrum, and SWIR bands can be used to map rocks containing carbonates and hydroxyl-bearing minerals. In most cases of geologic applications, the algorithms developed are generally composed of a series of successive steps leading to the product.

### A. USGS Tetracorder

The USGS Tetracorder (Clark et al., 2003) is an expert system approach based on spectroscopic analyses and feature shape comparison between image and library reference spectra. The results are highly dependent on the spectral library that has to be fine-tuned to the study site to get better results. This algorithm follows a spectroscopic approach. Mineral identification is based on position and shape of diagnostic absorption features that are unique to particular materials. The unknown image spectrum is quantitatively compared to a large library of spectra of well-characterized materials, mainly powders of different grain sizes, of different minerals. In order to be faster and efficient, the algorithm focuses on portions of the spectrum that are known to be diagnostic (spectrally active and inactive regions), for example the VNIR for iron oxides or the SWIR2 for clays. The rules are implemented in a decision tree derived from scientific knowledge and expertise.

The first step aims at isolating the features of interest. In order to do so, the continuum is removed (Clark and Roush, 1984a) using the same wavelength interval on each side of the absorption feature to be mapped from both the library reference spectra and the unknown. The continuum here is the global envelope of the spectrum and is modeled by straight lines between high points of the reflectance spectrum. Identifying minerals relies on diagnostic absorption features. But remotely sensed spectra generally correspond to a mixture of minerals and are also influenced by parameters other than composition, such as grain size variations, lighting geometry, topography, surface roughness, etc. These parameters have mostly an influence on the general shape of the spectrum or the reflectance level. Removing the continuum normalizes the spectra and allows focusing on absorption only. This is only done for active regions of the spectrum in order to save time.
The second step is a spectral matching algorithm. Prior to try and match the unknown spectrum to reference spectra from the library, the spectral feature intensity of the unknown spectrum has to be normalized. That is because diagnostic spectral features of materials measured remotely are almost always weaker that those of pure reference materials. This can be due to mixtures or effect of lighting or topography, etc. After normalization using a simple additive constant, it is then possible to derive correlation coefficients between the unknown spectrum and the reference spectra using an iterative least-squares fit. The best fit gives the highest correlation coefficient that will allow a first identification.

The last step adds some additional constraints to take into account the possibility for similar principal features and different materials. Most minerals present ancillary spectral features that can also be diagnostic and help discriminate among them. Additional constraints can be added to take into account parameters that were discarded after continuum removal such as reflectance level or continuum slope.

This algorithm seems rather simple and efficient. However, various effects can introduce some uncertainty in the identification process. For example, the different types of mixture (rocks are mixtures of minerals), if they are not included in the reference library, could lead to misidentifications. Grain size variation can also introduce some changes in the shape and width of absorption features, so variations in grain size have to be included in the reference spectral library for better accuracy.

Examples

- (Dalton et al., 2004) Identification of spectrally similar materials using the USGS Tetracorder algorithm: the calcite-epidote-chlorite problem.
- (Rockwell, 2004) Spectral variations in rocks and soils containing ferric iron hydroxide and (or) sulfate minerals as seen by AVIRIS and laboratory spectroscopy.
- (Swayze et al., 2004) Preliminary report in using imaging spectroscopy to map ultramafic rocks, serpentinites, and tremolite-actinolite-bearing rocks in California.
- (Livo and Clark, 2001) Mapped minerals at Questa, New Mexico, using Airborne Visible-Infrared Imaging Spectrometer (AVIRIS) data.

B. Analytical Imaging and Geophysics (AIG) Algorithm

The AIG algorithm (Kruse, 1998) is available off the shelf (implemented in the commercial software ENVI®, ITT Visual) and allows identifying pure pixels that can be considered as reference endmembers. These spectra can then be used as reference for mineral mapping with different algorithms. This kind of approach is partly statistical and does not require a priori knowledge of the possible minerals present in the area.
The first step is the **Minimum Noise Fraction (MNF) Transform**. This is a modified Principal Components (PC) transform that orders the output components by decreasing signal to noise ratio. First, the noise is estimated in each image band using the spatial variation of brightness values. Then two successive PC transforms are performed: 1) Noise estimates are used to transform the dataset to a coordinate system in which noise is uncorrelated and equal in each component. 2) A standard PC transform is performed. The result is a set of components with increasing noise content which allows separating information from noise and focusing on the most informative components.

The next step is the calculation of the **Pixel Purity Index (PPI)**. The objective here is to locate pure pixels on the image that will correspond to spectra from which the entire image can be modeled. Working in the MNF space and keeping only the most significant (information rich) components, spectral points are projected onto random vectors or skewers, and the extreme points that should correspond to pure pixels are flagged. The process tallies the number of times a pixel is found to be extreme and the pixels with the highest values at the end of the process are considered pure. They are then identified as endmembers.

In order to visualize these endmembers and identify their composition, the n-D visualizer is used either to manually select the extreme points in the n-dimension plot or letting the algorithm automatically select the extreme clusters. These clusters, corresponding to specific pixels, are then used to access the spectral response of these pure pixels. These endmember spectra can then be used for mineral mapping using various algorithms, including MTMF, SAM and others.

**Examples**

- (Kruse et al., 2003) Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping.
- (Kruse, 1998) Advances in hyperspectral remote sensing for geologic mapping and exploration.

**C. Mixture Tuned Matched Filtering (MTMF)**

Automated spectral unmixing is a convex geometry problem. We are given the scatterplot of points. Determining its inherent dimensionality tells us the number of mixing endmembers. There are n+1 endmembers if the data is inherently n-dimensional, assuming the endmembers are spectrally distinct in terms of the observing instrument. Estimating the spectral signatures of the mixing endmembers is
done by finding the *best-fitting* simplex that contains the scatterplot. The vertices of this simplex represent the n+1 mixing endmembers. Estimation of the abundances of each endmember, at each pixel, corresponds to a simple transformation of the data to barycentric coordinates. Since the simplex contains all the data points, the derived fractions will be positive and sum to unity, as desired.

This MTMF algorithm performs unmixing of the image pixels based on user’s chosen endmembers. It is therefore not necessary to find all possible endmembers in the scene. The algorithm maximizes the response of a known endmember while suppressing the response of the background. The result is a grayscale image with values corresponding to the relative degree of match (Boardman, 1993).

**Examples**

- (Kratt et al., 2006) Geothermal exploration with Hymap hyperspectral data at Brady-Desert Peak, Nevada.
- (Ellis and Scott, 2004) Evaluation of hyperspectral remote sensing as a means of environmental monitoring in the St. Austell China clay (kaolin) region, Cornwall, UK.
- (Boardman et al., 1995) Mapping target signatures via partial unmixing of AVIRIS data.

**D. Spectral Angle Mapping (SAM)**

For more details see [Spectral angle mapper (SAM)](a) in a previous section of this document.

**Examples**

- (Crosta et al., 1998) Hydrothermal alteration mapping at Bodie, California, using AVIRIS Hyperspectral Data.
- (van der Meer et al., 1997) Spectral characterization of ophiolite lithologies in the Troodos ophiolite complex of Cyprus and its potential in prospecting for massive sulphide deposits.
- (Boardman and Kruse, 1994) Automated spectral analysis: A geological example using AVIRIS data, northern Grapevine Mountains, Nevada.

**E. Spectral Mixture Analysis (SMA)**
For more details see **Spectral mixture analysis (SMA)** in a previous section.

### Examples

- (Penn and Livo, 2001) Using AVIRIS data to map geologic signatures of Copper Flat porphyry copper deposit, Hillsboro, New Mexico.
- (Kruse, 1996) Geologic mapping using combined analysis of Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and SIR-C/X-SAR data.

### F. Polytopic Vector Analysis (PVA)

Polytopic vector analysis (PVA) is a multivariate method developed in the earth sciences for analysis of mixtures. PVA has been used extensively for analysis of multivariate mixtures in the earth and environmental sciences. (Johnson and Hulen, 2002) describe its development and formalism. PVA is an unsupervised classification method that is also referred to as a *self-training classifier*. With PVA, the analyst does not need to assume the composition of contributing end-member spectra. This is advantageous because it allows feasible end-members to be inferred in areas where few ground-truth data exist. Given samples from a mixed system, PVA is used to resolve three parameters of interest: (1) the number of end-members contributing to the mixture; (2) the spectral composition of each end-member (i.e. the spectral fingerprint) and (3) the relative proportion of each end-member in each sample. PVA includes a Principal Component Analysis step.

#### Example

- (Johnson et al., 2002) Principal components analysis and receptor models in environmental forensics.

### G. Absorption band modeling

The approaches to absorption feature mapping such as the USGS Tetracorder rely on correlations between the parameterization of features observed in remotely acquired data to those in spectral libraries or ground-truth data sets. Therefore, these are excellent methods for applications in which
spectral libraries are representative of the materials in the scene exist. They also include a certain
degree of flexibility to handle noisy data and complexities introduced by mixtures or particle size effects.
However, when presented with mineralogical compositions that are different from those found in
spectral libraries and/or with several minerals in a mixed surface, simple matching of absorption
features can prove to be inadequate. Mineral identification is particularly difficult when overlapping
absorption combine, obscuring the minima of one or more of the diagnostic absorption bands. Models
that account more completely for the processes of absorption by focusing on individual absorption
bands rather than combined absorption features are more easily extended beyond the scope of
preexisting libraries.

**Derivatives**
The locations of individual absorptions can be identified by examining the various higher order
derivatives of spectral signatures. This method assumes that each absorption is symmetric around its
band center, but unlike other approaches, does not require that absorptions have a specific shape.
Band centers are identified where the second derivative of the spectrum is negative, the fourth
derivative is positive, and the fifth derivative is zero. Like any derivative analysis, this method is highly
sensitive to noise. This approach is capable of resolving overlapping band centers separated by as little
as 0.3 to 1.0 of the full width at half maximum (assuming Gaussian shaped absorptions).

**Gaussian absorption model**
Much work has been done to approximate absorption bands with Gaussian distribution of the form

\[ g(x) = s \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right) \]

where \( x \) is the wavenumber (energy), \( s \) the strength or amplitude of the distribution, \( \mu \) the center
(mean) wavenumber, and \( \sigma \) the width. However, simple Gaussian approximation has been shown to be
an inadequate physical model for absorption bands (Sunshine et al., 1990). This is particularly well
illustrated for the absorption bands in the mineral orthopyroxene. The 0.9 \( \mu m \) absorption is due to
electronic transitions from a single site in the orthopyroxene structure (Burns, 1970) and thus should be
modeled by a single Gaussian, if the absorption is in fact Gaussian in form. However, two Gaussian
distributions are required to fit the feature adequately.

The underlying assumption of the Gaussian model is that energy (wavenumber) is a random variable.
However, the center of a band, the location of the maximum probability of absorption, is a function of
the average crystallographic and molecular properties of the material. Similarly, the width of the absorptions is primarily a consequence of perturbations to these average properties. Fundamentally, the energy of absorptions is determined by the average bond length of the atoms defining the absorption site. Thus, to the first order the random variable should be variations in bond length (Sunshine et al., 1990).

Since the average bond length is randomly or Gaussian distributed, the key determining factor for the shape of absorption is how bond length variations are mapped into wavelength or energy. According to the crystal field theory of electronic absorptions (Burns, 1993; Marfunin, 1979), absorption energy (e.g. wavelength) and bond length are related to a power law. Sunshine and co-workers (1990) empirically determined the optimum mapping of the bond length to energy by fitting absorption due to a single absorption, the 0.9 µm orthopyroxene absorption. This analysis revealed that energy and bond length are inversely proportional, suggesting that Coulombic potential energy dominates the absorption site. This inverse mapping of bond length to energy leads to a modified Gaussian distribution, \( m(x) \), of the form

\[
m(x) = s \exp \left[ - \frac{(x^{-\frac{1}{2}} - \mu^{-\frac{1}{2}})^2}{2\sigma^2} \right]
\]

The difference between Gaussian and modified Gaussian distribution is reflected in the symmetry of the band models.

Having formulated a physically based mathematical description for isolated absorption bands, Sunshine and co-workers (1990) were able to model spectra accurately as a series of modified Gaussian distributions superimposed on a continuum. In the Modified Gaussian Model (MGM), the continuum is modeled as a straight line in energy (cm\(^{-1}\)) that sits above the general spectral curve. It should be noted that continua used in most other methods use reflectance maxima to define the continuum. This difference is critically important for the case of overlapping absorptions, where the true reflectance maxima are obscured. Under the MGM each absorption is characterized by three model parameters: a band center (\( \mu \)), a bandwidth (\( \sigma \)) and a band strength (\( s \)), while the continuum is characterized by the slope and offset of a straight line in energy. The model calculations are performed in natural log reflectance and energy, and all parameters, including the continuum, are fit simultaneously using a stochastic least-squares method (Tarantola and Valette, 1982).

For example, for minerals that exhibit systematic changes in vibrational absorptions as a function of composition and/or environment of formation, the MGM may be an effective tool for characterizing these absorptions and relating them to composition and environment of formation.
It should be noted that in contrast to other methods, the MGM neither relies on a library of spectra that may or may not reproduce all permutations and combinations of absorption bands, nor requires knowledge of end-member spectra as is needed for both linear and non-linear mixing. Instead, the MGM derives compositional information directly from the measured spectrum. This is particularly useful in applications where the exact mineralogy, major and minor element chemistry, and physical state (particle size, alteration state, etc.) of materials is unknown or may not exist in library spectra.

Examples

- (Verpoorter et al., 2007) Retrieval of physical properties of mudflat sediments from hyperspectral data using the modified Gaussian model and spectral curve fitting.
- (Combe et al., 2005) Mapping microphytobenthos biomass by non-linear inversion of visible-infrared hyperspectral images.
- (Sunshine, 1994) Inferring the composition of mafic lithologies from multi- and hyperspectral data sets: implications for remote sensing of terrestrial bodies.
- (Sunshine and Pieters, 1993) Estimating modal abundance from the spectra of natural and laboratory pyroxene mixtures using the modified Gaussian model.
- (Sunshine et al., 1993) Reflectance spectra of the Elephant Moraine A79001 meteorite: implications for remote sensing of planetary bodies.
- (Mustard, 1992) Chemical composition of actinolite from reflectance spectra.

H. Nonlinear Mixture Modeling

The fact that at the scale of most observing systems, the pixel contains a mixture of materials is generally treated as a second-order effect. It is accommodated through adding complexity to the approaches, such as mineral combinations or simply considered a consequence of working with natural scenes. The concept of the mixed pixel has certainly been recognized for many years. However, the idea that mixing at the subpixel scale is a natural consequence of earth processes and an inherent feature of remote sensing data sets that can be quantitatively exploited has only evolved over the last two decades.

The basic premise of mixture modeling is that within a given scene, the surface is dominated by a small number of common materials that have relatively constant spectral properties. If most of the spectral variability within the scene is a result of varying proportions of these common components (endmembers), it follows that the spectral variability captured by the remote sensing system can be
modeled by mixtures of these components. It has been documented that the reflectance spectrum of a mixture is a systematic combination of the reflectance spectra of the components in the mixture. The systematics are basically linear if the components are arranged in spatially distinct patterns, analogous to the squares on a checkerboard. In this case, the scattering and absorption of electromagnetic radiation is dominated by a single component on the surface, and thus the spectrum of a mixed pixel is a linear combination of the endmember spectra weighted by the areal coverage of each endmember in the pixel. If, however, the components of interest are in an intimate association, like sand grains of different composition in a beach deposit, the mixing systematics between these different components are nonlinear. The spectral properties of the different endmembers become convolved in this case, because the electromagnetic radiation interacts with more than one endmember as it is multiply scattered in the surface.

The question of whether linear or nonlinear processes dominate the spectral signatures of mixed pixels is still an unresolved question. It probably depends on a number of factors and conditions of the scene.

A fundamental limitation of the linear model is that it is strictly only valid for the situation where the endmembers are arranged in discrete patches on the surface. This condition is almost never met in nature and many constituents of interest for earth science investigations exist in soils or at small scales in intimate association with one another. When the endmember materials are intimately mixed on spatial scales smaller than the path length of photons through the medium, light typically interacts with more than one component and the measured spectrum is a complex convolution of the endmember spectra rather than a simple linear mixture.

The nonlinear effects in spectra of particulate mineral mixtures have been recognized for many years (e.g., Nash and Conel, 1974). A variety of methods have been developed to treat this situation, including Hapke's model for particulate surfaces (Hapke, 1981, 1993). Hapke's photometric model has been shown to be a powerful and useful model for linearizing the mixture systematics has been demonstrated in laboratory studies of directional-hemispherical reflectance (Johnson et al., 1992; Johnson et al., 1983) and bidirectional reflectance (Mustard and Pieters, 1987, 1989) and shown to be accurate to approximately 5% absolute abundance. The technique has also been applied successfully to imaging spectrometer data for desert soils in Utah (Mustard and Pieters, 1987).

The effect of nonlinear mixing on reflectance spectra can be quite dramatic. There are some important implications of these differences for spectral mixture analysis. If a linear mixing model is used on data where the systematics are nonlinear, the fractions calculated will be significantly erroneous. In tests of linear versus nonlinear mixing on the laboratory data, the fractions calculated may have an absolute error of as much as 30%. In addition, the linear model can cause considerable ambiguity and false fractions when used on nonlinear mixtures.
Absorption bands and continua in nonlinear mixtures cannot be fit adequately with a linear model. However, the least-squares approach will minimize fitting errors using any of the endmembers in the equation. Thus endmembers not present in a mixture will be calculated to be present simply to minimize the error.

Despite the obvious advantages of using a nonlinear approach for intimate mixtures, this has not been widely applied to remotely acquired data of particulate surfaces. There are several reasons for this. It is generally thought that the detailed photometric properties of all the endmembers and surfaces are required to perform the calculations. However, many common particulate materials exhibit quasi-Lambertian behavior when viewed at nadir for incidence angles up to 40°. Therefore, a Lambertian approximation should be adequate for a first-order solution. The calculation of single-scattering albedo from reflectance requires knowledge of the incidence and emergence angles of the endmembers and remotely acquired spectra.

The most important obstacle to the application is that the particle size, composition and alteration state of the endmembers are very important controlling parameters of the solutions. Nevertheless, the nature of the remote data and as spectral libraries become more well endowed with data, many of the obstacles may be removed.

**Examples**

- (Johnson et al., 1992) Simple algorithms for remote determination of mineral abundances and particle sizes from reflectance spectra.
- (Mustard and Pieters, 1989) Photometric phase function of common geologic minerals and applications to quantitative analysis of mineral mixture reflectance spectra.
- (Johnson et al., 1983) A semiempirical method for analysis of the reflectance spectra of binary mineral mixtures.
- (Nash and Conel, 1974) Spectral reflectance systematics for mixtures of powdered hypersthene, labradorite, and ilmenite.

### 2.4.2 Environmental geology

#### 2.4.2.1 Acid Mine drainage – Soil contamination
The main subject of geologic environment monitoring is related to mining activities. Several recent incidents have shown that mining activities can seriously impact the environment. More specifically, entire mining districts have been abandoned around the world with often little consideration to environmental impacts.

In most cases, the processes involved are related to oxidation of pyrite at the surface of mine tailings. This oxidation produces acidic water that is gradually neutralized as it drains away from the waste. Pyrite oxidation leads to deposits of Fe-bearing secondary minerals in roughly concentric zones that emanate from the mine waste piles. Depending on hydration and pH conditions, Pyrite will be replaced first by sulfates such as Copiapite or some other sulfate (Schwertmannite, Ferrihydrite, etc.). When pH decreases, Jarosite becomes stable and is progressively replaced by iron oxides and hydroxides such as Goethite and Hematite. These minerals can be used as indicators of geochemical conditions such as pH or metal leachability. This last information is particularly important since Fe-bearing secondary minerals often contain heavy metals such as arsenic (As) or lead (Pb) that substitute for iron (Fe) in the mineral structure and can be released in the environment depending on leachability conditions.

Interestingly enough, each of these Fe-bearing minerals is spectrally unique. Consequently, imaging spectroscopy can be used to rapidly screen entire mining districts for potential sources of surface acid drainage and detect acid producing minerals in mine wastes or unmined rock outcrops. The US Environmental Protection Agency (EPA) estimated that mineral maps made from AVIRIS data accelerated remediation efforts by 2 years, saving over $2 millions in investigation costs.

A. USGS Tetracorder

The reader is referred to **USGS Tetracorder** in the previous section.

Examples

- (Montero et al., 2005) Characterization of waste rock associated with acid drainage at the Penn Mine, California, by ground-based visible to short-wave infrared reflectance spectroscopy assisted by digital mapping.
- (Ong et al., 2003) An investigation of the use of Tetracorder expert system for multi-temporal mapping of acid drainage-related minerals using airborne hyperspectral data.
- (Swayze et al., 2000) Using imaging spectroscopy to map acidic mine waste.

B. Analytical Imaging and Geophysics (AIG) algorithm
The reader is referred to **Analytical Imaging and Geophysics (AIG) Algorithm** in the previous section.

### Examples
- (Vaughan and Calvin, 2004) Synthesis of high-spatial resolution hyperspectral VNIR/SWIR and TIR image data for mapping weathering and alteration minerals in Virginia City, Nevada.
- (Ferrier, 1999) Application of imaging spectrometer data in identifying environmental pollution caused by mining at Rodaquilar, Spain.

### C. Stepwise Multiple Linear Regression (MLR)

To observe the quantitative relationship between spectral parameter values and metal levels, the stepwise multiple linear regression can be used. At each step, the independent variable not in the equation and with the smallest probability of F is entered, provided that the probability is sufficiently small. Variables in the regression equation are removed if their probability of F becomes sufficiently large. The method terminates when no more variables are eligible for inclusion or removal. For more information on MLR the reader is also referred to **Regression analysis** in a previous section of this document.

### Examples
- (Choe et al., 2008) Mapping heavy metal pollution in stream sediments using combined geochemistry, field spectroscopy, and hyperspectral remote sensing: a case study of the Rodalquil mining area, SE Spain.
- (Garcia-Haro et al., 2005) A new tool for variable multiple endmember spectral mixture analysis (VMESMA).
- (Rockwell et al., 2004) Remote sensing for environmental site screening and watershed evaluation in Utah Mine lands – East Tintic Mountains, Oquirrh Mountains, and Tushar Mountains.
- (Kemper and Sommer, 2003) Mapping and monitoring of residual heavy metal contamination and acidification risk after the Aznalcollar mining accident (Andalucia, Spain) using field and airborne hyperspectral data.
- (Kemper and Sommer, 2002) Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy.
D. Artificial Neural Networks (ANN)

ANNs are model-free estimators in that they do not rely on an assumed form of the underlying data. Rather, based on some observed data they attempt to obtain an approximation of the underlying system that generated the observed data. In general, one can look at computation in neural networks from the perspective of estimating an unknown function based on these observations. This general framework of learning from examples makes it possible to use neural networks in situations where exact cause-and-effect relationships are not known. However, this does not imply that neural networks by themselves are heuristic techniques. Far from it: neural networks have a rich theory underlying them and are based on sound mathematical principles.

Underlying a neural network are three main ingredients: (1) the model of a single neuron that is assumed; (2) how these simple neurons are interconnected to form a network; and (3) how the strengths of the interconnections between the neurons are determined. Most models of neural networks arise by varying these constituents so as to ensure that the resulting network has properties that are useful and that can be guaranteed. A neuron accepts an input $\xi^k$ and produces an output $y^k$. Each component of $\xi^k$ is carried to the neuron through adjustable weights. The two common forms of neurons are the so-called linear basis function neurons and the radial basis function neurons. The more popular forms of neural networks are based on a network of either of these two models of neurons. In most models, a numerical value called bias is added to the net input. The activation function can be linear or non linear.

Examples

- (Garcia-Haro et al., 2005) A new tool for variable multiple endmember spectral mixture analysis (VMESMA).
- (Kemper and Sommer, 2003) Mapping and monitoring of residual heavy metal contamination and acidification risk after the Aznalcollar mining accident (Andalucia, Spain) using field and airborne hyperspectral data.
- (Kemper and Sommer, 2002) Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy.

E. Variable Multiple Endmember Mixture Analysis (VMESMA)

This algorithm is based on unmixing and is a modified version of MESMA. First, the study area has to be partitioned based on spectral information to allow the application of different sub-models to the
selected areas. Based on an iterative feedback process, the unmixing performance can be improved in each stage until an optimum level is reached. This technique allowed for example the semi-quantitative estimate of sludge respectively of residual pyrite bearing material to be transformed into quantitative information for an assessment of acidification risk and the distribution of residual heavy metal contamination based on an artificial mixture experiment and to derive simple stochiometric relationships. As a result, a sludge abundance map can be rescaled to quantities of residual pyrite sludge, associated heavy metals, acidification potential and counteracting calcite buffering need. This information may then be used as input for further remediation planning and as basis for monitoring the success of the remediation measures (Garcia-Haro et al., 2005).

Examples

- (Garcia-Haro et al., 2005) A new tool for variable multiple endmember spectral mixture analysis (VMESMA).
- (Wu et al., 2005) Feasibility of reflectance spectroscopy for the assessment of soil mercury contamination.

F. Spectral Angle Mapper (SAM)

The reader is referred to Spectral angle mapper (SAM) in a previous section of the document.

Examples

- (Ferrier et al., 2009) Characterization of the environmental impact of the Rodalquilar mine, Spain by ground-based reflectance spectroscopy.
- (Riaza et al., 2006) Dehydration and oxidation of pyrite mud and potential acid mine drainage using hyperspectral DAIS 7915 data (Aznalcollar, Spain).

2.4.2.2 Production of pH maps

pH maps, showing the evolution of pH through time, can also be derived from mineral maps after some specific geochemical modeling is performed to establish clear relationships between stable mineral and pH conditions. These relationships are generally specific to a site since they depend on parameters that regulate the chemical reactions such as hydrology or climate conditions. For example, CSIRO in Australia managed to derive pH maps that are more readily usable by mining companies or environmental agency.
Examples

- (Ong and Cudahy, 2002) Deriving quantitative monitoring data related to acid drainage using multitemporal hyperspectral data.

2.4.3 Geologic hazards

2.4.3.1 Expansive clay soils

Expansive clay soils constitute a major geologic hazard. Field observation and engineering analyses have shown that Smectite has the highest swelling potential and is responsible for most of the severe damages observed. Cost of post-construction mitigation and standard engineering soil tests to create regional hazard maps is important.

NIR reflectance spectroscopy can be used to discriminate among pure Smectite (high swelling potential), mixed-layer illite/smectite (intermediate swelling potential) and kaolinite (low swelling potential) based on characteristic absorption bands. The main limitation can be in some cases heavy vegetation cover such as forests or green grass and when soil reflectance is < 10 %. In the case of dark soil, the SNR is low so the absorption features will not be evident enough, limiting identification. The construction of clay hazard maps can be achieved with IS data provided that soils are adequately exposed and instrument SNR good enough.

A. Analytical Imaging and Geophysics (AIG) Algorithm

The reader is referred to Analytical Imaging and Geophysics (AIG) Algorithm in the previous section.

Examples

- (Chabrillat et al., 2002) Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution.
2.4.3.2 Oil seeps

Oil seeps are another major geologic hazard. For environmental and economic reasons, future petroleum exploration needs to be non-invasive and cost effective. Explorationists evaluating remote terrain can now consider using airborne hyperspectral imaging for detecting onshore oil seeps. Previous to this new technology the spatial and spectral resolution of sensors was too limited for successful detection of onshore oil seeps.

Hydrocarbon microseepage theory establishes a cause-and-effect relation between oil and gas reservoirs and some special surface anomalies, which mainly include surface hydrocarbon microseepage and related alterations. Therefore, it is possible to explore for oil and gas by determining reflectance spectra of surface anomalies. Significant gas fields are found in subtle stratigraphic traps in addition to structural traps. Both may be detectable by measuring surface hydrocarbon microseepage and related anomalies. In order to decrease cost of exploration and improve wildcat success rates, various combinations of surface methods have been used to locate those wildcat tests.

Remote sensing for hydrocarbon exploration generally focuses on the identification of indirect evidence of hydrocarbons. It is possible to map indirect evidence of seeps, using the spectra of stressed vegetation or through the identification of alteration minerals associated with seeps. In rare cases, it may even be possible to identify hydrocarbons directly, if they accumulate on the surface in sufficient quantities. Diagnostic spectral features of methane and crude oil consist of four distinct hydrocarbon absorption bands. Two bands, 1.18 and 1.38 μm are narrow and sharply defined, though fairly weak. The other two, 1.68-1.72 and 2.3-2.45 μm are broader, but much stronger. Hydrocarbon induced alterations associated with indirect detection can express themselves in a variety of forms, such as mineralogical changes (calcium carbonate mineralization near 2.35 μm), bleaching of red beds (near 1.0 μm), clay minerals alterations (near 2.2 μm), electrochemical changes, radiation anomalies, microbial anomalies, as well as biogeochemical and geobotanical anomalies.

A. Analytical Imaging and Geophysics (AIG) Algorithm

For methodological details, see Analytical Imaging and Geophysics (AIG) Algorithm in the previous section.

Examples

- (Ellis et al., 2001) Exploring for onshore oil seeps with Hyperspectral imaging.
- (Ellis et al., 2000) Airborne hyperspectral imagery for the petroleum industry.

© HYPER-I-NET February 09
B. Spectral Angle Mapper (SAM)

For more information see Spectral angle mapper (SAM) in a previous section of this document.

Examples

- (Xu et al., 2008) Exploring for natural gas using reflectance spectra of surface soils.
- (Feeman, 2003) Evaluation of the use of Hyperspectral imagery for identification of microseeps near Santa Barbara, California.

C. Spectral Feature Fitting (SFF)

Spectral Feature Fitting (SFF) is another spectral library matching technique for classifying unknown image spectra (Clark et al., 1990). A particular strength of SFF is that it isolates individual absorption features for comparison and only the shapes of the features are compared, not their depth.

First the continuum is removed from the image and the reference library spectra. The continuum is formed by connecting the local maxima of the spectrum with straight-line segments. Without removing the continuum, it is difficult to define distinct absorption features because illumination and particle size differences tend to dominate the spectra. Therefore, the image and reference spectra are normalised by dividing the reflectance values by the estimated continuum values.

A constant is added to the library continuum removed spectrum to provide a scaling factor in comparing the library and image spectra. This scaling is needed because the absorption features in the library data typically have greater depth than in the image spectra.

Next, the least squares fit (LSF) is calculated band-by-band between each reference endmember and the unknown spectrum using standard statistical methods. SFF does not produce a color-coded map so post classification is required to generalize the classes, for instance with the rule classifier in ENVI software based on thresholds from the histograms for each endmember.

Examples

- (Feeman, 2003) Evaluation of the use of Hyperspectral imagery for identification of microseeps near Santa Barbara, California.
- (Clark et al., 1990) Material absorption band depth mapping of imaging spectrometer data using a complete band shape least-squares fit with library reference spectra.

D. Principal Component Analysis (PCA) based on Wavelet Transform
PCA has been widely used for decades to reduce dimensionality, and it is a useful technique if the spectral class structure of the transformed data is such that it is distributed along the few axes. Otherwise, the transformed data may be similar to the original data. In such a case, the wavelet decomposition technique might be a better approach. Wavelet decomposition can reduce hyperspectral data in the spectral domain. This will not only reduce the data volume, but also preserve the distinction among spectral signatures. The characteristic is related to the intrinsic property of wavelet transforms that preserve high-frequency and low-frequency features during the signal decomposition and therefore preserve peaks and valleys found in typical spectra. In general, most classification errors occur at the boundary between classes. Since wavelet decomposition is applied to each local test data, a wavelet-based reduction might not well-differentiate classes among neighboring test data in the spatial domain. PCA, however, can provide more local spatial information among neighboring classes than wavelets.

Examples

- (Xu et al., 2008) Exploring for natural gas using reflectance spectra of surface soils.

E. Feature characteristics

Absorption band parameters such as the position, depth, width, area, and asymmetry can be used to quantitatively estimate composition of samples from reflectance data (van der Meer, 2004). The spectral absorption features of two absorption bands of 2.1-2.26 µm, related to clays, and 2.24-2.398 µm, related to carbonates, are extracted and studied, respectively. The absorption-band depth in the region from 2.3 to 2.6 µm and the overall reflectance slope in this range can be related to hydrocarbon concentrations.

Examples

- (Xu et al., 2008) Exploring for natural gas using reflectance spectra of surface soils.

F. Spectral Mixture Analysis (SMA)

For more details see Spectral mixture analysis (SMA) in the NPV section of this document.

Examples

- (Xu et al., 2008) Exploring for natural gas using reflectance spectra of surface soils.

© HYPER-I-NET February 09
2.4.4 Regolith mapping

The regolith is the part of the alteration profile between the soil which contains organic matter and the unaltered parent rock. Cudahy (1992) has developed a regolith model showing changes in the abundance, chemistry and crystallinity (disorder) of various minerals, especially the kaolins (kaolinite, dickite and halloysite), iron oxides (hematite, goethite) and primary minerals such as white micas (muscovite, paragonite and phengite), chlorite and talc.

2.4.4.1 Kaolin disorder

The nature of the kaolin in the regolith was further examined which led to the development of a refined regolith model (Cudahy, 1997). This model predicts that kaolin physicochemistry, including lattice structural order such as dickite-kaolinite interlayering cation substitution effects such as Fe for Al, can be used to map parent rock composition and regolith stratigraphy. In particular, transported materials are characterized by poorly ordered kaolin whereas saprock/saprolite developed over mafic and felsic rocks are characterized by more relatively well ordered kaolin. Weathered ultramafic rocks have poor development of kaolin and what kaolin is developed is poorly ordered and affected by abundant cation substitution. Therefore, kaolin disorder is one of the parameters of interest to locate transported versus in situ materials. Kaolin disorder changes the shape of the famous doublet. The depth of the 2.16 μm absorption appears to be very sensitive to disorder. A relationship between the relative depth of the 2.16 μm absorption and kaolin disorder can be established using X-ray diffraction (XRD) data.

A. Kaolin disorder mapping algorithm

The methodology used includes several steps:

- Focus on mineral-diagnostic features only.
- Use of continuum-band normalization procedures to both provide a measure of the absorption-band-depth and to cancel unnecessary wavelength-independent, multiplicative effects like illumination geometry, surface scattering and albedo.
- Avoid the use of any image or column/row statistics.
- If a diagnostic mineral absorption is not present then that mineral is not present.
- Mask out those minerals with features that spectrally overlap/complicate with the minerals of interest.
Absolute mineral abundances should be achievable using the above combined with an appropriate mixing model.

Examples

- (Cudahy et al., 2006a) Mapping alteration through the regolith at Kalgoorlie using airborne hyperspectral data.

### 2.4.4.2 Micas composition and metasomatic hydrothermal alteration

Cudahy also developed a spectral-mineralogical model for metasomatic hydrothermal gold (Au) alteration superimposed on different Archean host rocks. This model predicts that white mica progressively changes its cation composition from Al-rich Muscovite to Al-poor phengite with increasing alteration. Spectrally, this produces an apparent increase in wavelength of the white mica absorption from shorter (~2.185 µm) to longer (~2.215 µm) wavelengths. The spectral shift is related to the level of substitution.

For example, the large domains of short wavelength paragonite-rich mica and large domains of longer wavelength muscovite/phengite have not been previously recognized in public geoscience data and potentially represent big footprint signatures of different hydrothermal fluids responsible for Archean Au mineralization of the Kalgoorlie region in Australia. White mica is relatively resistant to erosion and so is possibly a mineral constituent of the soils that cover much of the region as pediment cover.

These mineralogical signatures have implications for the nature of the mineralizing fluids, structural architecture and water-water and water-rock ratios. Previously unrecognized alteration zonation has been revealed using airborne hyperspectral data in a very well mapped area that arguably had been considered inaccessible to this technology because of extensive regolith cover.

**A. Multiple logistic regression**

Multiple regression can be used to find regression equations for the prediction of the presence of white micas and their absorption wavelength measured with field spectroscopy (predicted variables) from the airborne band ratios (predictor variables). Multiple logistic regression is first used to determine the presence of mica followed by a multiple linear regression for estimating their absorption wavelengths. Stepwise procedures are used to obtain a selection of optimal predictor variables from the set of airborne bands.
Multiple logistic regression is a technique for the prediction of presence or absence of a characteristic or outcome based on values of a set of predictor variables. It is similar to linear regression but is suited to models where the dependent variable is dichotomous. The following regression model is used to obtain probability values for the presence of white micas from a set of airborne band ratios:

\[
p_i = \frac{1}{1 + e^{-(b_0 + b_1 x_{i1} + \ldots + b_n x_{in})}}
\]

where \( p_i \) the estimated white mica probability of the pixel that spatially coincides with the \( i \) th rock sample, ranging from 0 to 1, \( x_{ij} \) the \( j \) th predictor (airborne band ratio) value of the pixel that spatially coincides with the \( i \) th rock sample, \( b_j \) the \( j \) th regression coefficient, \( j = 1, \ldots, n \) and \( n \) the number of predictors.

The coefficients of the logistic regression model are estimated using a maximum likelihood method, where the difference between measured white mica probability in rock samples and estimated white mica probability in spatially coinciding airborne pixels was minimized.

**Examples**

- (Cudahy et al., 2006a) Mapping alteration through the regolith at Kalgoorlie using airborne hyperspectral data.
- (van Ruitenbeek et al., 2006) Mapping white micas and their absorption wavelengths using hyperspectral band ratios.

**B. Linear-regression model for estimation of absorption wavelength**

Multiple linear regression is used for the prediction of continuous variables based on the values of a set of predictor variables. Linear regression is used to estimate the absorption wavelengths of white micas from a set of airborne ratios:

\[
y_i = c_0 + c_j x_{ij} + \ldots + c_n x_{in}
\]
where \( y_i \) is the estimated absorption wavelength of the pixel that spatially coincides with the \( i \) th rock sample, \( x_j \) the \( j \) th predictor (airborne band ratio) value of the pixel that spatially coincides with the \( k \)th rock sample, \( c_j \) the value of the \( j \)th regression coefficient, \( j = 1, \ldots, n \) the number of predictors.

The coefficients of the linear regression model are estimated using the least squares method, where the difference between measured absorption wavelength in rocks and estimated absorption wavelengths in spatially coinciding airborne pixels is minimized.

**Examples**

- (van Ruitenbeek et al., 2006) Mapping white micas and their absorption wavelengths using hyperspectral band ratios.

### 2.4.4.3 Mapping non magnetic, talc-bearing ultramafic rocks

Mapping non-magnetic, talc-bearing ultramafic rocks is another important subject. Locating ultramafic rocks is important for ore exploration. In most cases, this is done through the use of airborne magnetics to locate magnetite-bearing ultramafic rocks. This magnetite is the product of hydration processes that alter olivine-pyroxene to serpentine and magnetite. However, not all ultramafic rocks bear magnetite because they either did not experience hydration reactions, or were later metamorphosed at sufficiently higher temperatures to reform olivine-pyroxene again, or were affected by CO\(_2\) flushing which converted serpentine and magnetite to talc and carbonate.

Exploration for these magnetite-poor ultramafic rocks which can potentially host significant massive Ni mineralization is thus a challenge using conventional mapping methods. Talc is a mineral relatively resistant to weathering that can be spectrally detected if it is exposed at the surface in sufficient abundance. Talc presents absorptions at 2.308 and 2.39 \( \mu \)m that can be used to create talc maps.

**Examples**

- (Cudahy et al., 2006b) Mapping non-magnetic, talc-bearing ultramafic rocks in the Kalgoorlie region using airborne hyperspectral data.

## 2.5 HYDROSPHERE
Remote sensing of inland and coastal water provides a synoptic view of the spatial distribution of different biological, chemical and physical variables of both the water column and, if visible, the substrate. This knowledge of the distribution is essential in environmental water studies as well as for resource management. Many inland and coastal waters are highly affected by anthropogenic influences. In combination with the complex hydrological situation, highly contrasted structures evolve in time and space in these aquatic environments. It is obvious that a water system with different optically active substances with temporal and spatial variations is by far more complex and requires more sophisticated models for remote sensing and separation of the water constituents than a system containing one component only like the ocean waters. It is for this reason that airborne imaging spectrometry mainly gets applied to coastal and inland water environments and not to oceans. The vast dimension of oceans necessitates the use of ocean colour sensors on satellite platforms. An extensive discussion of imaging spectroscopy of water is given in (Dekker et al., 2001).

### 2.5.1 Bathymetry, bottom types (albedo, sand, silt, macrophytes) and water column optical properties

Quantification of benthic community structure is central to understanding coral reef ecosystem function. Community structure determines rates of reef metabolism and indicates reef status. Different bottom types are important in life history strategies of reef-dwelling organisms, e.g., as recruitment sites for coral larvae and juvenile fish, and as habitat for adult fish. Reef community structure exhibits tremendous spatial heterogeneity over scales of centimeters to hundreds of meters and, in contrast to phytoplankton and macrophyte communities, is inherently stable on time scales of months to years (Hochberg et al., 2003). Remote sensing techniques allow large reef areas to be viewed and are the only practical solution for repeated monitoring of reef systems with sufficient frequency to detect major ecological phenomena on different scales. With imaging spectroscopy identification of spectral features indicative of degradation in reefs is feasible and this can lead to better ecological assessment and forecasting of detrimental conditions of the reef environment worldwide ([http://geo.arc.nasa.gov/sge/coral-health/index.html](http://geo.arc.nasa.gov/sge/coral-health/index.html)). For further reading and examples we refer to Brando et al. (2009), Hochberg and Atkinson (2000), Hochberg et al. (2003, 2004), Joyce and Phinn (2003), Louchard et al. (2003), Mumby et al. (2004, 2008) and Sandidge and Holyer (1998).

### 2.5.2 Phytoplankton, suspended sediments, dissolved organic material (DOM), fCover of bottom green macrophytes
Determination of the abundance of various phytoplankton species in coastal waters is important in studies of biological productivity and water quality. The problem is daunting, because the level of the reflectance signal from phytoplankton is usually close to the noise level, and it is often masked in coastal waters by the presence of colored dissolved organic matter (CDOM). To extract and reliably identify the low-level reflectance from phytoplankton, we need hyperspectral remote sensing data, data taken over the entire visible and near-infrared spectrum that can be used to isolate the effects and the spectral signatures of various constituents contributing to the reflectance from ocean water. The majority of organic carbon in the oceans is present as dissolved organic matter (DOM); therefore understanding the distribution and dynamics of DOM is central to understanding global carbon cycles. Describing the time-space variability in colored dissolved organic matter (CDOM) has been difficult, as standard spectrophotometric methods for CDOM determination are laborious and susceptible to methodological biases. For further reading and specific examples we refer to Andrefouet et al. (2004), Cipollini et al. (2001), Doxaran et al. (2005), Kirkpatrick et al. (2003) and Sydor (2006).

2.6 PEDOSPHERE

Soil is a complex material that is extremely variable in its physical and chemical composition. It is formed from exposed masses of partially weathered rocks and minerals of the earth’s crust. Soil formation or genesis is strongly dependent on the environmental conditions of both the atmosphere and the lithosphere. The soil body is a product of five factors: climate, time, organisms, topography, and parent materials. The great variability of soils is the result of interactions of these factors and their influence on the formation of different soil profiles. In general, the soil profile, composed of several horizons, typically refers to A (the upper, termed as alluvial horizon), B (the intermediate, termed as illuvial horizon), and C (the lowest, a parent material horizon) (Ben-Dor et al., 2008). Careful observation of the spectra of more than 400 American and some Brazilian soils (in the VNIR) revealed only five spectral categories for describing all soil groups (Stoner and Baumgardner, 1981). Analyzing visually many soil spectra shows that the soil spectrum is even more general than Baumgardner and Stoner’s criteria. A generic soil spectrum is characterized by a monotonous spectral increase in the VIS with or without a smear absorption feature of iron and a changing slope value, a function of its organic matter content (Ben-Dor et al., 1997). Far in the NIR-SWIR, two strong absorption features at 1400 and 1900 nm, with a slight decrease toward the TIR region can usually be visible based on the specific surface area (SSA) content of the solid soil phase that eventually controls
the hygroscopic moisture content (MC) in the sample (Banin and Amiel, 1970). This general view may be followed by moderate or weak absorption features in the SWIR region due to clay minerals (2200 nm), carbonate (2330 nm), salt or primary minerals, and organic compounds in the soil.

The complexity of soil materials does not often permit utilization of simple spectral analysis routines or spectral matching, such as the match filter approach with a known spectral library database. Therefore, multivariate analysis appeared to be more suitable to unravel the complex mixtures of soil constituents, as has been shown in numerous studies (e.g., Leone and Sommer, 2000). More recently, it has been demonstrated that a wide range of soil constituents such as the total metal content, water content, carbonate, and organic carbon can be derived through reflectance spectroscopy by using advanced methods such as artificial neural networks (ANNs) and partial least squares regression analysis (e.g., Udelhoven et al., 2003a). This seems to open up additional perspectives not only for mapping visible patterns (e.g., erosion) but also for deducing the soil constituents for the use in precision farming (e.g., the mineral content).

### 2.6.1 Soil degradation (salinity, erosion, deposition)

Soil degradation is frequently related to change in climate and global land use dynamics, leading to the intensification of agriculture, deforestation, and desertification (Eswaran et al., 2001). Any process-related evidence needs to be deduced from the spectral properties of the surface, implying that the spectral analysis has to focus on optically active tracers that can be connected to the process under question. Frequently, an optically active tracer itself might be significant for a process, such as surface salts in the case of salinization. But spectral indicators might also result from secondary effects such as mineralogical and structural changes that occur as a result of the problem.

For mapping erosion-affected soils with IS-remote sensing systems, a framework that has proven useful is based on concepts that consider soil development to be either progressive or regressive with time (Birkeland, 1990). Under progressive development, soils become better differentiated by horizons and horizon contrast becomes stronger. Pedogenetic processes involve the formation of clay-size particles by the weathering of larger grains, the alteration of clay minerals to other clay-mineral species, and the release and accumulation of iron by weathering. Some solids (silt, clays and CaCO$_3$) and ions (Ca$^{2+}$, Na$^{+}$, etc.) dissolved in rain water are added from the atmosphere, and as a result, the content of the topsoil organic matter increases with the decomposition of the plant and animal residues. Material transferred within the soil profiles result in the accumulation of silt and clay, Fe, Al, CaCO$_3$, gypsum, or halite in the B horizon, or due to bioturbation processes, they accumulate on the soil surface. In
contrast, regressive pedogenesis refers to the addition of material to the surface at a rate that suppresses soil formation (e.g. aeolian dunes, sand encroachment, glacial moraines, and distal fans) or to the suppression/interruption of pedogenetic processes by surface erosion, which, if continued, might produce truncated soil profiles along linear flow structures (channel erosion) or across larger areas (sheet erosion). Both progressive and regressive pedogenetic processes cause alterations of the soil surface that to a certain extent, they are spectrally detectable (e.g., Hill et al., 1995).

The remote sensing approach thus comprises two parts: first, the identification of optically active natural tracers associated with specific soil processes (e.g. erosion) and second, the design of efficient approaches to map the natural tracers in their spatial context.

### 2.6.1.1 Soil characterization

There is a recognized need to develop rapid and inexpensive techniques for soil characterization to support applications such as quantitative soil landscape modeling, precision agriculture and global soil C monitoring. Soil-landscape models and soil maps have been constructed largely on the basis of field observations – including Munsell colors, hand texturing, pH indicators and acid reaction. Recent research has suggested that proximal VNIR diffuse reflectance could provide inexpensive predictions of soil physical, chemical and biological properties. VNIR soil reflectance is a function of soil composition – organic materials, primary minerals, clay minerals, salts, and poorly-crystalline andic materials. VNIR reflectance has been proven, based on a broad sampling of soils over the world, to have strong relationships with the following important soil properties: relative kaolinite content, relative montmorillonite content, clay content, cation exchange capacity (CEC) (NH4 at pH = 7.0), soil organic carbon (SOC), inorganic carbon (IC) (HCl with electronic manometer) and citrate-dithionite extractable Fe. This demonstrates the fundamental viability of VNIR reflectance for global soil characterization.

- **Boosted regression trees**

  Following Friedman (2001), boosted models can be expressed in the general form:

  \[
  F(x; \{\beta_m a_m \}_{m=1}^M) = \sum_{m=1}^M \beta_m h(x; a_m)
  \]

  where \( h(x; a) \) represents a simple classification function or base learner with parameters \( a \) and input variables \( x \), \( m \) represents the model step, and \( \beta_m \) is a weighting coefficient for step \( m \). With the well known AdaBoost algorithm (Freund and Schapire, 1997), the base learner is applied sequentially to reweighted calibration datasets such that observations with larger residuals receive proportionally
greater weights in subsequent iterations. The final classification is computed with a weighted vote as shown in the equation above. (Friedman, 2001, 2002) has developed a new approach to fitting additive models of the form shown in the equation above, termed a Gradient Boosting Machine (GBM). With this approach, a numerical solution is found that involves sequentially fitting the base learner (using least squares) to pseudo-residuals computed from the gradient of a differentiable, prescribed loss function (lack of fit) – with respect to the predicted value for each calibration observation for the current step. In a further development, Friedman (2002) has found that the use of bagging or random subsampling from the calibration set in conjunction with boosting improves on boosting alone. Boosted regression trees (BRT) essentially add boosting to a regression tree routine like CART® (Breiman, 1984). The primary advantage of BRTR include (i) the ability to include a large number of weak relationships in a predictive model; (ii) insensitivity to outliers in the calibration dataset; (iii) the elimination of the need to uniform data transformations; and (iv) the relative immunity to overfitting.

BRT have many desirable qualities such as (i) the ability to model interactions and non-linear relationships; (ii) relative immunity to overfitting; and (iii) the ability to utilize a large number of weak classifiers and thereby make maximum use of the entire VNIR spectrum.

Examples

- (Brown et al., 2006) Global soil characterization with VNIR diffuse reflectance spectroscopy.

- **Partial Least Squares regression**

For soil characterization, Partial Least Squares (PLS) regression using the 1st derivatives of soil reflectance is commonly used to reduce high-dimensional spectral data obtained from NIR detectors. PLS regression, a chemometric technique, is similar to the Principle Components Regression (PCR) in that both employ statistical rotations to overcome the problem of high-dimensional, correlated predictors, except that in PLS the X and Y variables are rotated relative to the response variables to maximize predictive power.

A key limitation for both black box empirical modeling is that results cannot be reliably extended beyond the range of materials used in the calibration phase (Dardenne et al., 2000).

Examples

- (Brown et al., 2006) Global soil characterization with VNIR diffuse reflectance spectroscopy.
- (Dardenne et al., 2000) Multivariate calibration and chemometrics for near infrared spectroscopy: which method?
- **Artificial Neural Networks (ANNs)**

For more details the reader is referred to **Artificial Neural Network (ANN)** in a previous section of this document.

**Examples**

- (Chang and Islam, 2000) Estimation of soil physical properties using remote sensing and artificial neural network.

- **Near Infrared Reflectance Analysis (NIRA)**

The Near Infrared Reflectance Analysis (NIRA) method assumes that a concentration of a given constituent is proportional to the linear combination of several absorption features. The method is empirical, and although no physical or chemical assumptions are made, the method has a strong spectroscopy foundation. The NIRA approach has two stages: (1) the calibration stage, where a prediction equation for evaluation of a property is developed; and (2) the validation stage, where the preceding stage is validated. The calibration stage uses training samples, which represent the study population in terms of spectral and physicochemical properties. Then a prediction equation is generated based on multiple regression analysis between the soil chemistry data (determined in the laboratory) and selected spectral bands. This calibration equation is further validated in stage 2 against unknown samples and is examined statistically for its prediction performance.

Diagnostic absorption features of clay around 2.2 μm and calcium carbonate around 2.34 μm are used to estimate concentrations. After continuum removal, the depth or strength of these absorptions can be directly related to concentrations obtained by laboratory analyses. Linear regressions are derived between clay concentration and depth of the 2.2 μm absorption and calcium carbonate concentration and depth of the 2.34 μm feature. Intensity of absorption features for bands centered at 1.4, 1.9, 2.2, 2.265, 2.365 μm can also be used to estimate water content and mineral influences.

**Examples**

- (Lagacherie et al., 2008) Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements.

- (Demattê et al., 2004) Visible-NIR reflectance: a new approach on soil evaluation.
Feature characteristics (band shape analysis – band shape parameters)

For a more detailed description, see Feature characteristics in a previous section of this document.

Generally, bands associated with soil characteristics and centered at 450, 500, 600, and 1600 nm are characterized by angle measurements. Intervals between 450 and 600 nm and 780 and 1000 nm, chosen for iron oxide influence are also characterized by their concavity.

Examples

- (Farifteh et al., 2008) Spectral characteristics of salt-affected soils: A laboratory experiment.
- (Demattê et al., 2004) Visible-NIR reflectance: a new approach on soil evaluation.

Derivative analysis

For more methodological details the reader is referred to Derivative analysis in the pigments section of this document.

Examples

- (Demattê et al., 2004) Visible-NIR reflectance: a new approach on soil evaluation.

2.6.1.2 Soil salinity

Salinization and alkalization are the most common land degradation processes, particularly occurring in arid and semi-arid regions, where precipitation is too low to maintain a regular percolation of rainwater through the soil. Soil salinity is one of the major factors affecting biomass production and is the principal cause of soil degradation. Salinity is usually due to rising water tables, either induced by land clearing alone, referred to as dryland salinity, or by irrigation-induced salinity. The major soil components are inorganic solids consisting primarily of crystalline minerals and non-crystalline substances, organic matter, air, and water or a solution containing a variety of dissolved compounds (Farifteh et al., 2006). Salts in soils are reported to be Na₂CO₃, NaHCO₃, and NaCl, which are very soluble and mobile components in the soil environment. Typically, saline soils are highly erosive and have poor structure, low fertility, low microbial activity, and other attributes not conducive to plant growth.

There is a high correlation between soil reflectance and several soil properties such as: mineralogy, organic matter content, moisture content, particle size distribution, iron oxide content and surface
conditions. Ground observations and radiometric measurements indicate that the main factors affecting the reflectance are quantity and mineralogy of salts, moisture content, color, and surface roughness. Most of the features seen in spectrum of saline minerals (400–2500 nm) can be attributed to internal vibration modes of certain molecular groups and neutral water molecules. The reflectance of these salts can quantitatively be measured by laboratory or field spectroradiometers. Presence of white salt crusts on the surface, puffy soil surface, dark greasy surface of pure alkali soils, dehydration cracks of 1-2 cm wide and coarser topsoil texture are soil related indicators. Performance-oriented indicators used to distinguish saline/alkaline soils are spotty growth of crops, presence of dead trees, a blue-green tinge and moisture stress conditions.

The IS approach has shown a promising capability to further enlarge the information spectrally. Spectrally based studies of many researchers (e.g., Ben-Dor, 2002; Metternicht and Zinck, 1997, 2003; Taylor et al., 1994) have shown that detailed spectral information can directly and indirectly pinpoint saline-infected areas. Taylor et al. (2001) showed that dryland salinity in the Dicks Creek catchment of central New South Wales, Australia, could be characterized by the occurrence of spectrally distinct smectitic clays around surface salt scalds.

In many areas, it is obvious that the surface salts are ephemeral and can be removed during a rain event. Consequently, work commenced to investigate the nature of the mineralogical and structural changes that occur when soils are affected by salinity. Taylor (2004) described how the shape of the hydroxyl absorption feature at 2200 nm (depth and width) changes with increased salinity (hydroxyl depth versus width). An asymmetry to shorter wavelengths, due to kaolinitic minerals, is replaced by an asymmetry to longer wavelengths, characteristic of smectites. This corresponds to the observed increased frequency of swelling clays around saline discharges, which often leads to dispersive soils and subsequent soil erosion.

### A. Feature characteristics (band shape analysis – band shape parameters)

For a more detailed description, see Feature characteristics in a previous section of this document.

**Examples**

- (Farifteh et al., 2008) Spectral characteristics of salt-affected soils: a laboratory experiment.
- (Farifteh et al., 2006) Assessing salt-affected soils using remote sensing, solute modeling, and geophysics.
- (Taylor et al., 2006) Hyperspectral imagery as a tool for the assessment of stubble content in a fallow soil.
B. Salinization indicators

Continuum removal is the normalization of reflectance spectra using common spectral reference points to create a continuum spectrum, then continuum-normalizing the spectrum by dividing the original by the continuum. Continuum removal is a decorrelation technique that maximizes the effects of spectral absorption features (for a more detailed description see Continuum removal analysis in a previous section of this document).

Salinized soils have distinctive spectral features in the VNIR parts of the spectrum, related to combined water in hydrated evaporite minerals. They show absorption features at 505, 920, 1415, 1915 and 2205 nm. The salt scald and highly salinized soil show additional distinct absorption features at 680, 1180 and 1780 nm. These features allow the recognition of minerals such as gypsum, bassanite, and polyhalite that can be used as indicators. Additionally, hydroxyl features at approximately 2200 nm are becoming less developed as samples become more salinized. The reduction of the 2200 nm absorption feature intensity seen in the more salinized soils may be the result of a loss of crystallinity in the clay minerals. The overall slope of the curve between 800 and 1300 nm decreases as samples become more salinized. This decrease in slope parallels changes in soil color from red to grey and could be due to either changes in oxidation state of the iron present within the soils or due to variations in the evaporite mineral content within the soils. The deeper uncombined water features at 1415 and 1915 nm broaden and become more asymmetrical as salinization increases, suggesting an increase in soil moisture content with increasing salinization. The increase in depth and complexity of the absorption features has been described as being due to hydroxyl ions and water within the lattice of various hydrated evaporite minerals.

Examples

- (Dehaan and Taylor, 2003) Image-derived spectral endmembers as indicators of salinisation.
- (Dehaan and Taylor, 2002) Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced salinization.

C. Spectral Angle Mapping (SAM)

The reader is referred to Spectral angle mapper (SAM) in a previous section for more details.
Examples

- (Dehaan and Taylor, 2003) Image-derived spectral endmembers as indicators of salinisation.
- (Dehaan and Taylor, 2002) Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced salinization.

D. Matched Filtering (MF)

Matched Filtering (MF) allows the abundance of endmembers to be determined without any knowledge of the remaining endmembers. For a full description, see section **Mixture Tuned Matched Filtering (MTMF)**.

Examples

- (Dehaan and Taylor, 2003) Image-derived spectral endmembers as indicators of salinisation.
- (Dehaan and Taylor, 2002) Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced salinization.

E. Spectral Feature Fitting

Spectral Feature Fitting (SFF) is an absorption feature based method that matches image spectra to reference spectra using a least-squares technique and continuum-removal analysis. Absorption feature depth, relating to material abundance, is output in one image, while in a second image, the root mean square error is calculated using a least squares technique. For more details see **Spectral Feature Fitting (SFF)** in the Geologic hazard section.

Examples

- (Dehaan and Taylor, 2003) Image-derived spectral endmembers as indicators of salinisation.
- (Dehaan and Taylor, 2002) Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced salinization.

2.6.1.3 Soil erosion and deposition

Detecting and mapping areas that are affected by erosion is important for monitoring the process as well as for drafting and implementing suitable management plans. Ben-Dor (2002) and Goldshleger et
al. (2001) introduced a spectral empirically based idea that soil areas with potential to erode (low and high) can be dealt with long before the actual erosion takes place. A recent study demonstrated that in the field, erodability of a dust-producing playa in Australia can be detected via analyzing digital images and multi-angular reflectance measurements of spectroscopy before and after in situ rain simulator and wind tunnel abrasion have been applied to specific soil plots (Chappell et al., 2007).

Although empirical statistical models are quite powerful (e.g., Udelhoven et al., 2003b), numerous applications do not necessarily require a full quantitative assessment of soil properties. Areas affected by soil erosion, for example, may be successfully identified by mapping simple spectral indicators that are conceptually related to the eroded processes and provide a sufficient sensitivity range. Undisturbed and eroded soils exhibit different spectral characteristics because the top soil might be lost due to erosion and finally the parent material is becoming more abundant at the surface or is subsequently changed.

Organic carbon content in topsoils is an indicator for soil quality and soil erosion processes. In the context of landscape evolution, the spatial variability of soil organic matter concentrations might be understood as an indicator which emphasizes areas of accumulation and relative stability (sinks) in comparison to active erosion and transport cells (source areas) (Hill and Schuett, 2000). Various important soil characteristics concerning the nutrient budget are determined through soil organic matter, such as cation exchange capacity, soil acidity, the capacity of binding organic chemicals (pesticides, herbicides, fertilizer, etc.), the release and sequestration of N, P, and S during decomposition of soil organic matter, and nutrient availability for soil microorganisms. Besides the obvious impact of organic matter on chemical soil properties which directly affect the quality of the soil for agricultural use, organic matter also influences the physical characteristics of a soil with regard to accelerated soil erosion processes, as, for example, hydraulic conductivity (percolation and retention of soil water), soil structure (type and size of soil aggregates), and soil color which influences albedo and soil temperature.

General models related to organic matter dynamics indicate that the physical characteristics of a soil reflect the balance between regenerative and degenerative processes of soil structure development. Stable macroaggregates (> 250 µm) formed by inorganic compounds are, in general, held together by iron and aluminum oxides and occur predominantly in oxisols. Otherwise, stable macroaggregates are formed by soil organic matter such as fungal hyphae, fibrous roots, and polysaccharides. They rather occur under pasture or where organic residues have been added to the soil. The stability of macroaggregates increases with the amount of stable microaggregates. Since soil erodability (i.e. the resistance to mechanical detachment and breakdown of soil aggregates, and the ability to accept rainfall as infiltration) is greatly influenced by the structural organization of the soil, it is obvious that
highly calcareous silty Mediterranean soils become, in the absence of clay, highly dependent on organic matter. Increased organic matter concentrations render these soils more favorable because of better aggregation, higher infiltration and water retention, and increased nutrient availability. Also, better infiltration of rainfall and the formation of macroaggregates are among the most important soil properties to increase their resistance to accelerated erosion.

Organic matter content is also a soil property that significantly influences the bidirectional reflectance characteristics of a soil substrate.

A. Bi-directional reflectance model

Pinty et al. (1989) extended the work of Hapke (1963, 1981) to describe the range of soil surfaces on Earth where individual particles have non-uniform angular distributions. Jacquemoud et al. (1992) extended the model to explain backward and forward scattering (the specular effect) of light by smooth soils of different types. This model and its parameters are useful for characterizing prepared soil surfaces and changes to them induced by laboratory rainfall simulation and wind tunnel abrasion (Chappell et al., 2007). It is also used here because of its appropriateness to the smooth playa soils. The parameters of the model include: (i) the single scattering albedo (SSA) (= the ratio of the scattered energy to the total energy either scattered or absorbed by the particle), (ii) a roughness parameter which appears to be related to particle size distribution, porosity and the gradient compaction with depth (Hapke, 1963) and is caused by small shadows behind individual particles and micro-aggregates (Cierniewski, 1987). The remaining parameters are diagnostic of the type of scattering of light and have been demonstrated by Jacquemoud et al. (1992) and used by (Chappell et al., 2007), to identify backscattering (towards the direction of illumination) and forward scattering (away from the direction of illumination) and mixed scattering (backwards and forwards relative to the direction of illumination). The retrieved values should depend on the soil surface condition, not the illumination and viewing geometry. Jacquemoud et al. (1992) concluded that the retrieved parameters were invariant for a given soil, except for smooth soils with a large clay content, which exhibited a large specular effect near soil moisture saturation, but with a decreased specular effect and increased backscatter with drying.

Since the radiometric properties of a bare soil surface can be described by the model of Jacquemoud et al. (1992) the challenge is to find the values of the six parameters such that the computed value of the reflectance best approximates the actual observations. Following Pinty et al. (1989) and Jacquemoud et al. (1992) a non-linear least squares fitting procedure was used to solve that inverse problem. Details of the approach can be found in those original papers and the practical implementation and sensitivity of the approach used here is the same as described by Chappell and co-workers (2007).
Examples

- (Chappell et al., 2007) Detecting induced in situ erodability of dust-producing playa in Australia using a bi-directional soil spectral reflectance model.

B. Spectral Mixture Analysis (SMA)

For more details see Spectral Mixture Analysis (SMA) in a previous section.

Examples

- (Hill et al., 1995) Land degradation, soil erosion and desertification monitoring in Mediterranean ecosystems.

C. Multivariate pattern recognition

Detection methods may include multivariate pattern recognition algorithms that operate either on direct or transformed spectral variables. For example, polynomial coefficients (Hill and Schuett, 2000) that describe the curvature of the spectral continuum between 0.4 and 1.6 \( \mu \)m to derive organic carbon content of soils from their spectral continuum. Parametrizing soil reflectance spectra with variables related to specific shape characteristics of the spectral profile permits organic carbon concentrations in soils to be estimated on the basis of regionally validated regression models.

Examples


D. Spectral shape (organic matter)

Since the specific influence of organic matter on spectral reflectance is not expressed in narrow absorption bands, as is for example the case for iron oxides or carbonates, but determines the overall shape of the reflectance between 0.35 \( \mu \)m and 1.4 \( \mu \)m, it is proposed to parametrize the shape of the reflectance continuum in this wavelength range through the coefficients of the following third-order polynomial:

\[
p_\lambda = b_0 + b_1 \lambda + b_2 \lambda^2 + b_3 \lambda^3
\]
This polynomial fitted to the spectra can usually approximate the reflectance continuum of soils and rocks quite accurately. Whereas the constant term $b_0$ of the polynomial relates to the intersection of the curve with the ordinate, coefficients $b_1$-$b_3$ describe the slope and curvature of the function. It results from a detailed statistical analysis of spectral data that coefficients $b_1$ and $b_3$ are the most important variables for the characterization of the reflectance continuum, as a function of the organic carbon concentration. Then the coefficients $b_1$ and $b_3$ of the resulting equations are introduced as independent variables into a multiple linear regression analysis to determine the corresponding soil organic matter concentration. Accounting for nonlinearities in the relationship, the logarithm of soil organic matter concentration ($\ln C_{org}[\text{wt\%}]$) can be modeled from the continuous spectra with the following model:

$$\ln C_{org}[\text{wt\%}] = -1.885 + 1.591*b_1 - 6.552*b_3 .$$

**Examples**


**E. Regression analysis**

As part of a research project associated to the European RTD Project DeSurvey: A surveillance system for assessing and monitoring desertification (http://www.desurvey.net), regression functions were developed built on CIE color coordinates and specific absorption features to derive corresponding maps from atmospherically corrected HyMap imagery (Jarmer et al., 2007). For more methodological details the reader is referred to Regression analysis in a previous section of the document.

**Examples**

- (Jarmer et al., 2007) The use of hyperspectral remote sensing data for the assessment of chemical properties of dryland soils in SE Spain.

**F. USGS Tetracorder**

The reader is referred to USGS Tetracorder of a previous section for more methodological details.

**Examples**

- (Kokaly et al., 2007) Characterization of post-fire surface cover, soils, and burn severity at the Cerro Grande Fire, New Mexico, using hyperspectral and multispectral remote sensing.
G. Soil color spectral indices

Color indices are based on three bands representing the red (R, 693 nm), the green (G, 565 nm) and the blue (B, 447 nm) spectral range (Mathieu et al., 1998):

- Brightness Index, \( BI = \sqrt{\frac{B^2 + G^2 + R^2}{3}} \)
- Coloration Index, \( CI = \frac{R - G}{R + G} \)
- Hue Index, \( HI = \frac{2 * R - G - B}{G - B} \)
- Redness Index, \( RI = \frac{R^2}{B * G^3} \)
- Saturation Index, \( SI = \frac{R - B}{R + B} \)

Examples

- (Zhuo et al., 2008) Quantitative mapping of soil organic material using field spectrometer and hyperspectral remote sensing.
- (Mathieu et al., 1998) Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typic soil color of an arid environment.

2.6.3 Soil genesis and formation

One innovative field IS can add to its credit is its ability to track chemical and physical properties that can serve as tracers to monitor soil genesis and formation. For example, increasing redness is frequently associated with increasing age of sediments and has been used to suggest different source materials. It has also been linked to regional climatic gradients and to degrees of aeolian and fluvial abrasion.

One of the chemical tracers used in this direction is the free iron oxides (species and content) that may represent, for example, soil rubification process. Rubification is defined as a stage in pedogenesis in which iron is released from primary minerals to form free iron oxides that coat quartz particles in soils with a thin reddish film. The free iron oxides coat the quartz particles and provide a reddish color to the matrix as well as stability (Ben-Dor and Singer, 1987). The accurate assessment of redness can
therefore provide surrogate information about the concentration of iron oxides in sediments. The quantification of iron oxides is possible by means of spectral radiometric measurements in the laboratory. Quantification of iron oxides was also found possible in the field (Ben-Dor et al., 2006; Clark and Swayze, 1995; Palacios-Orueta et al., 1999).

Soil formation is followed not only by the release of iron oxides but also by the production of clay minerals. (Ben-Dor et al., 2006) plotted the absorption of the free iron oxides a 0.5 µm against the absorption of OH in clay lattice at 2.2 µm (as measured with an ASD) and found a significant relationship between the two traces. This shows that clay can also serve as an indicator of pedogenesis process over sand dune areas.

Surface reflectance is also controlled by two main superficial processes, which are used as relative age criteria (Crouvi et al., 2006b): the degree of desert pavement development (gravel coverage %) controls the absorption feature depths, while the rock coating development influences significantly the overall reflectance of the surface, but its effect on the absorption feature depths is limited. The predominant features are associated to carbonates around 2.33 µm, ferric features in the VNIR region and a weak Al-OH feature at 2.21 µm. The variations in the spectral signature of alluvial surface was found to be controlled by two main processes: the desert pavement development controls mainly the location and depth of the different absorption features, while the rock coating accretion controls mainly the reflectance of the surface, through the reflective region (0.4-2.5 µm). As the desert pavement develops, the absorption feature of the dominant gravels (in this case carbonate) is increased, while the absorption features of the fines (in this case Al-OH and ferric) are decreased.

A. Soil color spectral indices

For a more detailed description of the indices, see Soil color spectral indices (section G above). These spectral indices are used in a correlation analysis with Fe-dithionite-citrate-bicarbonate (DCB) to describe and quantify for instance rubification. The Redness index (RI) has proven to be correlated with Fe-DCB.

Examples

- (Ben-Dor et al., 2006) Quantitative mapping of the soil rubification process on sand dunes using an airborne hyperspectral sensor.
- (Mathieu et al., 1998) Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typic soil color of an arid environment.
B. Feature characteristics (band position, depth, area, etc.)

Scaled absorption band depth is used as a spectral parameter for soil rubification estimates. After continuum removal, two absorption depths are used. The first depth is based on the wavelengths of 0.489, 0.556 and 0.59 µm and the second is based on the wavelengths of 0.477, 0.534 and 716 µm (Ben-Dor et al., 2006). These spectral parameters are then used in a correlation analysis with Fe-dithionite-citrate-bicarbonate (DCB) to describe soil rubification. Carbonate absorption depth (2.33 µm) and Al-OH absorption depth (2.21 µm) as well as ferric absorption slope are also used. For more details on continuum removal analysis, the reader is referred to Feature characteristics in a previous section of this document.

Examples

- (Richter et al., 2009) Free iron oxide determination in Mediterranean soils using diffuse reflectance spectroscopy.
- (Ben-Dor et al., 2006) Quantitative mapping of the soil rubification process on sand dunes using an airborne hyperspectral sensor.
- (Crouvi et al., 2006a) Quantitative mapping of arid alluvial fan surfaces using field spectrometer and hyperspectral remote sensing.
- (Bullard and White, 2002) Quantifying iron oxide coatings on dune sands using spectrometric measurements: an example from the Simpson-Strzelecki Desert, Australia.

C. Linear regression analysis

For more methodological details, see Regression analysis in a previous section of this document.

Additionally, feature characteristics as determined in paragraph B (above) are then used in regression analysis for quantitative estimations of iron content, clay content, carbonate content, % gravel coverage and others.

Examples
• (Ben-Dor et al., 2006) Quantitative mapping of the soil rubification process on sand dunes using an airborne hyperspectral sensor.

• (Bullard and White, 2002) Quantifying iron oxide coatings on dune sands using spectrometric measurements: an example from the Simpson-Strzelecki Desert, Australia.


D. Multivariate analysis

For full description, see Multivariate pattern recognition in a previous section of this document.

Examples


2.6.4 Soil moisture

Moisture content is another interesting parameter that can give insights on the hydrological cycle for example. Soil moisture plays a very important role in the exchange of mass and energy between the surface and the atmosphere as part of the water cycle.

Spectral information related to moisture content in soils or sediment will appear in two ways. First, the strength of the specific liquid water absorptions (1.4 and 1.9 µm for example) will increase with moisture content. Second, moisture content influences the level of reflectance in the VNIR: the higher the moisture content, the lower the reflectance, and in the SWIR where the consequence of higher moisture content is lower reflectance with increasing wavelength. Reflectance decreases when moisture content increases but this is true only for low moisture levels. For high moisture levels, reflectance can increase with moisture, which is rather surprising. This is tentatively explained by some scattering effects due to water molecules in the soil structure. This inversion happens passed a critical point which seems to be related to soil hydrodynamic properties such as hygroscopic limit, field holding capacity and others.

A. Feature characteristics (band position, depth, area, etc.)
For soil moisture determination, the absorption features of interest are the 1.4 and 1.9 µm main water absorption bands (in the case of laboratory spectral measurements), the secondary water absorption bands at 0.96 and 1.16 µm and the 2.2 µm OH absorption feature related to clay content. For more methodological details see Feature characteristics in a previous section of this document.

Examples

- (Demattê et al., 2006) Determining soil water status and other soil characteristics by spectral proximal sensing.
- (Lobell and Asner, 2002a) Moisture effects on soil reflectance.

B. Stepwise linear regression

A stepwise linear regression is generally performed to reduce the number of bands and focus on the wavelengths that are more sensitive to water changes. This kind of approach can be justified because reflectance spectra of soils are usually rather smooth. The change in shape is often gradual with moisture so there is a strong correlation between bands. To emphasize the changes due to moisture, soil reflectance spectra can be normalized by the spectrum corresponding to the driest conditions. This also helps minimizing the effects of soil type and other multiplicative factors such as roughness or measurement configuration. Multiple regressions are also used to establish quantitative relationships between absorption band characteristics and soil moisture content as measured in the laboratory. More information on regression analysis can be found in a previous section of the document: Regression analysis.

Examples

- (Demattê et al., 2006) Determining soil water status and other soil characteristics by spectral proximal sensing.
- (Liu et al., 2002) Relating soil surface moisture to reflectance.

C. Exponential model

The effect of soil moisture on reflectance can also be summarized for each sample by determining the best-fit coefficients of an exponential model relating moisture and reflectance:

\[ R = R_{sat} + (R_{dry} - R_{sat}) \times \exp(-c \times wc) \]
Where $R_{sat}$ is the reflectance of saturated soil, $R_{dry}$ is the reflectance of dry soil, $c$ describes the rate of change because of soil moisture, $wc$ is the water content expressed as $\theta$ or $s$, and all values except $wc$ are wavelength dependent.

It has been shown that there is a clear exponential relationship between soil reflectance at 0.6, 1.2 and 2.2 μm and moisture content for different kind of soils (Lobell and Asner, 2002b).

**Examples**

- (Lobell and Asner, 2002b) Moisture effects on soil reflectance.

**D. Soil Moisture Gaussian Model (SMGM)**

The method used here (Whiting et al., 2004) is based on modeling of the shape of the fundamental water absorption feature centered at 2.8 μm. Convex hull points are selected in the SWIR part of the spectrum. The idea is then to use these points to fit an inverted gaussian function that will model the water absorption feature following the equation:

$$g(\lambda) = R_{sl} + (R_{\lambda o} - R_{\lambda i}) \exp\left(\frac{-(\lambda - \lambda_o)^2}{2\sigma^2}\right)$$

The Gaussian function is fitted to the SWIR region from the maximum reflectance wavelength to the function’s center. The wavelength of the functional center us beyond the range of laboratory or field instruments so the fitting algorithm extrapolates the function to the point of minimum reflectance, or the maximum absorption, $\lambda_o$, here the position of the fundamental water absorption at 2.8 μm.

The inverted amplitude ($R_d = R_{\lambda o} - R_{\lambda i}$) and area $A$ of this Gaussian function is then estimated using the following equation:

$$A = R_d \sigma \sqrt{\frac{\pi}{2}} \text{erf}\left(\frac{\lambda_o - \lambda_i}{\sqrt{2} \sigma}\right)$$

where $R_d = R_{\lambda o} - R_{\lambda i}$ and erf denotes the Gaussian error function.

Regressions are then developed between soil moisture content and amplitude and area of the Gaussian function. A general, universal, relationship can be established for all analyzed soil types but specific relationships, more precise, can be established for each soil type. These regressions can potentially be
used to correct the effects of soil moisture in hyperspectral images and increase precision on extraction of other parameters which are spectrally active in the SWIR such as presence of clays or carbonates.

**Examples**

- (Whiting et al., 2004) Predicting water content using Gaussian model on soil spectra.

**E. Normalized Soil Moisture Index (NSMI)**

Haubrock et al., (2004, 200) have developed and successfully tested a new model for determining soil moisture by means of remote sensing techniques. It is based on a combination of multi-temporal high spatial resolution IS observations with field and laboratory spectral studies, along with hydrological measurements. The method is called Normalized Soil Moisture Index, and it was tested for the best spectral prediction of soil moisture content (MC) in the field using the 400–2500 nm region. The coefficient of determination ($R^2$) ranges from 0.61 up to 0.71 for natural field samples taking into account the influence of different environmental factors: heterogeneous soil types and related MC, variable soil water profiles, and the presence of soil crust and vegetation cover. This index allows the production of surface soil moisture maps from IS images which were found to be highly correlated with the field MC measured at the time of the overflight. This index could be used to develop global, regional or local soil moisture maps neede for different applications (erosion modeling, desertification) because it does not need a priori knowledge that usually is requirement, such as soil type or dry reflectance. The authors suggested that this index or similar ones are potentially widely applicable as long as soils are well exposed at the surface.

**Examples**

- (Haubrock et al., 2008) Surface soil moisture quantification models from reflectance data under field conditions.
- (Haubrock et al., 2004) Application of hyperspectral imaging and laser scanning for the monitoring and assessment of soil erosion in a recultivation mining area.

**F. Reconstruction of soil spectral signatures through the use of various water film depths**

Bach and Mauser (1994) were able to simulate the reflectance change of the soil spectra from dry to moist. They combined the model by Lekner and Dorf (1988) for internal reflectance with the absorption
coefficients from (Palmer and Williams, 1974) into Beer-Lambert’s law to simulate the moist reflectance \( R \) from the dry reflectance \( R_0 \) by the exponential of the absorption coefficient and an empirically determined active thickness \( I \) in the following equation:

\[
R = R_0 e^{-\alpha*I}
\]

**Examples**

- (Bach and Mauser, 1994) Modeling and model verification of the spectral reflectance of soils under varying moisture conditions.

**G. SOILSPEC**

Jacquemoud et al. (1992) using a modification of Hapke’s single scattering albedo model (Hapke, 1981) separated the surface geometry component in a radiative transfer model for soil reflectance, called SOILSPEC. They also noticed quasi-homothetic variations in the VNIR and SWIR with the moisture content, though the moisture data set was limited. Future investigations to account for this albedo decline may help resolve this modeling problem (Pinty et al., 1998).

**Examples**

- (Jacquemoud et al., 1992) Modeling spectral and bidirectional soil reflectance.

### 2.6.5 Soil grain size

Most of the published studies deal with grain size variations in desert environment, mainly sandy aeolian deposits. In general, it is well known that reflectance of soil surface gradually increases when grain size decreases due to increasing scattering. In some cases, when soil composition is homogeneous, the depth of specific absorption features can also be used to estimate grain size. In this case, absorption increases when grain size increases because of increasing path length of light through the grains.

**A. Spectral indices**
(Xiao et al., 2005) compared the correlations between several spectral indices and soil grain composition and found a new index that correlated best with topsoil fine sand content, the Grain Size Index (GSI):

\[ GSI = (R-B)\times(R+B+G) \]

Where \( R \), \( B \), and \( G \) are red, blue and green bands of remote sensing data. This index showed its potential to detect fine sand area and help monitoring desertification process in arid regions.

**Examples**

- (Xiao et al., 2005) Detection of land desertification and topsoil grain size using remote sensing.

**B. Radiative transfer (RT) modeling**

Calculations of scattering in desert soils are based on the assumption that scattering by irregularly shaped grains can be matched by that of soil spheres in each other’s far field. Arguments for assuming sphericity are presented by Mugnai and Wiscombe (1989); however, the size parameters for the particles discussed are much smaller than those in soils. The greatest potential error form the sphericity assumption lies in its effect on the structure of the single scattering phase function and in turn, in the structure of the hemispherical-directional reflectance distribution function of soils. The radiative transfer model results were used by (Okin and Painter, 2004) to validate the change in reflectance with a change in particle size and because they did not attempt to accurately model the directional reflectance of soils.

They used the Mie theory to calculate the single-scattering properties of desert grains: scattering phase function, single-scattering albedo, and extinction efficiency. Each sphere was assumed to have a core of quartz with radii ranging from 0.05 to 0.75 mm and a rind of montmorillonite and hematite with a thickness of 0.005 mm. The complex refractive indices for quartz, montmorillonite and hematite came from data compiled by (Sokolik and Toon, 1999). With the single scattering properties, it is then possible to model the spectral reflectance of desert soils in this case with the discrete ordinates solution to the radiative transfer equation. The irradiance components (diffuse and direct) may be modeled using a standard atmospheric model in an atmospheric transmission model such as DISORT (Ricchiazi et al., 1998) or MODTRAN (Berk et al., 1989). With these boundary conditions and the single scattering properties described before, it is possible to model the radiance field reflected from a homogeneous
sandy layer. The radiance field allows then the calculation of the hemispherical-directional reflectance distribution function (HDRDF) and the hemispherical-directional reflectance factor (HDRF).

Examples


C. Feature characteristics (absorption band depth)

For a full description, see Feature characteristics in a previous section of this document.

Examples

- (Ghrafat et al., 2007) Modeling grain size variations of Aeolian gypsum deposits at White Sands, New Mexico, using AVIRIS imagery.

2.7 URBAN SPHERE

Hyperspectral data are very useful for the characterization of urban and periurban environments with respect to many of the elements affecting their ecological behaviour. We mention sealed parts, terrain and soil composition, dangerous materials used in built up blocks or left in uncontrolled areas, trees as well as lower vegetation and other biotopes. Some of these elements may be studied by means of high spatial as well as spectral resolution data and all of them require at least a spectral knowledge of the studied area. Indeed, in the following we will list a few examples of the realization of this kind of analysis by using hyperspectral data. In particular, we will refer here to urban material characterization, vegetation canopies and tree analysis, as well as to land cover and sealed parts detection. In all of these papers the advantages of using hyperspectral data for these tasks are clearly acknowledged.

However, some of these urban ecology analyses could give more refined results (and so better inputs to ecological models) by using not hyperspectral measurements alone, but rather a combination of data coming from different sensors. This could be done, for instance, by complementing hyperspectral mapping with 3D terrain models coming from interferometric SAR or by combining multispectral, radar and LIDAR measurements. In other words, to make more effective the proposed analysis, a multi-sensor urban research area is required. For further reading about urban remote sensing we refer to the textbook of (Weng, 2007).
2.7.1 Urban materials and biotopes

The urban landscape is the manifestation of both natural and human processes expressed in intricate geometry and corresponding land use and cover mixtures that have relatively unpredictable spatial patterns. In many ways, our understanding of the urban structure, environmental parameters and function of natural systems exceeds our comprehension of the structures and dynamics of urban areas. One focus of necessary research recognized by several recent global change reports involves the need for regional scale assessments of human vulnerability and ecosystem sustainability to activities concentrated in and near urban areas. Researchers in hyperspectral data analysis for urban material characterization are proposing to develop a methodology to specify, monitor and understand the urban landscape. Indeed, with hyperspectral data we are able not only to characterize the urban geometry, but also to individuate the interactions among soil composition, vegetation characteristics and generic natural/artificial materials used in urbanized areas. The output of hyperspectral data analysis are many of the elements of the urban landscape that affect ecological models and that are mandatory, for instance, for flood risk assessment and loss estimation models that consider urban asset vulnerability and exposure.

A. Spectral similarity

The idea of spectral distance originates in the need of many hyperspectral applications to use measures for assessing the similarity (or distance) between spectra, or between a spectrum and a group of spectra. Distance measures are frequently used in classification as well as in target detection. Most commonly distance-based metrics are used to measure the distance between spectral signatures and to quantify the degree of spectral similarity and discriminability, such as the Euclidean (Robila, 2005) or Bhattacharyya Distance (Kailath, 1967). Other techniques to assess similarity are based on orthogonal projection measures, such as the frequently used spectral angle mapper (SAM) (Robila, 2005) or statistical spectral measures, for instance the spectral information measure (SIM) or the Hidden Markov Model (HMM)-based measure. More details can be found in the textbook of (Chang, 2003, 2007).

Examples

- (Herold et al., 2004) Spectrometry for urban area remote sensing - Development and analysis of a spectral library from 350 to 2400 nm.

B. Spectral mixture analysis (SMA)
For methodological details see **Spectral mixture analysis (SMA)** in the NPV section.

**Examples**

- (Segl et al., 2000) Differentiation of urban surfaces based on hyperspectral image data and a multi-technique approach.
- (Xiao et al., 1999) Characterization of the structure and species composition of urban trees using high resolution AVIRIS data.

**C. Continuum removal analysis**

For methodological details we refer to **Continuum removal analysis** in the pigments section.

**Examples**

- (Bassani et al., 2007) Deterioration status of asbestos-cement roofing sheets assessed by analyzing hyperspectral data.

### 2.7.2 Mapping and cartographic feature extraction

Hyperspectral data are an excellent input for urban remote sensing. However, classification of remote sensing data with very high spatial resolution in urban areas is a challenging problem. It is not only necessary to use the available spectral information but also to exploit the spatial information. Land cover mapping of urban hyperspectral data has been considered in some recent works. Approaches based on spectral signature recognition have been proposed. In addition, papers address spectral mixture analysis of hyperspectral data or consider the identification of spectral signatures in mixed pixels. Even though spatial analysis has been attempted for hyperspectral data in these urban environments, it remains difficult because of the commonly low ground resolution. As a consequence, the availability of fine spatial resolution hyperspectral data requires the development of new algorithms or the improvement of already proposed techniques to use the complete spectral and spatial characteristics of a pixel.

**A. Spectral mixture analysis (SMA)**

For methodological details see **Spectral mixture analysis (SMA)** under the NPV section.
Examples

- (Ben-Dor et al., 2001) A spectral based recognition of the urban environment using the visible and near-infrared spectral region (0.4-1.1 \(\mu\)m). A case study over Tel-Aviv, Israel.

**B. Data Fusion**

The urban environment is so complex and challenging that it is optimistic to assume that a single sensor can provide all the information that one may require for its characterization. Multiple data sets are becoming available routinely on urban test sites, with ground resolutions that may range from several tens of meters to less than one meter. The very first challenge posed by such large amounts of data is the establishment of suitable application-driven criteria for usable data selection. An associated challenge is the development of automatic or semi-automatic techniques capable of extracting vector layers from raster data, in order to integrate them in a Geographic Information System (GIS). As a matter of fact, the use of remotely sensed information in geographic application is increasing, and notwithstanding the fact that urban applications often require data at the finest ground resolutions, many interesting characterizations are feasible using even lesser resolution data from existing sensors (Gamba et al., 2005). Data fusion aims at achieving the highest potential of the fused data based on the spectral, spatial and temporal properties. The literature of data fusion covers a wide range of data, applications and techniques (Gamba et al., 2005; Pohl and Van Genderen, 1998).

Examples

- (Madhok and Landgrebe, 1999) Supplementing hyperspectral data with digital elevation.
- (McKeown et al., 1999) Fusion of HYDICE hyperspectral data with panchromatic imagery for cartographic feature extraction.
- (Gamba and Houshmand, 1999) Integration of high resolution multispectral imagery with LIDAR and IFSAR data for urban analysis applications.

**A. Spectro-spatial analysis (Mathematical morphology)**

High-resolution hyperspectral remote sensing data provide both detailed structural and spectral information about urban scenes. The classification of hyperspectral urban data is a challenging problem for two main reasons. First, the hyperspectral data contain a lot of information about the spectral properties of the land cover in the data, but no spatial information is inherent in the spectral data. Second, the individual images from an urban scene contain spatial information but very limited
information about the spectral nature of the data. Consequently, a joint spectral/spatial classifier is needed for classification of urban hyperspectral data, meaning a morphological method that is based on making use of both the spectral and spatial information for classification.

Mathematical morphology (MM) is a theory for the analysis of spatial structures which was initiated in the late sixties (Serra, 1982). It stems from the study of the geometry of binary porous media such as sandstones. They are binary in the sense that they are made up of two phases: the pores embedded in a matrix. It is called morphology since it aims at analysing the shape and form of an object. A wide variety of morphological transformations are available for extracting structural information in spatial data. Morphological operations are grounded on set operators such as union, intersection, complementation and translation (Soille and Pesaresi, 2002). For more information on mathematical morphology we refer to (Serra, 1982; Soille, 1999) and http://web.ukonline.co.uk/soille/.

Examples

- (Benediktsson et al., 2005) Classification of hyperspectral data from urban areas based on extended morphological profiles.
- (Dell'Acqua et al., 2004) Exploiting spectral and spatial information in hyperspectral urban data with high resolution.

REFERENCES


Gamba, P., & Houshmand, B. 1999. *Integration of high resolution multispectral imagery with LIDAR and IFSAR data for urban analysis applications*. Procs. of ISPRS Workshop Mapping Surface Structure and Topography by Airborne and Spaceborne Lasers, La Jolla, USA.


<table>
<thead>
<tr>
<th>Algorithm Theoretical Basis Document</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doc. N.: D6.1</td>
</tr>
<tr>
<td>Issue: 1.0</td>
</tr>
<tr>
<td>Date: FEB 2009</td>
</tr>
<tr>
<td>Page: 109 of 117</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Algorithm Theoretical Basis Document</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doc. N.: D6.1</td>
</tr>
<tr>
<td>Issue: 1.0</td>
</tr>
<tr>
<td>Date: FEB 2009</td>
</tr>
<tr>
<td>Page: 111 of 117</td>
</tr>
</tbody>
</table>


© HYPER-I-NET February 09


