1. Introduction

Soil reflectance spectroscopy is a well-known technique to assess soil properties rapidly and quantitatively in both point (spectroscopy) and spatial (imaging spectroscopy (IS)) domains. The quantitative approach has been developed in the past two decades by many researchers, with much literature on the topic. Basically, this approach was adopted from other disciplines (e.g. food, textiles), whereas the mapping approach was incorporated with spectral-based remote sensing means. Soil contamination is an ever-growing concern, and a reliable and environmentally friendly method is needed to rapidly detect and monitor soil contaminants at any stage for the diagnosis of suspected contaminated areas, as well as for controlling rehabilitation processes. Conventional methods for investigating soil contamination based on point sampling and chemical analysis are time-consuming, costly and sometimes incorporate the use of environmentally damaging chemicals. Soil reflectance spectroscopy is a promising tool for the efficient detection and monitoring of soil contaminants. Using reflectance spectroscopy measurements, several direct and indirect soil properties, as well as soil contamination characteristics, can be extracted. Due to the vast amount of data in soil spectroscopy, be it point or image, data-modeling processes are required. This chapter reviews soil spectroscopy theory and its quantitative capabilities, as well as data-modeling methods for soil spectra that are related to soil contamination. Discussed contaminants include petroleum hydrocarbons, heavy metals such as Ni, Cr, Cu, Cd, Hg, Pb, Zn and As, acid mine drainage and pesticides. Limitations, obstacles and problems in recording soil spectra in the field and in the laboratory are discussed, along with IS technology. The future potential of spectral technology and IS missions used for soil contamination monitoring is presented, based on ongoing worldwide initiatives. This chapter also provides some notes on how to merge this technology with other systems, such as a penetrating spectral device for underground measurements utilizing the catheterization approach.

1.1 Energy matter interactions

Most, if not all matter interacts with electromagnetic radiation in one way or another. Electromagnetic radiation that is radiated onto any kind of material may be reflected, absorbed or transmitted. For most materials, specific light interactions of a certain
wavelength can be observed in the same range as solar radiation (Hunt 1980). For that reason, this chapter focuses on this radiation range (VNIR-SWIR, 400 to 2,500 nm). This range is divided into two sub-ranges: VNIR (400 to 1,100 nm) and SWIR (1,100 to 2,500 nm). Interactions between light and matter are wavelength-dependent. While electromagnetic radiation radiated onto a specific material may be reflected at one wavelength, it might be absorbed at another. This means that the reflected radiation of a material is a function of wavelength, surface material absorption, emissivity and reflectance characteristics. Therefore, energy-matter interactions in the VNIR, SWIR and IR (2,500 to 12,000 nm) wavelength regions can be used for qualitative and quantitative chemical analysis of a material.

1.2 Spectroscopy
VNIR-SWIR and IR spectroscopy has been used for qualitative and quantitative analysis for several decades (Taylor et al. 1970; Stark et al. 1986). Both are based on characteristic vibrations of chemical bonds in molecules, as well as electronic processes (the latter affects only radiation at 400 to 1,000 nm) (Ben-Dor et al. 1999). In general, incident IR is absorbed by the chemical-bond vibrations in the molecules (3N-6 fundamental possibilities where N is the number of atoms in a given molecule). The combination modes and overtones of the fundamental vibrations in the IR region are observed in the VNIR-SWIR region, and can be detected if the incident radiation in this latter region interacts with the matter in question. Whereas using the IR region in the laboratory requires sample preparation and caution during the measurements (e.g. not to exceed the saturation point), the VNIR-SWIR region has the advantage of allowing direct measurement of the material without complicated pretreatment and reagents. Another advantage of VNIR-SWIR over IR spectroscopy is that it is a non-destructive and non-contact method: it allows real-time measurements based on a simple illumination source (e.g. sun, tungsten halogen bulb). It is therefore suitable for online, in-situ monitoring and analysis of many types of compounds, mixtures and materials. In contrast, IR spectroscopy is a laboratory analysis method that requires considerable sample preparation. Its advantage over VNIR-SWIR lies in the fact that all chemical bonds within the molecule are active in this region, whereas in the VNIR-SWIR region only some of the bonds are active (e.g. OH-, NH-). Another benefit of IR is that it can be interpreted directly based on known chemical bonds, while in the VNIR-SWIR region, the spectra must be studied in advance to evaluate which wavelengths will best represent a given property (this is due to the possible overlap between some combination modes and overtones in the studied material). As a result, complex chemometric processes are used for the interpretation of VNIR-SWIR spectra into a quantitative model, in a process termed NIRA or NIRS (near infrared analysis). In this chapter, we will focus on hyperspectral data (continuous spectral data, amounting to hundreds of spectral bands), in the field of soil-contamination monitoring by reflectance spectroscopy and NIRS.

1.3 Soil spectroscopy
Today, the NIRS technique is well-known and utilized in many disciplines, such as food science, pharmacology, textile, tobacco, and oil industries, among others. Since the mid-1960s, when Bower and Hanks showed that soil moisture is highly correlated with soil spectra (Bowers & Hanks 1965), soil has captured the attention of many researchers who have realized that despite its complexity, soil spectroscopy can provide remarkable amounts of quantitative information. Stoner et al. (1980) were the first scientists to systematically
gather soil spectral information and publish it in the form of a soil spectral atlas. Their soil spectral library very soon became a classic tool that soil scientists came to rely on. Later, when laboratory and portable field spectrometers were introduced into the market (around 1993), more scientists realized the potential of soil spectroscopy, and consequently more spectral libraries were assembled and new quantitative, chemometric applications, such as NIRA, were developed and implemented for various soil materials. A summary of soil reflectance theory and its applications can be found in various publications (Irons et al. 1989; Ben-Dor et al. 1999; Ben-Dor 2002). A study by Brown et al. (2006) showed that NIRA can work successfully under a generic global view rather than relying on a specific soil population. An extensive study with over 1,100 soil samples, using several data-mining algorithms to model and interpret diffuse reflectance spectra of soils and predict a number of soil attributes was recently performed by Rossel & Behrens (2010).

1.4 Imaging spectroscopy and remote sensing
Imaging spectroscopy (IS), or hyperspectral remote sensing, is an advanced tool that generates data of high spectral resolution, with the aim of providing near-laboratory-quality reflectance (or emittance) for each single picture element (pixel) from a far distance (Vane et al. 1984). This information enables the identification of objects based on the spectral absorption features of chromophores and has been found very useful in many terrestrial and marine applications (Clark & Roush 1984; Dekker et al. 2001). IS brings a new dimension to the field of remote sensing by expanding the envelope of point spectrometry to the spatial domain. It provides a tangible perspective by adding spatial detail to spectral information, thereby enhancing the thematic application of spectral recognition algorithms. This capability exists for both far and close distances, such as data acquired by satellites or by microscopic sensors, respectively. Whereas the former is used for mapping the earth from space, the latter is used for mapping micro targets, such as microorganisms and cell bodies, to account for their biochemical processes in a spatial domain (Soenksen et al. 1996; Levenson & Farkas 1997). It is interesting to note, however, that although soil scientists have recognized the potential of reflectance spectroscopy and in fact termed it a novel technology, in many ways, the use of IS for soil applications remains undeveloped and is seldom reported. Though the IS approach is a cost-effective method, its adoption is limited because the data is difficult to process, only a few sensors are operated worldwide, and it has not yet been recognized by many end users. Hence, the journey from point spectroscopy to a cognitive (imaging) spectral view of soils has not yet been completed, although there is no doubt that it may open up new frontiers in the field of soil science. Thus far, only exclusive and select groups around the world have been able to use IS for soil applications. Nevertheless, over the past 10 years, these groups have demonstrated remarkable achievements and have documented its significant capability. For further information refer to the chapter on “Optical remote sensing techniques for soil contamination monitoring and vulnerability assessment” in this book.

1.5 Acquiring reflectance information from soil
To acquire reflectance information from the VNIR-SWIR region, several sensors, methods, protocols and platforms are used. Basically, the VNIR-SWIR sensors consist of an apparatus (with filters, grating, prisms, interferometer, etc.) that splits the measured radiation into individual wavelengths (or regions), several detectors (mostly CCDs) which are sensitive to
different regions), and various other optional optical components (lenses, collimators, beam-splitters, integrating spheres, optical fibers, etc.). The radiation source is usually a tungsten halogen light bulb when recording spectra under laboratory conditions, and the sun when recording spectra in the field. These sensors are generally periodically calibrated to provide the radiance values but in practice, the reflectance is calculated against a white reference target (Spectralon) that is assumed to reflect 100% of the radiation in the sample geometric configuration. There are point and image sensors. The point sensor records a single spectrum of the target (with a line-array detector) while averaging the mixed information of all components within the sensor’s field of view. The size of the single pixel being measured depends upon the optics and the distance from the target. The image sensor provides data similar to the point sensor except that an area-array detector is used that enables the acquisition of spectral-cube information constructed from spatial and spectral views of the area (target) in question. Image sensors usually have lower spectral resolution as well as lower signal-to-noise ratio, as the integration time over the targets is less than in the point spectrometer measurements. Laboratory sensors are more accurate, with higher spectral resolution and a better signal-to-noise ratio: the integration time is long and the geometric scheme is constant. They usually include a sample holder and a stable light source. Field sensors are generally battery-operated and are less accurate than laboratory sensors. They rely on the sun's radiation and the geometric measurement is not fixed. Nonetheless, they provide flexibility and in-situ measurement capabilities that cannot be achieved under laboratory conditions. The field sensors are a valuable tool for recording reference spectra for space- or airborne remote sensing imagery which will be later used for calibration, validation and quality control (Brook & Ben-Dor 2011). Image sensors can be operated from both ground and air (space) domains. They can be integrated with the aircraft’s GPS/INS system and produce multi or hyperspectral data cubes that can be georeferenced; thus each pixel in the image corresponds to a single area unit on the ground and is represented by a spectrum. The signal-to-noise ratios of air- and spaceborne sensors are lower than those operated on the ground and are governed by many factors (e.g. pixel size, integration time, frame rate, electronic noise, IFOV and atmosphere attenuation). A comprehensive review of image-based sensors for soil application can be found in (Ben-Dor et al. 2008).

2. Quantitative applications of spectroscopic data
Spectroscopic data are multivariate in nature. There are two approaches: 1) the chromophore absorption band in question is known and under saturation (supervised), and 2) the chromophore absorption band is unknown and is affected by several factors (unsupervised). Whereas the first approach is valid for the IR region by using selected wavebands and pretreatment to prevent saturation, the second represents the case of reflectance spectroscopy across the VNIR-SWIR region. For that purpose, multivariate statistical techniques (also called chemometrics) are required to extract the information about the quality attributes that is hidden within the spectral information. Essentially, this involves regression techniques coupled with spectral preprocessing.

2.1 Preprocessing
Spectral preprocessing techniques are used to remove any irrelevant information which cannot be handled properly by the modeling techniques. The preprocessing techniques include averaging, centering, smoothing, standardization, normalization and
transformations, among others. Some of the most common preprocessing techniques are presented here. For a full review of preprocessing methods, please refer to Nicolaï et al. (2007).

2.1.1 Multiplicative scatter correction
Multiplicative scatter correction (MSC) is one of the most commonly used normalization techniques. In MSC, the light scattering is estimated for each sample relative to an ideal sample obtained by averaging the complete wavelength range of the data set. Each spectrum is then corrected such that all samples appear to have the same scatter level as the reference spectrum. For a further description of this method, see (Isaksson & Naes 1988; Naes et al. 1990; Vohland et al. 2009).

2.1.2 Absorbance
Spectroscopic measurements performed in transmission mode can be quantified using Beer’s Law (concentration $\sim \log(1/T)$ where $T$ is the transmitted radiation). Accordingly, reflectance measurements are frequently converted to $\log(1/R)$ values, which are then used in a manner similar to optical density readings (Nicolaï et al. 2007).

2.1.3 Kubelka-Munck
The Kubelka-Munck transformation, $(1-R^2)/2R$, is similar to absorbance but accounts for scattering (Nicolaï et al. 2007).

2.1.4 Continuum removal
The continuum is the background absorption on which other absorption features are superimposed. The spectrum is divided by a convex hull that is fitted over it (Clark & Roush 1984; Clark & others 1999).

2.1.5 Derivatives
Derivation is often used to remove baseline shifts and superposed peaks. Second-derivative spectra can correct for both additive and multiplicative effects (like MSC). They are usually calculated according to the Savitzky-Golay algorithm (Naes et al. 2002). The parameters of the algorithm (interval width, polynomial order) should be carefully selected to avoid amplification of spectral noise (Nicolaï et al. 2007).

2.1.6 Spectral mixture analysis
Spectral mixture analysis (SMA) is a widely used method to determine the sub-pixel materials that fundamentally contribute to the spectral signal of mixed pixels. This is of particular importance for obtaining quantitative estimates of distinct materials, a typical application of remote sensing hyperspectral data. SMA aims to decompose the measured reflectance spectrum of each pixel into the proportional spectral contribution of so-called endmembers (EMs). In recent years, many authors have proposed and used a more complex model, in which both the number and the set of EMs vary dynamically on a per-pixel basis; this has become known as multiple EM SMA (MESMA). The idea consists of restricting the large set of possible EMs to a small set of better suited EMs, which can be different for each pixel, thereby allowing an accurate decomposition using a virtually unlimited number of
EMs. An improved strategy, called variable MESMA (VMESMA), allows segmentation of the image to increase flexibility and accuracy (Kemper & Sommer 2003).

2.2 Modeling

Modeling refers to relating a set of spectral parameters that are derived from the spectral information (before or after the aforementioned preprocessing treatment), to the real chemistry of the material in question. This is done by using a set of well-known samples as a training group. The data are divided into three groups: training, validation and test. The relationship between the chemistry and the spectroscopy data is found via the training group and simultaneously cross-validated by the validation group. Finally, the model is applied to the test group, independent of the training and validation process. Multivariate regression techniques are modeling methods that search for the relationship between two matrices: the spectral data matrix that can be very complex due to large amounts of data (X variables, the independent data), and a specific chemical reference value data matrix (Y variables, the dependent data). The common multivariate regression techniques are presented herein. For an in depth reference, please read further at (Esbensen et al. 2002; Nicolaï et al. 2007).

2.2.1 Multiple linear regression

Multiple linear regression (MLR) is a classical method that creates a linear combination of the spectral values at every single wavelength to correlate as closely as possible to the dependent reference values. The regression coefficients are estimated by minimizing the error between predicted and observed response values in a least squares sense. MLR models typically do not perform well with spectral data because spectral data usually exhibit high co-linearity, noise and more variables, i.e. more spectral bands, than measured samples (Esbensen et al. 2002; Nicolaï et al. 2007).

2.2.2 Principal component regression

Principal component regression (PCR) is a combination of principle component analysis (PCA) and MLR. The independent data matrix (spectral data) is transformed by PCA, and the first few principal components (PCs), which represent most of the independent data variance, are used as inputs for the MLR model instead of the original spectral data. The advantage over standard MLR is that PCs are uncorrelated, and the noise is filtered. The first few PCs are usually sufficient for a robust model and over-fitting issues can be eliminated. Although PCR consists of the two most studied multivariate methods (PCA and MLR), the major criticism against it is that the PCs (several first PCs) selected for the MLR input are not necessarily the best predictors for the reference data. There is no guarantee that the first PCs will include the spectral data related to the specific dependent variable that needs to be modeled (Esbensen et al. 2002; Nicolaï et al. 2007).

2.2.3 Partial least squares regression

Introduced in 1983 by Wold et al., partial least squares regression (PLS) is similar to PCR, but in PLS the PCs are constructed such that they include the chemical reference (Y variables, dependent data) in the calculation process. This technique orders the PCs according to their relevance for predicting the dependent variables, rather than to their description of the most variance of the spectral data. This method excels when the
dependent data (X variables) express common information, as usually happens in spectral
data. The required number of PCs is typically smaller than that in a PCR calibration model
for similar model performance (Wold et al. 1983; Esbensen et al. 2002; Nicolaï et al. 2007).

2.2.4 Artificial neural networks
Artificial neural networks (ANNs) are based on their ability to “learn” during a training
process where they are presented with inputs and a set of expected outputs. The basic
structure of an ANN consists of three layers of “nodes” or “neurons”: an input layer (i.e.
spectral data or PCs), a hidden layer (which can consist of one of more nodes), and an
output layer (which combines the outputs of the hidden layer into a single output). The
node is a fundamental processing unit; each node has a series of weighted inputs, either
from an external source or the output from other nodes. The inputs to the node are
analogous to synapses, and the weights correspond to the strength of the synaptic
connection. The sum of the weighted inputs is transformed with linear or nonlinear transfer
functions, a popular nonlinear transformation function being the sigmoid function. The
learning (or training) is an iterative process in which the resultant output is compared to the
provided expected output, and an algorithm adjusts the weights accordingly. This method
was first tested in the field of spectroscopy on simulated data (Long et al. 1990). It was
proven to provide the best results in modeling soil constituents in a recent large-scale study
that included over 1,100 soil samples (Rossel & Behrens 2010).

3. Heavy metals
3.1 Introduction
Heavy metals are released into the environment—the air (e.g. during combustion, extraction
and processing), surface water (via runoff and release from storage and transport) and soil
(and hence into the groundwater and crops). Although the adverse health effects of heavy
metals have long been known, exposure to heavy metals continues and is even increasing in
some areas, especially in less developed countries. The main threats to human health from
heavy metals are associated with exposure to Pb, Cd, Hg and As. Heavy metal exposure
may cause headaches, irritability, abdominal pain, kidney damage, skeletal damage, acute
pulmonary effects, cardiovascular disease, chronic renal failure, cancer, lung damage,
neurological and psychological symptoms, nervous system damage and much more (Järup
2003). Heavy metals as soil contaminants typically occur in low concentrations (10,000 mg
kg\(^{-1}\) dry soil). Inorganic material in general does not exhibit characteristic absorption
features in the VNIR-SWIR wavelength region. It is therefore considered impossible to
directly detect inorganic soil contaminants in general, and heavy metals in particular
(Winkelmann 2005). The conventional method of estimating the spatial distribution of heavy
metals is by raster sampling and a time-consuming laboratory analysis, followed by
geospatial interpolation (Kemper & Sommer 2002). While pure metals do not absorb in
the VNIR-SWIR region, they may be detected indirectly via their complexing with organic
matter (OM), association with moieties such as hydroxides, sulfides, carbonates, or oxides
which are detectable, or adsorption to clays that absorb light in this wavelength range
(Malley & Williams 1997). To explain possible variations in spectral signals due to heavy
metals bound to minerals, it is necessary to consider the binding reaction of the metal onto
the mineral surface. This approach starts from the premise that the spectrally assigned
position of minerals can change with chemical composition and surface activity (Ben-Dor et
Soil Contamination

al. 1999). Despite the occurrence of otherwise similar minerals in different samples, variations in spectral features (e.g., shifts in peak wavelength) may occur according to the nature of highly enriched cations in the mineral. The surface complexation model of the binding reactions of heavy metals describes the binding of metal ions to the mineral surface functional group to form a more stable molecular unit (Christl & Kretzschmar 1999; Zachara & Westall 1999). The main surface functional groups are inorganic hydroxyl groups that bind to surface Al, Fe, Mn, or Si on oxides or Al and Si on the edges of clay minerals (Sparks 1995; Zachara & Westall 1999; Choe et al. 2008).

3.2 Reflectance spectroscopy of heavy metals

The first report of quantitative prediction of heavy metals in sediments by NIRS was published in 1997 by Malley and Williams. They showed the feasibility of predicting six of the seven metals examined (Cd, Cu, Pb, Zn, Ni, Mn, Fe) with NIRS in a highly variable set of sediment samples, collected at Precambrian Shield Lake in northwestern Ontario, Canada. The metals were modeled by both MLR and PLS, and OM was found to be responsible for the prediction of sediment heavy metal concentrations (Malley & Williams 1997). A study published in 2002 by Kemper and Sommer investigated the ability to predict heavy metals following a mining accident that caused an area to be flooded with pyritic sludge contaminated with high concentrations of heavy metals: 214 soil samples were collected, and prediction of heavy metals was achieved by MLR and ANN approaches. It was possible to predict six out of nine elements (As, Fe, Hg, Pb, S, Sb) with high accuracy. Correlation analysis revealed that most of the wavelengths that were important for the prediction could be attributed to absorption features of Fe and Fe oxides (Kemper & Sommer 2002). The study by Kemper and Sommer continued with the collection of an additional 132 soil samples and acquisition of hyperspectral data with the HyMap airborne sensor (Cocks et al. 1998). A spectral mixture modeling approach was applied to the field and airborne hyperspectral data. VMESMA was used to estimate the quantities and distribution of the remaining tailing material. The semi-quantitatively determined abundance of residual pyritic material in the sludge could be transformed into quantitative information for an assessment of acidification risk and distribution of residual heavy metal contamination based on an artificial mixture experiment prepared with three different soils and pure sludge. Unmixing of the HyMap images allowed identification of secondary minerals of pyrite as indicators of pyrite oxidation and associated acidification (Kemper & Sommer 2003). The first study on NIRS of heavy metal contamination in agricultural soils was reported by Wu et al. (2005): 120 soil samples were collected in the Nanjing region, and Hg concentration was modeled by PCR. Correlation analysis revealed that Hg concentration is negatively correlated with soil reflectance, while positively correlated with the absorption depths of goethite at 496 nm and clay minerals at 2,210 nm, suggesting Hg sorption by clay-size mineral assemblages as the mechanism by which to predict spectrally featureless Hg in soils (Wu et al. 2005). In later work, Wu et al. (2007) also explored a physicochemical mechanism that allows estimation of heavy metals with the reflectance spectroscopy method, and concluded that correlation with total Fe (including active and residual Fe) is the major mechanism (Wu et al. 2007). Choe et al. (2008) explored the possibility of extending the use of spectral variations associated with heavy metal parameters to map the distribution of areas affected by heavy metals on HyMAP data in the Rodalquilar gold-mining area in southeast Spain. The chosen spectral parameters showed significant correlations with concentrations of Pb, Zn and As. Later, Choe et al. (2009) examined the
potential use of spectral absorption feature parameters to quantify heavy metals using stepwise MLR (SMLR) and enter MLR (EMLR).

Contradictory conclusions on the major absorption feature mechanisms that can enable indirect prediction of heavy metals were presented by the abovementioned studies. However, their results cannot be generalized, as soil properties (e.g., soil C<sub>org</sub> content) were distinctly different in the different studies (Vohland et al. 2009). Vohland et al. (2009) used soil samples taken from grassland plots that showed a very wide range of both C<sub>org</sub> (0.37%–9.24%) and clay contents (0%–37.8%), with the aim of investigating whether the heavy metal contents of these samples can actually be assessed by NIRS. In addition, they explored statistically whether a valid prediction mechanism can be identified. According to the correlation spectra and PLS coefficients, it was concluded that the active C<sub>org</sub> fraction is the key variable for an indirect spectroscopic approach (Vohland et al. 2009). Another study investigating the binding form of heavy metals with agricultural soil constituents was conducted by JiA et al. (2010). The results showed that Cr, Cu, Zn and As have stronger negative correlation coefficients with the absorption features of Fe oxides, clays and OM, However, Cd, Pb and Hg only displayed a significant correlation with the spectral region related to organic carbon (JiA et al. 2010). Pandit et al. (2010) showed very high correlation coefficients between laboratory-determined abundance and the abundance predicted from PLS calibration for several heavy metals, with the correlation coefficient for Pb being the highest of all metals (R=0.992). Soil Pb contamination exhibited a weak correlation (0.473) with percentage of OM, suggesting that while OM is a factor in Pb retention, it cannot be used to determine the quantity of Pb because only a little OM is needed to retain Pb (Pandit et al. 2010).

3.3 Conclusion

Heavy metal elements cannot be detected with reflectance spectroscopy at concentrations under 1000 mg kg<sup>-1</sup> (Wu et al. 2007). The presence of Fe in ferric or ferrous forms results in absorption features at wavelengths in the 0.4 to 1.3 μm region. A doublet absorption feature near 496 nm is caused by goethite. Two distinct absorption bands at 1,410 and 1,910 nm are attributable to vibrational frequencies of OH groups in the adsorbed water, and the absorption feature at around 2,210 nm is related to OH groups in the crystal lattice water. For pure goethite, two absorption bands, at 670 and 950 nm, are also characteristic and even stronger than the 496 nm band. However, the soil matrix effect prevents these two bands from being displayed in soil spectra (Wu et al. 2005). The intercorrelation between heavy metals and total Fe (including active and residual Fe) is a major predictive mechanism (Wu et al. 2005). Later, Wu et al. (2007) showed that all heavy metals that have a high correlation with Fe also have high cross-validation statistics. The three siderophile elements Ni, Cr, and Co had the highest correlation coefficients with Fe, and prediction accuracy for these elements was also the highest among the eight elements studied. The most poorly predicted element, Cd, showed the lowest correlation with Fe. Another important soil constituent, OM, could absorb electromagnetic waves and adsorb heavy metals; however, the correlation between OM and heavy metals was very poor in the study area (Wu et al. 2007). This is different from results obtained using reflectance spectroscopy in food and forage crop research, where the content of inorganic elements is low relative to the organic matrix. Thus, successful prediction of inorganic components in forage crops can be based on the correlations between inorganic components and organic constituents (Siebielec et al., 2004). Choe et al. (2008) showed that despite the requirement for further corrections, the
considered spectral absorption feature parameters had the potential to detect heavy metals. Moreover, the image-derived spectral parameters themselves showed a capacity to screen areas affected by heavy metals as a preliminary observation in determining sampling strategies and precise analyses for the investigation of environmental pollution (Choe et al. 2008). While weak prediction performance might result from the occurrence of many overlapping bands and the limited number of molecules that respond spectrally in the VNIR-SWIR range, a spatial distribution map of heavy metals by EMLR shows a distribution pattern similar to that on a map of measured values (Choe et al. 2009). However, the partial trade-off for time saved is reduced accuracy. Thus, spectral assessment of soil samples cannot replace, but rather complements the classical chemical analysis. The benefits are the practicable processing of a large number of samples and the savings on chemicals, lengthy and tedious processes and manpower. However, a calibration using information obtained from chemical analyses is mandatory (Vohland et al. 2009). In terms of spatial analysis, an airborne or space borne hyperspectral sensor may be useful for the screening of large areas and reproduction of the spatial distribution patterns.

### 3.4 Summary table

The results of the previously mentioned NIRS prediction models of heavy metals are summarized in Table 1. Different modeling techniques as well as various preprocessing methods were employed in the development of these models, for an in depth description of the models please refers to the proper references.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Contaminant ($R^2$)</th>
<th># of Samples</th>
<th>Modeling Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malley &amp; Williams (1997)</td>
<td>Cd(0.63), Cu(0.91), Fe(0.86), Mn(0.93), Ni(0.88), Pb(0.81), Zn(0.93)</td>
<td>169</td>
<td>MLR / PLS</td>
</tr>
<tr>
<td>Kemper &amp; Sommer (2002)</td>
<td>As(0.84), Cd(0.51), Cu(0.43), Fe(0.72), Hg(0.96), Pb(0.95), S(0.87), Sb(0.93),</td>
<td>214</td>
<td>MLR / ANN</td>
</tr>
<tr>
<td></td>
<td>Zn(0.24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kemper &amp; Sommer (2003)</td>
<td>As(0.73), Pb(0.73)</td>
<td>346</td>
<td>polynomial fit</td>
</tr>
<tr>
<td>Siebielec et al. (2004)</td>
<td>Cd(0.54), Cu(0.61), Fe(0.87), Ni(0.84), Pb(0.45), Zn(0.67)</td>
<td>70</td>
<td>PLS</td>
</tr>
<tr>
<td>Wu et al. (2005)</td>
<td>Hg(0.48)</td>
<td>120</td>
<td>PCR</td>
</tr>
<tr>
<td>Wu et al. (2007)</td>
<td>As(0.72), Cd(0.20), Co(0.80), Cr(0.85), Cu(0.67), Ni(0.81), Pb(0.55), Zn(0.56)</td>
<td>61</td>
<td>PLS</td>
</tr>
<tr>
<td>Choe et al. (2008)</td>
<td>As(0.88), Pb(0.61), Zn(0.60)</td>
<td>49</td>
<td>EMLR</td>
</tr>
<tr>
<td>Choe et al. (2009)</td>
<td>As(0.60), Cu(0.81)</td>
<td>22</td>
<td>EMLR</td>
</tr>
<tr>
<td>Ren et al. (2009)</td>
<td>As(0.62), Cu(0.41), Fe(0.78)</td>
<td>33</td>
<td>PLS</td>
</tr>
<tr>
<td>Vohland et al. (2009)</td>
<td>Cu(0.75), Fe(0.84), Mn(0.71), Pb(0.76), Zn(0.81)</td>
<td>149</td>
<td>PLS</td>
</tr>
<tr>
<td>Jia et al. (2010)</td>
<td>As(0.30), Cd(0.10), Cr(0.68), Cu(0.46), Hg(0.15), Pb(0.26), Zn(0.40)</td>
<td>122</td>
<td>Linear Regression</td>
</tr>
<tr>
<td>Pandit et al. (2010)</td>
<td>Cd(0.43), Cu(0.81), Mn(0.81), Pb(0.75), Zn(0.76)</td>
<td>8</td>
<td>PLS</td>
</tr>
</tbody>
</table>

Table 1. Heavy metals prediction via NIRS.
4. Petroleum hydrocarbons

4.1 Introduction
Petroleum hydrocarbons (PHCs) are environmental contaminants of great significance. With production rates of 13.42 million cubic meters of crude oil per day (Energy Information Administration 2009), their potential as soil and water contaminants is apparent. Crude oil and refined crude oil products therefore appear in almost every location on earth at the different stages of crude oil production, handling, transport, refining and consumption. The widespread use of PHCs for transportation, heating and industry has led to the release of these fuels into the environment through accidental spills, long-term leakage, and operational failures. Consequently, many soil and water areas are contaminated with PHCs.

PHCs are well known to be neurotoxic to humans and animals. They have been found to affect brain activity and development, and cause many other symptoms including, but not limited to: nausea, disorientation, mental confusion, speech slurring and memory disorders. Exposure to higher levels can cause extreme debilitation, loss of righting, coma, seizures and death. Long-term exposure has been shown to cause changes in neurophysiological or psychological capacity, and to increase the risk of lung, skin and bladder cancers, as well as having other carcinogenic effects (Hutcheson et al. 1996); (Boffetta et al. 1997); (Ritchie et al. 2001). Due to its toxicity, several environmental regulations have been enforced worldwide to protect the environment from PHCs contamination. As a result, attention has also been devoted to rehabilitating PHCs-contaminated soils. For both the diagnosis of suspected areas and the possibility of controlling the rehabilitation process, there is a great need to develop and implement a method that can rapidly detect small amounts of PHCs in soils as well as identify PHCs type, for better soil management and maintenance as a source of life and food.

4.1.1 Total PHCs
Various fuel products are the outcome of crude-oil refinement. These products contain mostly C (83-97%), H (10-14%), N (0.1-2%), O (0.0005-1.5%), S (0.05-6%) and various types of metals such as V, Fe, Cu and so on (<0.1%). These fuel products are classified by two main criteria: structure of the hydrocarbon (aliphat ic or aromatic), and number of C in the chemical structure. These two criteria affect several of the products' characteristics, such as boiling point, solubility, absorption and toxicity. There are some overlapping characteristics among the various groups of hydrocarbons but in general, aromatic compounds are more toxic than aliphatic structures and there is a connection between the length of an aliphatic molecule and its toxicity (Tetra Tech EM Inc. for EPA 2000). Total PHCs (TPH) is a commonly used measurement index for quantifying environmental contamination originating from fuel. Choosing an analytical method for assessing TPH in soil samples is a problematic task due to the complex nature of the fuel ingredients and the fact that those ingredients change with time. Today, the most commonly used analytical method for assessing TPH in soil samples is based on the no-longer approved Environmental Protection Agency (EPA) method 418.1. This method was originally intended for assessment of TPH in water (EPA 1978), but was later adjusted to assess TPH in soil samples (EPA 1983). Not only has this previously approved method been withdrawn by the EPA, but it is also problematic for various other reasons.
4.1.2 Analytical methods

4.1.2.1 EPA method 418.1

The most common method in use today, which is based on the EPA's method 418.1 which was adjusted for soils, is costly and time-consuming. Briefly, the sample is brought to the laboratory and a specific extraction technique is applied using Freon 113 (trichlorotrifluoroethane). The TPH level is then determined in the solvent by Fourier transform infrared (FTIR) spectrophotometer. This method necessitates careful shipping of the samples to the laboratory (e.g., cooling conditions and use of special quartz tubes) and several replications to precisely establish the chemical analysis. Measuring one sample can cost about $50 and the results cannot provide information in a near-real-time domain. Consequently, in obtaining a reliable spatial pattern of the contaminated area, many samples are required to generate useful information that will facilitate the rehabilitation or precautionary processes. This method can only be used for quantification of TPH and not for identifying PHCs type.

4.1.2.2 Gravimetric

Hydrocarbons are removed from the sample by an extraction solvent, the solvent is weighed before and after the extraction, and the difference in weight expresses the TPH. This method is analogous to weighing a battleship with its captain, then subtracting the weight of the battleship in order to find out the weight of the captain: while it is relatively inexpensive, it is highly inaccurate.

4.1.2.3 EPA method 8440

Hydrocarbons are extracted from the sample with CO₂ at very high pressure (340 atm), at 80°C for 30 min. After extraction, TPH is determined by FTIR spectrometry and compared to a calibration model (similar to method 418.1). This method is not only expensive but it is also dangerous due to the high pressures involved.

4.1.3 TPH field methods

The abovementioned laboratory methods are costly and time-consuming, and usually cannot be implemented in a real-time domain at the contaminated site. Samples need to be transported to the laboratory in a cooled container and there is a waiting period before results are reported. In addition to the laboratory tests, a few field-measurement kits are available. The seven field-measurement devices listed below were tested at the Navy Base Ventura County site in Port Hueneme, California by Tetra Tech EM Inc. (2000):

- CHEMetrics, Inc.: RemediAid™ Total Petroleum Hydrocarbon Starter Kit
- Wilks Enterprise, Inc.: InfraCal® TOG/TPH Analyzer, Models CVH and HATR-T
- Horiba Instruments, Inc.: OCM-350 Oil Content Analyzer
- Dexsil® Corp.: PetroFLAG™ Hydrocarbon Test Kit for Soil
- Environmental Systems Corp.: Synchronous Scanning Luminoscope
- siteLAB® Corp.: Analytical Test Kit UVF-3100A
- Strategic Diagnostics, Inc.: EnSys Petro Test System

These field tests, making use of several technologies, are summarized in Table 2.
Reflectance Spectroscopy as a Tool for Monitoring Contaminated Soils

<table>
<thead>
<tr>
<th>Technology</th>
<th>Measurement Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friedel-Crafts alkylation reaction and colorimetry</td>
<td>RemediAid™ Total Petroleum Hydrocarbon Starter Kit</td>
</tr>
<tr>
<td>Infrared analysis</td>
<td>Infracal® TOG/TPH Analyzer, Models CVH and HATR-T; OCMA-350 Oil Content Analyzer</td>
</tr>
<tr>
<td>Emulsion turbidimetry</td>
<td>PetroFLAG™ Hydrocarbon Test Kit for Soil</td>
</tr>
<tr>
<td>Ultraviolet fluorescence spectroscopy</td>
<td>Synchronous Scanning Luminoscope; siteLAB® Analytical Test Kit UVF-3100A</td>
</tr>
<tr>
<td>Immunoassay and colorimetry</td>
<td>EnSys Petro Test System</td>
</tr>
</tbody>
</table>

Table 2. Technology of TPH field-measurement kits

These technologies are thoroughly discussed in the report from the EPA field test conducted in June 2000: while some of these tests produced good results, the kits were found to be expensive and to require a knowledgeable operator (Tetra Tech EM Inc. for EPA 2000).

All analytical methods, whether conducted in the laboratory or in situ, that are available for the determination of TPH in soil samples have individual drawbacks, for example: shipping of the sample to a distant laboratory (no in-situ availability), long waiting periods for results, high cost per sample, inaccuracy, and the requirement of a skilled staff to perform the measurement. Due to these drawbacks, a new low-cost in-situ method for assessing PHC contamination, particularly at a new site, is greatly needed. Emphasized desired properties of this new method are cost-effectiveness and its ability to be implemented by not only highly trained personnel.

4.2 TPH and reflectance spectroscopy

4.2.1 Absorption features of organic compounds

Carbon-hydrogen bonds (C-H, C-H₂, C-H₃), hydroxy groups (O-H), double and triple bonds of aliphatics and aromatics, carboxyl groups (C=O), ester groups (C-O-C), amino groups (N-H) and other structural groups of organic chemicals exhibit characteristic fundamental vibrations that are evident in the 2,500 to 6,670 nm wavelength region of the IR spectrum. Furthermore, these compounds exhibit overtone and combination bands in the IR and VNIR-SWIR spectrum between 500 and 6,670 nm. These fundamental vibrations can be determined by IR spectroscopy and some of the overtone and combination bands can be determined by VNIR-SWIR spectroscopy (Winkelmann 2005).

4.2.2 VNIR-SWIR spectroscopy of organic chemicals

As IR spectroscopy is the better choice for qualitative analysis, VNIR-SWIR spectroscopy is primarily applied for quantitative analysis in many disciplines (termed “chemometrics”). However, despite the fact that VNIR-SWIR spectra of organic chemicals are often highly complex due to many overlapping overtone and combination absorption features, they still contain information on the functional and structural groups of the chemicals under investigation. During the past 10 years, a new quantitative methodology has been widely developed that utilizes the reflected radiation in solids. Basically, this technology was adopted 40 years ago from a strategy developed in the food science discipline (Ben-Gera & Norris 1968a; Ben-Gera & Norris 1968b), whereas today it is widely utilized in many industrial and scientific applications. In this approach, the radiation reflected across the
Soil Contamination

80

VNIR-SWIR region is modeled against constituents determined by traditional chemical analysis and then used to predict unknown samples. This technology permits rapid and cost-effective measurements on many samples at a given location and therefore significantly reduces chemical analysis costs. Despite the heterogeneity of the methodology, it was demonstrated by many (e.g: Nanni & Demattê 2006; Ben-Dor, et al. 2008; Rossel & Behrens 2010) that soil constituents can be extracted from a simple reflectance measurement (laboratory and field) if an appropriate calibration model has been previously developed. Since Ben-Dor et al. (1995)’s pioneering study, many other workers have explored this promising technology for soils and a significant breakthrough in this area was its replacement of wet chemistry in many scenarios. The adoption of this approach to evaluate hydrocarbon contamination in soil is obvious and a few workers have partially demonstrated this application, as described in the next section. As shown above, VNIR-SWIR spectroscopy holds great potential for quantitative chemical analysis. If calibration standards are used, the identification of known substances in a mixture and the determination of their concentrations are possible. Furthermore, it seems likely that the identification of substances based on their VNIR-SWIR spectra will become possible for classification purposes.

4.3 Reflectance spectroscopy of PHCs

The spectral properties of hydrocarbons were identified in the late 1980s, although it has been argued that these properties were only visible at concentrations of 4% wt (by weight) and above (Cloutis 1989). In the mid 1990s, a VNIR-SWIR reflectance sensor was developed as a proof of concept for the detection of OM in soil based on the same spectral properties identified by Cloutis. The sensor was designed for the detection of benzene in soil at a minimal concentration of 4.4% wt: several configurations were tested and minimal information was provided (Schneider et al. 1995). Soon after, the U.S. Department of Energy contracted a private company to investigate the application of reflectance spectroscopy to determine motor oil contamination in sandy loam. A schematic design for a field instrument was suggested, although only one contaminant and one type of soil were tested, using very few samples with a very limited range of contamination (Stallard et al. 1996). A more inclusive study was conducted shortly thereafter, using three types of soil contaminated in the laboratory with diesel and gasoline. A 0.1% wt and 0.5% wt minimum detection limit was achieved, respectively (Zwanziger & Heidrun 1998). The first study utilizing field-collected samples was not able to produce robust models, resulting in a very low correlation coefficient ($R^2 = 0.46$) and large errors, probably due to very low sample count and problems with the chemical measurements performed in the laboratory which produced inconsistent results (Malley et al. 1999). Attempts at mapping hydrocarbons using Landsat and Daedalus in 1994 and 1995 failed, probably due to the limited spectral resolution of the sensors (multispectral sensors); on the other hand, the higher spatial and spectral resolution, as well as the very high signal-to-noise ratio of the airborne hyperspectral sensor used (HyMap) (Cocks et al. 1998) yielded successful identification of hydrocarbon- and oil-contaminated soils, but only for high contaminant concentrations (2.5% wt) (Hörig et al. 2001). Based on the HyMap mission, a Hydrocarbon Index was developed for the mapping of hydrocarbon-bearing materials. This index is limited to very high signal-to-noise ratio sensors, as well as by other issues such as problems with land cover, vegetation and high concentration detection levels (Kühn et al. 2004). The most comprehensive work on reflectance properties

www.intechopen.com
Reflectance Spectroscopy as a Tool for Monitoring Contaminated Soils

of hydrocarbons was conducted by Winkelmann (2005): several types of hydrocarbons were mixed with several types of soil under laboratory conditions. They were measured spectrally and an attempt was made to separate them into hydrocarbon groups using the reflectance spectra; hyperspectral airborne remote sensing was also applied to identifying hydrocarbon contamination. No quantitative models were tested, although this was mentioned as an avenue of further study (Winkelmann 2005). A recent study by Chakraborty et al. (2010) on the prediction accuracy of VNIR-SWIR reflectance spectroscopy of petroleum contaminated soil, showed fair validation results ($R^2=0.64$). The study included 46 field collected samples that were preprocessed and modeled by several techniques. While the abovementioned studies addressed only concentration levels >0.1%, today’s environmental regulations require identification levels that are an order of magnitude lower. Moreover, none of the above studies dealt with identifying PHCs type and mixtures in soil using reflectance spectroscopy. A comprehensive study including several types of soil with several types of PHCs at a large variety of contamination levels has not yet been published. Most of the studies were performed with lab-prepared samples and almost none collected in the field. Furthermore, no generic models were developed, especially combining laboratory and field samples, and no real quantitative operation model was presented for real-life applications.

A long-term study conducted by Schwartz (2012) over the past 3 years included the development of NIRS models for the prediction of TPH in soil samples at high precision for low contamination levels, as well as PHCs classification. Five types of soil were mixed with three types of PHCs at 50 concentration levels, giving 750 laboratory samples, along with more than 200 other laboratory-prepared samples with various mixtures to study the soil properties' effects on spectral measurements and chemical analysis. The samples were evaluated for TPH level with EPA method 418.1, as well as measured spectrally by an ASD Fieldspec Pro. An “all possibilities” approach was taken for the creation of robust NIRS models, including the evaluation of many preprocessing techniques (smoothing, absorbance, first and second derivatives, and continuum removal), as well as PLS and ANN modeling methods. Forty-one soil samples were then collected in a field campaign and TPH level was predicted by both the PLS and ANN models, as well as being determined by the traditional EPA 418.1 method. A detection limit of 1,000 ppm was set, and samples with concentrations of 1,000 ppm or lower (27 samples) were successfully predicted as non-contaminated or with very low contamination. The rest of the predicted samples gave a correlation between the predicted and measured values of $R^2 > 0.9$ (this study will be published soon as a Ph.D. thesis, and as several papers in peer-reviewed journals).

5. Acid mine drainage

Acid mine drainage (AMD) from mine wastes is considered a major problem in mining areas, because of the risk of heavy metal contamination of soils and water, Reflectance spectroscopy and NIRS is also used in the field of AMD. AMD is a byproduct of the oxidation of iron disulfide mineral (pyrite). While AMD reacts with the surrounding rocks and soil, heavy metals become more soluble and they can leach into the environment. This process is also substantially enhanced by bacterial action. In mining areas, the release of heavy metals is accelerated due to the increased oxidation rates, which are caused by the greater accessibility to air through mine workings, waste rock, and tailings by mineral processing. Furthermore, there is the added risk of mining accidents such as in Aznalcóllar.
Soil contamination by agricultural pesticides is a matter of increasing concern. Several authors have reported accurate predictions of pesticide sorption to soil using NIRS. Bengtsson et al. (2007) used NIRS to predict lindane and linuron sorption to soil samples. The pesticide-content predictions were established by PLS of the VNIR-SWIR spectra and sorption data obtained using the batch equilibrium technique. PLS could predict 85% of the soil sorption (Kd) for lindane and 84% for linuron, respectively. Later, Brunet et al. (2009) tested the feasibility of determining chlordecone contamination in soils. Chlordecone, also known as kepone, is a chlorinated polycyclic ketone pesticide. It is toxic, persistent, and induces cumulative and delayed toxicity. The regular method to measure chlordecone involves extraction with two solvents and quantification by gas chromatography–mass spectroscopy. A total of 236 samples were collected at 0–30 and 30–60 cm depths in Andosols, Nitisols and Ferralsols from all over the eastern part of Martinique (French West Indies). Chlordecone content was determined by wet chemistry and the reflectance spectra were recorded. Modified PLS (mPLS) regression was used to correlate spectral data to conventional values. While conventional measures and NIRS predictions correlated poorly at high chlordecone contents (>12 mg kg\(^{-1}\)), when the validation set was divided into four classes based on chlordecone content, NIRS prediction properly classified 80% of the samples (Brunet et al. 2009).

7. Vegetation stress as an indicator for soil contamination

The concept of using plants as indicators for soil contamination has been tested for many years, for many good reasons: 1) inorganics in general do not exhibit characteristic absorption features in the VNIR-SWIR wavelength region; 2) the plant's root system is susceptible to contaminated material, and extracts and transports it to the aboveground plant parts, and 3) soil is far from always being bare; it is often covered with plants. Free metal ions in the soil solution and exchangeable ions of clays and OM are mostly available for plant uptake (Lasat 2000). For their development, plants require not only macronutrients
(N, P, K, S, Ca, and Mg), but also essential micronutrients such as Fe, Zn, Mn, Ni, Cu, and Mo. Most of these micronutrients accumulate in the plant tissues for their metabolic needs, but they never exceed 10 ppm. Higher levels of these metals (and other heavy metals) in plant tissue might have phytotoxic effects, sometimes resulting in death (Winkelmann 2005). Vegetation stress caused by heavy metal toxicity can show different symptoms, depending on the heavy metal's type and concentration, and on the plant species' sensitivity. The phototoxic effects typically cause physiological, morphological and ecological changes, expressed in many different ways, from chlorosis (reduced production of chlorophyll and hence reduced photosynthetic activity), leaf wilting, needle retention, branch density reduction, defoliation, growth inhibition, flowering and fruiting changes, dwarf growth and gigantism, to changes in plant species distribution, the establishment of adapted species, and the dying-off of whole plants or communities (Winkelmann 2005). However, stress can be induced in vegetation by a large number of other factors, including water deficiency, poor soil drainage, poor soil aeration, soil salinity, weed competition, pest infestation, nutrient deficiency, or nutrient poisoning (Levitt 1980; Lichtenthaler 1996).

A large number of studies have reported on spectra of metal-stressed vegetation being clearly different from those of unstressed vegetation. These differences usually include decreases in both NDVI and RVI with increasing plant stress (Davids & Tyler 2003; Dunagan et al. 2007), signs of reduced biomass and a shift of the red edge position (REP) (Kooistra et al. 2004; Dunagan et al. 2007) or red edge slope (Zhou et al. 2010), signs of reduced photosynthetic activity due to chlorosis and decreased reflectance in the 700 nm to 2,500 nm wavelength region (Kooistra et al. 2003). Some heavy metals which are considered contaminants are also essential micronutrients for all vegetation species, in particular Cu and Zn. Therefore, in some cases, even positive effects in reflectance spectra can be observed when one of the elements under investigation is clearly a micronutrient with limited availability. Such effects include a red-shift of the red edge and decreased reflection in the VIS wavelength region indicating increased photosynthetic activity (Horler et al. 1980). As no vegetation stress symptoms and corresponding spectral characteristic could be related specifically to heavy metal stress or any other contaminant stress, the sources causing the observed stress need to be carefully identified and separated.

8. Limitations, obstacles and problems

8.1 Laboratory

A wide range of factors can affect soil reflectance spectra in both laboratory and field domains. In the lab, different spectrometers, or even repeating a specific sample's measurements in the same spectrometer may result in variations. Such variations might include subtle or strong alterations in wavelength location, peak absorption shape or radiometric intensity. In addition to the instrumentation itself, internal electronic noise can affect the measurements and mechanical noise factors (e.g. homogeneity and purity of the white reference panel, or subtle movement when holding the fiber optic) can strongly affect their consistency. In soil samples, where very weak spectral features are monitored for chemometric purposes, these noise factors can alter the robust use of a selected spectral model for a wide range of spectrometers and users. Recently, Pimstein et al. (2011) examined the variation stemming from the above sources using three ASD FieldSpec Pro spectrometers and developed a standard protocol for laboratory spectral operation. They also suggested using the same internal standard worldwide in order to correct the spectra of
any spectrometer to a selected master spectrometer. They demonstrated that chemometric analytical results are more coherent and precise. A global soil spectral library, not just as with minerals, has to be followed by chemical information. Recently, Viscarra Rossel (Rossel, Soil Spectroscopy Group 2009) initiated a spectral archive containing more than 10,000 spectra taken worldwide and is in the process of integrating them into this library with some major soil attributes measured in the alternative “wet chemistry” way.

8.2 Field

In the field, there are more uncertainties than in the laboratory, such as atmosphere attenuation, sun angle, aspect and slope of the sample area, large pixel sample, BRDF effects and most of all, soil crusting and sealing which can emerge on any soil surface. The latter prevent representation of the real soil body in the measurement. In the laboratory, the soil is crushed to pass a 2 mm sieves, and all stones and litter debris are consequently removed before the measurement. In the field, those materials are present, as are physical and biogenic crusts and possible dust contamination. Another problem that might emerge in field soil measurement is the adjacency effect and the mixed pixel. The first can occur when the pixel in question is surrounded by bright reflective targets. The mixed pixel problem occurs when a pixel is composed of several chromophores. It is simple when the mixture is binary and complex when it is not. The sun's angle, slope and aspect might not be factors in the field if artificial illumination is used in a standard procedure. A contact probe equipped with tungsten-halogen illumination can be used, with caution. This is because the soil surface measured by such an instrument may consist of a very narrow field of view which might not represent the soil in question, but rather debris, stones or even soil aggregation.

8.3 Airborne and spaceborne

Soil reflectance can also be measured from aircrafts and satellites, using either a point spectroscopy (Karnieli et al. 2001) or IS (Ben-Dor et al. 2009) sensor. However, more difficulties arise during such measurements to extract the correct reflectance, such as: small integration time (less photons), strong atmosphere effects, large pixel size and varying quality of the sensor’s stability and sensitivity. Brook & Ben-Dor (2011) have recently developed a more moderate method (SVC—Supervised Vicarious Calibration) to standardize all sensors' radiometric readings, with the aim of deriving the optimal soil reflectance from the airborne IS sensor. This method uses artificial net targets with varying densities placed on a bright background area. It is easy to use and performs well. The method has been recently validated in a European campaign over southern France, using three different sensors simultaneously onboard two different airplanes. The preliminary results were relatively good as compared to the alternative ways of deriving the reflectance from the airborne sensor (unpublished data). The artificial target and the suggested method help assess the atmospheric attenuation, and minimize sensor instability while correcting for the systematic noise. Another limitation using airborne IS for soil is its high cost, low availability, and the complexity involved in processing the raw data into a final reflectance product (Ben-Dor et al. 2009). These factors actually prevent ordinary users from using this technology, classifying it as an exclusive method. Vegetation coverage (partial or total) is also a problem for deriving correct soil reflectance from afar when it is mixed in the sensor’s field of view.
As a further step, the retrieval of quantitative mapping using spectral parameters derived from hyperspectral images should take into account differences in the spectral and spatial resolution between ground spectrometer and image data (Choe et al. 2008).

8.4 General issues
In all of the above domains, it must be remembered that reflectance spectroscopy is strongly affected by water content, particle size distribution and the measurement protocol. In addition, ways of calculating the reflectance from radiometric readings, i.e. against a white standard panel or using the radiance-to-irradiance ratio, may change the final product. Furthermore, the reflectance represents only the surface and cannot provide information on the soil profile (unless a penetrating fiber optic is used such as that described by (Ben-Dor et al. 2008)). When obtaining a soil spectrum from a user, a meta-data set which characterizes all of these factors (moisture content, grain size and method of reflectance calculation) is strongly needed, because they have a significant impact on the final spectra's behavior.

8.5 Summary
Soil reflectance is an inherent property of the soil but many factors can affect its performance. Internal standards, a standard protocol and controlled conditions are a few of the things that can assist in sharing and comparing soil spectra (and chemometric models) worldwide (e.g. the Global Soil Library by Rossel and Soil Spectroscopy Group (2009)). It is obvious that uncertainties in the laboratory are smaller than those in the field, and the latter are smaller than those obtained from air or space. Use of radiometric data acquired from remote sensing domains to measure soil reflectance information should therefore be undertaken with caution.

9. Future potential of remote sensing technology for assessing soil contamination
Soil reflectance has become a very useful tool over the past 20 years in the laboratory, in the field and from air and space. As the sensitivity of portable field spectrometers increases, field soil spectroscopy is expected to become a basic tool for rapid point-by-point monitoring of the soil environment. The commercial development, operation and use of air- and spaceborne image spectrometers can provide near-laboratory-quality spectra of every pixel in an image and very soon, will permit remote sensing of soils with high standards. Information about soils from reflectance spectra in the VNIR and SWIR spectral regions represents almost all of the data passive solar sensors can provide. It is anticipated that the thermal IR regions will also become part of the soil spectral arena as they contain diagnostic information on some soil attributes that the reflective spectral region does not. The development of a thermal spectrometer (either point or imaging) will enable obtaining more spectral information with better accuracy. Another future insight is the capability to measure the soil profile's spectral distribution using a fiber optic assembly and small boreholes. Today, the major limitation of this technique is the fiber optics' length (as the fiber length increases, the signal across the SWIR region decreases—today, the length is limited to 1.5 m). Miniaturizing the spectrometer will enable placing it on the penetrating optical head, thus bypassing the limitation imposed by fiber length. This, in turn, will enable spectral measurement of any profile depth which, when combined with the NIRS model, might be able to provide in-situ soil contamination attributes of the soil profile. Combining
the penetrating optical device information with the hyperspectral remote sensing data will also enable extracting the spatial distribution of the contaminant in question in 3D view. Data merging of soil reflectance information with GIS layers and other potential sensors that can be used simultaneously is also a key vision for the future. Time-series spectral measurements of large areas are also very important. Future IS from orbit will enable global coverage of every point on the globe with very good signal-to-noise ratio, such as the PRISMA (Labate et al. 2009), HyspIRI (Knox et al. 2010) and EnMAP (Stuffler et al. 2007) programs. This will enable monitoring soil surface changes in short- and long-term scenarios. Another future activity related to soil spectroscopy and soil contamination assessment is the development of better computing systems in which better models to derive highly accurate soil attributes will be achieved. To that end, the “all possibilities” approach in which all statistical and data-manipulation techniques can be applied automatically to a set of data is strongly required. Computing power and simple operation algorithms are key factors to that end.

10. General conclusions and summary

Soil reflectance spectroscopy can assess soil properties rapidly and quantitatively in both point and spatial domains. Direct and indirect soil properties, as well as soil contamination characteristics, can be extracted efficiently at low cost in situ. To that end, the VNIR-SWIR spectra must be preprocessed and modeled against reference data obtained by traditional methods. While some soil contaminants are featureless in the VNIR-SWIR region, their detection and quantification is possible, as they may be detected indirectly based on their association with other detectable materials. Although a wide range of factors can affect soil reflectance spectra in both laboratory and field domains thus strongly affecting the consistency of the resulting measurements, recent developments and proper protocols are allowing for more consistent and accurate results. While using NIRS to predict soil contaminants can save time, some cases involve the trade-off of reduced accuracy. Thus, the spectral assessment of soil samples cannot completely replace, but rather complements the classical chemical analysis in these and other cases. The benefits of using NIRS can result in the practicable processing of a large number of samples and savings in chemicals, lengthy tedious processes and manpower. In terms of spatial analysis, an airborne or spaceborne hyperspectral sensor can be useful for the screening of large areas and the reproduction of the spatial distribution patterns of contaminated soil areas. Nevertheless, the field of reflectance spectroscopy as a tool for monitoring contaminated soils still requires further study toward increased accuracy and the development of practical real-life applications.

11. References


Zhou, G. et al., 2010. Heavy Metal Stress to P. Tomentosa in Coal Mining Area and Its Spectral Feature. In Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on. pp. 1–4.

Soil contamination has severely increased over the last decades, mainly due to petroleum hydrocarbons, solvents, pesticides, lead and other heavy metals from industrial wastes and human activities. The critical point regarding contaminated soil monitoring is the intrinsic difficulty in defining fixed monitoring variables and indicators as the establishment of any a priori criterion and threshold for soil quality can be still considered subjective. This book is organized into eight chapters and presents the state-of-the art and new research highlights in the context of contaminated soil monitoring and remediation strategies, including examples from South America, Europe and Asia. The chapters deal with the following topics: - monitoring of dioxin, furan, hydrocarbons and heavy metals level in soils - bioindicators and biomarkers for the assessment of soil toxicity - use of reflectance spectroscopy for soil contaminants and waste material detection - remediation technologies and strategies.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: