

SCALE-WAVELENGTH DECOMPOSITION OF HYPERSPECTRAL SIGNALS  
- USE FOR MINERAL CLASSIFICATION & QUANTIFICATION -

Zur Erlangung des akademischen Grades eines  
DOKTOR-INGENIEURS  
von der Fakultät für  
Bauingenieur-, Geo- und Umweltwissenschaften

des Karlsruher Instituts für Technologie (KIT)  
genehmigte  
DISSERTATION

von  
Dipl.-Geoinf. Christoph Ehrler  
aus Reichenbach/Vogtland

Tag der mündlichen  
Prüfung: 13. Februar 2014

Referent: Prof. Dr.-Ing. Stefan Hinz  
Korreferent: Prof. Dr. Thomas Neumann

Karlsruhe 2014



## Abstract

This research aims at the spatial mapping of minerals. Data acquisition encompassed a hyperspectral airborne survey in the reflective (HyMap) and a multispectral survey in the thermal (AHS) domains, coupled with spectroscopic reference measurements and laboratory analysis of collected samples. Intense lignite mining led to several environmental problems at the demonstration site Sokolov. One is acid mine drainage (AMD), the acidification of soils and waters due to sulphide mineral weathering and mobilisation of heavy metals. Mineral mapping candidates were limited by climate, geology and available data to the weathering stable goethite as indicator for potentially AMD affected sites, siderite as fast neutralising carbonate, kaolinite as slowly reacting acidity buffer and quartz as inert mineral not participating in buffering reactions.

Reproducible classification & quantification require a standardised and quality controlled pre-processing, a robust parameterization of spectral absorptions and a validation of all processing.

Pre-processing of the thermal domain data lacked an in-flight calibration step. A generic module adaptable to any multi- or hyperspectral thermal sensor was developed. The skin temperatures of the calibration targets are estimated from the imagery, hence relieving field teams from measuring temperature during sensor overflight. The target's spectral emissivity is typically more stable and for several materials pre-known.

Adaptation of the thermal atmospheric model and destripping in summary resolved the issue of noisy and strongly deviating emissivity spectra. The thermal data was brought to a similar pre-processing quality as is standard for the reflective data. It allowed its use for quantification.

Quantitative mineral fractions were derived from the samples' elemental oxide concentrations by use of the normative model NORMA, adapted to the Sokolov site. The feature based quantitative analysis allowed to link reflective and thermal domain through validation of the map products.

Overlaying spectral absorption features in mixtures can result in very complex spectral signatures. Classical spectral feature analysis methods are based on simple geometric primitives like triangles or on few feature describing parameters. Lots of information is lost during feature parameterization. A multi-scale analysis of spectra at several levels-of-detail and more versatile feature describing parameters would be beneficial.

This can be realised by the continuous wavelet transform (CWT). It linearly decomposes a spectrum in frequency components (scales). Noise is separated into the smallest scale. Narrow absorptions are emphasised in small to medium scales. Wide absorptions and continuum appear in high scales. CWT can be interpreted as cross-correlation of the spectrum with the wavelet, or as multi-scale derivative of the smoothed spectrum. The Second Derivative of Gaussian ( $\text{DOG}_{m=2}$ ) mother wavelet was used as its shape is similar to a generalised absorption feature. The transform is the scaled second derivative of the Gaussian filter smoothed spectrum. Maxima and minima in the spectra leave extrema in the transform, inflection points form zero crossings. The development of these transform features over the scales characterises the shape of the absorption feature and its spectral neighbourhood. Effects of noise, illumination, moisture, grain size, viewing geometry, additive offsets and linear bias are normalised or separated in specific scales. The thesis develops two spectral analysis methods on this multi-scale decomposition.

The *Wavelet Transform Feature Analysis* is a quantification method. It traces related transform features over the scales and chains them into vectors. The vectors convey information about the inducing spectral feature. A versatile parameterization is possible by length, curvature, wavelength and transform amplitude development of the vectors. This drives a multivariate feature based quantification, shown to be a generalisation of the classical Spectral Feature Analysis and Spectral Derivative Analysis. This approach exploits the multi-scale decomposition of the input signal with its benefits of selective accentuation of spectral components, suppression of noise, normalisation of the input signal and selective focus on varying levels of spectral detail and neighbourhood. No previous knowledge of shape and location of characteristic features and no methodological differentiation between reflectance or emissivity nor absorption minima or maxima is required. Complex spectral shapes can be conveyed into the regression as no geometric primitives are needed for the approximation of signal structures.

The *Wavelet Coherence Mapper* is a classification method. It makes use of wavelet coherence, a localised spectral similarity measure which returns a whole similarity spectrum instead of an aggregated value. Together with the multiple wavelet scales, it allows to interrogate the unknown spectrum at selectable wavelength and level of spectral detail. The local similarity around a specific narrow absorption up to the general similarity of the whole spectral continuum shape can be compared. The wavelet coherence is independent of multiplicative factors and amplitude differences. It is very sensitive to curvature direction differences, less to curvature strength. Pronounced deep absorptions are equally weighted as subtle shallow absorptions and it is invariant to inter-class amplitude differences. These features are beneficial for identification of possibly weak features within signatures of mineral mixtures.

Classification of siderite, goethite and clay minerals as well as quantification of goethite/ $\text{Fe}_2\text{O}_3$  and kaolinite from the hyperspectral reflective data were facilitated by the wavelet based methods and normative modelling. Derived mineral distribution patterns were plausible and followed the expectations. To complement and validate the reflective domain results, a band-ratio based quantification of kaolinite and quartz was performed on the multispectral thermal data.

Validation showed a general agreement of the distribution patterns and relative abundance trends between reflective and thermal domain for areas matching the mineralogic assemblage of the training data, albeit an offset existed. Limited transferability of the reflective domain quantitative kaolinite model was apparent for coal-clay mixtures. Pre-processing was successful. Normalised image, field and laboratory spectra matched well. Deviations in amplitude and spectral contrast of the not normalised spectra were partially compensated by the normalising character of CWT. The influence on the quantitative models hence was small.

Future research aspects are border effects in the CWT. Masking of possibly affected bands led to classification errors at higher scales. Fast filter-bank implementations of CWT can reduce processing times and the change to a spline biorthogonal wavelet can confine the border effects. Future availability of hyperspectral airborne sensors (HySpex, TASI, SEBASS) and satellite missions (EnMAP, HypSIPI) with largely varying spatial and spectral resolution will rise issues of spectral mixing and be a promising testbed for the developed methods.

# Zusammenfassung

Das übergeordnete Ziel vorliegender Arbeit ist die Ableitung räumlicher Verteilungsmuster von Mineralen. Als Datenbasis stehen hierfür eine Befliegung im reflektiven Spektralbereich des elektromagnetischen Spektrums mit dem hyperspektralen Sensor „HyMap“ sowie eine Befliegung im thermalen Spektralbereich mit dem multispektralen Sensor „AHS“ zur Verfügung. Weiterhin wurden umfangreiche spektroskopische Referenzmessungen durchgeführt sowie Mineralproben geochemisch analysiert.

Intensiver Braunkohleabbau hat im böhmischen Tagebaugebiet Sokolov zu beträchtlichen Umweltauswirkungen geführt. Eines der Hauptprobleme ist die Entstehung saurer Wässer durch die Sulfidverwitterung, englisch „Acid Mine Drainage“ (AMD) genannt. Die Versauerung von Gewässern und Böden geht einher mit der Mobilisierung von Schwermetallen. AMD wird von drei Prozessen bestimmt, i) der Verwitterung von vornehmlich Pyrit unter Freisetzung saurer Wässer, ii) der Neutralisation durch Karbonate, Hydroxide und Silikate, sowie iii) der Ausfällung von Sekundärmineralen. Aus der Kenntnis der räumlichen Verteilung dieser Schlüsselminerale kann auf die geochemischen Prozesse geschlossen werden. Die genannten Minerale weisen charakteristische spektrale Absorptionsmerkmale auf. Eine Identifizierung und Klassifizierung der Minerale ist über ihre Reflektions- und Emissionsspektren möglich. Somit ist die abbildende Spektroskopie ein adäquates Mittel zur räumlichen und zeitlichen Erfassung dieser.

Bedingt durch die lokalen geologischen und klimatischen Gegebenheiten sowie die verfügbaren Daten sind nur einzelne Minerale für die Auswertungen in Sokolov relevant. Dies sind der verwitterungsresistente Goethit als Indikator für potentiell AMD beeinflusste Flächen, das schnell neutralisierende Karbonat Siderit, das Tonmineral Kaolinit als langsam reagierender Säurepuffer sowie Quarz als inaktives Mineral, welches zu keiner Pufferung beiträgt.

Voraussetzungen für eine reproduzierbare Klassifizierung und insbesondere Quantifizierung sind eine standardisierte und qualitätskontrollierte Vorprozessierung, eine robuste Parametrisierung der spektralen Charakteristika sowie die Validation aller Arbeitsschritte. Dies sind die drei methodischen Schwerpunkte der vorliegenden Arbeit.

Auf Seiten der Vorprozessierung thermaler Bilddaten fehlte der wichtige Schritt der In-Flight Kalibrierung zur Kontrolle und Verbesserung der Laborkalibration des Sensors mit Hilfe von Referenzmessungen. Im Zuge dieser Arbeit wurde deshalb ein generisches In-Flight Kalibrationsmodul entwickelt, welches sich leicht an beliebige multi- und hyperspektrale thermale Sensoren adaptieren lässt. Das Modul integriert sich durch die gemeinsame Nutzung tabellierter Atmosphärenparameter und der Verwendung eines kompatiblen Strahlungstransfermodelles nahtlos in die Atmosphärenkorrektursoftware „ATCOR“. Die Oberflächentemperaturen gewählter Kalibrationsflächen werden aus den Bilddaten abgeleitet. Dies befreit von der Notwendigkeit, eine genaue und repräsentative Temperatur zum Zeitpunkt des Sensorüberfluges bestimmen zu müssen. Die spektrale Emissivität der Kalibrationsflächen als zweite notwendige Eingangsgröße für die Kalibration ist hingegen stabiler gegenüber Umwelteinflüssen und für viele Materialien bereits bekannt.

Bedingt durch Sensor und Aufnahmezeit waren die Anpassung des Atmosphärenmodells sowie die Entwicklung einer „Destriping“-Methode nötig. Zusammenfassend konnte die Problematik stark abfallender Emissionsspektren zu den Rändern des thermalen atmosphärischen Fensters behoben werden. Für die Thermaldaten konnte eine Vorverarbeitungsqualität, vergleichbar mit dem Standard für reflektive Daten, erreicht werden. Die nun realistischen Emissivitätsverläufe ermöglichten im Rahmen der Auswertungen die Quantifizierung ausgewählter Mineralfraktionen.

Ein wichtiger zusätzlicher Schritt war die Überführung der quantitativen Elementgehalte der Bodenproben in Mineralgehalte. Denn nicht die chemischen Elemente, sondern deren Verbindungen bestimmen die spektralen Absorptionseigenschaften. Dies wurde mit dem an das Testgebiet Sokolov angepassten normativen Modell „NORMA“ erreicht. Somit war die featurebasierte quantitative Analyse der Bilddaten möglich sowie die Verknüpfung von thermalem und reflektivem Spektralbereich über die Validation der entsprechenden Kartenprodukte.

Spektrale Absorptionseigenschaften im thermalen und reflektiven Spektralbereich weisen eine breite Formenvielfalt auf. Aufgrund von Mischungen und Überlagerungen können sehr komplexe spektrale Signaturen entstehen. Demgegenüber stehen relativ einfach gehaltene featurebasierte Auswertungsverfahren, wie z.B. die Näherung über geometrische Primitive (Dreiecke) oder über wenige Parameter (Absorptionstiefe, Fläche, Asymmetrie). Viele Informationen gehen somit für die nachgeschalteten Analysen verloren. Vor diesem Hintergrund ist es vorteilhaft, die Spektren auf unabhängigen Detaillierungsgraden analysieren zu können und flexiblere absorptionsbeschreibende Parameter zur Verfügung zu haben.

Dies kann die „Kontinuierliche Wavelet Transformation“ (CWT) leisten. Sie ist eine lineare Zerlegung eines Spektrums in seine Frequenzkomponenten (Skalen). Rauschstörungen werden in die kleinste Skale separiert. Schmalbandige Absorptionen werden in kleinen und mittleren Skalen hervorgehoben, weite Absorptionen bis hin zum Kontinuum treten in hohen Skalen hervor. CWT kann als Kreuzkorrelation des Spektrums mit dem Wavelet und als Multiskalen-Ableitung des geglätteten Spektrums interpretiert werden. Hier wurde das „Second Derivative of Gaussian“ ( $\text{DOG}_{m=2}$ ) Wavelet genutzt, da es der generellen Form einer Absorptionsbande ähnelt. Die Transformation ist damit die skalierte zweite Ableitung des mit einem Gauß-Filter geglätteten Spektrums. Die Filterbreite ist proportional zur Skale. Spektrale Absorptionsfeatures wie Maxima und Minima zeichnen sich durch Extrema in der Transformation aus, Wendepunkte erzeugen Nulldurchgänge. Das Verhalten der Transformationsfeatures zu höheren Skalen charakterisiert die Form und die spektrale Umgebung des Absorptionsfeatures. Effekte durch Rauschen, Beleuchtung, Feuchte, Korngrößenverteilung, Betrachtungswinkel sowie additive und lineare Einflüsse werden normalisiert oder in bestimmte Skalen separiert. Die vorliegende Arbeit entwickelt zwei spektrale Analyseverfahren, die auf dieser Multiskalen-Zerlegung beruhen.

Die „Wavelet Transform Feature Analyse“ ist eine Quantifizierungsmethode. Sie analysiert die Entwicklung von Extrema und Nulldurchgängen in der Transformation über die Skalen hinweg. Verwandte Features werden zu Vektoren verknüpft, deren Eigenschaften wiederum die Charakteristika der sie auslösenden spektralen Absorptionen beschreiben. Damit können die Absorptionen flexibel parametrisiert werden, z.B. über Länge, Krümmung, Wellenlänge und Entwicklung der Amplitude der Transformation über die Skalen der Vektoren. Die Parameter fin-

den Einzug in eine multivariate featurebasierte Quantifizierung, welche eine Verallgemeinerung der klassischen „Spektralen Feature Analyse“ und der „Spektralen Ableitungs Analyse“ ist. Der vorgestellte Ansatz nutzt die Multiskalen-Zerlegung des Eingangssignals mit den Vorteilen der selektiven Hervorhebung spektraler Komponenten, der Rauschunterdrückung, der Normalisierung sowie des selektiven Fokussierens auf verschiedene spektrale Details. Die „Wavelet Transform Feature Analyse“ benötigt kein Vorwissen über die spektrale Form und Lokalisierung charakteristischer Absorptionen. Auch ist keine methodische Unterscheidung zwischen Reflektions-, Emissions- oder Absorptionsspektren sowie Maxima oder Minima vorzunehmen. Komplexe spektrale Absorptionsformen können an das Regressionsmodell weitergereicht werden, da keine geometrischen Primitive zur Näherung der Signalstrukturen benötigt werden.

Der „*Wavelet Coherence Mapper*“ ist eine Klassifizierungsmethode. Sie verwendet die Wavelet Kohärenz, ein lokalisiertes Ähnlichkeitsmaß, so dass anstelle nur eines aggregierten Wertes ein gesamtes Ähnlichkeitsspektrum erzeugt wird. In Kombination mit mehreren Skalen erlaubt dies, ein Spektrum in wählbaren Detaillierungsgraden und Wellenlängenbereichen zu untersuchen. Die lokale Ähnlichkeit in der Umgebung einer schmalen Absorptionsbande bis hin zur globalen Ähnlichkeit des spektralen Kontinuums können verglichen werden. Die Wavelet Kohärenz ist ein normalisiertes lineares Korrelationsmaß und daher invariant gegenüber multiplikativen Faktoren und Amplitudendifferenzen. Die Kohärenz reagiert sensitiv auf Unterschiede in der Krümmungsrichtung, weniger auf die Stärke der Krümmung. Damit werden ausgeprägte tiefe und subtile flache Absorptionen gleich stark gewichtet. Amplitudendifferenzen zwischen verschiedenen Klassen werden ignoriert. Diese Eigenschaften machen den „Wavelet Coherence Mapper“ für die Identifikation potentiell schwach ausgeprägter Features oder bei charakteristischen Absorptionen verschiedener Stärke in spektralen Mischsignalen besonders geeignet.

Mit den waveletbasierten Methoden konnten in den hyperspektralen reflektiven Bilddaten Siderit, Goethit und Mischungen von Tonmineralfraktionen erfolgreich klassifiziert werden. In Verbindung mit der normativen Modellierung konnten Kaolinit und Goethit/ $\text{Fe}_2\text{O}_3$  quantifiziert werden. Die Regressionsmodelle erreichten mittlere Qualität. Die abgeleiteten Mineralverteilungsmuster waren plausibel und folgten den Erwartungen aus Feldbegehungen und Stratigraphie. Die praktische Anwendbarkeit der CWT auf einen anwendungsnahen hyperspektralen Datensatz konnte demonstriert werden. Zur Komplementierung und Validierung der aus den reflektiven Daten abgeleiteten Ergebnisse wurden die multispektralen thermalen Daten herangezogen. Über Bandratios und die normative Modellierung konnten Quarz und Kaolinit erfolgreich quantifiziert werden.

Die Validation zeigte eine generelle Übereinstimmung der Mineralverteilungsmuster und Abundanztrends für Flächen, die der mineralogischen Zusammensetzung der Trainingsdaten entsprachen. Es existierte jedoch ein Unterschied in den Absolutwerten der abgeleiteten Abundanzen. Die eingeschränkte Übertragbarkeit des reflektiven Kaolinit-Abundanzmodelles wurde an Mischungen von Kohle und Tonmineralen ersichtlich.

Die normalisierten Bild-, Feld- und Laborspektren zeigten eine gute Übereinstimmung. Jedoch existierten für die nicht normalisierten Spektren große Unterschiede in Amplitude und spektralem Kontrast. Diese Abweichungen wurden durch die normalisierende Eigenschaft von CWT

abgemindert. Im Ergebnis waren die Einflüsse auf die quantitativen Modelle eher gering. Eine weiterführende Evaluierung dieses Sachverhaltes war durch den limitierten Umfang der Referenzdaten jedoch nur eingeschränkt möglich.

Ein wichtiger Ansatz für zukünftige Arbeiten sind das Auftreten von Randeffekte in der CWT. Die Möglichkeiten zu deren Unterdrückung sollten weiter untersucht werden, so dass die eventuell betroffenen randnahen Transformationswerte mit höherer Konfidenz genutzt werden können. Die hier verwendete Maskierung potenziell betroffener Bänder in der „Wavelet Coherence Mapper“ Methode führte zu Klassifikationsfehlern in höheren Skalen. Mit zunehmender Skale steigt auch die Zahl der potentiell betroffenen Bänder. Dies führte zum Ausschluss diskriminativer spektraler Eigenschaften bei der Klassifikation.

Die Nutzung schneller CWT Implementierungen als Filterbank kann zukünftige Prozessierungszeiten signifikant verkürzen. Der induzierte Wechsel zu einem „Spline Biorthogonal“ Wavelet kann die Zahl potenziell von Randeffekten betroffener Bänder weiter einschränken. Die meisten vorteilhaften Eigenschaften des hier verwendeten „Second Derivative of Gaussian“ ( $\text{DOG}_{m=2}$ ) Wavelets bleiben dann erhalten.

Die entwickelten Verfahren sind generisch. Sie sollten im Hinblick auf eine zukünftige Verfügbarkeit weiterer flugzeuggetragener hyperspektraler Sensorsysteme (HySpex, TASI, SEBASS) und Satellitenmissionen (EnMAP, HypIRI) geprüft werden. Insbesondere die unterschiedlichen geometrischen und spektralen Auflösungen und die damit verbundene Mischpixelproblematik spielen eine wichtige Rolle. Daraus ergeben sich vielversprechende Anwendungsfälle für die entwickelten Methoden.



# Contents

<b>Abstract</b>	<b>I</b>
<b>Zusammenfassung</b>	<b>III</b>
<b>Table of Contents</b>	<b>VII</b>
List of Tables . . . . .	X
List of Figures . . . . .	XI
<b>Nomenclature</b>	<b>XIV</b>
Abbreviations . . . . .	XIV
Symbols . . . . .	XVI
<b>1. Introduction and Objectives</b>	<b>1</b>
1.1. Specific Study Objectives . . . . .	3
1.2. Thesis Structure . . . . .	4
<b>2. Basics</b>	<b>5</b>
2.1. Study Area . . . . .	5
2.1.1. Geologic Setting . . . . .	5
2.1.2. Coal Mining and Environmental Problems . . . . .	6
2.2. Application Objective & Mineral Mapping Candidates . . . . .	7
2.3. Passive Optical Remote Sensing for Mineral Quantification . . . . .	9
2.3.1. Electromagnetic Radiation as Signal . . . . .	9
2.3.2. Signal Interaction with Atmosphere . . . . .	10
2.3.3. Signal Interaction with Earth's Surface . . . . .	11
2.3.4. Spectral Signature Retrieval . . . . .	12
2.3.5. Spectral Characteristics of relevant Minerals . . . . .	13
2.4. Decomposition of Signals . . . . .	18
2.4.1. Continuous Wavelet Transform . . . . .	20
2.4.2. Wavelet Choice . . . . .	24
2.4.3. Cone Of Influence . . . . .	27
2.4.4. Wavelet Transform Properties . . . . .	28
2.4.5. Hyperspectral Signature Transform . . . . .	31
<b>3. Mineral Classification and Quantification</b>	<b>34</b>
3.1. State-of-the-Art & Open Issues . . . . .	34
3.2. Continuous Wavelet Transform based Developments . . . . .	38
3.2.1. Signal Similarity Measure . . . . .	39
3.2.2. Signal Feature Reconstruction . . . . .	43
3.3. Application to Hyperspectral Signatures . . . . .	47
3.3.1. Implementation . . . . .	47

3.3.2.	Wavelet Coherence Mapper . . . . .	49
3.3.3.	Wavelet Transform Feature Analysis . . . . .	52
<b>4.</b>	<b>Thermal Imagery Pre-Processing</b>	<b>59</b>
4.1.	State-of-the-Art & Open Issues . . . . .	59
4.2.	In-Flight Calibration Module Developments . . . . .	60
4.2.1.	Atmospheric Modelling . . . . .	60
4.2.2.	In-Flight Calibration Module . . . . .	63
4.2.3.	Calibration Target Selection . . . . .	65
4.2.4.	Temperature Estimation . . . . .	67
4.2.5.	At-Ground Calibration . . . . .	67
<b>5.</b>	<b>Workflow &amp; Product Validation</b>	<b>68</b>
5.1.	State-of-the-Art & Open Issues . . . . .	68
5.2.	Validation Concept Developments . . . . .	69
5.2.1.	Empirical Validation Options . . . . .	70
5.2.2.	Two-Tier Workflow Model & Validation Concept . . . . .	71
<b>6.</b>	<b>Application Scenario Mining Environment</b>	<b>73</b>
6.1.	Overall Workflow Concept . . . . .	74
6.2.	Data Sets and Measurement Protocols . . . . .	76
6.3.	Spectroscopic Data Pre-Processing . . . . .	78
6.3.1.	Field & Laboratory Measurement Corrections . . . . .	78
6.3.2.	Airborne Imagery Corrections . . . . .	79
6.4.	Normative Modelling . . . . .	89
6.5.	Mineral Classification . . . . .	92
6.5.1.	Wavelet based Classification of Siderite, Goethite & Kaolinite . . . . .	92
6.5.2.	Wavelet based Classification of Clays . . . . .	104
6.6.	Mineral Quantification . . . . .	113
6.6.1.	Wavelet based Quantification of Goethite & Kaolinite . . . . .	113
6.6.2.	Band-Ratio based Quantification of Quartz & Kaolinite . . . . .	127
6.7.	Evaluation of Processing Workflow & Results . . . . .	134
6.7.1.	Validation of Processing Steps . . . . .	134
6.7.2.	Validation of Quantification Results . . . . .	146
<b>7.</b>	<b>Discussion &amp; Summary</b>	<b>153</b>
7.1.	Critical Discussion of Developments . . . . .	153
7.1.1.	Review of Specific Research Objectives . . . . .	153
7.1.2.	Wavelet based Classification & Quantification Concepts Review . . . . .	154
7.1.3.	Thermal Pre-Processing Concept Review . . . . .	157
7.1.4.	Workflow Validation Concept Review . . . . .	158
7.2.	Critical Discussion of Developed Concepts' Application . . . . .	159
7.2.1.	Review of Processing Results . . . . .	159
7.2.2.	Major Uncertainty Sources . . . . .	161

7.2.3. Review of Developed Concepts' Performance . . . . .	161
7.3. Recommended Application Scenario for Developed Concepts . . . . .	164
<b>8. Outlook</b>	<b>166</b>
<b>References</b>	<b>167</b>
<b>Appendix A. Acid Mine Drainage Supplements</b>	<b>i</b>
<b>Appendix B. Normative Modelling Supplements</b>	<b>vi</b>
<b>Appendix C. Sensor Descriptions</b>	<b>xvi</b>
<b>Appendix D. Mathematical Supplements</b>	<b>xvii</b>
<b>Appendix E. Validation Supplements</b>	<b>xxi</b>
<b>Erklärung</b>	<b>xxvi</b>
<b>Acknowledgements</b>	<b>xxvii</b>
<b>Curriculum Vitae</b>	<b>xxviii</b>
Scientific Publications . . . . .	xxix
Related Unpublished Reports . . . . .	xxix

# List of Tables

3.1. Analogy of Spectral Feature Analysis & Wavelet Transform Feature Analysis . . .	56
3.2. Analogy of Spectral Derivative Analysis & Wavelet Transform Feature Analysis .	57
4.1. Atmospheric modelling parameters for MODTRAN . . . . .	61
6.1. Spectra pre-processing parameters reflective domain . . . . .	78
6.2. Spectra pre-processing parameters thermal domain . . . . .	79
6.3. Image pre-processing parameters reflective domain . . . . .	82
6.4. Image pre-processing parameters thermal domain . . . . .	88
6.5. Size of COI area depending on scale . . . . .	94
6.6. Quantiles of wavelet coherence for kaolinite, goethite & siderite classification . .	94
6.7. Clays classification approaches and thresholds . . . . .	105
6.8. Clays classification confusion matrix WCM vs. SAM (2,000-2,500 nm) . . . . .	107
6.9. Clays classification confusion matrix WCM vs. SAM (2,102-2,242 nm) . . . . .	112
6.10. NORMA versus ASD correlations . . . . .	115
6.11. Parameters used for the clustering into common features . . . . .	117
6.12. Feature parameters of elevated correlation to kaolinite abundance . . . . .	120
6.13. Feature parameters of elevated correlation to goethite abundance . . . . .	120
6.14. Feature parameters of elevated correlation to Fe <sub>2</sub> O <sub>3</sub> abundance . . . . .	120
6.15. Final feature parameters used for kaolinite, goethite & Fe <sub>2</sub> O <sub>3</sub> MLR models . . .	121
6.16. Correlations between NORMA abundances & AHS band combinations . . . . .	128
6.17. Kaolinite MLR model validation on library spectra . . . . .	142
6.18. Goethite & Fe <sub>2</sub> O <sub>3</sub> MLR model validation on library spectra . . . . .	142
6.19. Summary of validation results for processing steps . . . . .	145
6.20. Deviations of kaolinite quantitative maps in reflective & thermal domain . . . . .	151
A.1. Reactivity of acidity neutralising minerals . . . . .	ii
A.2. pH buffering reactions . . . . .	iv
B.1. SEDNORM mineral calculation sequence . . . . .	ix
B.2. Laboratory analysis methods . . . . .	xi
B.3. XRD results per sample . . . . .	xii

# List of Figures

2.1. Geologic map of Sokolov Basin . . . . .	5
2.2. Radiation sources and atmospheric windows . . . . .	10
2.3. Reflectance & emissivity spectra of selected iron bearing minerals . . . . .	15
2.4. Reflectance & emissivity spectra of carbonates . . . . .	15
2.5. Reflectance & emissivity spectra of feldspars . . . . .	15
2.6. Reflectance & emissivity spectra of clays . . . . .	16
2.7. Reflectance & emissivity spectra of quartz & micas . . . . .	16
2.8. Reflectance & emissivity spectra of other materials . . . . .	16
2.9. Haar wavelet . . . . .	19
2.10. Morlet wavelet & Heisenberg boxes . . . . .	22
2.11. Schematic signal interrogation by a wavelet . . . . .	23
2.12. Gaussian derivatives . . . . .	26
2.13. Second Derivative of Gaussian wavelet . . . . .	27
2.14. Signal interrogation by a DOG wavelet . . . . .	29
2.15. Edge detection by a DOG wavelet . . . . .	30
2.16. CWT of an albite emissivity spectrum . . . . .	31
2.17. CWT of a jarosite-clay mixture reflectance spectrum . . . . .	32
3.1. Correlation scaleogram example . . . . .	36
3.2. Fingerprint in scale space example . . . . .	37
3.3. Modulus maxima lines example . . . . .	37
3.4. Wavelet coherence & cross wavelet transform for several synthetic signals . . . . .	41
3.5. Wavelet coherence between a jarosite-clay mixture & jarosite endmember . . . . .	42
3.6. Analytic $\text{DOG}_{m=2}$ wavelet transform of a Gaussian . . . . .	44
3.7. Test signals for modulus maxima line parameter retrieval . . . . .	46
3.8. Padding & CWT of a kaolinite reflectance spectrum . . . . .	48
3.9. Wavelet coherence between an unknown image spectrum & siderite endmember . . . . .	50
3.10. Workflow for the Wavelet Coherence Mapper method . . . . .	50
3.11. Workflow for the Wavelet Transform Feature Analysis method . . . . .	52
3.12. Transform feature vector extraction . . . . .	53
3.13. Transform feature vector clustering . . . . .	54
3.14. Common feature pattern extraction . . . . .	58
4.1. View geometry influence on path radiance and transmittance . . . . .	61
4.2. Atmospheric state influence on path radiance and transmittance . . . . .	62
4.3. Thermal in-flight calibration procedure . . . . .	64
4.4. In-flight calibration example for vegetation target . . . . .	66
4.5. In-flight calibration example . . . . .	66
5.1. Two-tier workflow model and validation concept . . . . .	72
6.1. Impressions of active coal mine and AMD site . . . . .	73

6.2. Processing concept flow chart . . . . .	75
6.3. Imagery pre-processing steps . . . . .	80
6.4. In-flight calibration determined per-band gains & offsets in the reflective domain . . . . .	81
6.5. Influence of water vapour & atmosphere model on emissivity spectra . . . . .	83
6.6. In-flight calibration determined per-band gains in the thermal domain . . . . .	84
6.7. Application of destriping filter to AHS imagery . . . . .	86
6.8. Artefacts after application of destriping filter . . . . .	86
6.9. Application of diffusion filter to AHS imagery . . . . .	87
6.10. Aggregated mineral occurrences within 50 samples according to XRD . . . . .	89
6.11. Active mines image subset used for showcasing the analysis . . . . .	92
6.12. Linear mixture series of kaolinite and siderite class templates . . . . .	93
6.13. Full spectral range kaolinite best match . . . . .	96
6.14. Wavelet Coherence Mapper classification of kaolinite . . . . .	97
6.15. Full spectral range goethite best match . . . . .	98
6.16. Wavelet Coherence Mapper classification of goethite . . . . .	99
6.17. Full spectral range siderite best match . . . . .	100
6.18. Wavelet Coherence Mapper classification of siderite . . . . .	101
6.19. Siderite classification false positives . . . . .	101
6.20. Clay classification template spectra . . . . .	104
6.21. Slope influence on SAM classification of clays . . . . .	106
6.22. Comparison of clays classification class assignment SAM vs. WCM . . . . .	108
6.23. Reflectance-based SAM classification of clays - 2,000-2,500 nm . . . . .	109
6.24. Reflectance-based SAM classification of clays - 2,102-2,242 nm . . . . .	109
6.25. Wavelet-based SAM classification of clays - 2,000-2,500 nm . . . . .	110
6.26. Wavelet-based SAM classification of clays - 2,102-2,242 nm . . . . .	110
6.27. Wavelet Coherence Mapper classification of clays - 2,000-2,500 nm . . . . .	111
6.28. Wavelet Coherence Mapper classification of clays - 2,173 nm . . . . .	111
6.29. Correlation scaleograms for NORMA goethite & kaolinite . . . . .	115
6.30. Scatter plot matrix of NORMA mineral inter-correlations . . . . .	116
6.31. Example of transform maxima feature vector clustering . . . . .	117
6.32. Wavelet Transform Feature Analysis quantification of kaolinite . . . . .	123
6.33. Wavelet Transform Feature Analysis quantification of $\text{Fe}_2\text{O}_3$ . . . . .	123
6.34. Kaolinite & goethite MLR regression quality . . . . .	124
6.35. Common feature patterns for kaolinite and $\text{Fe}_2\text{O}_3$ . . . . .	125
6.36. Impact of $\text{Fe}_2\text{O}_3$ common feature pattern matching on mapping result . . . . .	126
6.37. Match of image spectra to the common feature patterns of $\text{Fe}_2\text{O}_3$ . . . . .	126
6.38. Scatter plot of NORMA quartz vs. kaolinite and NORMA quartz vs. $\text{Fe}_2\text{O}_3$ . . . . .	128
6.39. Workflow for AHS data quantification . . . . .	129
6.40. AHS band combinations for TIR quantification . . . . .	130
6.41. Scatter plot NORMA kaolinite versus AHS band ratio #72/#74 . . . . .	131
6.42. Scatter plot NORMA kaolinite versus AHS band ratio #76/#77 . . . . .	131
6.43. Scatter plot NORMA quartz versus AHS band ratio #72/#78 . . . . .	132

6.44. Band ratio quantification of quartz . . . . .	133
6.45. Band ratio quantification of kaolinite . . . . .	133
6.46. Example of field vs. laboratory reflectance spectra (ASD vs. ASD) . . . . .	134
6.47. Example of field vs. laboratory emissivity spectra (FT-IR vs. SR-5000) . . . . .	136
6.48. Example of field vs. image reflectance spectra (ASD vs. HyMap) . . . . .	137
6.49. Example of field vs. image emissivity spectra (FT-IR vs. AHS) . . . . .	138
6.50. Cross-validation of individual VNIR-SWIR transform feature parameters . . . . .	140
6.51. Cross-validation of VNIR-SWIR MLR abundance models . . . . .	140
6.52. Field vs. laboratory spectra derived kaolinite & Fe <sub>2</sub> O <sub>3</sub> abundances . . . . .	141
6.53. Abundance sum constraint conformity of quartz, kaolinite & Fe <sub>2</sub> O <sub>3</sub> maps . . . . .	147
6.54. Siderite versus Fe <sub>2</sub> O <sub>3</sub> distribution comparison . . . . .	147
6.55. Comparison of kaolinite abundance maps versus clay classification . . . . .	148
6.56. Comparison of quantification of kaolinite - kaolinite extraction site . . . . .	149
6.57. Comparison of quantification of kaolinite - overburden dump site . . . . .	150
A.1. pH buffering sequence in tailings impoundments . . . . .	iv
A.2. AMD reaction paths and element circulations model . . . . .	v
B.1. Stratigraphy of Sokolov Basin . . . . .	vi
B.2. SEDNORM mineral calculation sequence . . . . .	vii
B.3. Wet chemistry elemental concentrations of samples . . . . .	xiii
B.4. SEDNORM mineral abundances of samples . . . . .	xiv
B.5. NORMA mineral abundances of samples . . . . .	xv
E.1. Mineral abundance difference between NORMA and SEDNORM . . . . .	xxi
E.2. Field versus laboratory reflectance spectra (ASD vs. ASD) . . . . .	xxii
E.3. Field versus laboratory emissivity spectra (FT-IR vs. SR-5000) . . . . .	xxiii
E.4. Field versus image reflectance spectra (ASD vs. HyMap) . . . . .	xxiv
E.5. Field versus image emissivity spectra (FT-IR vs. AHS) . . . . .	xxv

# Nomenclature

## Abbreviations

AHS	Airborne Hyperspectral Scanner (imaging spectrometer, 80 bands)
AMD	Acid Mine Drainage
ARD	Acid Rock Drainage
ASD	Analytical Spectral Devices spectroradiometer
ASU	Arizona State University
BRDF	Bidirectional Reflectance Distribution Function
cal	Calibration
COI	Cone Of Influence
CWT	Continuous Wavelet Transform
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
deriv	Derivative
DLR	Deutsches Zentrum für Luft- und Raumfahrt, Germany
DOG	Derivative Of Gaussian
DOG <sub>m=2</sub>	Second order Derivative Of Gaussian
DWT	Discrete Wavelet Transform
DZC	Downward Zero Crossings
FFT	Fast Fourier Transform
FOV	Field Of View
FWHM	Full Width at Half Maximum
FT-IR	Fourier Transform InfraRed spectroradiometer
GDBSCAN	Generalized version of DBSCAN
GIFOV	Ground-projected Instantaneous Field Of View
HyMap	Hyperspectral Mapper (imaging spectrometer, 128 bands)
IIFOV	Instantaneous Field Of View
INTA	Instituto Nacional de Técnica Aeroespacial, Spain
JHU	Johns Hopkins University
LUT	Look-Up-Table
MAX	MAXima
MIN	MINima



---

MIR	Mid InfraRed
MLR	Multiple Linear Regression
MNF	Maximum/minimum Noise Fraction
MP	Molar Proportion
NEM	Normalised Emissivity Method
NIR	Near InfraRed
OpAiRS	Optical Airborne Remote Sensing and Calibration Facility of DLR
PTB	Physikalisch-Technische Bundesanstalt
QC	Quality Control
RMS	Root Mean Square
RMSE	Root Mean Square Error
RPD	Relative Percent Difference
SAM	Spectral Angle Mapper
SCM	Spectral Correlation Mapper
SNR	Signal to Noise Ratio
SR-5000	CI Systems spectroradiometer
stddev	Standard deviation
SWIR	Short-Wave InfraRed
TES	Temperature-Emissivity-Separation
TIR	Thermal InfraRed
USGS	U.S. Geological Survey
UV	Ultra Violet
UZC	Upward Zero Crossings
val	Validation
VIS	VISible
VNIR	Visible Near InfraRed
WCM	Wavelet Coherence Mapper
WTC	WaveleT Coherence
WTFA	Wavelet Transform Feature Analysis
wvl	Wavelength
XRD	X-Ray Diffraction
XWT	Cross Wavelet Transform

## Symbols

### units

$\epsilon$	emissivity [ - ]	(ratio)
$\lambda$	wavelength [ nm or $\mu\text{m}$ ]	nanometre or micrometre
$L$	spectral radiance [ $\frac{W}{\text{sr m}^2 \text{m}}$ ]	Watts per steradian, area & wavelength
$\rho$	reflectance [ - ]	(ratio)
$\omega, \xi$	frequency [ Hz ]	Hertz
$s$	scale [ - ]	(unitless)
$t, u$	time [ s ]	seconds
$\tau$	transmittance [ - ]	(ratio)
$\omega_i$	mass fraction [ wt-% ]	weight percent

### operators

$m!$	factorial of $m$	
$\Gamma(m)$	gamma function of argument $m$	
$ f(t) $	absolute value or modulus of $f(t)$	
$\ f\ $	$L^2$ norm of $f(t)$	$\sqrt{\int_{-\infty}^{\infty}  f(t) ^2 dt}$
$f^{(m)}(t)$	$m$ -th order derivative of $f(t)$ in respect to $t$	$\frac{d^{(m)}}{dt^{(m)}} f(t)$
$(f * g)(t)$	continuous convolution of $f(t)$ with $g(t)$	$\int_{-\infty}^{\infty} f(t-u)g(u) du$
$\psi^*(t)$	complex conjugate of $\psi(t)$	
$\langle Wf(u, s) \rangle$	smoothing of $Wf(u, s)$ over time and scale	

### transforms

$\hat{f}(\omega)$	Fourier transform of $f(t)$	$\int_{-\infty}^{\infty} f(t) \cdot e^{-2\pi i t \omega} dt$
$\hat{f}^{-1}(\omega)$	inverse Fourier transform of $\hat{f}(\omega)$ yielding $f(t)$	$\int_{-\infty}^{\infty} \hat{f}(\omega) \cdot e^{2\pi i \omega t} d\omega$
$Sf(u, \xi)$	Gabor transform of $f(t)$ with window function $g(t)$	$\int_{-\infty}^{\infty} f(t) g(t-u) \cdot e^{-2\pi i t \xi} dt$
$Wf(u, s)$	CWT of $f(t)$ with mother wavelet $\psi(t)$	$\int_{-\infty}^{\infty} f(t) \cdot \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt$
$XWT(u, s)$	cross wavelet transform of signals $f(t)$ & $g(t)$	$Wf(u, s) \cdot W^*g(u, s)$
$WTC(u, s)$	wavelet coherence of signals $f(t)$ & $g(t)$	$\frac{\langle \frac{1}{s} XWT(u, s) \rangle}{\sqrt{\langle \frac{1}{s} Wf(u, s)^2 \rangle \cdot \langle \frac{1}{s} Wg(u, s)^2 \rangle}}$

# 1. Introduction and Objectives

The overarching objective of this research is to use hyperspectral reflective and multispectral thermal imagery for mapping and quantification of selected minerals in a mining environment. Benefits of the availability of information in both wavelength domains are to be shown.

The demonstration site Sokolov, Czech Republic, is affected by extensive lignite mining in past and present. Active and closing open pit mines, dump sites and ancient mining remnants are characteristic to this area. Reclamation problems due to low-quality effluents are observed in this particular area which led to an acidification of soils and of surface waters. The resulting decrease in pH and mobilisation of heavy metals has an adverse environmental impact on aquatic life, vegetation and soil fertility. Spatial mapping of the source of acidification, so called acid mine drainage (AMD), its pathways and possibly affected areas are of interest in this study area.

AMD occurs where mining operations have exposed sulphur rich materials to atmospheric oxygen and water. The intermixed sulphide minerals, predominantly pyrite, are readily oxidised and hydrolysed, whereupon considerable amounts of acidity and mobilized metal cations are generated. Reaction of the acid solution with neutralising minerals leads to the precipitation of characteristic secondary mineral assemblages of iron oxides, hydroxides and sulphates. As a result, zones of indicative minerals for the pH buffering regimes do form. Weathering under precipitation-controlled climates, as in Sokolov, transforms many of this meta-stable precipitates into more resilient minerals. Spatial abundance estimates of iron oxides, sulphides and sulphates on the one hand and carbonates, feldspars and clays on the other hand are of major importance for describing the scale of the problem. Former minerals indicate potential acid generation locations or areas affected by AMD. Latter minerals describe acid neutralisation zones or areas of sufficiently high buffering capacity. As a third group quartz dominated soils with a greatly reduced acid neutralisation capacity are candidates for not damped acidification when affected by AMD (Blowes *et al.*, 2003). In summary, mapping of the spatial distribution patterns of the various key minerals and mineral groups can support the evaluation of the soil quality, the risk assessment of possible contamination of surface and ground water bodies, the planning of specific remediation efforts or the assessment of the (future) economical usage potential of former mining sites. Typically a monitoring is necessary to detect and document changes in the mineral distributions which govern the acidification process over time.

Above minerals have been demonstrated to be distinguishable by laboratory spectroscopic methods in the reflective and thermal wavelength domains (Salisbury *et al.*, 1991; Clark *et al.*, 1993; Christensen *et al.*, 2000). Airborne imaging spectroscopy is therefore the tool of choice to assess spatial patterns of potential AMD generating and buffering areas. The employed sensors typically offer a spectral resolution well above 100 narrow bands in each the reflective (500 - 2,500 nm, VNIR-SWIR) and thermal (8,000 - 14,000 nm, TIR) atmospheric windows. Practical applicability of the prevailing VNIR-SWIR reflectance spectroscopy to this problem has been successfully demonstrated by numerous authors, e.g. Krüger *et al.* (1998); Swayze *et al.* (2000); Mars & Crowley (2003); Choe *et al.* (2008); Alfaro (2009); Richter (2010a); Rianza *et al.* (2011). However, shortcomings due to spectral similarities or indeterminacy of various key minerals such as carbonates, clays, feldspars, iron oxides and quartz exist when atmospheric absorp-

tions, limited sensor spectral resolution and mixing degrade the characteristic spectral features. TIR spectroscopy has the potential to close these gaps by exposing (additional) characteristic features for these minerals which was demonstrated by various authors, e.g. [Vaughan \*et al.\* \(2003\)](#); [Schodlok \(2004\)](#); [Vaughan \*et al.\* \(2005\)](#); [Hecker \(2013\)](#). In this sense the combination with TIR data allows for specification of the information derived from the VNIR-SWIR data and for adding important layers exclusively determinable in the thermal domain as shown by [Roberts \*et al.\* \(2011\)](#). Furthermore this brings up the possibility to cross-validate derived results with the information from the complementary wavelength domain. This additional validation of abundance models and spatial patterns is rarely done.

For the proposed study an airborne hyperspectral reflective data set (HyMap sensor) and a multispectral thermal data set (AHS sensor) were available. Together with the two airborne surveys, field reference spectral measurements in the reflective (ASD Field Spec instrument) and in the thermal (D&P FT-IR Model 102 instrument) domain for calibration and validation have been taken. Mineral soil samples have been collected in parallel that were analysed for mineralogical and geochemical parameters by XRD and wet chemistry.

Justified by the demonstration site's geologic and climatic situation as well as the available data this study derives quantitative abundance maps for goethite as potential AMD indicator, kaolinite as slowly reacting acidity buffering mineral and quartz as inert mineral. Furthermore, qualitative distribution of the fast buffering carbonate mineral siderite and mixtures of clay minerals are investigated.

These information layers can serve as input for further geo-spatial modelling tasks not covered by this research, such as risk analysis, guidance for remediation tasks, soil fertility determination and pH buffering capacity estimation. The qualitative and quantitative analysis of both reflective and thermal data sets allows a more complete picture of the relevant mineral's spatial patterns in the affected areas than either of the datasets can reveal.

Reproducible classification and especially quantification require a standardised and quality controlled pre-processing, a robust parameterization of spectral absorptions and a validation of all individual processing steps leading to the final result. This study therefore focuses on the methodological advancement of these aspects. Three specific study objectives are derived which are detailed in the following.

## 1.1. Specific Study Objectives

Mineral classification and quantification typically rely on characteristic absorption/emission features within the spectra that are caused by elemental and molecular constituents ( $\text{Fe}^{2+}$ ,  $\text{Fe}^{3+}$ ,  $\text{Al-OH}$ ,  $\text{CO}_3$ ,  $\text{Si-O}$ ,  $\text{Al-O}$ , etc.). These features vary in shape and scale, from steep narrow drops to subtle wide troughs which also can overlap. An interesting concept for hyperspectral data analysis are methods based on continuous wavelet transforms of the spectra. These methods can take advantage of normalising, enhancing and decomposing hyperspectral signals into a scale-wavelength representation. This scale-wavelength representation of a spectrum highlights its different features at differing scales while preserving their wavelength localisation and complex shapes. This allows for a separate analysis which is of benefit for material classification and constituent quantification. Especially for mixtures with possibly overlapping and weak features, the wavelet based methods have advantages over the traditional spectral feature analysis. Yet the wavelet based methods leave space for improvements and partially lack practical application.

THE FIRST OBJECTIVE of this study therefore is to improve wavelet transform based classification and quantification methods and to demonstrate their practical applicability not only to spectral libraries but also to hyperspectral imagery.

Classification and especially quantification need datasets to be of sufficiently high quality. This encompasses calibration, validation, atmospheric and illumination correction and geocoding of the imagery. For the reflective domain there exist well established processing chains. In the thermal domain however there persist open issues. Many studies report residual effects after atmospheric correction and temperature-emissivity-separation forcing the use of empirical corrections on top of the calculated emissivity spectra. A standardised pre-processing chain, completed by adding in-flight calibration with field reference data, as established for the reflective domain, would allow for reproducible and quality controlled data layers.

THE SECOND OBJECTIVE is the development of a generic in-flight calibration module that integrates with and complements the thermal pre-processing chains of the AHS sensor operator INTA and the processing and archiving facility at DLR. Incorporation of the FT-IR field reference measurements and MODTRAN radiative transfer model are crucial.

The calibrated and validated reflectance and emissivity spectral data can be used for mineral classification and quantification. It is essential to ensure the accuracy of the input data layers, models and obtained results, such that different information layers can be combined into a reliable and reproducible map product. This leads to the need of standardized measuring, pre-processing and analysis workflows which are calibrated and validated against reference data as well as cross-validated against additional information.

THE THIRD OBJECTIVE of this study is to provide a concept for the inclusion of validation and plausibility checks covering each of the individual processing steps in the workflow. This enables to identify sources of uncertainty and aims towards the needed accuracy assessment of input data used for further modelling scenarios beyond the scope of this research.

## 1.2. Thesis Structure

The thesis at hand is structured in the following way: Starting from the general introduction and the derived three specific study objectives outlined above, Chapter 2 - *Basics* introduces the fundamental concepts on which the subsequent chapters rely. Starting with the demonstration site Sokolov, its geologic setting and environmental problems are discussed. From there, the thematic objective of this research is derived and candidate minerals that realistically can be mapped are reasoned. Next, an introduction to passive optical remote sensing in the reflective and thermal domains with focus on mineral quantification is given. This includes concepts signal interaction with atmosphere and ground surface, and the spectral characteristics of AMD relevant minerals. Finally, an introduction to signal decomposition is given leading to continuous wavelet transform. The chapter concludes with a reasoning for the choice of the wavelet, the resulting properties of the transform and implications when applied to hyperspectral signatures.

Chapter 3 - *Mineral Classification and Quantification* tackles the specific study objective of spectral feature based mineral classification/quantification using continuous wavelet transform. It provides a literature review and identifies open aspects. Based on the latter the Wavelet Coherence Mapper and Wavelet Transform Feature Analysis are developed.

Chapter 4 - *Thermal Imagery Pre-Processing* targets the specific study objective of thermal data pre-processing. Starting from a literature review the current open aspects are highlighted and a generic in-flight calibration module for the thermal domain is devised.

Chapter 5 - *Workflow & Product Validation* elaborates the specific study objective of validating a processing workflow. After a literature review a concept for validating a multi-sensor processing workflow based on the intermediate and final products is presented.

In Chapter 6 - *Application Scenario Mining Environment* the practical applicability of the developed concepts to an AMD affected site is shown. It covers a typical processing workflow from data acquisition over pre-processing to analysis and validation. Special focus is set on the congruent handling of reflective and thermal domain (within the limits of the available data's spectral resolution) to assure comparable and validated results.

The thesis at hand is concluded by a critical discussion of the results and performance of the developed concepts in view of the initial research questions in Chapter 7 - *Discussion & Summary*. This leads to a summary and outlook given in the two subsequent chapters.

Five Appendices A - *Acid Mine Drainage Supplements*, B - *Normative Modelling Supplements*, C - *Sensor Descriptions*, D - *Mathematical Supplements* and E - *Validation Supplements* give additional information on the acidification process, the site's stratigraphic setting and the normative modelling of the collected mineral samples, the technical specifications of the spectrometers used for data acquisition, the chosen wavelet's mathematical background, as well as the full validation plots.

## 2. Basics

The chapter briefly introduces the study area and its environmental issues. The mapping goals are derived and candidate minerals reasoned. Concepts of passive optical remote sensing are presented. Finally, signal decomposition with continuous wavelet transform is introduced.

### 2.1. Study Area

The study area is located in the Sokolov Basin occupying about 200 km<sup>2</sup> of the Western Eger Rift in North-West Bohemia province, Czech Republic (see Figure 2.1). The hilly area is mostly covered by agricultural fields, forests and open casts due to the presence of lignite, kaolin and clay deposits that have been mined since historic age.

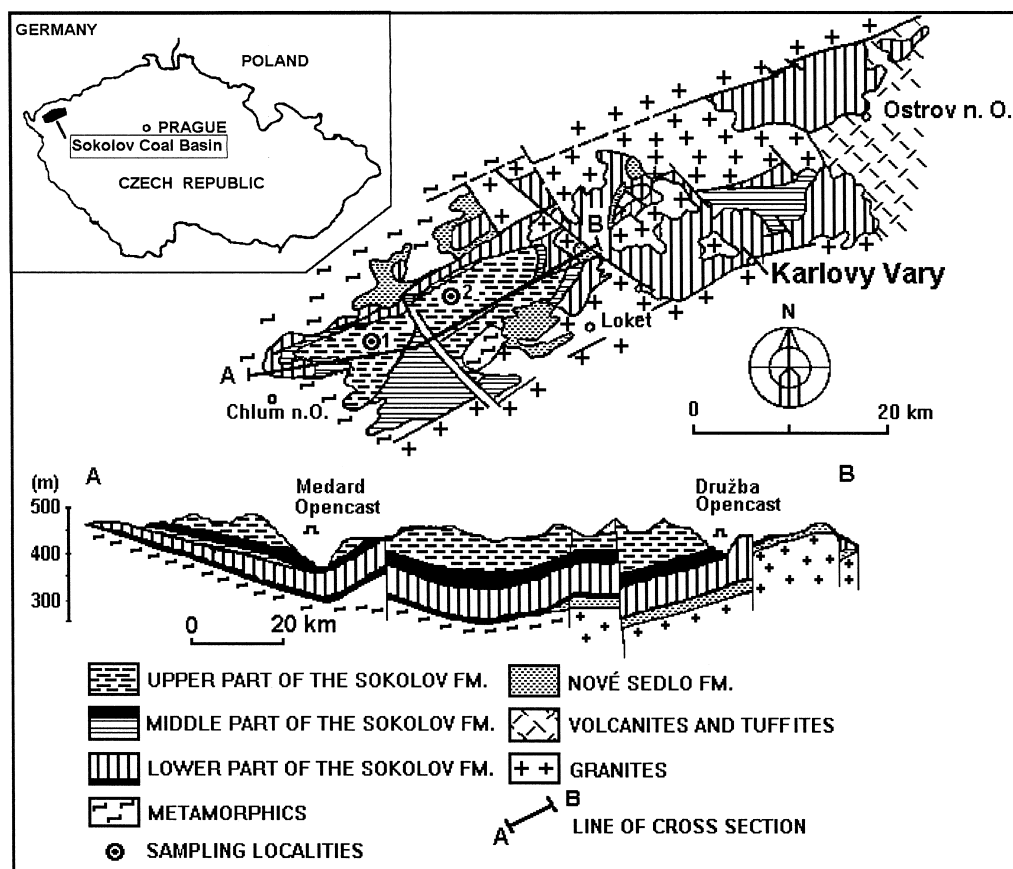


Figure 2.1.: Simplified geologic map and cross section of the Sokolov Basin from Krátek et al. (1998).

#### 2.1.1. Geologic Setting

Tectonic activities coupled with intra-plate volcanism, weathering and subsidence during warm climate shaped a basin that is framed by numerous faults. It bears deposits of characteristic properties. Rojčík (2004) proposes an updated stratigraphy of the Sokolov Basin (see Appendix B.1 - *Sokolov Stratigraphy*). Its base being formed by a complex set of metamorphic and granitic units of Variscan and pre-Variscan origin which upper layers have undergone strong weathering leaving kaolins and silica residues. On top of this weathered crystalline basement rests the Staré Sedlo formation of fluvial sandstones and conglomerates. During its deposition kaolinitic weathering

was still persistent resulting in the presence of re-deposited kaolinitic clays. The superseding Nové Sedlo formation is characterised by volcanic rocks with interbedded coal bearing sediments marking several episodes of synchronous tectonic subsidence and volcanism. This resulted in multiple alternating layers of effusive and pyroclastic rocks with kaolinized rocks and coal which are finally overlain by a massive layer of tuffs and tuffites marking the first culmination phase of volcanic and tectonic activity in the basin. The overlaying Sokolov formation is composed of coal bearing sediments with interfingering alluvial and volcanic deposits. The transition between organic and inorganic sedimentation coincides with phases of changing volcano-tectonic activity. A layer of pyroclastic and effusive rocks marks its second culmination phase. Superseding are the massive laminated clays and claystones of the Cypris formation consisting of kaolinite, illite and montmorillonite. They characterise a period of fresh-water lake sedimentation during declining volcanic and tectonic activity. The Cypris' lacustrine clays gradually change from illite-kaolinite claystones at the bottom into kaolinite-montmorillonite-illite claystones at the top reflected by a change in color from blue-brown to yellowish to brown-yellow. Krábek *et al.* (1998) confirm a variable carbonate content of up to 16 wt-% primarily made up of siderite.

### 2.1.2. Coal Mining and Environmental Problems

Coal mining in the Sokolov Basin is documented since 1642 starting with bituminous slates rich in pyrite and marcasite for production of alum, vitriol and sulphuric acid. In 1793 lignite mining for combustion started. From 1870 this expanded to large-scale underground and open cast operations which continue until today (Murad & Rojík, 2005). Three major coal seams exist. The Antonín seam comprises of high quality lignite of typically less than 1% sulphur. The underlying Anežka and Josef seams contain with 1.6% and 4.6% significantly higher sulphur contents (Bouška & Pešek, 1999). Elevated sulphur contents are additionally found at the basin's rims within mineralised hydrothermal veins (Murad & Rojík, 2005).

Extensive mining led to the exposure of sulphide rich material within numerous waste dumps, spoil heaps and abandoned mines leading to a large potential for acid mine drainage (AMD). Appendix A - *Acid Mine Drainage Supplements* gives a detailed summary on the AMD processes that lead to generation and seepage of acidic water enriched with metals due to the exposure of sulphur bearing minerals to oxygen and water, the precipitation of secondary minerals and the neutralisation reactions leading to an acidity buffering sequence. In Sokolov, the iron sulphides enriched in the coal seams, coal/clay boundaries, mineralised veins and faults generate acid waters which are largely neutralised by the carbonate-rich claystones. Nevertheless, field work undertaken by Šebestová *et al.* (1996), Murad & Rojík (2005), Kopačková *et al.* (2009), Kopačková *et al.* (2011) and during this study highlight the scale of the problem by documenting numerous acidic lakes, streams and ditches inside open pits and dumps displaying the characteristic ochre AMD precipitates. Named studies link AMD effluents to extreme acidity in lakes and streams (pH 2.3), enrichment of toxic elements in soil and water (arsenic, beryllium, copper) and remediation problems due to water and soil acidity coupled with a limited nutrient availability.



## 2.2. Application Objective & Mineral Mapping Candidates

The goal of the Sokolov case study performed in Chapter 6 - *Application Scenario Mining Environment* is two fold.

One is the practical application of the methodological developments corresponding to the three specific study objectives that are covered by Chapters 3, 4 and 5. Respectively to show i) the usability of the wavelet transform based classification and quantification methods to hyperspectral data in order to demonstrate their benefits and limitations, ii) the pre-processing of the thermal data such that it can be used for a supportive quantitative study revealing information not accessible to the reflective domain, and iii) the validation of the individual processing steps of this multi-sensor study to evaluate the product quality and sources of uncertainties.

The other aim is to derive qualitative and quantitative maps of minerals relevant for the assessment of the AMD problem in Sokolov. Knowledge of the distribution and abundance of the various key minerals may be used to i) relate back to the dispersion of AMD in the area, ii) to identify areas at risk of being affected by AMD, or iii) to assist in guiding remediation efforts.

The key minerals governing geochemical processes related to AMD are discussed in Appendix A - *Acid Mine Drainage Supplements*. They can be categorised into three classes. The first class encompasses minerals that indicate localities of acid generation potential or areas affected by AMD. The second class comprises of minerals that form potentially acidity neutralisation zones or areas of high acidity buffering capacity. The third class contains minerals with no neutralisation capacity which mark areas possibly prone to rapid acidification when affected by AMD.

The primary AMD source are sulphide minerals. During field work in Sokolov it was verified that the pyrite enriched in the coal is the relevant source material. Pyrite concentrations however are low (typically < 1 wt-%) and exposed pyrite is readily weathered. Its opaque and featureless character makes pyrite practically impossible to be determined from VNIR-SWIR airborne imagery (see Figure 2.3). In the TIR, spectral features are visible which however are highly variable depending on the oxidation state (pers. communication Dr. Martin Schodlok, CSIRO). Mapping pyrite from airborne imagery was found not feasible for the Sokolov site.

The typical secondary AMD precipitate jarosite was not found at the surface during field visits in Sokolov. It occurred either in ditches/streams covered by water or in deeper soil layers. The more weathering resistant iron oxyhydroxide goethite by contrast was found in larger abundances and spatial extent at AMD affected sites. Goethite is a mapping candidate given its characteristic VNIR-SWIR spectral signature dominated by the ferric ion transitions around 450, 670 and 930 nm (Richter, 2010a). Despite suggested by the plot in Figure 2.3 no consistent feature was found for goethite rich samples in the TIR.

Krřbek *et al.* (1998) found up to 16 wt-% siderite in the Cypris Formation claystones. This AMD buffering iron carbonate was present in almost 50% of the XRD analysed samples. Siderite shows a wide ferrous iron absorption at 1,000-1,300 nm and the characteristic narrower carbonate feature at 2,300-2,350 nm (see Figure 2.4). Gaffey (1987) and Morris *et al.* (2011) confirm the presence of a strong Fe<sup>2+</sup> doublet near 1,100 nm shaped by two individual absorptions at 1,060

and 1,300 nm. The carbonate vibration mode at 11,000 nm dominates the TIR. Being the only carbonate mineral of relevance in Sokolov siderite is a mapping candidate.

The slower neutralising micas, feldspars and clays are present at Sokolov primarily in the variants muscovite, k-feldspar and kaolinite. XRD analysis of the samples assigned kaolinite by far the highest fraction, hence it is expected to dominate. In the reflective domain no diagnostic feature exists for feldspars. Kaolinite exhibits the diagnostic Al–OH doublet at 2,165-2,200 nm while muscovite has the 2,200 nm Al–OH plus a 2,300 nm Mg–OH absorption (Clark, 1999). Characteristic features exist in the TIR for all three which however overlay in wavelength. Given the 10 band multispectral TIR imagery, most discriminative features are lost. Found in abundance in Sokolov, kaolinite with its 2,200 nm reflectance doublet and broad 9,500/11,000 nm emissivity minima is another mapping candidate (see Figure 2.6).

Based on the samples XRD analysis, quartz is the most abundant mineral in Sokolov. It is featureless in the reflective domain but exhibits the diagnostic broad 8,000-9,500 nm emissivity minimum in the TIR (see Figure 2.7). This makes quartz a candidate to be mapped exclusively from the TIR data.

Justified by the demonstration site's geologic and climatic situation as well as the available data, this study will derive quantitative abundance maps for goethite as potential AMD indicator, kaolinite as slowly reacting acidity buffering mineral and quartz as inert mineral. All three were found to occur in the field in abundances and with sufficient spatial coverage for an airborne mapping. Covered by only a limited number of field samples, too few for development of a reasonable regression model, the fast buffering carbonate mineral siderite and the differentiation of clay minerals are investigated by qualitative distribution maps.

The following section covers the principles of remote sensing, atmospheric radiative transfer and spectral features found in the minerals & mineral groups typically encountered in Sokolov. Later chapters will frequently refer back to this introduction.

## 2.3. Passive Optical Remote Sensing for Mineral Quantification

Remote sensing in its conventional sense describes the acquisition of information about the Earth’s surface by airborne and spaceborne platforms using electromagnetic radiation as carrier medium (Gupta, 1991). “optical” hereby covers the visible to thermal sub-region of the electromagnetic spectrum. “Passive” relates to the radiation’s source excluding active sensor concepts such as LIDAR. Passive optical remote sensing typically bases on the intensity modulation of reflected/emitted radiation when it interacts with matter. Compositional variations and changes in physical properties lead to characteristic modulation patterns in the electromagnetic signal over the wavelength ranges. These can be analysed to differentiate surface objects and to derive their properties. Hyperspectral sensors sample the electromagnetic signal using more than 20 “narrow, continuous and contiguous spectral bands” while multispectral sensors sample at “fewer, spectrally broader bands, which may be noncontiguously spaced” (Kramer, 2002).

Comprehensive overviews of remote sensing principles are given in Schowengerdt (1997) and Richards & Jia (2006). Kramer (2002) treats sensors and technology, while Gupta (1991) and Rencz (1999) focus on remote sensing methods for Earth sciences applications. This section summarises concepts of electromagnetic radiation, its interaction with Earth’s surface and atmosphere as well as mechanisms leading to mineral spectral features. Finally a summary is given on spectral characteristics of abundant rock forming minerals, minerals relevant for the assessment of AMD affected areas and common land covers.

### 2.3.1. Electromagnetic Radiation as Signal

Every object with a kinetic temperature above absolute zero emits electromagnetic radiation. According to quantum theory only discrete energy quanta are emitted at specific wavelengths depending on the atomic configuration. Overlapping quanta in diverse and densely packed media lead to quasi-continuous spectra. For an idealised perfect emitter, a so called blackbody, the Planck’s law relates spectral radiance  $L_{bb}$  [ $\frac{\text{W}}{\text{sr m}^2 \text{m}}$ ] to the kinetic temperature  $T$  [K] and wavelength  $\lambda$  [m] with

$$L_{bb}(\lambda, T) = \frac{2hc^2}{\lambda^5} \left( \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \right) \quad (2.1)$$

where  $h$  [J s] represents the Planck’s constant,  $k$  [ $\frac{\text{J}}{\text{K}}$ ] the Boltzmann’s constant and  $c$  [ $\frac{\text{m}}{\text{s}}$ ] the speed of light in vacuum. For passive optical sensors focusing on terrestrial applications two sources of electromagnetic radiation are of relevance. The Sun illuminates Earth with an intensity spectrum approximately that of a blackbody at 6000 K temperature when neglecting effects of the Solar atmosphere. The Earth’s surface itself also emits electromagnetic radiation which intensity spectrum can be approximated by a blackbody of its respective temperature, i.e. 300 K. According to Wien’s displacement law, the surface reflected solar radiation reaches maximum intensity in the visible wavelengths (480 nm) while the typical surface emitted radiation peaks in the thermal domain (9,700 nm). Both radiation distributions have little overlap of significant intensity (see Figure 2.2). Hence reflective and emissive domain can be used to independently derive surface information that complement each other.

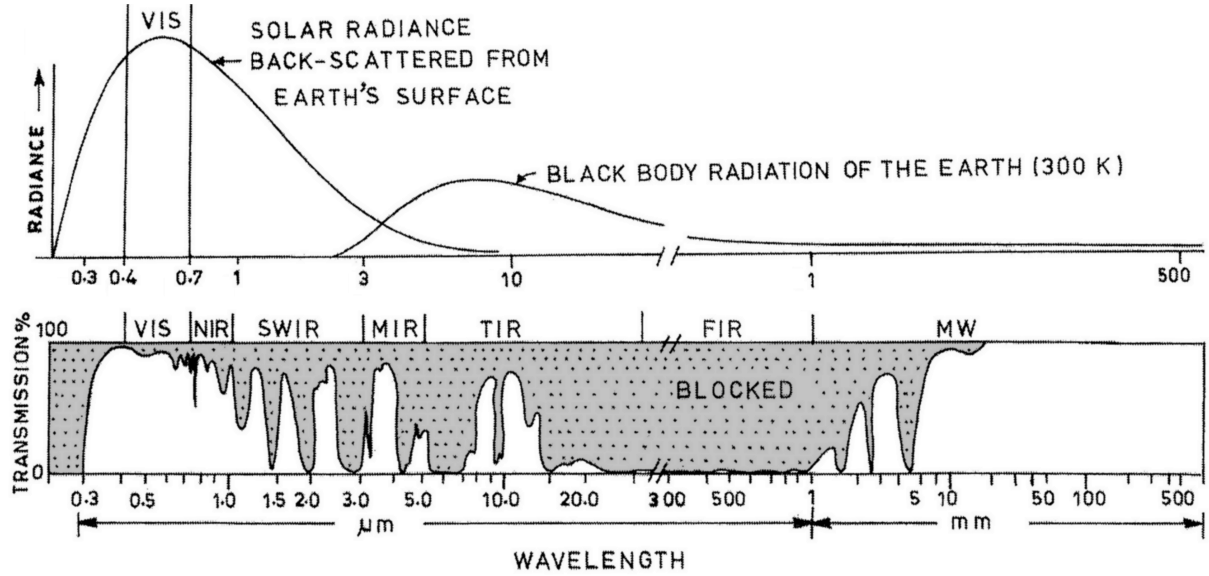


Figure 2.2.: Intensity distributions of electromagnetic radiation sources (top) and atmosphere's spectral transmission curve which determines atmospheric windows (bottom) used for remote sensing, modified after Gupta (1991).

### 2.3.2. Signal Interaction with Atmosphere

The electromagnetic signal is modified when interacting with gaseous, liquid or solid constituents of the atmosphere. Gupta (1991) summarises the main processes to be i) selective and non-selective atmospheric scattering by aerosols and dust, ii) selective absorption by water vapour, carbon dioxide, ozone and iii) atmospheric emission due to its thermal state and water vapour content. These processes lead to specific spectral regions, so called atmospheric windows, in which the atmosphere has high transmissivity and low emissivity. Clark (1999) gives an overview of the atmospheric windows used for optical remote sensing applications: visible (VIS: 400-700 nm), near infrared (NIR: 700-1,000 nm), short-wave infrared (SWIR: 1,000-2,500 nm), mid infrared (MIR: 3,500-5,000 nm) and thermal infrared (TIR: 8,000-14,000 nm) which are depicted in Figure 2.2.

Though relatively “transparent” in the atmospheric windows, remaining influences of the atmosphere need to be compensated, especially for quantitative applications (Liang, 2004). The inversion of radiative transfer models can be used to correct the recorded signal. In the reflective domain (VIS-NIR-SWIR) the basic radiative transfer equation is given in Lillesand & Kiefer (1994) and modified after Richter (1998) as

$$L_{as}^r(\lambda) = L_{path}^r(\lambda) + \tau(\lambda) \cdot L_{ag}^r(\lambda) + L_{adjacency}^r(\lambda) \quad (2.2)$$

with  $L_{as}^r(\lambda)$  the total radiance at the sensor,  $L_{path}^r(\lambda)$  the path radiance of photons scattered into the sensor,  $\tau^r(\lambda)$  the ground-to-sensor atmospheric transmittance,  $L_{ag}^r(\lambda)$  the ground reflected radiance and  $L_{adjacency}^r(\lambda)$  the reflected radiation of the neighbourhood scattered into the sensor. The main variable atmospheric parameters affecting above radiative transfer components are water vapour content, aerosol type and concentration of gases such as  $O_2$ ,  $CO_2$ ,  $O_3$  which all influence the atmosphere's optical depth.

In the thermal domain (TIR), Dash *et al.* (2002) give the basic radiative transfer equation

$$L_{as}^e(\lambda) = L_{path}^e(\lambda) + \tau^e(\lambda) \cdot L_{ag}^e(\lambda) + \tau^e(\lambda) \cdot [1 - \epsilon(\lambda)] \cdot \frac{F^e(\lambda)}{\pi} \quad (2.3)$$

with  $L_{as}^e(\lambda)$  the total thermal radiance at the sensor,  $L_{path}^e(\lambda)$  the thermal path radiance emitted by the atmospheric layer between ground and sensor,  $\tau^e(\lambda)$  the ground-to-sensor atmospheric thermal transmittance,  $L_{ag}^e(\lambda)$  the ground emitted radiance,  $\epsilon(\lambda)$  the ground surface emissivity and  $F^e(\lambda)$  the downwelling thermal sky flux at the ground (Richter & Coll, 2002). It assumes the local atmosphere in thermodynamic equilibrium with no scattering occurring and the Earth's surface a Lambertian reflector. The first term strongly depends on the atmosphere's vertical temperature and moisture profile. The second term is least affected inside an atmospheric window as long as the ground surface is warmer than the atmosphere. The third term, reflected downwelling radiation, is typically small given the emissivity [0.95...0.99] of natural materials.

### 2.3.3. Signal Interaction with Earth's Surface

The interaction of incident electromagnetic radiation with the ground surface is governed by reflection, transmission and absorption processes. Surface scattering reflects a part of the incident radiation. Specular reflection on plane surfaces and perfectly diffuse reflection on rough Lambertian surfaces exist as extreme ideal cases. For natural surfaces an intermediate semi-diffuse reflection occurs. Its dependence on the radiation's incident angle is described by the bidirectional reflectance distribution function (BRDF). Another part of the incident radiation is transmitted into the ground where inhomogeneities lead to volume scattering. A part of the volume scattered radiation is absorbed while the other part may exit the ground. This results in a mix of surface and volume scattering that sum up to a total back-scattered signal which can be received by a sensor operating in the reflective domain (VIS-NIR-SWIR, reflected solar radiation). The absorbed radiation leads to an increase of the kinetic temperature of the ground, which promotes the emission of electromagnetic radiation in the thermal domain (TIR) which a sensor operating in the proper spectral region can detect.

Modulated on the reflected and emitted signals are selective absorptions due to molecular and atomic rotational, vibrational and electronic excitation modes (Gupta, 1991) which are detailed in Section 2.3.5 - *Spectral Characteristics of relevant Minerals*. They characterise the material the signal interacted with. This leads to the material specific spectral properties transmittance  $\tau(\lambda)$ , absorptance  $\alpha(\lambda)$  and reflectance  $\rho(\lambda)$  which are the ratio of incoming to outgoing energy for each process. According to the energy conservation law they sum to unity.

$$\tau(\lambda) + \alpha(\lambda) + \rho(\lambda) = 1 \quad (2.4)$$

The macro-scale transmittance  $\tau(\lambda)$  is assumed to vanish for sufficiently thick and dense materials. Kirchhoff's law states absorptance equals emissivity  $\alpha(\lambda) = \epsilon(\lambda)$ , i.e. for a material in thermal equilibrium all absorbed energy is emitted again. This allows to convert between reflectance and emissivity.

$$\epsilon(\lambda) = 1 - \rho(\lambda) \quad (2.5)$$

Several idealised realisations exist. A blackbody as perfect emitter absorbs all incident radiation regardless of its wavelength, hence  $\epsilon = \alpha = 1, \rho = 0$ . A so called graybody emits a constant fraction of the blackbody radiation  $\epsilon = \alpha < 1, \rho > 0$  while a whitebody reflects all incident radiation  $\rho = 1, \epsilon = \alpha = 0$ . In reality only approximations to aforementioned perfect bodies exist which are typically used for calibration. The behaviour of natural materials however is a function of wavelength.

The Beer-Lambert law links transmittance  $\tau(\lambda)$  with the molar concentration  $c$  [mol/l] of the absorber and the path length  $l$  [cm] the signal travels through the medium

$$\tau(\lambda) = \frac{I}{I_0} = e^{-\epsilon(\lambda) \cdot c \cdot l} \quad (2.6)$$

where  $\epsilon(\lambda)$  [l/mol/cm] denotes the molar absorptivity of the specific absorber. Strictly, the Beer-Lambert law holds only for lower concentrations and in transparent, non-scattering media (CHP-Homepage, 2013). Ben-Dor *et al.* (1999) show it also approximates signal interactions with opaque soil volumes.

Above processes superimpose ground surface compositional information onto the sensor recorded signal in form of absorption/emission features leading to a so called spectral signature. Its characteristic troughs and peaks are the basis for material identification, compositional analysis, abundance estimation and physical parameter retrieval.

#### 2.3.4. Spectral Signature Retrieval

Retrieval of the spectral signature from the recorded signal is a fundamental task in remote sensing. In case of reflective domain data the Sun is the radiation source. The well known top of atmosphere signal travels the full path to the ground surface from where a fraction is scattered back and travels again through the atmosphere to the sensor. To retrieve the pristine spectral signature of the ground material, the influence of the atmospheric double pass must be removed. Using the radiative transfer model given in Equation 2.2 the ground reflected radiance  $L_{ag}^r(\lambda)$  can be computed. The desired ground surface reflectance spectrum  $\rho(\lambda)$  can then be isolated using the modelled global solar flux on the ground  $F^r(\lambda)$ .

$$L_{ag}^r(\lambda) = \rho(\lambda) \cdot \frac{F^r(\lambda)}{\pi} \quad (2.7)$$

For thermal domain data the Earth surface itself is the radiation source. The signal passes once through the atmosphere, which influence can be modelled and removed by the radiative transfer model given in Equation 2.3. This yields the ground emitted radiance  $L_{ag}^e(\lambda)$ . Isolation of the desired surface emissivity spectrum  $\epsilon(\lambda)$  however is difficult as the source  $L_{bb}^e(\lambda, T)$  is spatially and temporally highly variable.

$$L_{ag}^e(\lambda) = \epsilon(\lambda) \cdot L_{bb}^e(\lambda, T) \quad (2.8)$$

The surface's actual energy budget, described by the kinetic temperature  $T$ , is dependent on a variety of material specific and environmental parameters such as heat capacity, heat conduc-

tion, evaporation, long- and shortwave irradiation, etc. This imposes a major problem as for  $n$  spectral measurements (sensor bands) there are  $n+1$  unknowns (spectral emissivities plus temperature). This underdetermination has led to a variety of temperature-emissivity-separation (TES) algorithms. Quattrocchi & Luvall (2004) and Dash *et al.* (2002) give overviews.

### 2.3.5. Spectral Characteristics of relevant Minerals

Spectral absorption features are shaped due to energy transitions at the atomic and molecular level of the respective materials. The reviews of Gupta (1991), Clark (1999) and (Salisbury & Daria, 1992) give a summary highlighting electronic and vibrational processes pertinent to the VIS-TIR wavelength range. Electronic processes dominate the VIS and NIR, partially extending also into the SWIR. Three effects are of relevance for mineral spectra: i) charge transfer between ions or ion and ligands, ii) electronic transitions in transition metals and iii) crystal field effects, i.e. the modification of ion energy levels. Charge transfers and electronic transitions lead to absorption bands of low reflectance. Crystal field effects cause shifts of the absorption band's wavelength positions. The TIR is dominated by several strong fundamental molecular vibration bands (reststrahlen bands) where reflection is maximal and emission minimal. The so called Christiansen features describe wavelengths where a maximum emissivity is reached due to minimal scattering at the air-mineral surface. Such Christiansen features are typically located at a wavelength just prior to a vibration band. In the SWIR the combinations and overtones of the fundamental vibrations dominate the spectral signatures.

Spectral features are modified by a number of factors. In the reflective domain increasing grain size reduces spectral contrast (Clark, 1999). In the thermal domain this reverses as spectral contrast is reduced by less ordering of crystal surfaces, decrease in grain size and increase in pore space which eventually can result in a volume scattering effect manifesting as a broad emissivity drop (Salisbury & Daria, 1992). Next to this grain size distribution and grain packing effects, also mixing and viewing geometry have an influence on the apparent spectral characteristics. Clark (1983) and Johnson *et al.* (1992) demonstrate that intimate mixtures of materials exhibit non-linear spectral behaviour in the reflective domain, with the darker component dominating and finer grains partially covering coarser grains. Ramsey & Christensen (1998) show that in the thermal domain the validity of linear mixing holds also for the intimate mixing case and over a wide grain size range. The viewing geometry, i.e. the angles between incident radiation, material surface and the sensor, modifies the material's spectral response. In the reflective domain, the spectral intensity is modified while feature depth, shape and position are basically constant as long as no significant transition to specular reflection occurs (Clark & Roush, 1984). In the thermal domain Moroz & Schade (2011) show "significant changes in the shapes and positions of Christiansen features and the Reststrahlen bands in their vicinities" due to viewing geometry.

Overviews of spectral features for prevailing mineral groups can be found in Gupta (1991) and Clark (1999). Hunt & Salisbury (1971), Hunt *et al.* (1971a), Hunt *et al.* (1971b), Hunt (1977) and Adams & Goullaud (1978) investigate features in the VIS-NIR-SWIR. Salisbury & Daria (1992), Hook & Kahle (1996), Schodlok (2004) and Yitagesu *et al.* (2011) focus on TIR. Christensen *et al.* (2000); Clark *et al.* (1993); Salisbury *et al.* (1991) published mineral spectral libraries.

From the AMD processes scheme (Figure A.2) and the buffering sequence (Tables A.2 and A.1) in Appendix A - *Acid Mine Drainage Supplements* a number of key minerals and mineral groups were identified that govern the AMD phenomenon. In the following paragraphs the spectral characteristics of minerals relevant for acidity production, neutralisation and buffering at the Sokolov demonstration site are introduced. Their reflectance and emissivity spectra are displayed side by side in Figures 2.3 to 2.8. Spectra are given also for the abundant land covers vegetation and water which are of relevance for the vicarious calibration of the thermal image data.

### Iron oxides, sulphides, sulphates and hydroxides — Figures 2.3 and 2.8

In the reflective region the diagnostic iron electronic transition and charge transfer features can be seen. Absorptions between 1,000-1,300 nm for ferrous ion ( $\text{Fe}^{2+}$ ) are visible in siderite and, very weak, in pyrite. Jarosite, goethite and haematite exhibit the ferric ion ( $\text{Fe}^{3+}$ ) transition around 860 nm plus an absorption edge at 600 nm. Metal hydroxides or hydrated minerals (e.g. goethite) show vibrational absorptions in the SWIR at 1,400 nm and 1,900 nm (Gupta, 1991; Clark, 1999). In the TIR, sulphates such as jarosite and gypsum have an absorption around 9,000 nm. Clark (1999) reports a hydroxide fundamental for the metal–OH bending vibration at 10,000 nm. This seems to shift depending on the metal ion as Gupta (1991) reports the respective feature at 11,000 nm for Al–OH in kaolinite and Wang & Andrews (2006) at 22,200 nm for Fe–OH. The characteristic 11,200/12,400 nm doublet in goethite may be caused due to Fe–O stretching vibrations dependent on the lattice structure (Wang & Andrews, 2006). Iron oxides and hydroxides are generally fine grained which, together with a high absorption in the TIR, leads to a high emissivity and little spectral contrast. Iron oxides exhibit no distinct features as the reststrahlen bands occur beyond the TIR (Clark, 1999).

### Carbonates — Figure 2.4

The  $\text{CO}_3^{2-}$  bending vibration modes cause diagnostic absorption features in the spectra of carbonates. The most prominent are located at 2,300-2,350 nm (SWIR), 2,500-2,550 nm (SWIR) and 11,100-11,300 nm (TIR). Weaker bands at 1,860 nm, 1,985 nm and 2,140 nm exist in the SWIR. Replacement of Ca by Mg from calcite over dolomite to magnesite shifts the minima to slightly shorter wavelengths (Gupta, 1991; Clark, 1999). The replacement by Fe in siderite shows no apparent effect on the 11,100-11,300 nm absorption in the thermal reference spectra. The SWIR features at 1,860 nm and 2,300 nm interfere with a water and a clay feature respectively (Gupta, 1991), which may cause some indeterminacy.

### Silicates — Figures 2.5, 2.6 and 2.7

TIR spectra of silicates exhibit the strong asymmetric Si–O and Al–O stretching vibration fundamentals of the  $\text{SiO}_4$  and  $\text{AlO}_4$  tetrahedrons causing broad emissivity minima. The minima are located around 8,500 nm for tectosilicates (quartz, feldspar) and shift towards longer wavelengths for the increasing mafic phyllosilicates (mica, clay), inosilicates (pyroxene, amphiboles) to 9,400 nm for nesosilicates (garnet, olivine). Weaker symmetric Si–Si and Si–Al stretching fundamentals exist at longer wavelengths (12,200 nm for quartz) (Salisbury & Daria, 1992; Clark, 1999). In the VIS-NIR-SWIR, quartz and feldspars do not exhibit diagnostic features but lead to a general increase in overall reflectance (Gupta, 1991). Trace amounts of iron in plagioclase



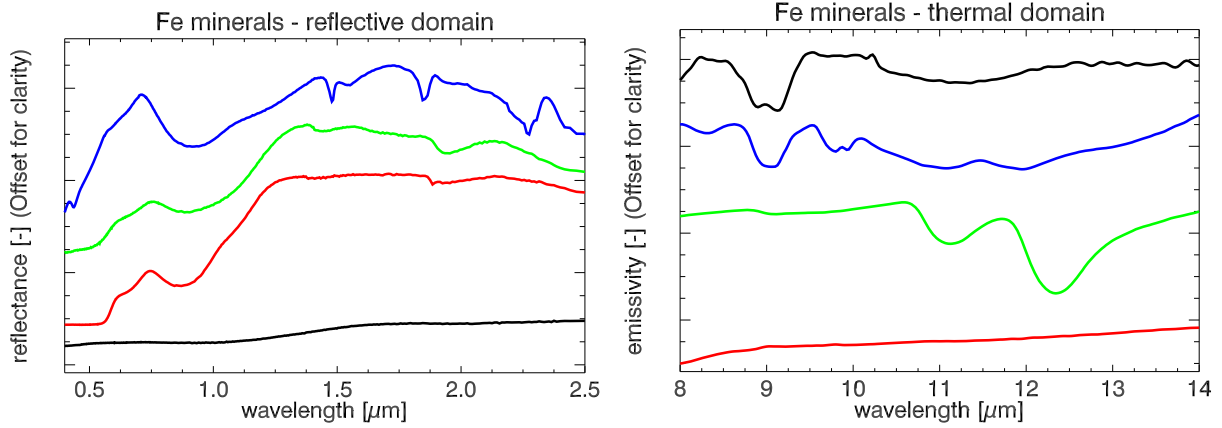


Figure 2.3.: Reflectance & emissivity spectra of selected iron bearing minerals: haematite (red), goethite (green), jarosite (blue) and pyrite (black). Plots from the USGS spectral library (Clark et al., 1993) & courtesy CSIRO Division of Earth Science & Resource Engineering.

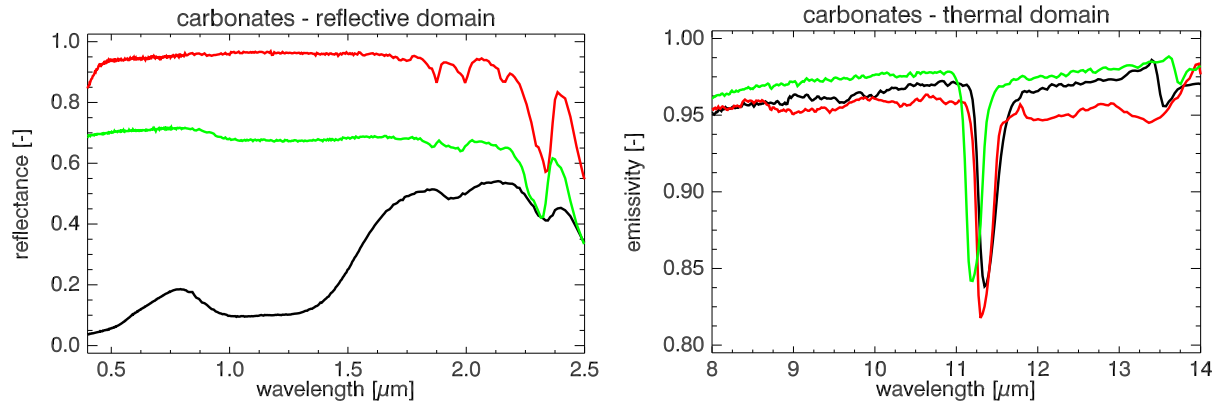


Figure 2.4.: Reflectance & emissivity spectra of common carbonates: calcite (red), dolomite (green) and siderite (black). Plots from the USGS (Clark et al., 1993) & ASU (Christensen et al., 2000) spectral libraries.

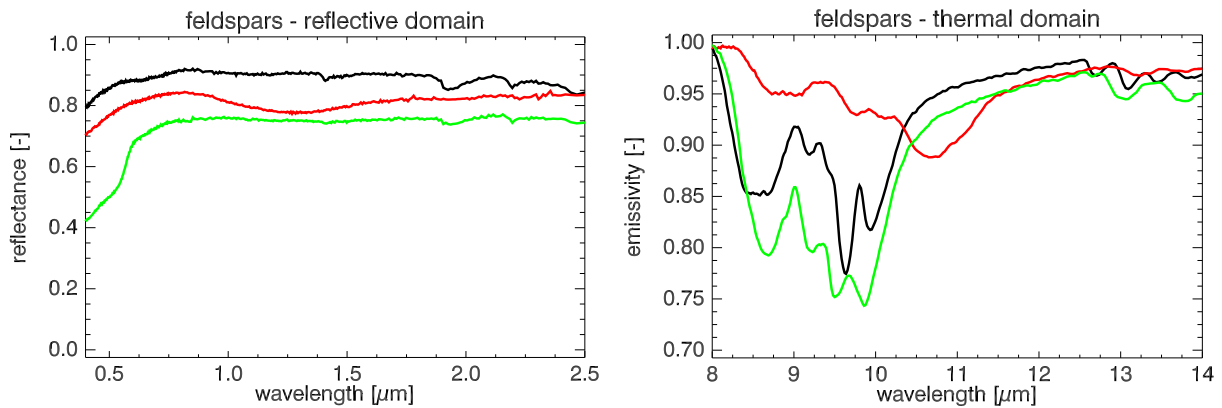


Figure 2.5.: Reflectance & emissivity spectra of feldspar endmembers: anorthite (red), microcline (green) and albite (black). Plots from the USGS (Clark et al., 1993) & ASU (Christensen et al., 2000) spectral libraries.

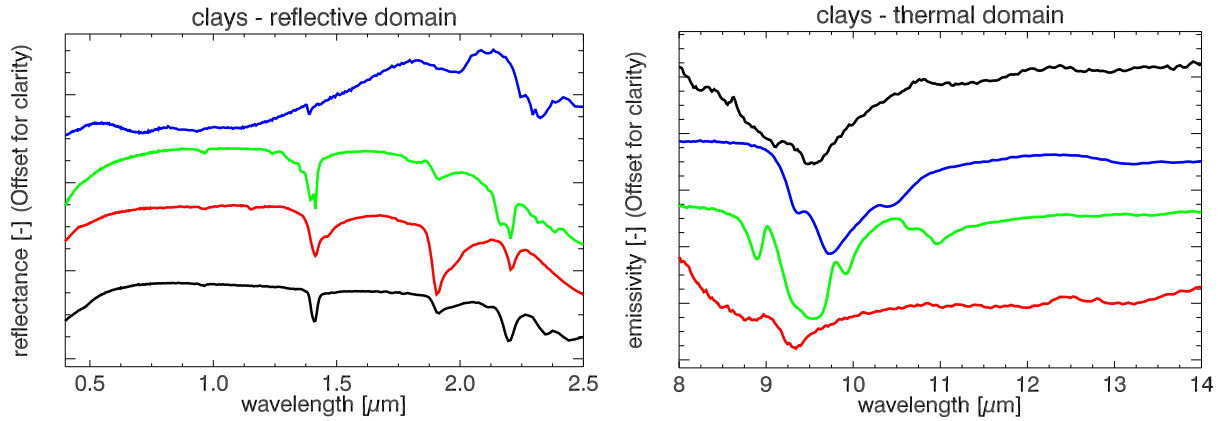


Figure 2.6.: Reflectance & emissivity spectra of clays: montmorillonite (red), kaolinite (green), chlorite (blue) and illite (black). Plots from the USGS (Clark et al., 1993) & ASU (Christensen et al., 2000) spectral libraries.

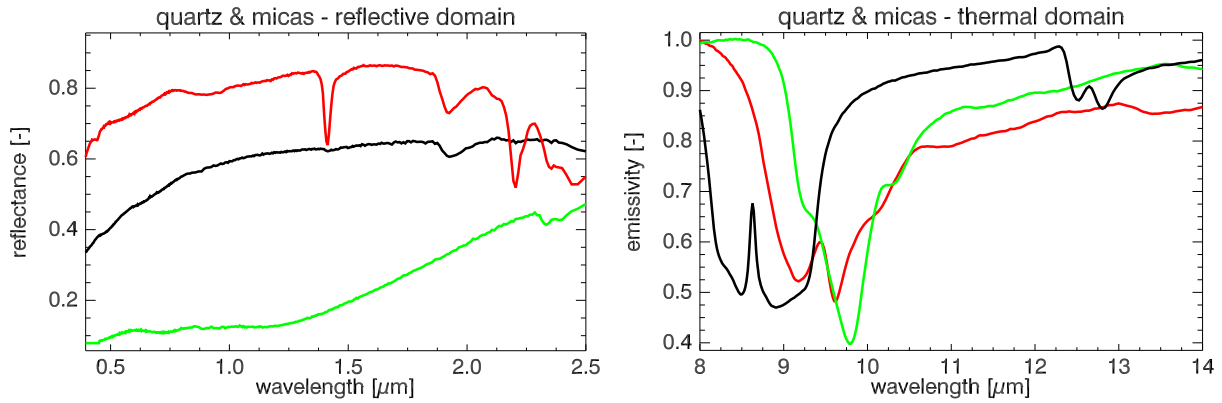


Figure 2.7.: Reflectance & emissivity spectra of quartz & micas: quartz (black), muscovite (red) and biotite (green). Plots from the USGS (Clark et al., 1993) & ASU (Christensen et al., 2000) spectral libraries.

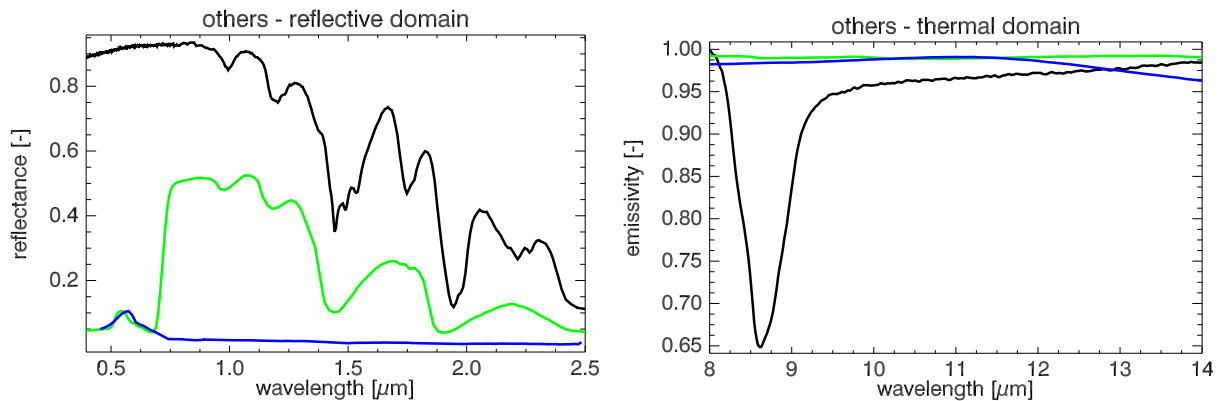


Figure 2.8.: Reflectance & emissivity spectra of other materials: gypsum (black), conifer (green) and sea water (blue). Plots from the USGS (Clark et al., 1993), ASU (Christensen et al., 2000) & JHU (Salisbury et al., 1991) spectral libraries.

feldspars may cause a  $\text{Fe}^{2+}$  electronic transition band between 1,100-1,300 nm. K-feldspars show the typical  $\text{Fe}^{3+}$  absorption band at 860 nm and absorption edge at 600 nm. However, both features are very weak and readily masked by other iron bearing minerals (Adams & Goullaud, 1978). Clays and micas show characteristic sharp absorptions at 2,200 nm due to Al–OH and at 2,300 nm due to Mg–OH vibrational modes. When both Al–OH and Mg–OH occur, an absorption doublet forms as can be seen for kaolinite. In case Fe substitutes the Al or Mg, the intensity of the respective absorption is reduced. For clays and micas also a sharp hydroxide vibrational overtone exists at 1,440 nm (Gupta, 1991; Clark, 1999).

### Water — Figure 2.8

Water is an effective absorber. In the reflective and thermal domains, suspended matter has little effect due to water's high absorption coefficient (except VIS region). Hence almost no reflection occurs in the NIR and SWIR for deep clear water bodies. Shallow or turbid water can reach considerably reflectance in the NIR and SWIR due to suspended matter and bottom reflection (Gupta, 1991). In the TIR, water bodies approximately behave like a graybody of  $\epsilon = 0.98$  but deviate considerably from this at wavelengths beyond 11,200 nm (Salisbury & Daria, 1992). Moisture causes a uniform feature contrast reduction in the thermal spectra of minerals (Salisbury & Daria, 1992). Dominant SWIR water molecule absorptions are visible at 1,400 nm and 1,900 nm. Introduced by structural and/or atmospheric water, they mask other features in the spectral surroundings.

### Vegetation — Figure 2.8

VIS-NIR-SWIR vegetation spectra are controlled by foliage type, leaf pigments, cell structure and water content. This leads to the characteristic spectral signature of photosynthetically active vegetation of low reflectance in the VIS, red-edge marking a steep rise to a NIR plateau with gradually decreasing reflectance in the SWIR overlain by two strong water absorptions (Gupta, 1991). TIR spectra of green foliage vary between virtually constant  $\epsilon = 0.98$  for conifers to well-defined troughs down to  $\epsilon = 0.95$  for deciduous plants. Senescent foliage spectra show a distinct cellulose spectrum with considerable contrast  $\epsilon = [0.85 \dots 0.95]$ , but conifers retain their high-emissivity featureless behaviour. Multiple scattering in the complex canopy of dense vegetation reduces spectral contrast and leads to a graybody behaviour in the TIR with  $\epsilon = 0.985$ . Fractional vegetation cover causes a uniform feature contrast reduction of the underlying (Salisbury & Daria, 1992).

In summary, it is possible relate back to minerals and their mixture from the characteristic features they leave in spectral signatures. A multitude of absorption shapes, simple to complex, narrow to broad exist. Standard feature based methods for their analysis and parameterization such as continuum removal, band ratios and the fitting of geometrical primitives are limited in describing the wide range of shapes. A lossless decomposition of the signatures into a simpler, multi-scale representation on which analysis methods can operate is desirable. The next section introduces the continuous wavelet transform which features such properties.

## 2.4. Decomposition of Signals

Representing complex shaped signals as superposition of simple base functions is a fundamental approach in signal processing. A compact representation of the signal can be found by decomposing it over appropriate base functions. Conclusions on the properties of the original signal can be drawn from studying the simpler base functions and decomposition coefficients. A transition is made from spectral signatures, a function of wavelength like  $\rho(\lambda)$ , to a more generic idea of a function of time  $f(t)$ . In Section 3.3 - *Application to Hyperspectral Signatures* the peculiarities in reverting back to a function of wavelength are described.

Mentioned base functions need to satisfy several requirements to be practically useful. A universal choice does not exist but rather specific functions tailored to the application are chosen. Following Băni (2002), the requirements can be condensed to three properties of the base functions desirable for most applications:

1. Superposition of the base functions should allow to represent a sufficiently large number of different signals. Calculation of decomposition coefficients and the signal reconstruction from the coefficients should be numerically fast and stable.
2. The time support of the base functions should be limited. That is the functions are (nearly) zero except for a finite time window. The base functions are then said to be well localised in time. Also their frequency support should be limited, implying their Fourier transform is well localised. Both properties allow to capture localised signal features in few and also localised decomposition coefficients.
3. The base functions should form an orthonormal basis. As such the decomposition coefficients are uniquely defined and free of redundancy.

In 1807, Joseph Fourier published his findings in representing periodic real functions as series of trigonometric functions. The Fourier transform is a well known example of a decomposition into an orthonormal basis. It breaks down a signal  $f(t)$  into a sum of sinusoidal functions  $e^{-2\pi it\omega}$ .

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) \cdot e^{-2\pi it\omega} dt \quad (2.9)$$

For discrete signals a fast algorithm for computing the Fourier transform plus its inverse exists (Fast Fourier Transform), making the Fourier transform an attractive tool widely applied in mathematics, physics and engineering. It is well suited for time-invariant, periodic and uniformly regular signals but performs poorly in representing time localised events such as sharp signal transitions or discontinuities. This characteristic attributes to the Fourier base functions. They are perfectly localised in frequency but exhibit no localisation in time, i.e. live over the whole infinite time domain. A signal feature localised in time therefore can not be represented by a local transform coefficient but is smeared into all coefficients making the interpretation difficult.

In 1910, Alfréd Haar introduced a piecewise constant function which values are different from zero only for a limited time interval, the atom. An orthonormal basis can be constructed by dilation and translation of the atom along the time axis (see Figure 2.9). This was the first wavelet base. Unlike Fourier transform, wavelets are well localised in time allowing to decom-

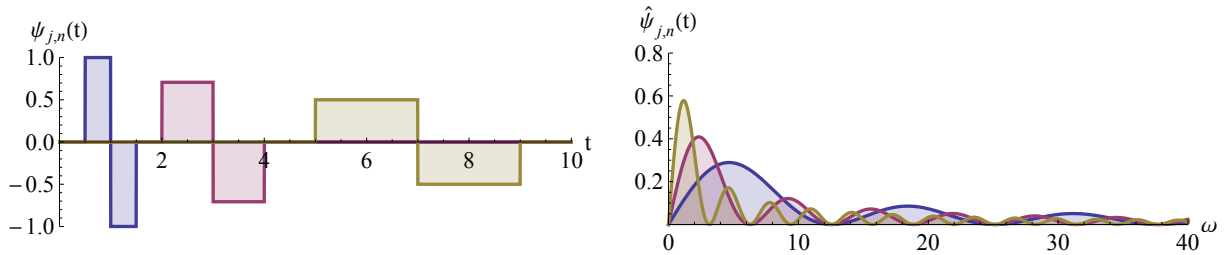


Figure 2.9.: Haar wavelet for different translation and dilation (left) together with respective Fourier transforms (right) limited to positive frequencies.

pose piecewise regular signals containing singularities and sharp transitions into few significant coefficients. The frequency localisation of the Haar wavelet however is inferior attributed to its jump discontinuities introducing oscillations in its Fourier transform. As such the Haar basis and the Fourier basis can be regarded contrary in respect to the localisation in time and frequency. Time and frequency localisation can not be chosen arbitrarily but follow the Heisenberg uncertainty principle. The better the localisation in time the worse in frequency, and vice versa.

In 1946, Dennis Gabor proposed the use of time-frequency atoms in signal decomposition imitating the human perception of sound and visual pattern recognition. Bounded by the Heisenberg uncertainty the atoms shall have minimal spread in time and frequency.

The Gabor transform, also known as windowed or short time Fourier transform, applies a windowing function  $g(t)$  to achieve time localisation. The signal  $f(t)$  is multiplied by a translated windowing function  $g(t - u)$  and then Fourier transformed. A highly redundant continuous windowed Fourier transform results.

$$Sf(u, \xi) = \int_{-\infty}^{\infty} f(t) g(t - u) \cdot e^{-2\pi i t \xi} dt \quad (2.10)$$

A one-dimensional signal  $f(t)$  is decomposed into a two-dimensional spectrogram  $(u, \xi)$  with a fixed time and frequency resolution. Obviously the windowing function's width governs the time resolution. Signal structures requiring finer time or frequency resolution can not be resolved. Discretisation of  $u$  and  $\xi$  aiming at constructing an orthonormal basis free of redundancy results in poor localisation either in frequency or time.

Wavelet bases avoid latter issues by decomposing into base functions of changing time and frequency resolution. In 1984 Jean Morlet and Alex Grossmann paved the way to continuous wavelet transform (CWT). CWT uses a wavelet mother function  $\psi(t)$  which is dilated by scale  $s$  and translated by  $u$ . Essentially the mother wavelet can be seen as windowing function. Rather than just translating it, the wavelet transform additionally rescales it. The continuous wavelet transform represents a one-dimensional signal  $f(t)$  thus as highly redundant time-scale-frequency cube  $(u, s, \xi)$ . Choosing the mother wavelet itself localised in frequency  $\xi_0$  the rescaling by  $s$  will change the time resolution but also the frequency resolution  $\xi_0/s$ . This enables to skip the factor  $e^{-2\pi i t \xi}$  and hence the frequency parameter  $\xi$ . The time-frequency image is replaced by a time-scale image.

$$Wf(u, s) = \int_{-\infty}^{\infty} f(t) \cdot \frac{1}{\sqrt{s}} \psi^* \left( \frac{t - u}{s} \right) dt \quad (2.11)$$

The esteem of wavelet transform is the possibility to construct bases which fulfil all three desirable properties mentioned above when an orthogonal mother wavelet is used, scale is discretised in dyadic steps  $s_j = 2^j$  and localised time follows the adapted discretisation  $u_n = 2^j n$  with  $j, n \in \mathbb{Z}$ . This discrete wavelet transform (DWT) then decomposes the signal into an orthonormal base. The resulting coefficients are free of redundancy and thus yield a compact signal representation. Fast algorithms implemented as filter banks exist that exploit the proportional decrease of the number of convolutions necessary to compute the discrete transform with increasing scale. DWT is a tiling of the time-frequency space with no overlaps. The resulting time-scale image consists of blocks linked to the time-frequency tiles where the transform value is constant for a discrete time and scale range. This leads to the time-variance of DWT which is a drawback when analysing signal patterns. The time-variance manifests in strongly changing transform coefficients despite only a minor shift of the signal in time. DWT is thus mainly applied in signal compression and denoising techniques. [Holschneider \*et al.\* \(1989\)](#) introduced the stationary wavelet transform, also known as “algorithme à trous”, which overcomes the lack of translation-invariance of DWT. It uses a fast recursive implementation as filter bank and produces a continuous transform in respect to time. The algorithm however requires use of an orthogonal or biorthogonal mother wavelet and is limited to dyadic scales.

The continuous wavelet transform (CWT) computes coefficients for each time and scale. The non-orthonormal basis results in a highly redundant signal representation. Especially for high scales, the transform is strongly correlated in time. This results from the large time support of the atoms but continuous sampling. Latter oversampling results in the translation-invariance of CWT at the cost of higher computational expense. Hence CWT is useful in searching a signal for a specific localised event which is exploited in pattern-matching, time-series analysis ([Torrence & Compo, 1998](#)) or study of signal transients with a zooming procedure ([Mallat, 2008](#)). CWT can be approximated for discrete signals by computing a convolution at each sampling point for densely discretised scale steps. Use of a Fast Fourier Transform algorithm allows an efficient calculation that simultaneously evaluates all convolutions per scale ([Torrence & Compo, 1998](#)).

Since the thesis at hand aims at hyperspectral signal analysis, the full wavelength resolution in all scales is necessary which leads to the use of CWT. The remainder of this chapter discusses CWT and its properties in respect to the chosen wavelet mother function. Chapter 3 - *Mineral Classification and Quantification* develops wavelet-based methods for the evaluation of signal similarity and the parameterization of signal features against the background of hyperspectral data analysis. Mathematical notations and terms follow the conventions found in [Mallat \(2008\)](#).

### 2.4.1. Continuous Wavelet Transform

Wavelet analysis uses localised waveforms - the wavelets - as base functions to decompose the signal. Translation and dilation operations are applied to construct time-frequency atoms of changing resolution in both time and frequency. The template for such an atom is called the mother wavelet  $\psi(t)$ . Besides localisation, the mother wavelet must satisfy several criteria to qualify for wavelet analysis.

It must be of finite energy described by the integral of its squared magnitude.

$$E = \|\psi(t)\| = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (2.12)$$

The Fourier transform of the mother wavelet  $\hat{\psi}(\omega)$  must fulfil the admissibility condition.

$$\int_0^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty \quad (2.13)$$

This implies  $\hat{\psi}(\omega)$  must be continuously differentiable and must have no zero frequency component  $\hat{\psi}(0) = 0$  leading to the zero average of wavelets.

$$\hat{\psi}(0) = \int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (2.14)$$

The mother wavelet is centred around  $t = 0$  and typically normalised. Latter ensures comparability of the transforms between scales and between different signals at the same scale (Torrence & Compo, 1998). Normalisation is achieved by choosing the mother wavelet parameters such that the wavelet has unit energy. The wavelet transform is thus weighted only by the signal amplitude.

$$E = \|\psi(t)\| = \int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1 \quad (2.15)$$

The wavelet base functions are generated from the mother wavelet  $\psi(t)$  when scaling by  $s$  and translating by  $u$ . The atoms  $\psi_{u,s}(t)$  remain normalised due to the factor  $\frac{1}{\sqrt{s}}$ .

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \quad \text{with } u \in \mathbb{R}, s \in \mathbb{R}^+ \quad (2.16)$$

From the admissibility condition (Equation 2.14) a mother wavelet can be interpreted as band-pass filter for which a passband centre frequency  $\xi_0$  is defined. The filter accentuates signal components of a finite frequency range around  $\xi_0$  whereas components of other frequencies are damped. The scaling of the mother wavelet dilates the bandpass filter and leads to a characteristic frequency of the wavelet at scale  $s$  given by  $\xi = \xi_0/s$ . Since the mother wavelet is centred at  $t = 0$  translation by  $u$  will centre the wavelet at  $t = u$ .

Translating this into spectral signature analysis, the time localisation parameter  $u$  allows to centre the wavelet over individual sampling wavelengths. It thus allows to selectively interrogate one absorption feature in the whole signature because the limited time support of the wavelet excludes information from neighbouring absorptions. Furthermore, selection of the scale parameter defines the size of the time support thus allowing to change the spectral neighbourhood which is used for the interrogation. With scale governing the filtering bandpass, the spectral features of a corresponding width are emphasised during the interrogation while narrower or wider spectral features are suppressed.

Figure 2.10 displays the tiling of the time-frequency plane by the Morlet wavelet. The wavelet's effective frequency and time support are proportional to its energy densities in the frequency  $|\hat{\psi}_{u,s}(\omega)|^2$  and time  $|\psi_{u,s}(t)|^2$  domains, which are plotted for three different translations  $u$  and

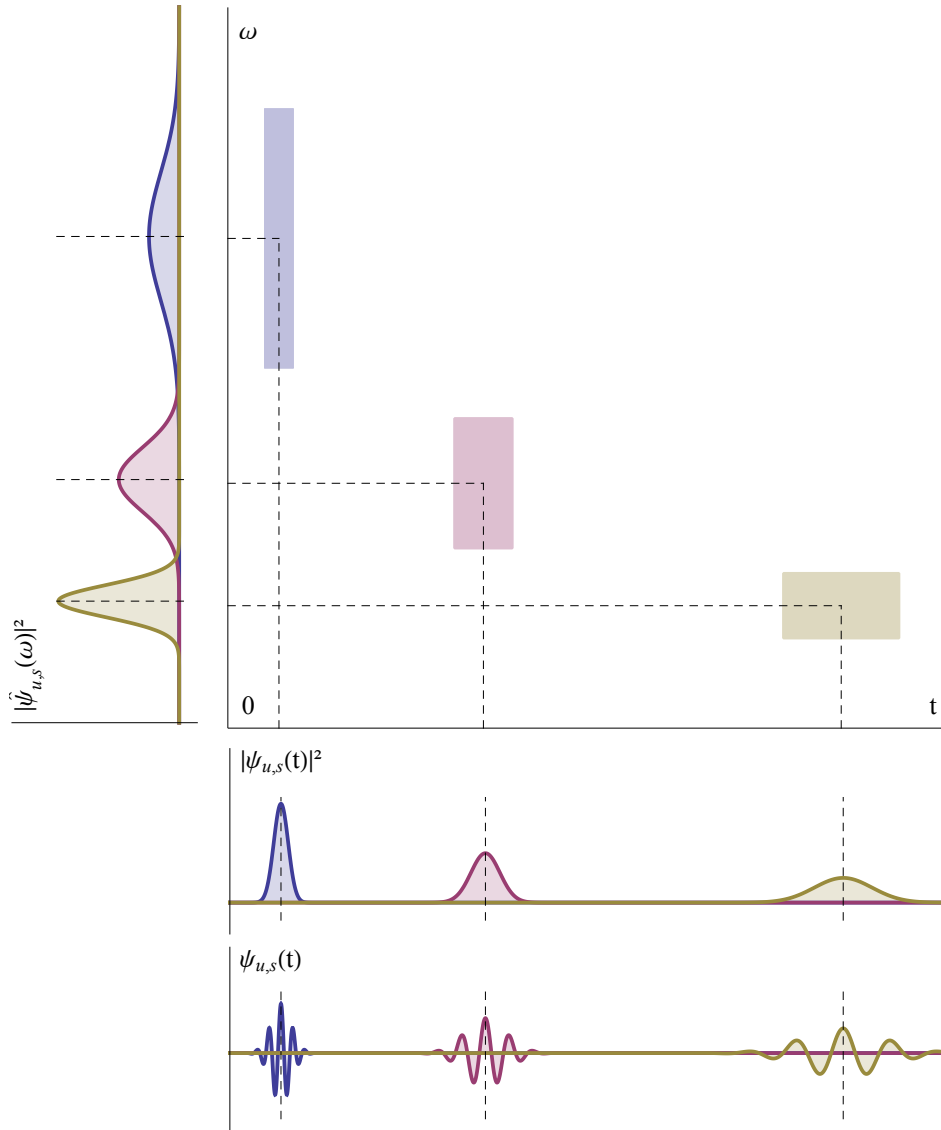


Figure 2.10.: Tiling of the time-frequency plane by the Morlet wavelet. The bottom plot shows the real part of the Morlet wavelet  $\psi_{u,s}(t)$  at three different scales and translations. Plotted above are the respective energy densities in the time domain  $|\psi_{u,s}(t)|^2$  merging the wavelet's real and imaginary parts. The left plot shows the respective energy densities in the frequency domain  $|\hat{\psi}_{u,s}(\omega)|^2$ . The resulting Heisenberg boxes are drawn in the centre part. Modified after Addison (2002).

scales  $s$  along the respective axes. In the time-frequency plane this forms so called Heisenberg boxes centred at the respective  $(u, \xi)$  coordinates. The box' sides shrink proportional to  $1/s$  in frequency and increase proportional to  $s$  in time. The area stays constant and has a lower bound according to the Heisenberg uncertainty.

This has the following implications for signal analysis. The combined frequency and time resolution cannot be chosen arbitrarily fine. It is dependent on the wavelet choice and a boundary exists. With increasing scale, the wavelet's time support expands but its frequency support contracts and the passband shifts to lower frequencies. It follows that at small scales high frequency signal components can be measured but with a low frequency resolution. Conversely, higher scales allow to interrogate low frequency signal components at a better frequency resolution but at the cost of time resolution.



Following the principles of the Fourier and Gabor transforms and considering the characteristics of the mother wavelet, Mallat (2008) gives the continuous wavelet transform of signal  $f(t)$  at localised time  $u$  and wavelet scale  $s$  by

$$Wf(u, s) = \int_{-\infty}^{\infty} f(t) \cdot \frac{1}{\sqrt{s}} \psi^* \left( \frac{t-u}{s} \right) dt \quad (2.17)$$

where  $\psi^*$  indicates the wavelet's complex conjugate. Alternatively Equation 2.17 can be rewritten as convolution product of the signal  $f(t)$  with the time-reversed wavelet  $\bar{\psi}_s(t)$ .

$$Wf(u, s) = (f * \bar{\psi}_s)(u) \quad \text{with} \quad \bar{\psi}_s(t) = \frac{1}{\sqrt{s}} \psi^* \left( \frac{-t}{s} \right) \quad (2.18)$$

From Equation 2.18 the continuous wavelet transform at scale  $s$  can be interpreted as cross correlation of the signal  $f(t)$  with the wavelet atom  $\psi_{u,s}(t)$ . The cross correlation measures the similarity of two waveforms. Hence the CWT value at time  $u$  quantifies the similarity between signal and the scaled wavelet centred over that time location  $u$ . Times where the wavelet atom and the signal are of similar shape give a positive contribution to the integral value of  $Wf(u, s)$ . Regions of opposite shape give a negative contribution.

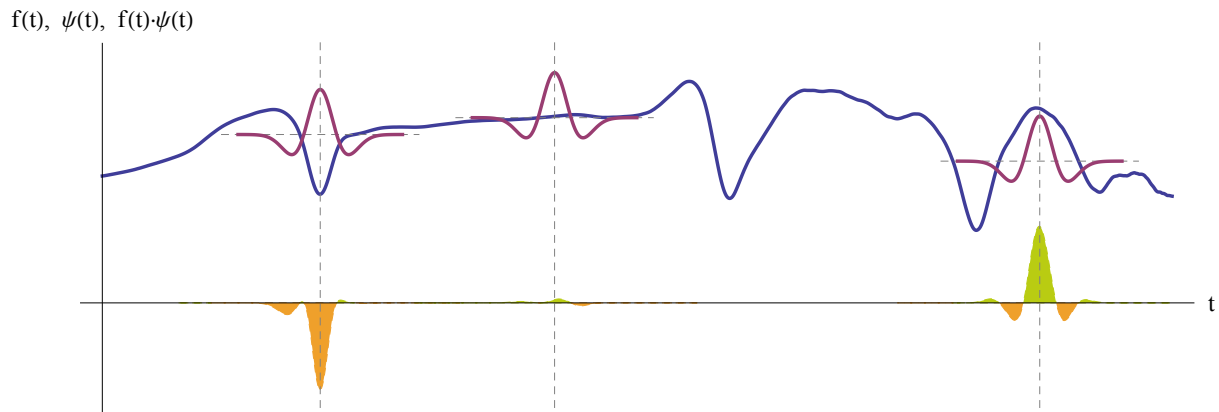


Figure 2.11.: Schematic interrogation of a signal  $f(t)$  by a wavelet  $\psi(t)$  for three selected translations (top) versus resulting product  $f(t) \cdot \psi(t)$  (bottom).

This is exemplified in Figure 2.11. It shows a signal  $f(t)$  with three Derivative Of Gaussian wavelet atoms  $\psi_{u,s}(t)$  centred at different times  $u$ . The plot below gives the product of signal and wavelet atom  $f(t) \cdot \psi_{u,s}(t)$ . Positive contributions are shaded in green while negative are shaded in orange. The CWT value at  $u$  equals the integral over the respective positive and negative contributions within the time support of the wavelet atom. At the first location, signal and wavelet atom are out of phase resulting in a large negative value of  $Wf(u, s)$ . At the second location, positive and negative contributions nearly cancel out yielding a  $Wf(u, s)$  close to zero. At the third location, signal and wavelet atom match well resulting in a large positive value for  $Wf(u, s)$ . For the analysis of spectral signatures this has the implication that a mother wavelet shaped similar to an absorption feature causes large transform values at the wavelength locations where such an absorption is present in the signature. Hence CWT can be seen as an absorption shape matching method.

Mallat (2008) demonstrates another interesting property of CWT. If the mother wavelet is chosen the  $m$ -th order derivative of a fast decaying function  $\theta(t)$  such that

$$\psi(t) = (-1)^m \frac{d^{(m)}}{dt^{(m)}} \theta(t) \quad \text{with} \quad \int_{-\infty}^{\infty} \theta(t) dt \neq 0 \quad (2.19)$$

the wavelet transform can be expressed as multi-scale differential operator.

$$Wf(u, s) = s^m \frac{d^{(m)}}{du^{(m)}} (f * \bar{\theta}_s)(u) \quad \text{with} \quad \bar{\theta}_s(t) = \frac{1}{\sqrt{s}} \theta\left(\frac{-t}{s}\right) \quad (2.20)$$

The convolution  $(f * \bar{\theta}_s)(t)$  is equivalent to a weighted average of  $f(t)$  with a kernel defined by the function  $\theta(t)$  of size proportional to  $s$ . This finding allows to construct a wavelet transform which is the  $m$ -th derivative of the smoothed spectral signature by choice of an appropriate mother wavelet. The smoothing filter shape is then governed by the wavelet, more precisely by  $\theta(t)$ , and the filter width by the scale.

### 2.4.2. Wavelet Choice

The choice of the wavelet mother function has a crucial influence on the characteristics of the wavelet transform. Depending on the signal properties and the desired type of analysis, various optimized classes of wavelets exist. Torrence & Compo (1998) and Mallat (2008) give important wavelet mother function properties which are summarised in the following.

ORTHOGONALITY & COMPLEXITY Orthogonal wavelets imply the use of DWT whereas non-orthogonal wavelets can be used in DWT and CWT. Complex valued wavelets capture the time evolution of oscillations in the signal. The transform consists of a real and an imaginary component allowing the separation of amplitude and phase. Real valued wavelets are used to detect sharp transitions in the signal such as isolated peaks and discontinuities. The regularity of the signal around the transitions can be measured. Real wavelet transforms return no imaginary part which leaves phase undefined.

WIDTH Limited by the Heisenberg uncertainty, time and frequency resolution of a wavelet can not be chosen arbitrarily high at the same time. A wavelet with broad time support will have a good frequency resolution and vice versa. Dealing with non-periodic finite length signals border effects arise when the wavelet time support overlaps with the beginning and the end of the signal beyond which an extrapolation (padding) is necessary (see Section 2.4.3 - *Cone Of Influence*).

SHAPE & REGULARITY Interpreting CWT as cross-correlation between signal and the wavelet atom, then a wavelet mother function reflecting the shape of the features expected in the signal will facilitate feature detection. Symmetric wavelets will not alter the symmetry of structures within the signal. Smooth wavelets will result in a smooth transform.

VANISHING MOMENTS A wavelet that has  $p$  vanishing moments is orthogonal to polynomials up to  $p - 1$  degree. If the signal is locally  $k$  times differentiable ( $\mathbf{C}^k$ ) it can be approximated by a Taylor polynomial of the same degree. A wavelet of  $p > k$  vanishing moments is orthogonal to that Taylor polynomial. Hence wavelet transform values close to zero will

result. The higher the number of vanishing moments the more complex shapes in the signal can be represented with only a limited number of significant wavelet transform coefficients. For orthogonal wavelets a higher number of vanishing moments implies a broader time support.

The thesis at hand aims at the analysis of spectral signatures of minerals. Their geochemical composition leaves characteristic absorption features. Also mixing, moisture, viewing geometry and illumination conditions during the data acquisition affect the apparent signal. The mother wavelet hence should be chosen such that it represents the general shape of absorption features and such that application of the transform counters disturbing influences found in the data. Absorption features should be exactly located in wavelength and their shape be preserved for later analysis. Information about frequency and phase are of less importance. Given the rather limited widths of the atmospheric windows, limitations in spectral resolution of the sensors and the occurrence of characteristic spectral features close to the margins of the atmospheric windows, the chosen wavelet should allow to interrogate the signal close to the margins of the spectral coverage.

From above requisites a real valued mother wavelet of small time support (width) is to be chosen. Its shape should be smooth, symmetric and in accordance with the expected absorption shapes.

Several authors have used Gaussian functions to model absorption features. [Clark & Roush \(1984\)](#) give an overview noting the good agreement between theoretical predictions of Kubelka-Munk theory and Gaussian models for apparent absorbance and reflectance spectra. The authors suggest to convert reflectance spectra to apparent absorbance. [Sunshine \*et al.\* \(1990\)](#) too propose conversion of reflectance into absorbance and investigate the Gaussian model performance when deconvolving pyroxene mixture spectra. They agree with the physical and practical applicability of the Gaussian model but suggest a modified version introducing a skew.

Spectral derivatives enhance subtle spectral variations and hence are frequently used to identify and locate features otherwise hardly visible. Second or higher order derivatives are regarded relatively insensitive to variations in illumination intensity ([Tsai & Philpot, 1998](#)). Despite, first and second order derivatives are mainly found in literature probably because they can be readily interpreted as local slope and local curvature of the continuous reflectance spectrum. Higher order derivatives require a more abstract interpretation of the results. Derivatives amplify high-frequency noise thus reduce the signal-to-noise ratio. This effect greatly increases the higher the order of the derivative. The noise problem is typically addressed by a low-pass filtering step prior to the calculation of the spectral derivative or a Savitzky-Golay least-square-fitting method ([Savitzky & Golay, 1964](#)). The noise characteristic however is changing with wavelength (e.g. atmospheric effects, detector transitions & sensitivity) making it difficult to select an appropriate smoothing which removes high-frequent noise but preserves subtle and narrow absorption features. Also characteristic signal features may get distorted or may occur at varying window widths ([Tsai & Philpot, 1998](#)). Determination of the optimal smoothing is often experimentally and based on the particular application.

The continuous wavelet transform in particular can provide solutions for these issues. Rather than limiting analysis to a single spectral scale, a multi-scale approach takes advantage of different information of the input spectrum being emphasised at the respective scales.

The localised wavelets “home in” on narrow signal features and accurately locate them at smallest scale. Recalling the interpretation of CWT as multi-scale differential operator (Equation 2.20), a wavelet can be chosen such that the desired derivative order is achieved and signal features are preserved by a smooth and symmetric filter. Smoothing and derivative calculation are efficiently joined (Bruce & Li, 2001) yielding transforms at multiple scales. Therefore a more general and application independent feature search and parameterization method is at hand.

The Derivative Of Gaussian (DOG) wavelet family fulfils most of the desired properties listed above. They result, as the name suggests, from applying a differentiation of arbitrary order to a Gaussian function (see Figure 2.12). The wavelets are perfectly smooth, i.e. infinitely differentiable. For odd order derivatives the wavelets are antisymmetric around zero, for even orders the wavelets are symmetric. A mathematical introduction to the Gaussian function and its derivatives is given in Appendix D - *Mathematical Supplements*. Haar & Bart (2003) point out the bandpass filter behaviour of Gaussian derivative kernels. The higher the differentiation order, the higher the number of oscillations represented by the wavelet. The centre frequency  $\omega_c$  of the bandpass filter grows exponentially with the derivative order  $m$  following  $\omega_c = \sqrt{m}$ .

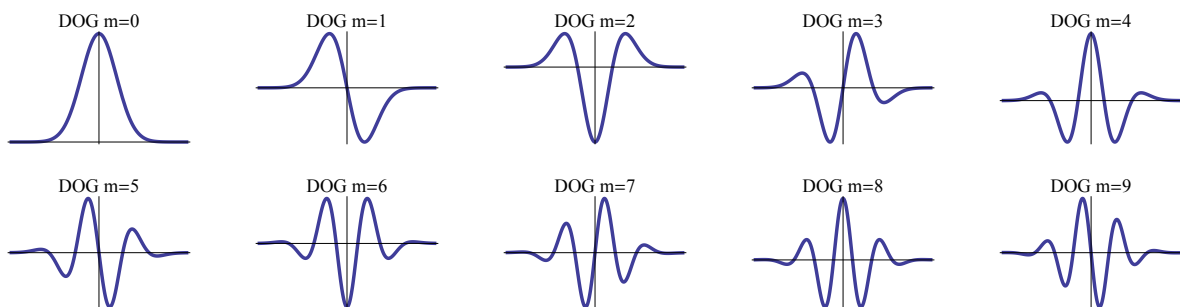


Figure 2.12.: Gaussian function ( $DOG_{m=0}$ ) together with its derivatives for numerous orders  $m$ , after Haar & Bart (2003)

Looking at Figure 2.12, the Second Order Derivative Of Gaussian ( $DOG_{m=2}$ ) is most similar in shape to a Gaussian ( $DOG_{m=0}$ ). The Gaussian function is actually a common element of all its deviates (cf. Hermite polynomials). The  $DOG_{m=2}$  wavelet is symmetric and by definition of zero average. Hence the areas under the side lobes equals the area under the main peak. Appendix D.4 - *DOG Wavelet Notations* lists notations for DOG wavelets found in the literature and derives the standard notation used here. The normalised  $DOG_{m=2}$  mother wavelet  $\psi(t)$  and its translated/dilated atoms  $\psi_{u,s}(t)$  are

$$\psi(t) = \frac{2}{\sqrt[4]{\pi}\sqrt{3}\sigma} \cdot \left( \frac{t^2}{\sigma^2} - 1 \right) \cdot e^{-\frac{t^2}{2\sigma^2}} \quad (2.21)$$

$$\psi_{u,s}(t) = \frac{2}{\sqrt[4]{\pi}\sqrt{3}\sigma s} \cdot \left( \frac{(t-u)^2}{\sigma^2 s^2} - 1 \right) \cdot e^{-\frac{(t-u)^2}{2\sigma^2 s^2}} \quad (2.22)$$

This definition is opposite in sign compared to the mathematical sense. The  $\text{DOG}_{m=2}$  wavelet used here is hence identical to the popular Mexican Hat or Sombrero wavelet. Figure 2.13 depicts three instances of a  $\text{DOG}_{m=2}$  wavelet  $\psi_{u,s}(t)$  of  $\sigma = 1$  at varying dilation  $s$  and translation  $u$  together with the respective Fourier transforms  $\hat{\psi}_{u,s}(t)$ . The wavelet atoms depicted from left to right  $\psi_{5,1}(t)$ ,  $\psi_{20,2}(t)$  and  $\psi_{45,4}(t)$  double their time support while the frequency support contracts proportionally. Compared to the Morlet wavelet (see Figure 2.10) the frequency localisation is inferior (big overlap between scales) and the resulting bandpass filter is asymmetric. The better time localisation of the  $\text{DOG}_{m=2}$  wavelet gives it advantages in highlighting small-scale signal variations and limiting border effects. The two zero crossings marking the transition from the main peak to the side lobes are at a distance  $\pm\sigma s$  from the centre of the atom.

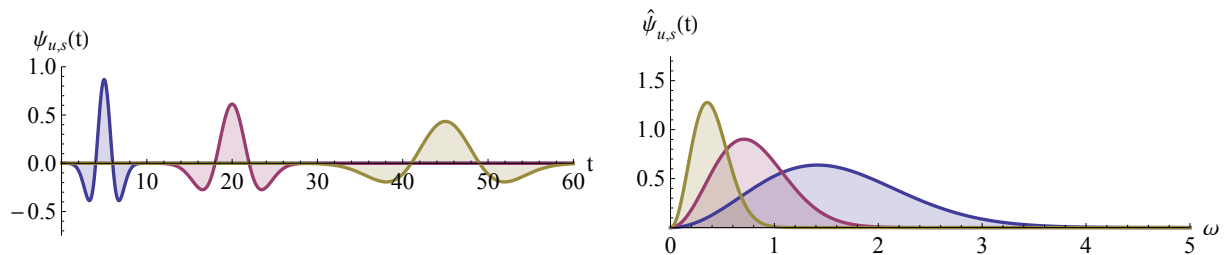


Figure 2.13.: Second Derivative of Gaussian wavelet for different translation and dilation (left) together with respective Fourier transforms (right) limited to positive frequencies (symmetry with respect to the  $y$ -axis exists, negative frequencies having no physical meaning are not shown).

### 2.4.3. Cone Of Influence

Dealing with non-periodic finite length signals border effects arise when the wavelet time support overlaps with the beginning and the end of the signal. Various signal extrapolation methods are listed in Addison (2002) and Su *et al.* (2011). The possible area of border effects grows linearly with the time support of the wavelet and hence with scale. The parts of the transform in which border effects become significant is described by the cone of influence (COI). Furthermore the COI gives the decorrelation time of a point feature in the signal (Torrence & Compo, 1998).

For a mother wavelet  $\psi$  of compact time support  $[-C, C]$  the dilated and translated time support of the atom  $\psi_{u,s}$  is given by  $[u - Cs, u + Cs]$  and thus the cone of influence centred on  $t = v$  encompasses all  $t = u$  that fulfil (Mallat, 2008)

$$|u - v| \leq Cs \quad (2.23)$$

For the DOG family no compact time support is defined, strictly it is infinite. Torrence & Compo (1998) propose an empirical COI definition using the  $e$ -folding time. The  $e$ -folding time marks the point at which the wavelet power of the edge discontinuity drops by a factor of  $e^{-2}$ . The cone of influence centred at  $t = v$  hence encompasses all  $t = u$  fulfilling

$$\frac{|Wf(u, s)|^2}{|Wf(v, s)|^2} \geq e^{-2} \quad (2.24)$$

An analytic solution for the wavelet transform of a Gaussian pulse of width  $\sigma_f$  at  $t = \mu_f$  is given in Equation 3.13. Realising the three leading terms are constant for a fixed scale and thus

cancel from the ratio plus the COI is centred on the pulse  $v = \mu_f$  rendering the denominator unity then above COI expression can be written as

$$\left| e^{-\frac{(u-v)^2}{2(\sigma_f^2 + s^2\sigma^2)}} \right|^2 \geq e^{-2} \quad (2.25)$$

$$\frac{(u-v)^2}{(\sigma_f^2 + s^2\sigma^2)} \leq 2 \quad (2.26)$$

$$\frac{(u-v)^2}{s^2\sigma^2} \cdot \frac{1}{1 + \frac{\sigma_f^2}{s^2\sigma^2}} \leq 2 \quad (2.27)$$

The second fraction on the left hand side can be expanded to a geometric power series

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots \quad \text{with } x = \frac{\sigma_f^2}{s^2\sigma^2} \quad (2.28)$$

For  $\sigma_f \rightarrow 0$  the Gaussian pulse in the signal approaches a Dirac delta function describing an idealised point discontinuity. As  $\sigma_f \ll \sigma, s \geq 1$  the geometric power series converges to unity and can be disregarded, thus the COI for the DOG family is given

$$\frac{(u-v)^2}{s^2\sigma^2} \leq 2 \quad (2.29)$$

$$|u-v| \leq \sqrt{2}\sigma s \quad (2.30)$$

resulting in an empirical COI constant  $\hat{C} = \sqrt{2}\sigma$  which is linearly dependent on the width  $\sigma$  of the DOG base function.

It must be noted that this COI definition proposed by [Torrence & Compo \(1998\)](#) limits the influence of a signal feature to about 14%, i.e. still a considerable contribution of a spike/discontinuity may persist. Hence a good padding of the signal ends is necessary to avoid severe discontinuities that propagate into and distort the signal's transform at higher scales. To be confident of a very limited border effect a COI constant of  $3\sigma$  should be chosen which ensures a damping to  $e^{-9}$  ( $\approx 0.01\%$ ). However this large COI area limits the number of meaningful transformation scales which is disadvantageous especially for short signals.

#### 2.4.4. Wavelet Transform Properties

Following the multi-scale differential operator interpretation of CWT in Equation 2.20, the choice of the  $\text{DOG}_{m=2}$  mother wavelet directly gives  $\psi = \theta^{(2)}$  with  $\theta$  a Gaussian of width  $\sigma$ . Since  $\theta$  is fast decaying, the  $\text{DOG}_{m=2}$  mother wavelet has two vanishing moments. Hence the continuous wavelet transform  $Wf(u, s)$  can be interpreted as scaled second derivative of  $f(t)$  averaged in the neighbourhood of  $u$  with a Gaussian kernel of width  $\sigma s$ . The wavelet transform thus equals the well studied heat diffusion equation from which several properties can be derived ([Perona & Malik, 1990](#)).

From the two vanishing moments it follows that local signal components up to polynomial order 1 are “not seen” by the wavelet. Constant offsets and a linear bias ( $a + bx$ ) in the local signal, described by the wavelet’s time support, cause no response in the wavelet coefficients and are therefore eliminated (see Figure 2.14). This is equivalent to the observation that the second derivative describes the curvature of the signal, not its slope nor amplitude. At higher scales the time support increases such that practically only the global linear signal component is eliminated while locally linear components are carried over (see Figure 2.15). Singularities in a signal result in high-amplitude wavelet coefficients limited to the neighbourhood of the singularities due to the localised wavelet support. The transform  $Wf(u, s)$  of a real wavelet measures the variation of the signal at time  $u$  in a neighbourhood of size proportional to  $s$ . The evolution of the wavelet transform amplitude in scale holds information on the type and characteristics of the singularity.

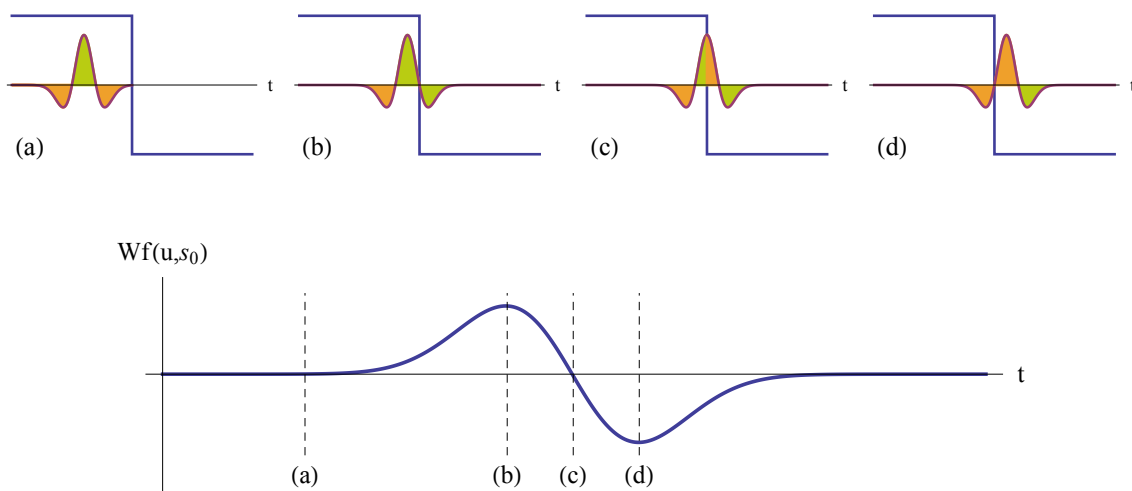


Figure 2.14.: Interrogation of a step discontinuity by a  $DOG_{m=2}$  wavelet for different translations (top) and the resulting wavelet transform (bottom). Green and orange coloured areas indicate positive and negative contribution to  $Wf(u, s_0)$  respectively. Modified after Addison (2002).

Figure 2.14 showcases the interrogation of a step discontinuity by a  $DOG_{m=2}$  wavelet of a fixed scale  $s_0$ . The upper part shows the wavelet atom at four translations moving from outside the discontinuity into it. Below the wavelet transform of the signal is given with markings for the translations. At location (a) the time support of the atom has no overlap with the discontinuity. Hence the convolution of the atom with the constant signal gives a value of zero since positive and negative contribution to the integral  $Wf(u, s_0)$ , indicated by the green and orange coloured areas, annihilate. At (b) the right side lobe of the wavelet atom fully traversed the discontinuity. Since signal and lobe are of equal sign the positive contribution produces a maximum positive transform value. Centring the atom on the discontinuity in (c) yields again a transform value of zero because the  $DOG_{m=2}$  wavelet is symmetric and the area enclosed by a side lobe equals half the area under the main peak. Finally, the transform value reaches a minimum when the atom traverses out of the discontinuity to (d) such that its zero crossing just before the left side lobe coincides with the discontinuity. The contributions of both side lobes cancel out leaving the maximum negative contribution of the main peak.

Stressing the interpretation as multi-scale differential operator (Equation 2.20), the transform features produced by an asymmetric triangular trough in a constant signal are readily explained.

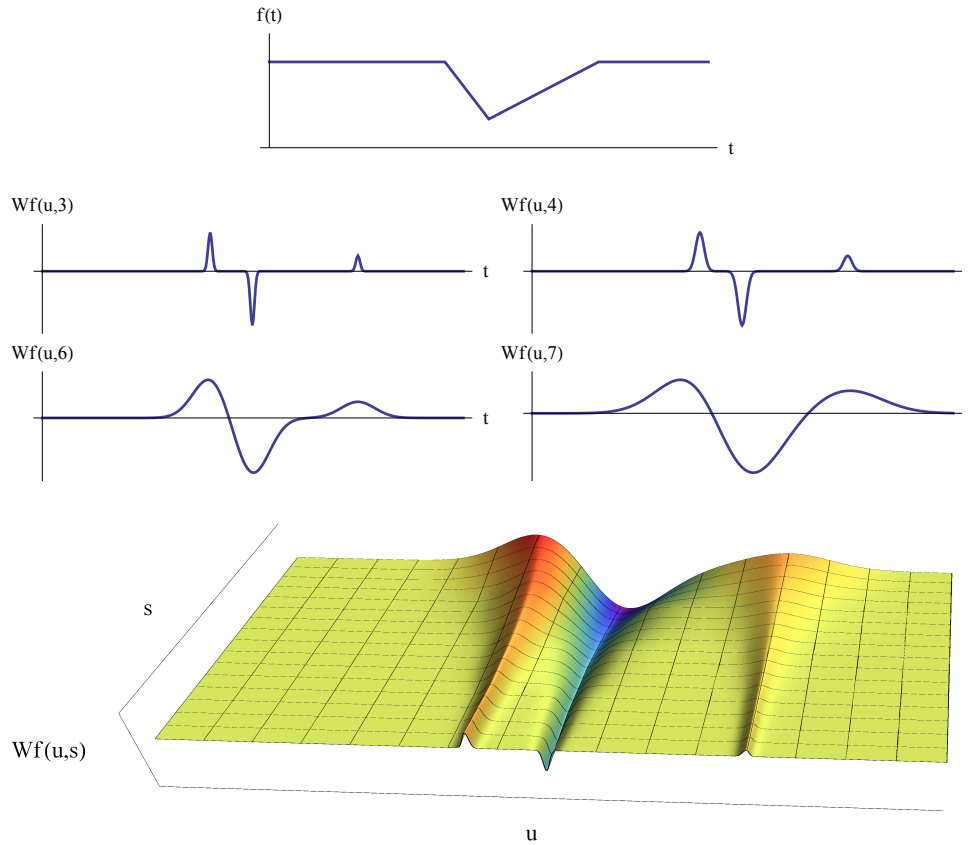


Figure 2.15.: Edge detection in a signal with  $DOG_{m=2}$  wavelet. Signal containing three singularities (top), its wavelet transforms at four fixed scales (middle) and 3d wavelet transform plot (bottom).

Figure 2.15 shows the synthetic signal with its wavelet transforms at four increasing scales. Each of the three discontinuities produces a ripple which width is proportional to the scale. At small scales the Gaussian kernel smoothing is local, accurately locating the discontinuities in time. With increasing scale the kernel approaches the width of the peak and the ripples interfere. A transform of similar shape like the mother wavelet but of noticeable asymmetry results.

The CWT of a signal exhibits four basic transform features: local maxima, local minima, upward zero crossings and downward zero crossings. The former two can be summed as modulus maxima when not differentiating their sign. The  $DOG_{m=2}$  transform can be interpreted the curvature of the signal. A positive transform value indicates the signal is locally concave down (“concave”) whereas a negative value indicates locally concave up (“convex”) behaviour. An upward zero crossing localises a change in the signal’s curvature from concave up to concave down. A downward zero crossing vice versa. Both mark inflection points in the signal. Transform modulus maxima locate maximum curvatures in the signal which are candidates for signal extrema.

The characteristics of a spectral signature are given by its absorption features and continuum. The former are readily described by local curvature changes which are picked up by the CWT at small to medium scales. The continuum contains broad absorptions plus influences by illumination, moisture and viewing geometry. The broad continuum curvatures are retrieved in high transform scales while the undifferentiated disturbing influences are potentially remedied by the cancelling of linear signature components. The CWT of spectral signatures into a multi-scale representation therefore is a promising approach for hyperspectral signal analysis.



### 2.4.5. Hyperspectral Signature Transform

On transition from the generic signal  $f(t)$  to a spectral signature  $\rho(\lambda)$  (reflectance) or  $\epsilon(\lambda)$  (emissivity) it is necessary to consider several peculiarities. Obviously, the interpretation of the transform changes to a scale-wavelength decomposition of the input spectral signature. This implies that i) the signal may be relatively short, ii) sampled at irregular wavelength intervals, with gaps or at changing bandwidths, and iii) noise characteristics are a function of wavelength. Technical aspects of the implementation used in this research to moderate these effects are discussed in Section 3.3 - *Application to Hyperspectral Signatures*.

In general, the CWT converts the 1-dimensional input spectrum into a 2-dimensional representation, which can be visualised as wavelet transform plot (also wavelet amplitude spectrum or scaleogram when plotting the squared magnitude). Each profile for a fixed scale gives a new spectrum. The absorption features, absorption shoulders and inflection points present in the input spectrum all leave characteristic local maxima, local minima and zero crossings in the transform. Figure 2.16 displays the CWT of an albite emissivity signature, showing complex shaped absorption structures around 9,000-10,000 nm.

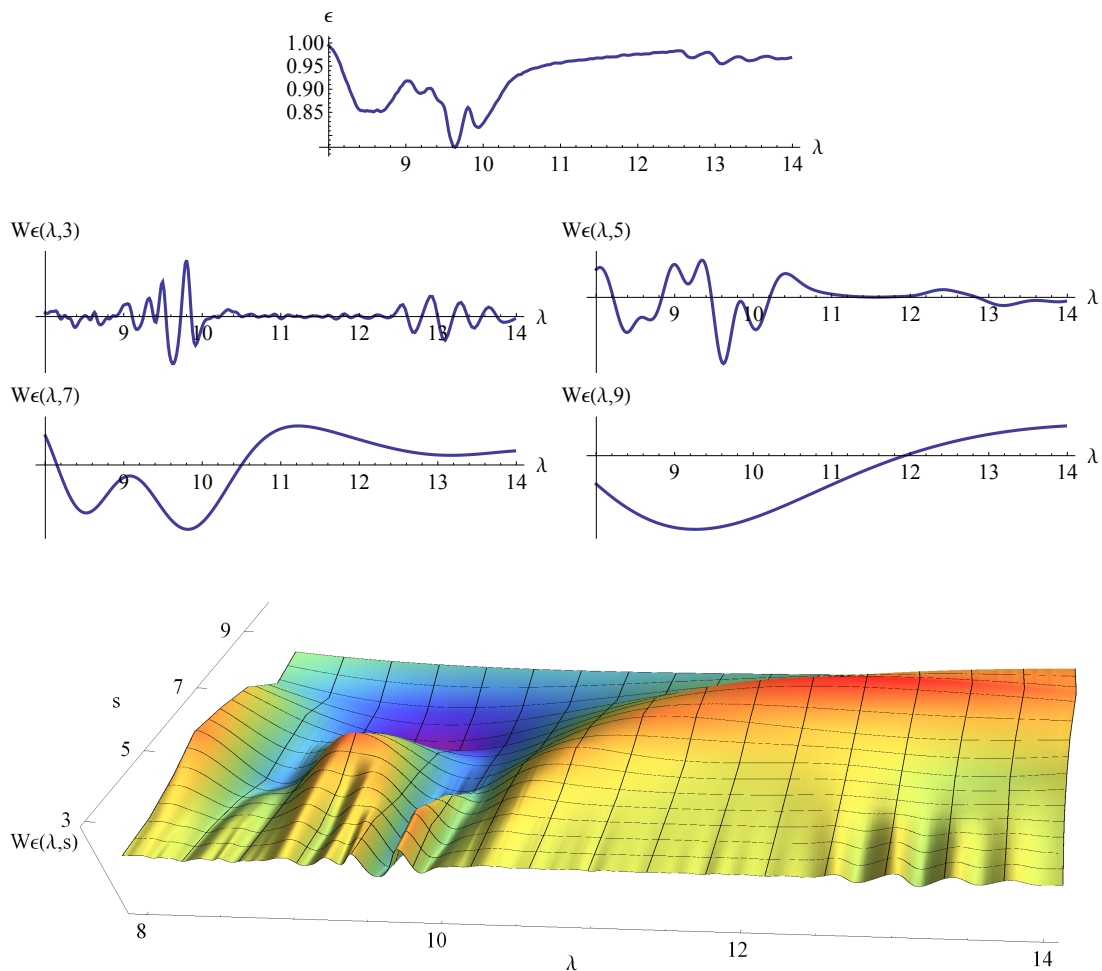


Figure 2.16.: Wavelet transform of a thermal signature from ASU spectral library (Christensen et al., 2000). Albite emissivity spectrum (top) together with its wavelet transforms at four fixed scales (middle) and 3d wavelet transform plot (bottom).

A spectral feature analysis approach that applies a continuum removal and fitting of geometric primitives could hardly describe such complex spectral details. Instead, CWT preserves all spectral details (except a global linear component) and makes them available for the subsequent analysis. The multi-scale decomposition allows to focus on a specific level-of-detail. At small scale narrow and subtle spectral curvature features are highlighted and accurately localised in wavelength. At higher scales the curvature of the spectral continuum is in focus. Information about the spectral feature shapes is conveyed by the development of the transform amplitude over the scales. Linear bias, i.e. the spectrum's local slope, is eliminated within the wavelength support ( $\approx 2 \cdot \text{COI}$ ). Small to medium scales appear therefore continuum removed while increased smoothing at large scales has the effect of emphasising the bias-corrected continuum.

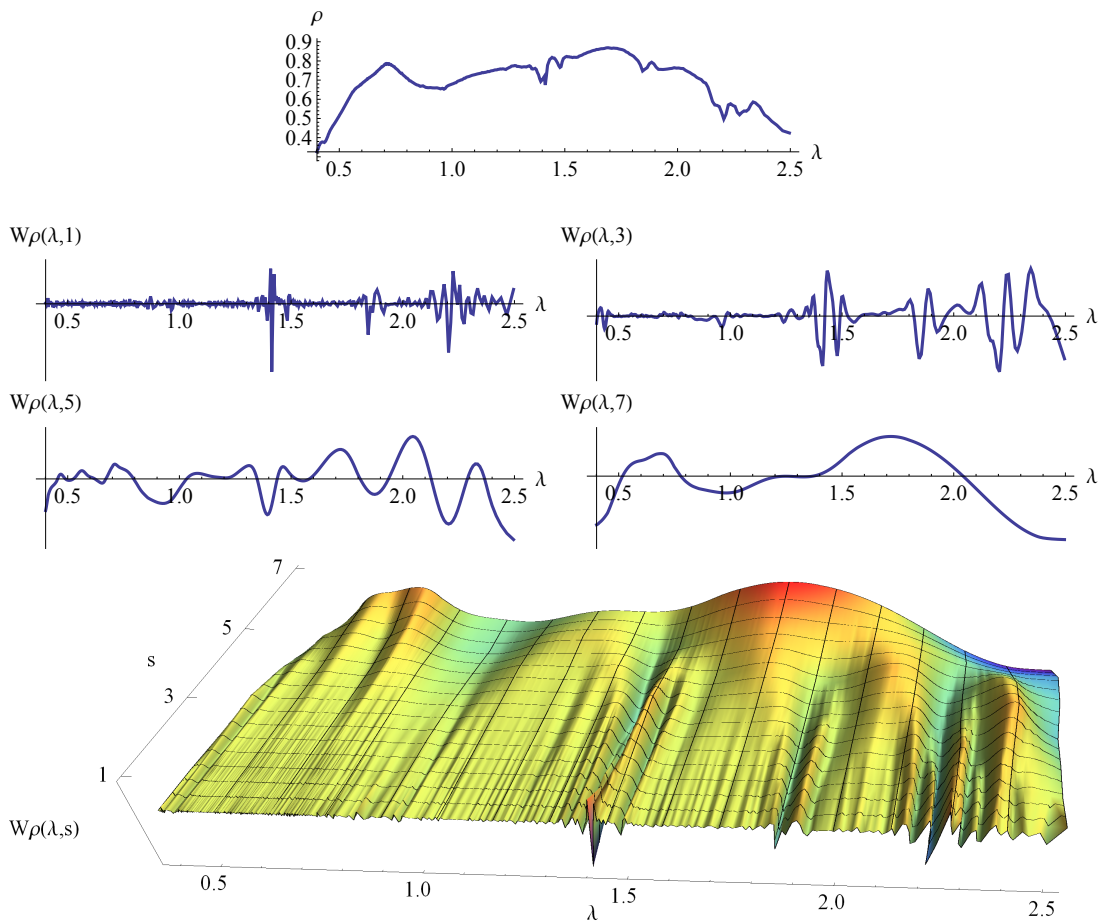


Figure 2.17.: Wavelet transforms (middle) of a synthetic reflectance spectrum (top) of a 50/50 linear mixture of jarosite and kaolinite from USGS spectral library (Clark et al., 1993) together with the 3d wavelet transform plot (bottom).

Figure 2.17 illustrates effects of noise, feature scale and border discontinuities on the CWT, exemplified for a synthetic reflectance signature from a 50/50 linear mixture of jarosite and kaolinite. In the smallest scale 1 spectral noise dominates. At scale 3 the noise is effectively eliminated by the inherent smoothing while subtle spectral features are preserved, e.g. the Al–OH overtone at 950 nm overlaying the  $\text{Fe}^{3+}$  electronic transition at 900 nm, or the shape of the clay doublet at 2,200 nm (Clark, 1999). Border effects can be seen for scale 7 where the slope of the transform flattens off in the COI towards the spectral margins. A custom-designed signal padding can reduce such effects (see Section 3.3 - *Application to Hyperspectral Signatures*).

In summary, various benefits can be exploited when basing spectral signature analysis on the CWT of a spectral signature:

MULTI-SCALE DECOMPOSITION with small scales emphasising local absorption features and large scales highlighting the spectral continuum. Subtle features superimposed on wider absorption flanks are retrieved and correctly localised. Multi-scale analysis and local slope normalisation allow feature separation from spectra where convex hull normalisation fails, e.g. emissivity, or the spectral features are too complex in shape to be parameterized with geometric primitives, e.g. mixtures.

LINEAR BIAS ELIMINATION corrects for additive offsets and linear gain shifts like viewing angle, illumination or moisture differences as well as shifts in emissivity spectra due to a wrong temperature estimate or atmospheric remnants (Hook & Kahle, 1996; Vaughan *et al.*, 2003). Depending on scale, this operation is local or global.

NOISE SEPARATION into the smallest scales since (random) noise is the highest-frequency signal component. The gradual increase of smoothing with scale allows to find a good compromise between smoothing and feature preservation. However, the wavelength dependency of the noise may lead to its leap into a higher scale and inaccurate global noise estimates.

GENERICITY & ROBUSTNESS of the decomposition being independent from the spectral shape characteristics, thus being applicable to reflectance and emissivity spectra, to absorption troughs and absorption shoulders without modification.

The preservation of spectral shape information in several detail levels is expected to improve feature based classification and quantification results for two reasons. Firstly, the multi-scale decomposition allows to focus on the relevant detail level of spectral features, or a combination thereof. This is especially advantageous for feature superpositions in mixtures. Secondly, characteristic features are emphasised and their complex shapes are retained for similarity comparison and feature parameterization. Thus possibly discriminative information is carried over into the classification and quantification models.

### 3. Mineral Classification and Quantification

Mineral classification necessitates a spectral similarity measure. It is used to define the membership of an unknown spectrum to a given mineral class template. Feature based mineral quantification requires a parameterization method. Compositional changes influence the shape and position of the spectral features. Their parameterization conveys this information into a compact representation on which regression models can be build.

CWT has the potential to improve existing classification and quantification methods, but also to devise new ones that benefit from the multi-scale transform, linearity and completeness of the decomposition. In the following sections the state-of-the-art is reviewed and open issues are highlighted. From there mathematical concepts are developed that advance the understanding of wavelet transform based similarity measures and feature parameterization. Finally, two methodological approaches for their application to hyperspectral signatures are elaborated.

#### 3.1. State-of-the-Art & Open Issues

Classification and quantification of minerals in hyperspectral imagery is a vast field of research and development. The high dimensionality of the data led to specialised methods that can operate efficiently on the data and take advantage of the high spectral resolution for improved pattern recognition. Amongst are the Spectral Angle Mapper, Spectral Feature Analysis and Spectral Derivative Analysis which are widely applied due to their simplicity and ease of interpretation.

Spectral shape based classification requires a similarity measure to compare the unknown image spectra against a set of known class templates. One of the widest applied is the Spectral Angle Mapper (SAM, [Kruse \*et al.\* \(1993\)](#)). SAM interprets an image spectrum  $\vec{I}$  and a reference spectrum  $\vec{R}$  as vectors in n-dimensional space, whose similarity is given by the enclosed angle  $\alpha$

$$\cos(\alpha) = \frac{\sum_{i=1}^n \vec{I}_i \cdot \vec{R}_i}{\sqrt{\sum_{i=1}^n \vec{I}_i^2 \cdot \sum_{i=1}^n \vec{R}_i^2}} \quad (3.1)$$

where  $n$  is the number of bands. Multiplicative constants do not alter the result. The spectral difference  $\alpha$  in radians is compared to a threshold to decide whether a class match is found. [Van der Meer \*et al.\* \(2008\)](#) showed that inclusion of irrelevant spectral information has an adverse effect on the matching result and that angles vary non-linearly despite linear mixing. [De Carvalho & Meneses \(2000\)](#) demonstrated that SAM is a constricted variant of the Person correlation coefficient resulting in its inability “of distinguishing between negative and positive correlations because only the absolute value is considered”. They propose a Pearson’s correlation measure  $r$  which mean-centres the spectra and dubbed it Spectral Correlation Mapper (SCM).

$$r = \frac{\sum_{i=1}^n [(\vec{I}_i - \bar{I}) \cdot (\vec{R}_i - \bar{R})]}{\sqrt{\sum_{i=1}^n (\vec{I}_i - \bar{I})^2 \cdot \sum_{i=1}^n (\vec{R}_i - \bar{R})^2}} \quad (3.2)$$

Both methods yield an aggregated similarity measure over the complete spectral range which varies between  $\cos(\alpha) \in [0 \dots 1]$  for SAM and  $r \in [-1 \dots 1]$  for SCM. Good matches score close to 1. The exclusion of irrelevant spectral information is crucial for both methods as they work on shape. To emphasise characteristic features a pre-treatment of the spectra might be necessary. This can be done by limiting to a spectral subset, continuum removal or derivative calculation. The optimal pre-treatment strategy however is typically not known from the beginning. This may result in multiple iterations.

Switching to spectral feature based approaches, the Spectral Derivative Analysis can be used for feature extraction from hyperspectral signatures, e.g. [Sunshine \*et al.\* \(1990\)](#); [Tsai & Philpot \(1998\)](#); [Smith \*et al.\* \(2004\)](#). As described earlier, this method is impeded by the inherent noise amplification and the requirement to select an appropriate smoothing filter. Another popular approach, the Spectral Feature Analysis, decomposes the signal into a background continuum plus superimposed local absorptions, e.g. [Kruse & Lefkoff \(1994\)](#); [Clark \*et al.\* \(2003\)](#); [van der Meer \(2004\)](#). [Rivard \*et al.\* \(2008\)](#) mention that selection of an appropriate continuum shape is an issue and the ratioing during continuum removal introduces non-linearity which impedes unmixing strategies. Also the characteristic features exist in various shapes which additionally may overlay each other. Classical spectral feature analysis has problems in addressing this variability as it relies on simple geometric primitives to describe continuum and absorptions. The information contained in complex shaped absorptions can hardly be conveyed into the quantification model.

Use of CWT is a possible solution to the issues highlighted above. The multi-scale derivative calculation is a linear operation that emphasises spectral features matching the wavelet scale while suppressing others as demonstrated in Section 2.4.5 - *Hyperspectral Signature Transform*. Most authors therefore use continuous wavelet transform to enhance spectral contrast, focus on selected feature scales, remove sensor noise or mitigate spectral differences due to measurement setup and sample treatment prior to the application of standard methods, e.g. [Bruce & Li \(2001\)](#); [Bruce \*et al.\* \(2001\)](#); [Koger \*et al.\* \(2003\)](#); [Kempeneers \*et al.\* \(2005\)](#); [Blackburn & Ferwerda \(2008\)](#). [Rivard \*et al.\* \(2008\)](#) used  $\text{DOG}_{m=2}$  wavelet transform on TIR mineral reflectance and emissivity spectral libraries. They found improved SAM classification performance for the summed smaller scale transform components that capture the compositional variability. Spectral variance due to illumination, viewing geometry and grain size variations were effectively separated into the higher scales. [Feng \*et al.\* \(2011\)](#) applied this methodology to TIR reflectance spectra of drill core logs. The SAM based rock type classification yielded an average accuracy of 72% using the low scale transform components comparing to 55% using reflectance data.

Yet CWT offers much more than a mere pre-treatment and feature extraction from hyperspectral signatures. [Torrence & Compo \(1998\)](#) presented two measures of similarity that operate in the time-scale space. The cross wavelet transform localises scales and times where two signals exhibit a large transform amplitude. The wavelet coherence instead is a measure of the cross-correlation between two time series irrespective of their transform amplitude. It was not until [Torrence & Webster \(1999\)](#) overcame a mathematical pitfall that it could be practically applied to study correlation between several time series of climatic data.

Grinsted *et al.* (2004) and Klein *et al.* (2006) extended the use of coherence to study geophysical and biomedical time series. The outstanding appeal of wavelet coherence is the combination of multi-scale feature enhancement with a localised similarity measure. Instead of an aggregated similarity value as for SAM and SCM, a dynamic measure is at hand that describes the change of similarity over time. To the knowledge of the author this has not yet been adapted and applied to hyperspectral signatures.

Also the time-scale image of the CWT itself can be exploited for the extraction of information. Lyder *et al.* (2010) used the two-dimensional representation of the wavelet transform in an integrated process for correlation screening and abundance estimation of bitumen content in reflectance spectra. The authors dubbed this method “correlation scaleogram analysis”. A correlation scaleogram is formed by calculating the CWT for each input spectrum and stacking the resulting wavelet transform plots into a cube. A profile through the cube lists the CWT values of all input spectra for a fixed scale and wavelength. The dependence between this CWT values and the targeted abundance can be assessed with standard correlation measures. This yields a correlation value for the selected scale-wavelength point which, repeated for all possible combinations, gives a two-dimensional representation of the correlation over the full scale-wavelength space, the correlation scaleogram (see Figure 3.1). Finally the authors located points that yielded highest correlations to establish a multivariate regression model using the promising scale-wavelength combinations as predictors.

All scale-wavelength points were tested for correlation independently. Information carried by hyperspectral signals however is primarily contained in the absorption features and may be non-static in respect of scale or wavelength.

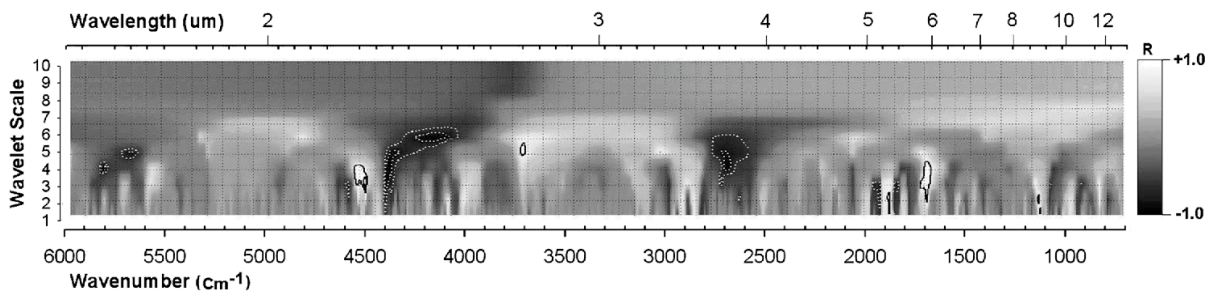


Figure 3.1.: Correlation scaleogram for total bitumen content versus wavelet power from Lyder *et al.* (2010). Contours mark scale-wavelength combinations that yield  $|r| > 0.9$ .

It would be of interest to complement the use of fixed scale-wavelength points by a more dynamic approach which extracts the characteristic patterns that spectral absorption features leave. Piech & Piech (1987) have extensively investigated the scale-wavelength space for the analysis of hyperspectral signals. The authors traced local points of inflection (zero crossings of the CWT in case of  $\text{DOG}_{m=2}$ ) leading to a “fingerprint” that relates to the spectral features (see Figure 3.2). A symbolic, compact, quantitative and hierarchical description of the hyperspectral signature resulted which was used to retrieve soil moisture content. Piech & Piech (1989) applied aforementioned method to airborne imaging spectrometer data to study effects of atmospheric absorption features and in Piech & Piech (1990) to discriminate landforms from terrain profiles.

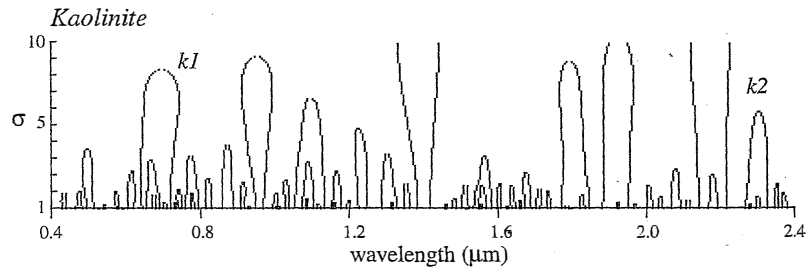


Figure 3.2.: Scale space fingerprint of kaolinite reflectance spectrum from [Schowengerdt \(1997\)](#). Contours trace inflection points which mark enclosed absorption features.

[Mallat & Hwang \(1992\)](#) and [Mallat \(2008\)](#) investigated the mathematical characterisation of signal singularities from the evolution of their wavelet transform modulus maxima in the scale space. Irregular structures in the signal, such as absorption features, leave local modulus maxima in the wavelet transform. Tracing the maxima over the transform scales and chaining them into modulus maxima lines, the authors showed that the evolution of the transform value along the maxima line bears information on the type of the singularity (see [Figure 3.3](#)). This relates closely to the discussed fingerprint analysis which traces zero crossings over the transform scales.

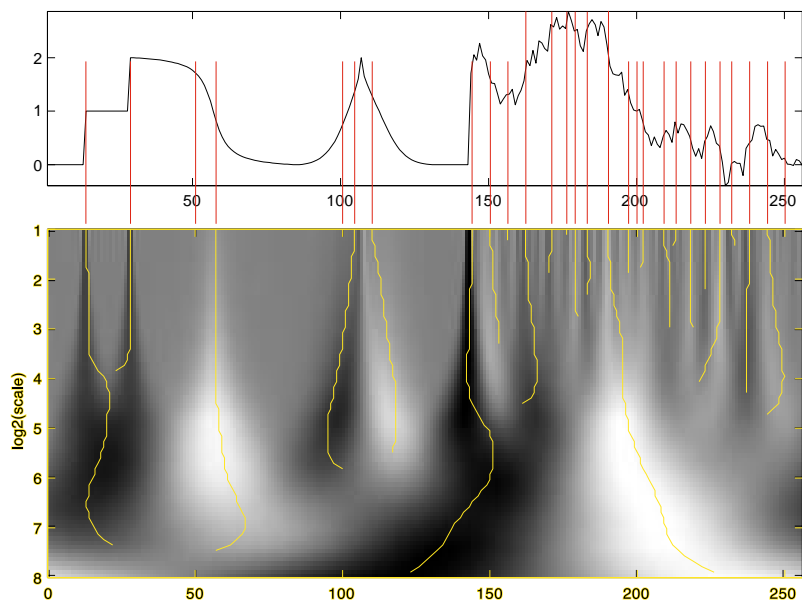


Figure 3.3.: Modulus maxima lines (yellow) extracted from time-scale image (bottom) of  $DOG_{m=1}$  CWT. Link between modulus maxima and signal singularities (red lines). From [Mallat \(2008\)](#).

[Bäni \(2002\)](#) further elaborated the approach for spikes in synthetic signals. He derived curve features and empirical relations between transform value plots along the maxima lines induced by spikes of varying area, amplitude, width and smoothing. Absorption features, absorption shoulders or inflection points in the spectral signatures therefore can be characterised based on the maxima, minima and zero crossings they leave in the scale-wavelength image.

[Hsu & Tseng \(2000\)](#) used a signal regularity measure based on the modulus maxima lines to distinguish between AVIRIS radiance spectra of grass, soybean and corn. [Liu et al. \(2011\)](#) assessed heavy-metal induced stress in crop VNIR-SWIR reflectance spectra. The authors used width, amplitude and a regularity measure based on the transform modulus maxima to quantify zinc levels in soybean, maize, cabbage and rice.

The aforementioned studies limit parameterization of the modulus maxima lines mainly to the regularity measure, i.e. its asymptotic decay, leaving other characteristics such as the amplitude development over scales, length or curvature untouched. The latter can be expected to convey information on shape and spectral neighbourhood of an absorption feature as shown by Bani (2002). Yet their analytic relations are still unknown. Also the studies presented above focus on individual spectra or spectral libraries. A practical demonstration of CWT based classification and quantification methods on hyperspectral imagery consisting of multiple flight lines is not known to the author.

## 3.2. Continuous Wavelet Transform based Developments

**Hypothesis and Approach** Wavelet coherence is a promising approach for quantifying the similarity of two spectral signatures. It brings together two highly desirable properties. Firstly, the linear decomposition of signals into multiple scales of their smoothed derivatives with the benefits listed above. Secondly, the localisation in wavelength of the similarity measure. The latter allows to investigate the signal's conformity over arbitrary spectral subsets and after calculating the similarity measure once. With the coherence a test for the concordance of narrow absorption features up to the broad continuum shapes can be conducted. Scale-wavelength ranges in which the two signals match are highlighted. This enables the user to compare two spectra at selectable spectral detail (scale) and chosen wavelength for matching spectral behaviour.

Section 3.2.1 - *Signal Similarity Measure* will develop the mathematical background of a wavelet coherence based spectral similarity measure and compare it to the Spectral Angle Mapper and Spectral Correlation Mapper. In Section 3.3.2 - *Wavelet Coherence Mapper* a methodological concept for the practical application of the coherence method for mineral identification and classification in hyperspectral data sets will be developed.

The presented fingerprints and modulus maxima lines in the scale-wavelength space are an avenue to compactly describe the spectral features in the input signal. It was shown that the development of modulus maxima and zero-crossings over the transform scales convey information about the shape of a feature and its surroundings. If a link between absorption shape parameters such as depth, asymmetry or width and quantitative descriptors of the modulus maxima and zero-crossing lines in the scale-wavelength space could be established, then Spectral Derivative and Spectral Feature Analysis could be generalized into one wavelet transform feature based method. The introduction of multiple derivative scales once more allows to selectively focus on varying levels of spectral detail and neighbourhood. Possibly an extended set of feature describing parameters, which might better suit the inherent characteristics of the data, could be used for quantification.

Section 3.2.2 - *Signal Feature Reconstruction* will focus on the mathematical link between the shape of absorption features and the induced modulus maxima / zero-crossing lines. In Section 3.3.3 - *Wavelet Transform Feature Analysis* the discovered relations are used to devise a spectral feature parameterization method operating in the scale-wavelength space. The wavelet transform based parameters are compared to the feature describing parameters used in the Spectral Derivative and Spectral Feature Analysis.



### 3.2.1. Signal Similarity Measure

This section introduces wavelet coherence as a localised multi-scale signal similarity measure. The wavelet transforms of two signals,  $Wf(u, s)$  and  $Wg(u, s)$ , can be tested for similarity by the cross wavelet transform defined in [Torrence & Compo \(1998\)](#) and by the wavelet coherence after [Torrence & Webster \(1999\)](#). The cross wavelet transform (XWT) highlights regions in time-frequency space where a high common signal power exists. This are scales and times where for both signals the transform amplitude is large (see Equation 3.3). If both signals are locally of opposing sign then the XWT is locally negative. For the remaining cases it is locally positive.

$$XWT(u, s) = Wf(u, s) \cdot W^*g(u, s) \quad (3.3)$$

The dependence on high common power is rather disadvantageous for comparing spectral signatures that may exhibit similar spectral curvature behaviour but of different amplitude due to brightness or moisture differences. A normalised measure would be of advantage. To find significant coherence between the two signals regardless of their common power the wavelet coherence (WTC) can be used. The WTC is a localised correlation measure in the time-frequency space highlighting “locally phase locked behaviour [...] at the price of slightly less localisation” ([Grinsted \*et al.\*, 2004](#)) due to the necessary application of a smoothing operator. Equation 3.4 defines the squared wavelet coherence where  $\langle \cdot \rangle$  represents the smoothing operation over time and scale.

$$WTC^2(u, s) = \frac{|\langle \frac{1}{s} XWT(u, s) \rangle|^2}{\langle \frac{1}{s} |Wf(u, s)|^2 \rangle \cdot \langle \frac{1}{s} |Wg(u, s)|^2 \rangle} \quad (3.4)$$

[Torrence & Webster \(1999\)](#) proposed a weighted moving average similar in size as the used wavelet’s footprint for smoothing in scale and time directions. For scale, a boxcar of the wavelet’s scale decorrelation length was used whereas for time a kernel defined by the wavelet’s absolute value and normalised to unity was used. This squared wavelet coherence varies between  $0 \leq WTC^2(u, s) \leq 1$ . [Grinsted \*et al.\* \(2004\)](#) pointed at the similarity of Equation 3.4 to a correlation coefficient. In fact the WTC can be seen a localised covariance measure between two signals since the wavelet transform conserves variance ([Torrence & Webster, 1999](#)). As such WTC can be interpreted as a time and frequency localised correlation coefficient. Values close to zero suggest the signals are uncorrelated at that time. Magnitudes close to one suggest that a linear relationship exists between both signals ([White & Boashash, 1990](#)).

$$WTC(u, s) = \frac{\langle \frac{1}{s} XWT(u, s) \rangle}{\sqrt{\langle \frac{1}{s} Wf(u, s)^2 \rangle \cdot \langle \frac{1}{s} Wg(u, s)^2 \rangle}} \quad (3.5)$$

Instead of the squared wavelet coherence, a modified wavelet coherence is defined here. It is better suited for quantifying the spectral similarity because it preserves the sign of the correlation. When using a real valued mother wavelet, e.g. DOG, taking the modulus  $|\cdot|$  can be omitted. Avoiding the squaring operation of the numerator leads to Equation 3.5. This preserves the sign of the XWT thus allowing to differentiate between locally positive and negative correlated signals. Hence the modified WTC’s domain is  $-1 \leq WTC(u, s) \leq 1$  with an interpretation similar to the traditional correlation coefficient.

Keeping in mind the large redundancy of the CWT and the possibility for a fast filter-bank implementation when using an orthogonal wavelet (algorithme à trous), it is proposed to use dyadic scales. This avoids calculation of the whole time-frequency representation by limiting on selected scales but preserving the full time resolution. This is justified by the aspired spectral signature analysis where the wavelength localisation is more important than the frequency resolution. Abandoning a quasi-continuous frequency resolution it is sufficient to perform the smoothing  $\langle \cdot \rangle$  only in time. Using a Gaussian kernel for the smoothing, its width  $\sigma_{\langle \cdot \rangle}$  defines the time localisation, i.e. “sharpness”, of the coherence. Choosing it half the width of the wavelet at the respective scale  $\sigma_{\langle \cdot \rangle} = \frac{1}{2}\sigma s$  and recalling the DOG wavelet’s COI of  $\sqrt{2}\sigma s$  then the time localisation is dilated to approximately  $(\frac{1}{2} + \sqrt{2})\sigma s \approx 2\sigma s$  when applying the coherence.

In Figure 3.4 four experiments with synthetic signals are displayed that demonstrate the properties of WTC (Equation 3.5) and its comparison to the XWT (Equation 3.3), SAM (Equation 3.1) and SCM (Equation 3.2).

- (A) The first experiment compares two opposite, triangular shaped, signals X and Y that are centred around  $f(t) = 5$ . Their  $XWT = X \cdot Y$  is positive, and highest at  $t = 3$  &  $t = 9$  where both signals intersect and their overlap is maximum. The smoothing operation  $\langle XWT \rangle$  has little effect on the shape. The wavelet coherence ( $WTC$ ) and squared coherence ( $WTC^2$ ) both indicate high similarity over the full range. A drop in coherence around  $t = 6$  localises the opposite peaks in the signals. The rather high overall similarity may seem contradicting, but is explained by the fact that calculation of the coherence is not mean centring the signals. The continuous wavelet transform (omitted here) however is, resulting in signal transforms that are typically oscillating around zero. This is comparable to SAM which indicates a rather high similarity for experiment 1 that could be arbitrarily increased by translating X and Y to higher ordinate values. The SCM instead mean centres the signals and detects the perfect negative correlation between both.
- (B) The second experiment is identical to the first, except that X and Y are mean centred as would be the outputs of a CWT. Therefore SAM and SCM both yield the same value indicating perfect negative linear correlation between both signals. The coherence gives a similarity measure over  $t$  indicating a perfect linear correlation for all data points. Note that  $WTC^2$  is not able to differentiate positive and negative correlation whereas the  $WTC$  used here does, hence yielding -1 for all data points.
- (C) In the third experiment two oscillating signals of opposite phase with a common large negative peak are given. The XWT localises the common peak with a high value while for the other data points it yields zero. The smoothing  $\langle XWT \rangle$  slightly degrades the time localisation which results in elevated  $\langle XWT \rangle$  values of the adjacent data points. The coherence is able to differentiate the positively and negatively correlated signal portions  $t \leq 3$  and  $t \geq 4$  respectively. Once again the squared coherence  $WTC^2$  is not differentiating positive and negative correlation showing misleading high overall similarity. SAM and SCM both report a high similarity because they are biased towards strong features in the data. The coherence instead disregards the feature strength and therefore is biased towards the sign of the transform, i.e. the signal curvature direction in case of  $DOG_{m=2}$ .

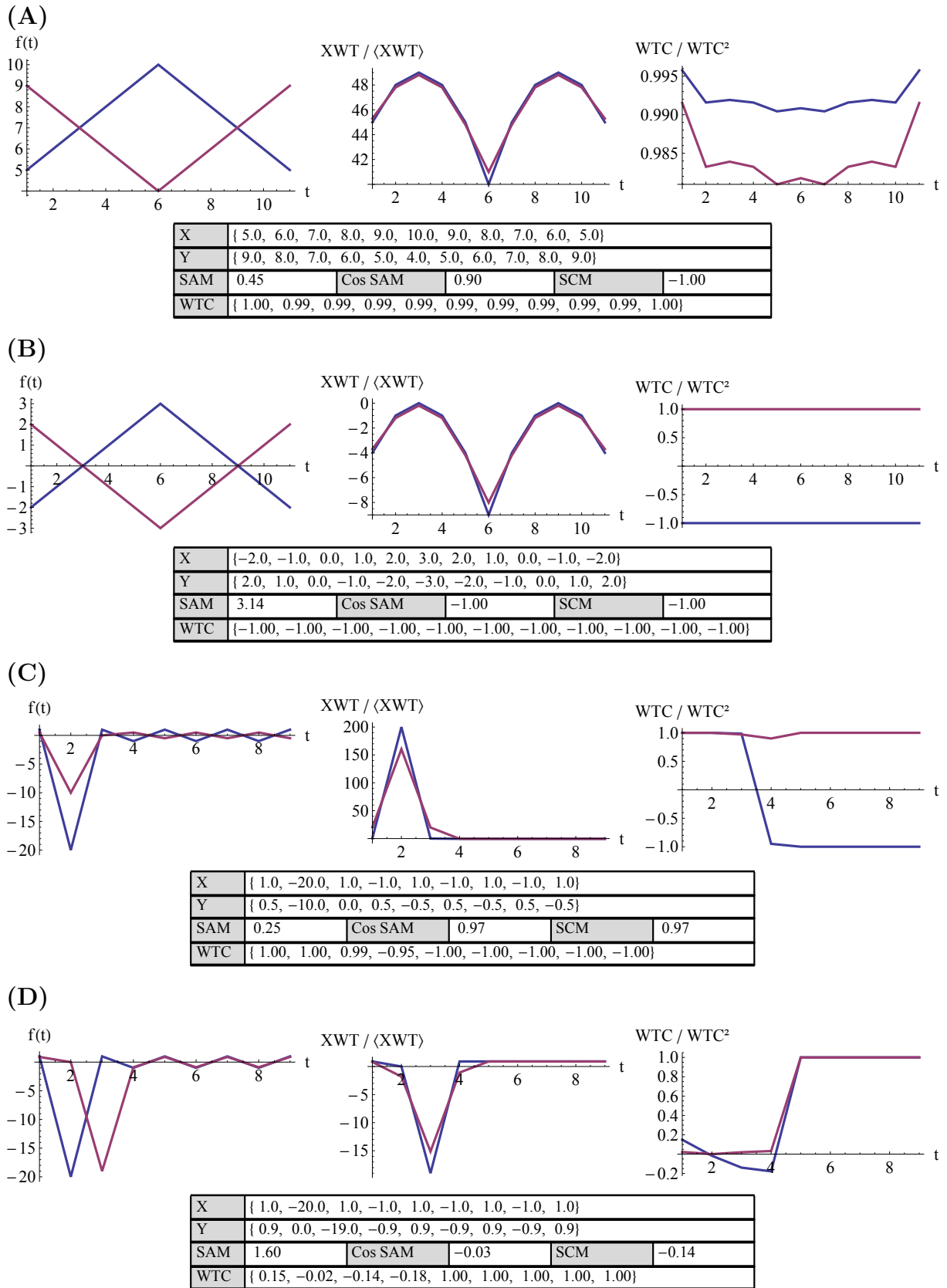


Figure 3.4.: Four similarity tests between synthetic signals  $X$  and  $Y$  (left). Their cross wavelet transform (middle), normal (blue) and smoothed (purple). Resulting wavelet coherence (right), as used in this research (blue) and squared variant from literature (purple). Numerical values and comparison to Spectral Angle Mapper & Spectral Correlation Mapper in table below.

(D) The fourth experiment simulates oscillating signals of identical phase but with a shifted large peak. The XWT localises the common large power at  $t = 3$  where the peaks overlap. The coherence localises the out of phase behaviour for  $t < 5$  and in-phase behaviour for  $t \geq 5$ . SAM and SCM indicate signal mismatch because of their focus on the main peaks.

In summary it was shown that the wavelet coherence is a potent method which complements SAM and SCM. Most notably it returns a similarity spectrum that allows a localised similarity test. The localisation is governed by the inherent smoothing operation and hence scale. Amplitude differences are neglected, as such the coherence accounts for deep and shallow signal features equally, in contrast to the shape driven SAM and SCM. This property make coherence ineffective when applied directly to reflectance/emissivity spectra. Instead it depends on the CWT that removes local additive and multiplicative signal components (in case of  $\text{DOG}_{m=2}$ ) such that a transform oscillating around zero is the input for the coherence calculation. Then its strength in localising phase locked behaviour can be put to optimal use. In conclusion,  $\text{DOG}_{m=2}$  CWT based wavelet coherence does localise wavelength ranges in which the changes in direction of the spectral curvature match. Deviations in the strength of the changes are causing much less penalty on the coherence than deviations in the direction of the curvature.

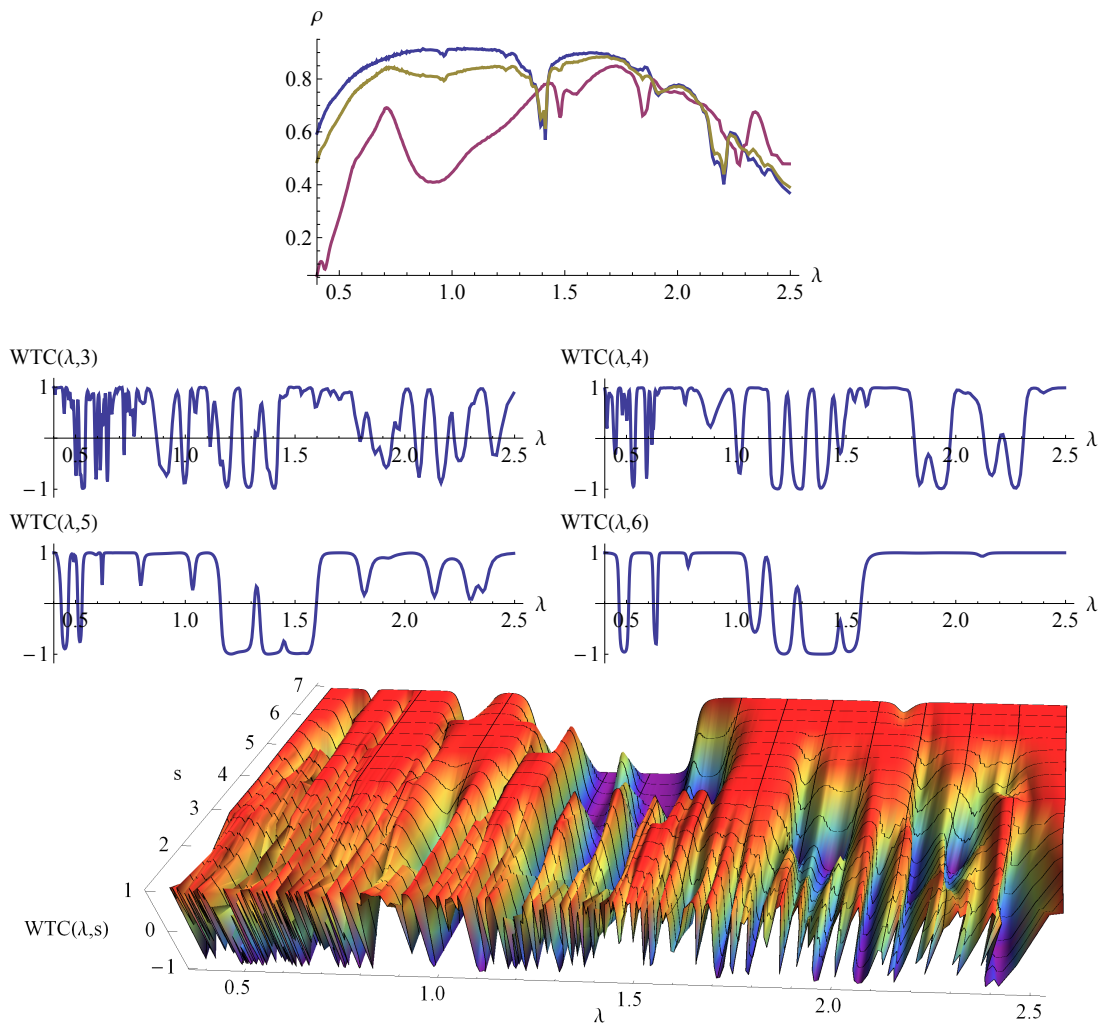


Figure 3.5.: Reflectance spectra (top) of a 20/80 linear mixture (yellow) of jarosite (purple) and kaolinite (blue) from USGS spectral library (Clark et al., 1993). Wavelet coherence spectra (middle) between mixture and jarosite for selected scales and full 3d wavelet coherence plot (bottom).

Figure 3.5 visualises this behaviour together with the multi-scale nature of the WTC for an experiment with a linear mixture of USGS library reflectance spectra (Clark *et al.*, 1993) “kaolini6.spc Kaolinite CM3” and “siderite.spc Siderite HS271.3B” (top). The coherence between a 20/80 % jarosite/kaolinite mixture and the jarosite endmember was calculated for 7 transform scales, visualised as 3d plot (bottom). Four coherence spectra of increasing scale 3, 4, 5 and 6 were picked (middle).

At scales 3 and 4 the detail differences due to narrow features in the kaolinite endmember are highlighted, e.g. 1,400/1,900 nm hydroxide absorptions and 2,200 nm Al–OH feature, or not superimposed narrow jarosite features, e.g. the 2,250 nm Fe–OH absorption (Clark, 1999). At scales 5 and 6 the smoothing (CWT plus WTC) becomes apparent and the similarity measure starts focusing on the curvature of the continuum. Both scales indicate a general good match, interrupted by a major mismatch from 1,100-1,600 nm. Thus even at 20 % mixing the strong curvature changes in the jarosite are sufficiently superimposed on the continuous kaolinite hull to be picked up by the WTC as match. This once more demonstrates that WTC works by curvature changes rather than by exact shape.

### 3.2.2. Signal Feature Reconstruction

This section introduces the reconstruction of signal features from the modulus maxima lines they leave in the scale-frequency space. In reverse, the signal features can be parameterized from the patterns they leave in the continuous wavelet transform.

Suppose the input signal  $f(t)$  consists of a single Gaussian-shaped peak, centred on time  $t = \mu_f$  with an amplitude  $A_f$  and a width defined by the standard deviation  $\sigma_f$ .

$$f(t) = A_f \cdot e^{-\frac{(t-\mu_f)^2}{2\sigma_f^2}} \quad (3.6)$$

The wavelet mother function  $\psi(t)$  is the normalised and sign reversed  $\text{DOG}_{m=2}$ , hence the wavelet base function  $\theta$  and its scaled variant  $\bar{\theta}_s$  are given by

$$\theta(t) = -\frac{2\sigma^2}{\sqrt[4]{\pi}\sqrt{3}\sigma} \cdot e^{-\frac{t^2}{2\sigma^2}} \quad (3.7)$$

$$\bar{\theta}_s(t) = \frac{1}{\sqrt{s}} \cdot \theta\left(\frac{-t}{s}\right) \quad (3.8)$$

$$\bar{\theta}_s(t) = -\frac{2\sigma^2}{\sqrt[4]{\pi}\sqrt{3s}\sigma} \cdot e^{-\frac{t^2}{2s^2\sigma^2}} \quad (3.9)$$

The convolution of the input signal  $f(t)$  with the scaled wavelet base function then gives, according to Equation D.17, another Gaussian function.

$$(f * \bar{\theta}_s)(t) = -\sqrt{2\pi} \cdot \frac{A_f \cdot 2\sigma^2 \cdot \sigma_f \cdot s\sigma}{\sqrt[4]{\pi}\sqrt{3s}\sigma \cdot \sqrt{\sigma_f^2 + s^2\sigma^2}} \cdot e^{-\frac{(t-\mu_f)^2}{2(\sigma_f^2 + s^2\sigma^2)}} \quad (3.10)$$

The second derivative of above Gaussian function is given by Equation D.6 and evaluates to

$$\frac{d^2}{dt^2} (f * \bar{\theta}_s)(t) = \frac{-(\sigma_f^2 + s^2\sigma^2) + (t - \mu_f)^2}{(\sigma_f^2 + s^2\sigma^2)^2} \cdot (f * \bar{\theta}_s)(t) \quad (3.11)$$

Recalling that the wavelet transform can be written as a multi-scale differential operator according to Equation 2.20, then the wavelet transform of the input signal  $f(t)$  can be assembled.

$$Wf(u, s) = s^2 \frac{d^2}{du^2} (f * \bar{\theta}_s)(u) \quad (3.12)$$

$$Wf(u, s) = \left( \frac{s^2}{\sigma_f^2 + s^2\sigma^2} - \frac{s(t - \mu_f)}{\sigma_f^2 + s^2\sigma^2} \right) \cdot \frac{\sqrt{2\pi} \cdot A_f \cdot 2\sigma^2 \cdot \sigma_f \cdot s\sigma}{\sqrt[4]{\pi} \sqrt{3s\sigma} \cdot \sqrt{\sigma_f^2 + s^2\sigma^2}} \cdot e^{-\frac{(u - \mu_f)^2}{2(\sigma_f^2 + s^2\sigma^2)}} \quad (3.13)$$

Equation 3.13 hence gives an analytic expression for the  $\text{DOG}_{m=2}$  CWT of a signal containing an isolated Gaussian peak. Figure 3.6 shows an example for an input signal  $f(t)$  with a Gaussian defined by  $\mu_f = 0$ ,  $\sigma_f = 2$  and  $A_f = 1.5$ . The wavelet base function  $\theta(t)$  was chosen with  $\sigma = 1$ . The convolution of the signal with the dilated wavelet base function,  $(f * \bar{\theta}_s)(t)$ , gives another Gaussian. The wavelet transform  $Wf(u, s)$  is then simply its second derivative. The right part of Figure 3.6 shows the overplot of  $Wf(u, s)$  for 15 increasing scales  $s$ . It can be seen that the positive Gaussian peak in the input signal causes three modulus maxima lines. The main modulus maxima line traces the points of highest positive amplitude of the wavelet transform. It is stationary at  $u = \mu_f = 0$  for all scales. These transform maxima arise from the maximum concave down curvature at the peak of the Gaussian signal. Two weaker side modulus maxima lines trace the transform minima located symmetrically to both sides of the main maxima. Since the local signal is asymmetric and the inherent smoothing increases with scale the lines are drifting apart. The two side minima arise from the maximum concave up curvatures in the side lobes of the Gaussian signal just before it decays to zero.

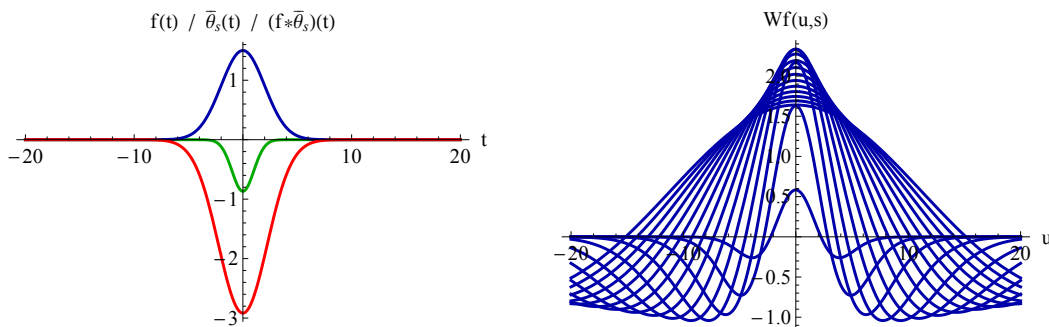


Figure 3.6.: Analytic  $\text{DOG}_{m=2}$  wavelet transform of a Gaussian peak. Comparison of input signal  $f(x)$  (blue), wavelet base function  $\bar{\theta}_s(t)$  (green) and their convolution  $(f * \bar{\theta}_s)(t)$  (red) for  $s = 1$  on the left. Plot of resulting wavelet transforms  $Wf(u, s)$  for scales  $1 \leq s \leq 15$  on the right.

The main peak at  $u = 0$  in the transform becomes wider with increasing scale. Its amplitude increases for the first four scales and then slowly decreases for the higher scales. Thus, focusing on the centre modulus maxima line, there exists a point  $(W_m, s_m)$  where the amplitude of the wavelet transform reaches its maximum along that line. In the following it is shown that the shape of the Gaussian input signal can be reconstructed from the parameters  $W_m$  and  $s_m$ .

Given that the centre modulus maxima line is caused by a Gaussian peak centred at  $t = \mu_f$ , then the transform values along that line can be found by evaluating Equation 3.13 at  $t = \mu_f$ .

$$Wf(\mu_f, s) = \frac{s^2}{\sigma_f^2 + s^2\sigma^2} \cdot \frac{\sqrt{2\pi} \cdot A_f \cdot 2\sigma^2 \cdot \sigma_f \cdot s\sigma}{\sqrt[4]{\pi} \sqrt{3s\sigma} \cdot \sqrt{\sigma_f^2 + s^2\sigma^2}} \quad (3.14)$$

Collecting all constant factors in respect to  $s$  in a constant  $A$

$$Wf(\mu_f, s) = A \cdot \frac{s^{5/2}}{(\sigma_f^2 + s^2\sigma^2)^{3/2}} \quad \text{with} \quad A = 2\sqrt{\frac{2}{3}} \cdot \sqrt[4]{\pi} \cdot A_f \cdot \sigma_f \cdot \sigma^{5/2} \quad (3.15)$$

then the first derivative yields

$$\frac{\partial}{\partial s} Wf(\mu_f, s) = A \cdot \frac{s^{3/2} \cdot (5\sigma_f^2 - s^2\sigma^2)}{2(\sigma_f^2 + s^2\sigma^2)^{5/2}} \quad (3.16)$$

To find the scale  $s_m$  at which the wavelet transform is extreme along the modulus maxima line the first derivative is set to zero

$$\frac{\partial}{\partial s} Wf(\mu_f, s_m) = 0 \quad (3.17)$$

Since  $s > 0$ ,  $\sigma_f > 0$  and  $\sigma > 0$  the second term of the numerator must be zero. It also implies that the found positive root is the only relevant one. Proof of the sufficient condition  $\frac{\partial^2}{\partial s^2} Wf(\mu_f, s_m) \geq 0$  is omitted as the presence of the maximum at  $\mu_f$  is known.

$$5\sigma_f^2 - s_m^2\sigma^2 = 0 \quad (3.18)$$

$$s_m = \sqrt{5} \cdot \frac{\sigma_f}{\sigma} \quad (3.19)$$

Hence the scale at which the extremum of the transform along a modulus maxima line occurs is determined only by the width of the Gaussian peak  $\sigma_f$  in the input signal and the width of the wavelet base function  $\sigma$ . In reverse, the width of the peak in the input signal can be determined from finding the scale  $s_m$  at which the transform amplitude along the line is highest.  $\sigma$  depends on the wavelet choice and is known from the beginning.

Inserting Relation 3.19 into Expression 3.15 gives an analytic relation for the extreme value of the wavelet transform  $W_m$  reached along the modulus maxima line.

$$W_m f(\mu_f, s_m) = A \cdot \frac{s_m^{5/2}}{(\sigma_f^2 + s_m^2\sigma^2)^{3/2}} \quad (3.20)$$

$$W_m f(\mu_f, s_m) = \frac{5}{9} \cdot \sqrt[4]{\pi} \cdot A_f \cdot \sqrt{s_m\sigma} \quad (3.21)$$

In summary it was shown that knowledge of the scale and transform value of the extremum reached along the modulus maxima line allows the retrieval of the width  $\sigma_f$  and amplitude  $A_f$ . Together with the time localisation  $\mu_f$  of the modulus maxima line, this allows the complete reconstruction of the Gaussian peak in the input signal.

From both amplitude  $A_f$  and width  $\sigma_f$  also the integral  $V_f$  of the Gaussian peak can be calculated. The latter allows an approximation of the enclosed area, e.g.  $\pm 2\sigma \approx 78\%$ .

$$\sigma_f = \frac{s_m \cdot \sigma}{\sqrt{5}}, \quad A_f = \frac{9 \cdot W_m}{5 \sqrt[4]{\pi} \cdot \sqrt{s_m \sigma}} \quad \text{and} \quad V_f = \sqrt{2\pi} \cdot A_f \cdot \sigma_f \quad (3.22)$$

Triggered by the work of [Sunshine et al. \(1990\)](#) suggesting a skewed Gaussian model for absorption features in absorbance spectra, several distributions introducing a skew were tested, including the suggested modified Gaussian, skew normal distribution, log-normal distribution and Weibull distribution. However, no analytic expressions for the convolution of these distributions with the wavelet base function could be found. Attempting the saddlepoint approximation ([Goutis & Casella, 1999](#); [DasGupta, 2008](#)) also led to no usable results.

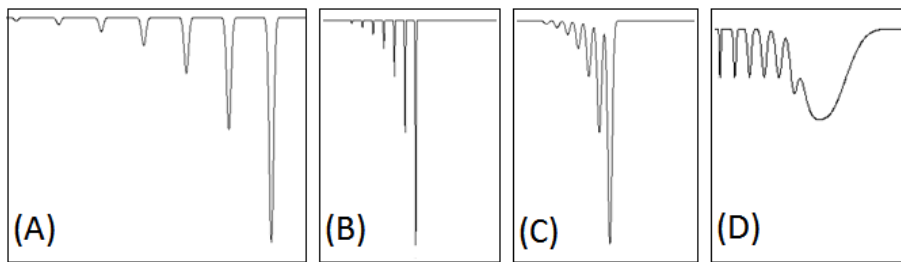


Figure 3.7.: Test signals for modulus maxima line parameter retrieval from seven isolated Gaussian pulses of varying amplitude (A), same but with overlapping COI (B), same but overlapping COI and pulses (C). 7 overlapping Gaussian pulses with varying width (D).

The matter is left here with above analytical relation for a Gaussian peak, the empirical observations of [Bäni \(2002\)](#) and own empirical tests from which important conclusions can be drawn. [Bäni \(2002\)](#) showed that an estimation of signal parameters is noise resistant and that the exact shape of a signal feature, triangular or Gaussian, has little effect on the modulus maxima line (as long as their amplitude and area are identical). Together with the use of discrete scales, the parameterization of signal features from the modulus maxima line is a generalisation. Interference of the wavelet transforms of neighbouring signal features (see Figure 3.7) has little effect on the parameter retrieval as long as the feature overlap is small (cases A to C). When multiple features overlap, then their modulus maxima lines converge into a single composite line. Features that merged (case D) hence can not be distinguished, but are treated as a single composite feature for the area and amplitude retrieval. At small scale the composite nature nevertheless is visible.

In summary,  $\text{DOG}_{m=2}$  CWT of spectral signatures localises the contained absorption features and causes characteristic patterns in the scale-wavelength image. These can be captured into modulus maxima lines which carry information for reconstructing properties of the absorption:

THE WIDTH OF A FEATURE is related to the scale where it causes a maximum transform value along its modulus maxima line.

THE AMPLITUDE OF A FEATURE is reflected by the transform value in relation to the feature's width, i.e. at a scale corresponding its width.

THE SKEW OF A FEATURE changes the apparent time localisation of the feature at higher scales, i.e. the modulus maxima line is curved towards the skew.



It suggests itself to use the scale-wavelength patterns, captured into modulus maxima and zero-crossing lines, for the parameterization of spectral features. Using the  $\text{DOG}_{m=2}$  CWT this can be expanded into a generalised, feature based quantification approach, which brings together Spectral Derivative Analysis and Spectral Feature Analysis.

### 3.3. Application to Hyperspectral Signatures

The method development so far was based on generic signals. The transition to spectral signatures and the application to extensive hyperspectral data sets (hyperspectral image cube) leads to practical and technical constraints. The conceptual benefits of CWT were summarised in Section 2.4.5 - *Hyperspectral Signature Transform*. This section covers the technical and practical implementation of the previously developed concepts from Sections 3.2.1 - *Signal Similarity Measure* and 3.2.2 - *Signal Feature Reconstruction* into applicable data analysis approaches.

#### 3.3.1. Implementation

**Software Environment** The technical realisation of all following analysis modules was made in the IDL<sup>®</sup> 8.0 scientific programming language. Together with ENVI<sup>®</sup> 4.8 it is a widely applied software environment for the processing, analysis and visualisation of multi- and hyperspectral imagery ([ExelisVIS-Homepage, 2013](#)). The CWT provided by the IDL Wavelet Toolkit (© 2002, Christopher Torrence) was used together with helper routines from the MIDL Library (© 2002, Mati Meron, [MIDL-Homepage \(2013\)](#)) for finding signal extrema and from Coyote Library (© 2008, David Fanning, [IDL-Coyote-Homepage \(2013\)](#)) for a linked list data structure.

**Wavelet Choice & COI** The  $\text{DOG}_{m=2}$  mother wavelet was chosen. Section 2.4.2 - *Wavelet Choice* gives the reasoning. The CWT therefore transforms a spectral signature into smoothed second derivatives (see Section 2.4.4 - *Wavelet Transform Properties*). Local maxima of the transform amplitude localise wavelengths where locally highest concave down curvatures are reached, e.g. absorption shoulders. Local transform minima correspond to extreme concave up curvatures within the spectra, e.g. absorption bands. Zero crossings of the transform localise inflection points where the spectral shape changes direction of the curvature. The scale defines the amount of inherent Gaussian smoothing. Here the width of the wavelet base function  $\theta$  (smoothing kernel) was set to its minimum  $\sigma_\theta = s$  and thus also the COI is minimum (for the given wavelet). The extend of apparent smoothing or border effects thus is  $\pm\sqrt{2} s$  which is approximately 1.5 times the scale number of bands to either side of a fixed wavelength sampling point (see Section 2.4.3 - *Cone Of Influence*).

**Signal Padding** Sensor technology currently allowing for some 1000 spectral bands for laboratory/field instruments and some 100 bands for airborne instruments plus the limited wavelength range in the atmospheric windows lead to spectra of rather limited length. Important spectral features may be found close to the margins of the spectral coverage. Border effects as discussed in Section 2.4.3 - *Cone Of Influence* become a concern. They have the potential to introduce artefacts that impair higher-scale analysis in the vicinity of the short ( $\lambda_1$ ) and long ( $\lambda_N$ ) wavelength margins. Addison (2002) and Su *et al.* (2011) list various signal extrapolation methods

and related studies. The difficulty in finding an appropriate padding is its dependence on the chosen wavelet, the signal and the intended analysis. All padding will de facto introduce some kind of artefact, hence masking of the affected areas is the safest.

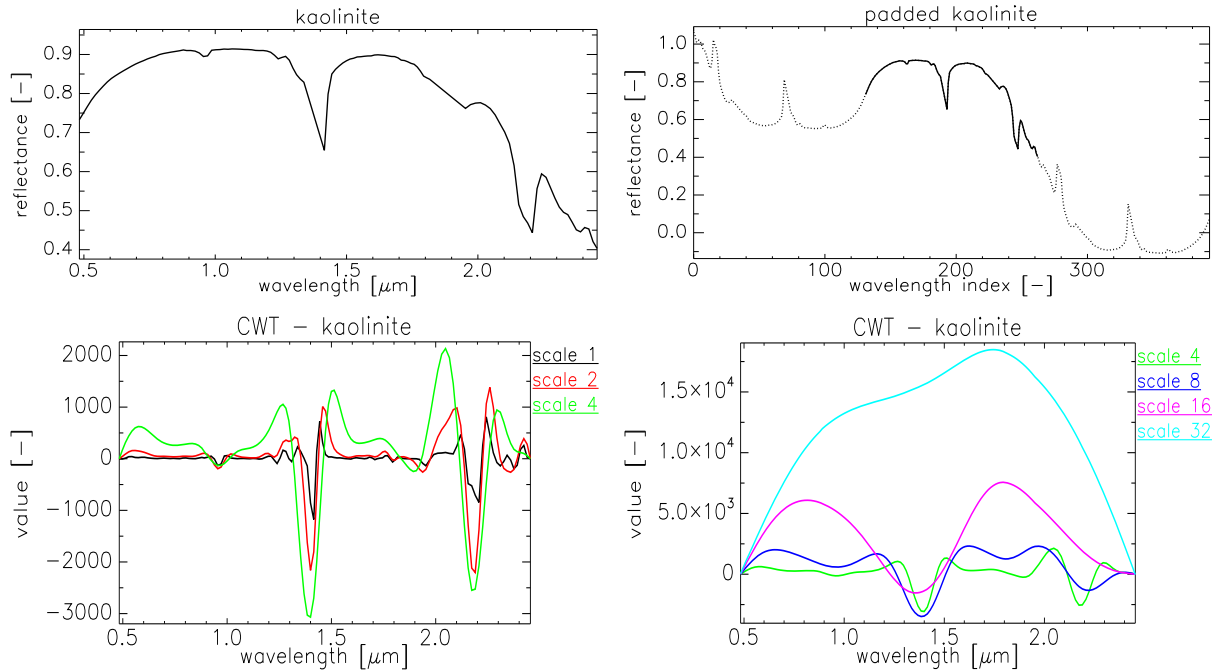


Figure 3.8.: 120 band resampled USGS library (Clark et al., 1993) reflectance spectrum of kaolinite (top left) together with its mirror padded analogon (top right) showing the extrapolation as dotted line. Below the  $\text{DOG}_{m=2}$  continuous wavelet transform showing smaller scales 1 to 4 (bottom left) and higher scales 4 to 32 (bottom right).

Here a point symmetry at the borders  $\lambda_1$  and  $\lambda_N$  was chosen and implemented by a “mirror padding”  $f_{pad}(\lambda)$ . It prevents a signal discontinuity by extrapolating the spectral signature  $f(\lambda)$  with its wavelength mirrored and value flipped analogon (see Figure 3.8).

$$f_{pad} : [\lambda_{-N+1}, \lambda_{2N}] = \begin{cases} f(\lambda) & \text{if } \lambda \in [\lambda_1, \lambda_N] \\ 2f(\lambda_1) - f(2\lambda_1 - \lambda) & \text{if } \lambda < \lambda_1 \\ 2f(\lambda_N) - f(2\lambda_N - \lambda) & \text{if } \lambda > \lambda_N \end{cases} \quad (3.23)$$

This ensures that i) the transform is governed only by the signal in the vicinity, ii) the influence of noisy bands with a low signal-to-noise ratio is limited, and iii) no elevated transform values are introduced that might be mistaken as absorption features. Possible drawbacks are the doubling, and hence overweighting, of signal features and the forced zero crossings in  $\lambda_1$  and  $\lambda_N$  in all scales due to the perfect symmetry. The latter decreases the transform variability in higher scales. This may lead to false positives when using a shape based classification method.

**Signal Resampling and Scales** Spectral resolution varies between different sensors and over wavelength. Also gaps in the wavelength sampling may exist as airborne sensors spare strong water absorption bands around 1,400 and 1,900 nm. It suggests itself to interpolate the spectral signatures to a fixed wavelength resolution of narrow equidistant spacing. This would allow direct comparison between sensors and the information found in individual scales. This comes however at the price of data volume. For a hyperspectral image cube the highly redundant CWT produces

for each scale an own data cube of size equal to the input cube. Data volumes beyond practical margin can be reached easily. For the thesis at hand dyadic scales were used. The signal is decomposed into few discrete scales but the full quasi-continuous wavelength resolution is kept. According to Mallat (2008) this dyadic decomposition is complete, i.e. no information is lost, and the possibility for later use of a fast filter-bank implementation (algorithme à trous) is retained. The application of CWT for spectral signature analysis where the wavelength localisation is more important than the frequency resolution justifies this choice of scales. No wavelength resampling was used, except for wavelength gaps in the signature that exceed 1.5 times the average sampling distance. Linear interpolation over the gaps avoided extreme transform values at the discontinuities. Spectral signatures were treated wavelength indexed during the CWT. No adjustment for changing spectral resolution was made.

### 3.3.2. Wavelet Coherence Mapper

The CWT offers with the wavelet coherence a versatile method to measure the similarity between two spectra. Its mathematical derivation was covered in Section 3.2.1 - *Signal Similarity Measure*. The multi-scale decomposition allows to selectively focus on narrow absorptions up to the continuum depending on the selected scale. Also the wavelength localisation allows to centre analysis on characteristic wavelength regions. Both properties are highly beneficial for material detection and discrimination when mixtures occur, absorption features overlay or are developed only partially.

Basically, the wavelet coherence between two spectra calculated using the  $\text{DOG}_{m=2}$  wavelet returns for each scale a correlation spectrum representing the match of their spectral curvature, highly sensitive to the curvature direction. CWT removes local additive and linear bias, while the coherence is a normalised correlation measure and thus independent of multiplicative factors and amplitude differences. In general the selected scale determines the “level of detail” and the localisation, i.e. the size of the spectral neighbourhood influencing the coherence at a given wavelength (its “sharpness”). The localisation of the coherence was chosen to be approximately  $2s$  bands to either side of a fixed wavelength.

To conserve data volume only the positive coherence values  $[0 \dots 1]$  were retained and mapped on the range of byte  $[0 \dots 250]$ . Negative coherence values were seen of no interest and mapped to a value of 0. The coherence in the COI area ( $\sqrt{2}s$ ) was also set to 0 to suppress possible border effects. This is a conservatory approach as influences can migrate up to  $3s$  bands from the signal boundaries (see Section 2.4.3 - *Cone Of Influence*).

Figure 3.9 gives an example for a siderite endmember from USGS library (Clark *et al.*, 1993) compared to an image spectrum rich in siderite and clay. Analysis of correlation over broad and featureless absorptions like the 1,000-1,300 nm  $\text{Fe}^{2+}$  trough is done best in a higher scale (here scale 8). Narrower absorptions rich in significant small scale variations like the 2,200 nm clay one should be investigated at smaller scales (here scale 4).

A local curvature mismatch is characterised by a coherence dropping towards zero. Local curvature matches manifest as plateaus or spikes reaching high coherence values. Averaging the coherence over a wavelength range gives the spectral similarity in that range decoupled from the

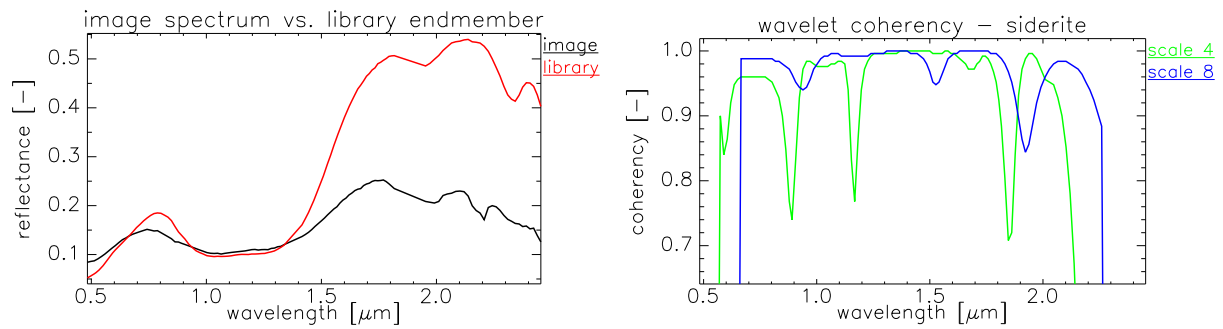


Figure 3.9.: 120 bands reflectance spectra (left) of siderite library endmember (red) and unknown image spectrum (black). Coherence spectra of both for scale 4 (green) and scale 8 (blue).

localisation. Therefore extended wavelength ranges can be investigated at a high level of detail while the reverse obviously is not possible, since low spectral detail and therefore a higher scale imply a wide smoothing kernel.

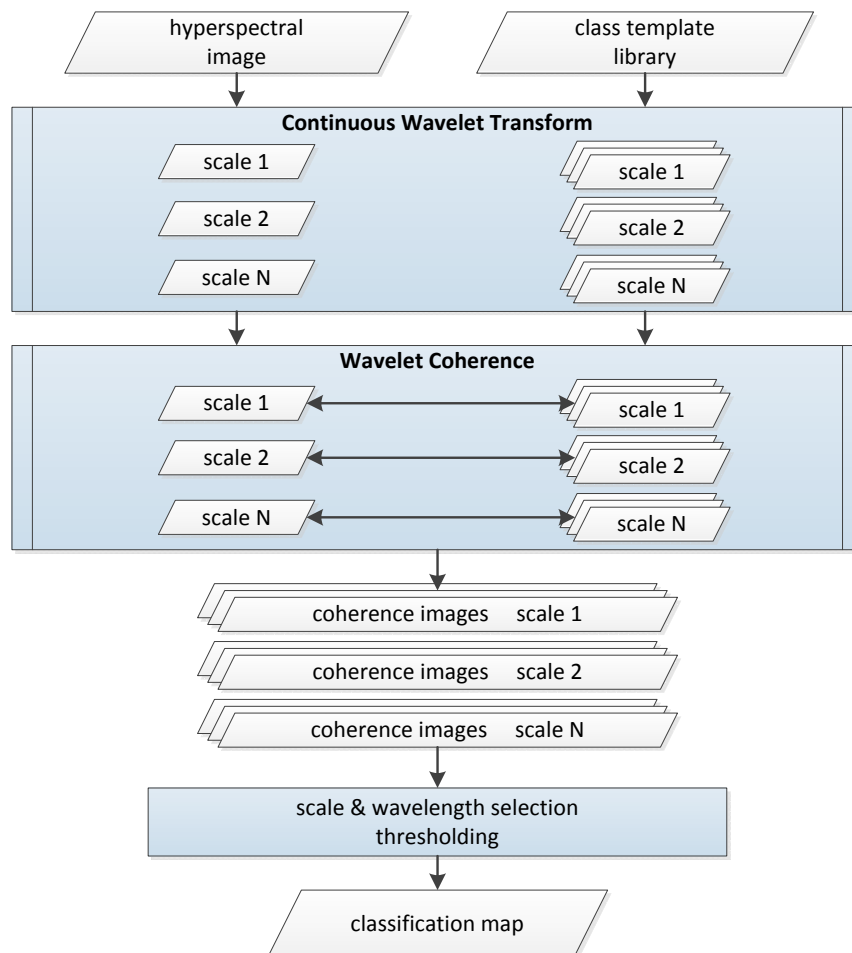


Figure 3.10.: Processing workflow for the application of Wavelet Coherence Mapper in a classification. Data layers depicted as white parallelograms, processing steps as blue rectangles.

The use of wavelet coherence for a pixel-based classification method, dubbed Wavelet Coherence Mapper, is depicted in Scheme 3.10. The conceptual application process is divided into four consecutive steps:

CONTINUOUS WAVELET TRANSFORM Starting from a hyperspectral image to be classified and a spectral library of class reference spectra, the CWT is used to decompose both inputs into a number of dyadic scales. For  $N_s$  scales and  $N_r$  reference spectra this results in  $N_s$  transformed image cubes and  $N_s \cdot N_r$  transformed reference spectra, one for each scale.

WAVELET COHERENCE CALCULATION For all scales individually, the coherence between a transformed image and each of the transformed reference spectra is computed. The calculation is done pixel-by-pixel resulting in a coherence image cube.  $N_s \cdot N_r$  coherence image cubes are generated. One coherence image cube holds the spectral similarity between its image pixels and a given reference spectrum at a certain scale, showing the similarity development over the wavelengths.

SCALE AND WAVELENGTH SELECTION For a classification a class discriminative must be found. An expert user may define a fixed wavelength or wavelength range that, in a certain scale, best discriminates the classes and is little affected by mixing. Rigorous classification geared towards exact matches may use the full wavelength range. The choice of scale gives the ability to selectively focus either on absorption features or the continuum. Also scale governs the spectral neighbourhood that influences the similarity measure. An advanced classification may use multiple wavelength-scale combinations and employ automated class separability optimisation.

THRESHOLDING Once a suitable scale-wavelength combination is found, the respective per-pixel coherence values are extracted, and possibly averaged over a spectral range. This forms a grayscale “rule image” that holds the aggregated similarity score for each image pixel. The process is repeated for every class such that  $N_r$  rule images are generated. A thresholding on the rule images then leads to a classification where individual pixels are assigned the class yielding the highest coherence for that pixel or no class when the maximum coherence is below the threshold.

The Wavelet Coherence Mapper generates, similar to SAM, “rule images” that quantify the similarity between each image pixel and the class template spectra. Class assignment is facilitated by selection of an appropriate coherence threshold. Threshold selection is a major point of criticism of SAM, as the SAM angle is nonlinear and neither a physical nor a statistical measure (van der Meer *et al.*, 2008). In case of Wavelet Coherence Mapper the threshold selection is simplified as i) the normalising character of coherence make it independent from intra- and inter-class amplitude differences, and ii) coherence has a statistical meaning (linear relation between the two signals). Nevertheless a dependence on the scene, sensor, scale and selected wavelength range exists. The latter two are intended and a result from the coherence’s localisation in scale and wavelength. This gives the user advanced options in focusing on discriminative feature scales or using a rule-based approach that combines multiple scales and wavelengths for the classification. This advanced flexibility introduces three degrees of freedom: scale, wavelength and spectral range over which to average. The implications of these choices are discussed in Sections 6.5.1 - *Wavelet based Classification of Siderite, Goethite & Kaolinite* and 6.5.2 - *Wavelet based Classification of Clays* where the Wavelet Coherence Mapper is practically applied and a comparison to the Spectral Angle Mapper is made.

### 3.3.3. Wavelet Transform Feature Analysis

The features of a spectral signature leave distinct patterns in the CWT, namely amplitude extrema and zero crossings which evolve over the transform scales (see Section 2.4.4 - *Wavelet Transform Properties*). Use of the  $\text{DOG}_{m=2}$  ensures this extrema and crossings persist from higher down to the smallest scale (Mallat & Hwang, 1992). When chaining related extrema/crossings over all scales, then so called transform feature vectors do form. They bear information on type and shape of the spectral feature they originate from. Section 3.2.2 - *Signal Feature Reconstruction* demonstrated the retrieval of feature width, amplitude, area and skew from a maximum transform feature vector (modulus maxima line).

The parameterization of this feature vectors which were extracted from the scale-wavelength space therefore leads to a spectral feature based quantification approach. Given that changes in mineral composition and biological, chemical or physical properties have a distinct impact on the spectral signature, this is reflected in the feature vectors. Parameters extracted from this feature vectors consequently can be used to test for correlation with these changes. On success, a regression model using the significant feature vector parameters can be constructed. This quantification method is dubbed Wavelet Transform Feature Analysis. Given the second derivative interpretation and reconstruction abilities when using the  $\text{DOG}_{m=2}$  CWT, the Wavelet Transform Feature Analysis can be seen a generalisation of the Spectral Derivative Analysis and Spectral Feature Analysis to a unified multi-scale approach.

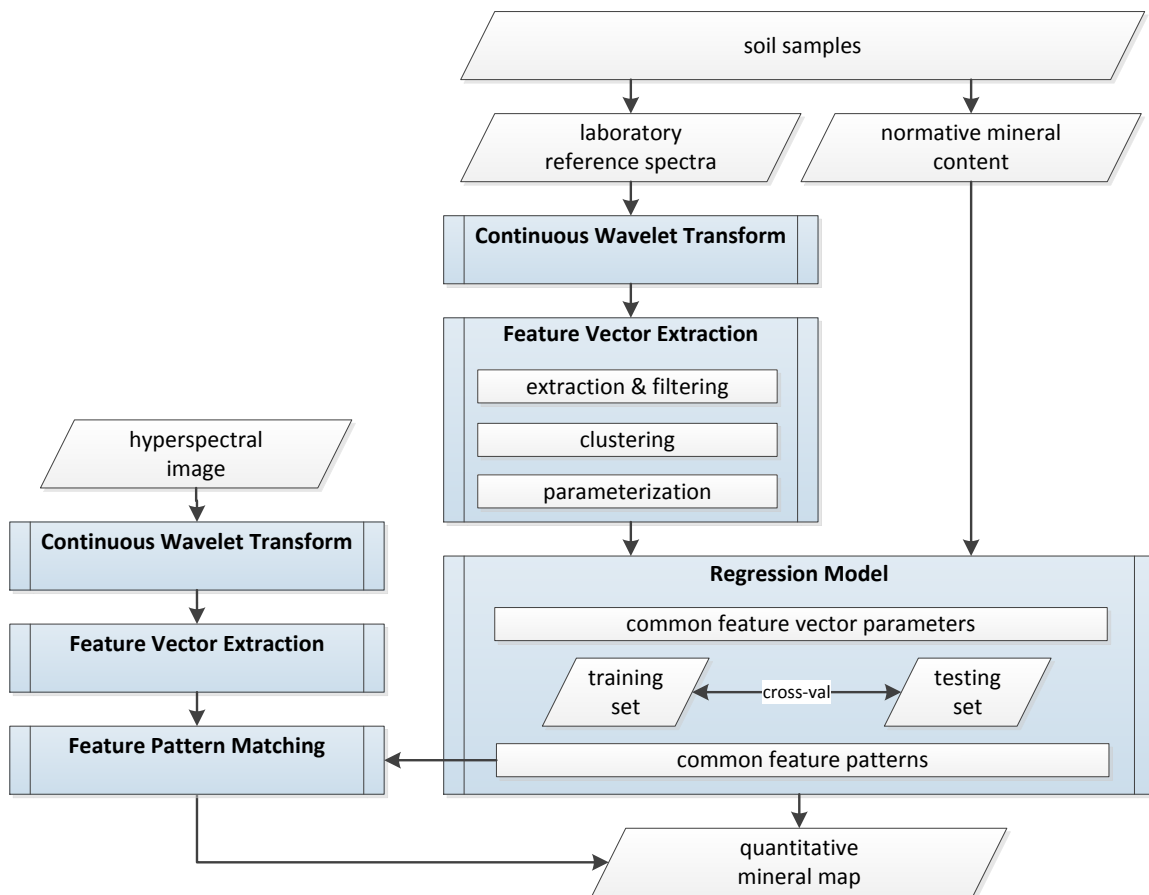


Figure 3.11.: Processing workflow for the application of Wavelet Transform Feature Analysis in a quantification. Data layers depicted as white parallelograms, processing steps as blue rectangles.

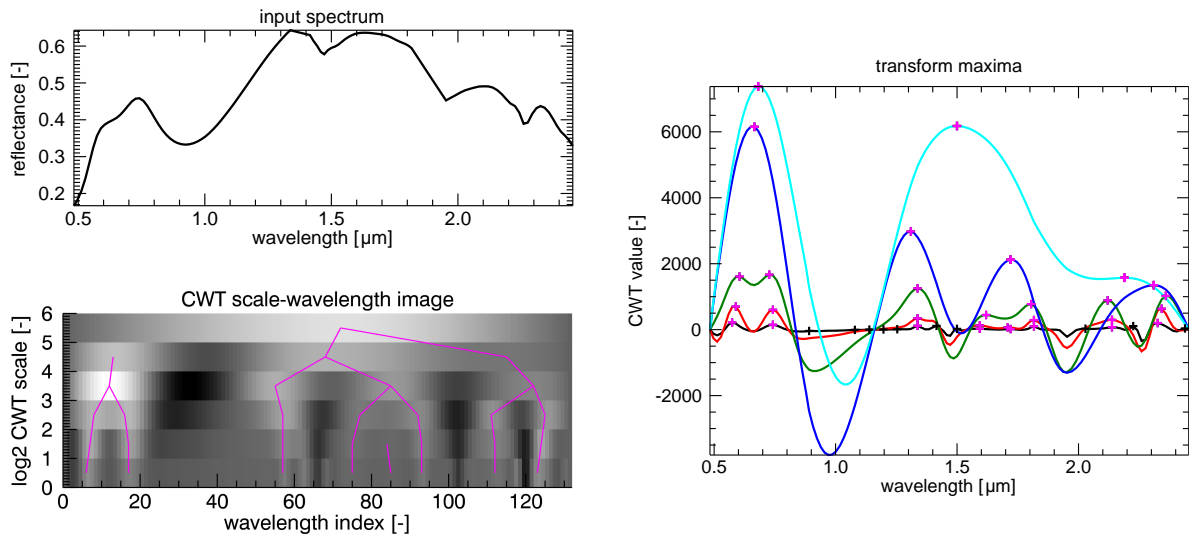


Figure 3.12.: Transform feature vector extraction from example spectrum (left top) showing its scale-wavelength image (left bottom) with extracted maximum transform feature vectors (magenta). Grayscale color ramp spanning from high positive transform values (white) through values close to zero (gray) to large negative values (black). Localisation of extracted transform maxima for multiple scales (right). Maxima belonging to a vector that passed sieving marked as magenta cross. Spurious maxima treated as noise marked black.

Application of the Wavelet Transform Feature Analysis necessitates two stages: i) model development on a set of training spectra and ii) application of the quantification model to image data. Scheme 3.11 presents the individual steps in both stages which are discussed in the following. The model development breaks down into five steps to extract the feature vectors, test them for significance and to derive parameters in order to correlate them with the targeted property:

FEATURE VECTOR EXTRACTION Each spectrum of the training set is wavelet transformed.

From each scale the wavelength localisations of the transform feature types local maxima, local minima as well as upward and downward zero crossings are extracted. Differentiating between this four types rather than treating only modulus maxima and crossings simplifies the following feature chaining and clustering. Also information about extremum type and zero crossing direction may be of importance for the subsequent analysis. For the four types each individual feature is chained into a transform feature vector by tracing its wavelength position from smallest to higher scales according to a search window, i.e. the COI. In case no position can be found in the search window the feature vector is terminated. If multiple features of one kind appear within the search window the one closest to the feature location in the previous scale is chosen (see Figure 3.12).

FEATURE VECTOR SIEVING To reduce the vast number of extracted feature vectors to a meaningful subset several sieving methods were implemented. A proper sieving can exclude spurious feature vectors resulting from numeric inaccuracies or sensor noise (see Figure 3.12). Problematic wavelength regions like water bands can be excluded, or focus on specific regions of interest can be set by considering only feature vectors that start inside a preset wavelength range. Setting a minimum vector length is a means to exclude short feature vectors that exist only in the noisy small scales. Presumably insignificant features that do not persist over the set minimum number of signal smoothing steps are excluded. A noise

threshold after Donoho (1995) can be used to exclude feature vectors of low amplitude. Latter criterion should be used carefully as the noise level varies with wavelength due to water bands, detector transitions, detector sensitivity and incoming radiance level.

**FEATURE VECTOR CLUSTERING** Identification of common features present in a majority of the spectra and distinction from nearby features requires a clustering algorithm. Here the *Generalized Density-Based Spatial Clustering of Applications with Noise* (GDBSCAN, Sander *et al.* (1998)) was chosen because the number of clusters is not required to be known a priori and the clustering is insensitive to its initial point.

A three-dimensional clustering space is proposed (see Figure 3.13). The primary criterion of distinction of subsequent features is their wavelength of occurrence. Common features are required to exist in the majority of spectra. Hence, the secondary distinction criterion is the completeness of the feature occurring over the input spectra indices. A tertiary distinction criterion for feature vectors close in respect to above criteria might be necessary. A ratio of the transform value at different scales or the vector length can be used.

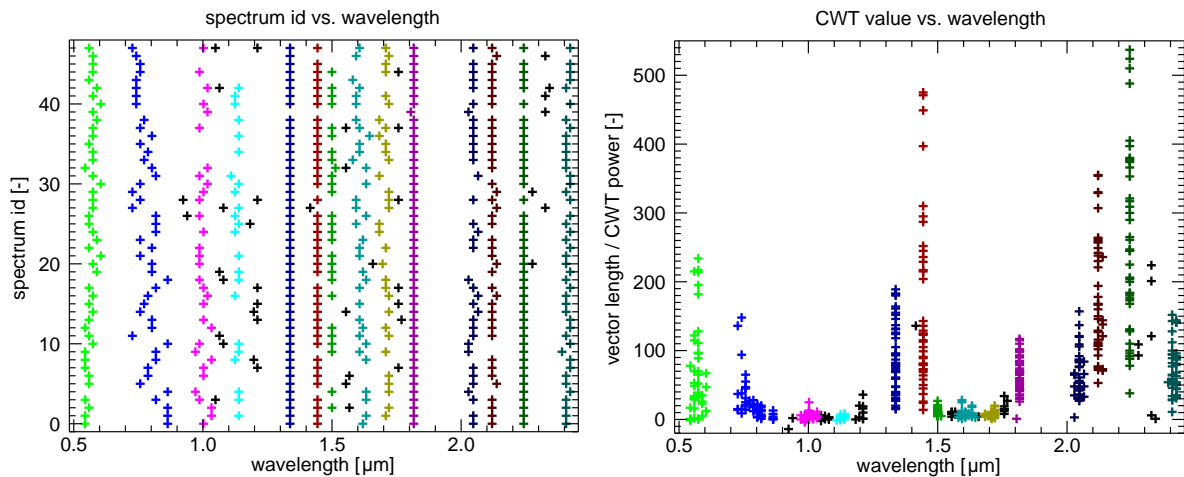


Figure 3.13.: Clustering of extracted maximum transform feature vectors from 50 samples. Left plots primary wavelength/spectrum-id clustering space. Each sample spectrum is represented by a row. Crosses mark the wavelength of occurrence of a transform maximum. Vertical clusters (columns) form that represent common transform features. As spectra are ordered by targeted abundance a systematic wavelength shift as seen for the blue cluster (700-900 nm) may indicate a correlation. Right plots secondary wavelength/value clustering space which in this case is of little distinction (several con-compact clusters forming columns). Colors represent the clustering result. Features separated as noise coloured black.

GDBSCAN defines a cluster by “density”, i.e. the number of points in the neighbourhood. The neighbourhood was implemented as ellipsoid spanned by three  $\epsilon$ -thresholds for wavelength, spectrum-id and wavelet transform value ( $\epsilon_{wvl}$ ,  $\epsilon_{sid}$ ,  $\epsilon_{val}$ ). The neighbourhood predicate of two points  $p_1$  and  $p_2$  results from the logical conjunction of distance tests.

$$nPred = |wvl_1 - wvl_2| \leq \epsilon_{wvl} \quad \wedge \quad |sid_1 - sid_2| \leq \epsilon_{sid} \quad \wedge \quad |val_1 - val_2| \leq \epsilon_{val}$$

A point is assigned a cluster core point when its neighbourhood contains at least a minimum number  $minCard$  of points. The cluster is expanded around the core point as long as more points are density reachable in respect to the neighbourhood predicate and  $minCard$ .



Four parameters  $\epsilon_{wvl}$ ,  $\epsilon_{sid}$ ,  $\epsilon_{val}$  and  $minCard$  need to be chosen. An educated choice can be realised. The spectrum-id are of integer type. A common feature is expected to occur in almost every spectrum. Hence  $\epsilon_{sid}$  can be used to define the maximum expectable count of neighbours, an upper bound for  $minCard$ . For  $\epsilon_{sid} = m$  the maximum count would be  $2m + 1$ . A too small  $\epsilon_{sid}$  may result in many individual clusters that need to be merged.  $minCard$  is to be chosen to allow for a margin of input spectra having no such feature. The closer  $minCard$  to the maximum count the fewer missing features are tolerated. A percentage threshold can be used, e.g.  $minCard = 0.75 \cdot (2\epsilon_{sid} + 1)$ . Obviously the remaining parameters  $\epsilon_{wvl}$  and  $\epsilon_{val}$  should be chosen such that neighbouring features are excluded.  $\epsilon_{val}$  is least distinctive and can be selected rather weak to reject only extreme outliers. Instead  $\epsilon_{wvl}$  is most suitable to describe cluster membership. Sander *et al.* (1998) propose a heuristic for its estimation. Calculating the distance of each point in the cluster set to its nearest neighbour a threshold which separates clusters from noise can be found.  $\epsilon_{wvl}$  is set the distance where the maximum concavity occurs in the sorted distance plot.

Typically the clustering result requires some reconciliation. That includes the merge of splinter clusters which are centred at the same wavelength as well as the split of overfull clusters that contain too many points or multiple points per spectrum. Common features are then represented by clusters exceeding a certain size. A percentage threshold can be used to reject clusters that do not contain a minimum count of transform features in respect to the number of input spectra.

FEATURE VECTOR PARAMETERIZATION Once the common features are found, each of the contained feature vectors can be parameterized. Several numerical descriptors for an individual transform feature vector were considered:

- **feature wavelength at scale** The wavelength  $\lambda_s$  where the respective feature vector is localised in scale  $s$ , i.e. the wavelength localisation of a spectral curvature feature (maximum/minimum curvature or inflection point) in the smoothed input spectrum. It describes the wavelength localisation of absorptions (maximum concave up curvature), absorption shoulders (maximum concave down curvature) and absorption flanks (inflection point).
- **feature wavelength drift at scale** The feature's wavelength drift  $\Delta\lambda_s$  describes the shift in transform feature localisation at a selected scale  $s$  in relation to its occurrence at smallest scale  $\lambda_{s=0}$ . The drift is related to the asymmetry of the spectral feature and its spectral neighbourhood.
- **feature significance** The feature length gives the number of scales the transform feature persists, i.e. its significance in respect to surviving smoothing. It relates to the distinctiveness of a spectral feature in respect to other features in its neighbourhood.

Previous three parameters are applicable to all four transform feature types. Four more parameters can be retrieved for maxima and minima feature vectors only. Section 3.2.2 - *Signal Feature Reconstruction* gives the mathematical background of the below mentioned parameters.

- **feature power at scale** The feature power is the modulus of the wavelet transform value  $|W\rho(\lambda, s)|$  along the vector and at a selected scale  $s$ . It relates to the spectral feature's absorption depth and width at the scope of the scale's respective wavelength support (width of the wavelet in that scale).
- **feature width** The width of the spectral feature is proportional to the scale  $s_m$  at which the transform power reaches its maximum along the feature vector.
- **feature amplitude** The absorption depth / amplitude of the spectral feature is proportional to the power  $W_m$  divided by the square root of the scale  $s_m$  at which the transform value reaches its maximum along the feature vector.
- **feature area** The area of an absorption feature is proportional to the feature width and amplitude given above.

In summary Wavelet Transform Feature Analysis is able to parameterize absorption feature properties as known from the Spectral Feature Analysis in a generalised way (see comparison in Table 3.1). It can also partially generalise continuum properties (see Table 3.2). Besides reflectance/emissivity minima (e.g. absorptions) also maxima (e.g. absorption shoulders) can be parameterized without changing the methodology. Together with the zero-crossings a generalised Spectral Derivative Analysis locating spectral curvature maxima, minima as well as inflection points is facilitated. The significance of a feature in the sense of its persistence over multiple smoothing steps is reflected in the length of the corresponding feature vector which can be used as additional parameter.

Table 3.1.: Analogy between feature descriptors of Spectral Feature Analysis and Wavelet Transform Feature Analysis.

Spectral Feature Parameters		Wavelet Transform Feature Analysis	
name	description	name	description
depth	reflectance difference of absorption minimum to continuum	normalised transform value	transform value normalised by the scale at which the minimum is reached along the feature vector
wavelength	wavelength of absorption minimum	wavelength	wavelength localisation of corresponding minimum feature vector at smallest scale, i.e. the occurrence of maximum curvature
width	wavelength range from left to right absorption shoulder or generalised FWHM	scale	scale at which the minimum transform value occurs along the minimum feature vector
area	area enclosed by the absorption and the local continuum	transform value and scale	product of scale and transform value of extremum along the feature vector
asymmetry	ratio of absorption area to left and right of absorption minimum wavelength	curvature of vector	for higher scales the feature vector's wavelength location shifts in direction of the skew

Table 3.2.: Analogy between continuum descriptors of Spectral Derivative Analysis and Wavelet Transform Feature Analysis.

Spectral Continuum Parameters		Wavelet Transform Feature Analysis	
name	description	name	description
absolute reflectance	reflectance of the local continuum or its average over a certain wavelength range	none	local constant offset is removed during the transform
slope	slope of the local continuum or its average over a certain wavelength range	none	local slope is removed during the transform
curvature	curvature of the local continuum or its average over a certain wavelength range	transform value	transform value represents spectral curvature smoothed by a Gaussian filter of width proportional to the scale

REGRESSION MODELLING For each of the common features, all vector parameters are individually iterated and tested for correlation with the targeted quantity. That means that from all spectra in the training set the value of the parameter currently under investigation is fed as predictor variable in a correlation test. Parameters achieving a significant correlation are combined into a multivariate regression model.

Application of the established regression model to unknown image spectra requires the feature vector extraction and parameterization for each single pixel spectrum. Depending on the common features and vector parameters used by the multivariate model it might be sufficient to limit the transform to a subset of scales or wavelengths. Three more processing steps are required:

PATTERN EXTRACTION For each common feature, the contained feature vectors span a common feature pattern. This pattern can be described by the mean, minimum and maximum wavelength where the contained transform vectors occur plus the minimum and maximum vector lengths. The pattern represents an area in the scale-wavelength image. It envelopes the variability of the feature vectors in the training data, i.e. the variability of the spectral shape characteristics the regression model was trained on (see Figure 3.14).

VECTOR EXTRACTION AND PATTERN MATCHING The spectrum of each individual image pixel is wavelet transformed, the transform features extracted and chained into vectors. The feature vectors are then compared to the established common feature patterns. Only when a pixel spectrum matches all required common feature patterns, then the multivariate model for quantification is applied. This filtering step skips spectra that do not conform to the training data. Hence a model application to materials not trained on or an extrapolation to parameter values not backed by the training data is avoided.

MULTIVARIATE QUANTIFICATION Finally, the required vector parameters of the conforming image spectra are retrieved. The abundances are calculated according to the multivariate model and are aggregated into an output quantitative map.

This feature based quantification approach exploits the multi-scale decomposition of the input signal with its benefit of selective accentuation of spectral components, suppression of noise and normalisation of the input signal. Limiting the regression's predictor variables to parameters

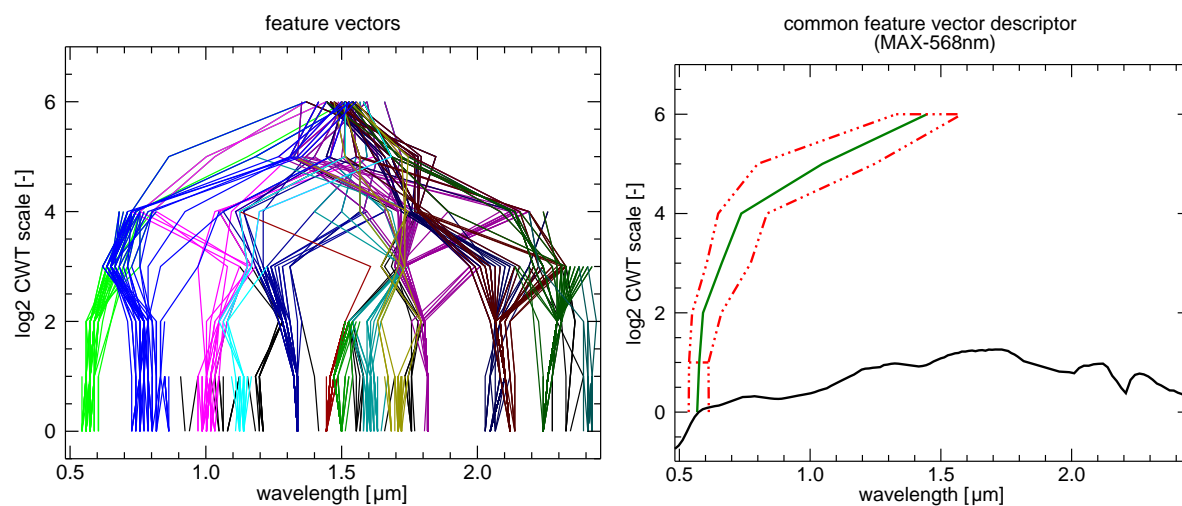


Figure 3.14.: Common feature pattern (right) extracted from the clustered maximum transform feature vectors (left) of 50 samples. Colours identify vectors belonging to same feature. Pattern describes the variability of the lime coloured common maximum around 568 nm in scale-wavelength space (a sample spectrum overplotted for comparison).

derived from spectral features, their interpretation is eased (curvature of smoothed spectrum) and a validation against literature sources (significant wavelength ranges, absorption form factor) is facilitated. The Wavelet Transform Feature Analysis requires no previous knowledge of shape and location of characteristic features and no methodological differentiation between reflectance, absorbance or emissivity nor absorption minima or maxima. Complex spectral feature shapes can be conveyed into the regression model as no geometric primitives were needed for the approximation of signal structures.

A practical application of the Wavelet Transform Feature Analysis for mineral quantification is conducted in Section 6.6.1 - *Wavelet based Quantification of Goethite & Kaolinite*.

## 4. Thermal Imagery Pre-Processing

Calibrated and validated datasets are needed for a meaningful classification or quantification which is comparable and repeatable. It must be ensured that effects due to sensor, illumination and viewing geometry as well as the changing conditions of the atmosphere are corrected for. This is accomplished by a so called pre-processing chain that encompasses the sensor calibration, validation, atmospheric and illumination correction and geocoding of the imagery. In the reflective domain well established processing chains exist. Here the DLR Optical Airborne Remote Sensing and Calibration Facility (OpAiRS) implementation (detailed in Section 6.3.2 - *Airborne Imagery Corrections*) is relied on. Yet, open issues persist in the thermal domain.

### 4.1. State-of-the-Art & Open Issues

Modifications of the TIR signal on its path through the atmosphere can be modelled with a radiative transfer model, e.g. MODTRAN (Berk *et al.*, 2006). Vertical profiles of various atmospheric parameters are required to scale the model. In the TIR, the most important are temperature, water vapour content, pressure and trace gas concentrations. These accurate atmospheric data however is not available to most studies. A number of atmospheric correction approaches that simplify the problem have been proposed. Most notably are Emissive Empirical Line Method (EELM, DiStasio & Resmini (2010)), In-Scene Atmospheric Compensation (ISAC, Young *et al.* (2002)), Autonomous Atmospheric Compensation (AAC, Gu *et al.* (2000)) and use of pre-calculated look-up-tables based on MODTRAN (e.g. ATCOR-4, Richter (2010b)).

The first four correction methods use in-scene data to establish current-state values for the atmospheric parameters. EELM and ISAC are dependent on reference targets/pixels. Typically these are not distributed sufficiently dense to establish a signal path length dependent solution which could accommodate influences of sensor view angle or topography. Instead the radiative transfer equation parameters are held constant for the complete scene. Also mathematically optimized yet physically not feasible atmospheric parameters and at-surface radiances may result. Especially for sensors with few channels or scenes of low contrast this is problematic. Performance of AAC depends on the spectral and spatial resolution of the sensor. It requires two narrow bands, one inside and one outside of an atmospheric water absorption. The radiative transfer equation parameters can be estimated from the ratio of these two bands. This calculation is done for image pixel blocks of assumed quasi-isotropic atmosphere. The look-up-table driven ATCOR interpolates required atmospheric parameters for each image pixel based on the individual viewing geometry. It requires few inputs (e.g. visibility, total water vapour) to select the appropriate radiative transfer parameters from the look-up-table. Obviously they can only be an approximation to the current-state conditions. Residual atmospheric effects may remain.

Above issues, together with uncertainties introduced by temperature-emissivity-separation, lead to unscaled emissivity spectra that are difficult to use for quantitative estimations. The AHS multispectral TIR data available to this study did not offer the necessary spectral resolution to apply AAC. ISAC and EELM were seen critical because of their empirical character and omission of view angle effects. The decision fell on using the physical model based ATCOR.

Strong atmospheric residuals in the emissivity spectra were apparent in the AHS data after pre-processing. This even hampered material identification. Also other researchers like [Cudahy \*et al.\* \(2000\)](#), [Vaughan \*et al.\* \(2005\)](#) and [Hecker \*et al.\* \(2011\)](#) reported such residual effects after atmospheric correction and temperature-emissivity-separation forcing the use of empirical corrections on top of the calculated emissivity spectra. In addition, no complete pre-processing chain existed at DLR OpAiRS which would ensure the quality of the processed thermal imagery.

## 4.2. In-Flight Calibration Module Developments

**Hypothesis and Approach** Pure in-scene data based atmospheric corrections have difficulties addressing spatial variability of radiative transfer equation parameters or need to omit certain terms. Radiative transfer model based methods on the other hand need to fall back on standard atmospheres, as accurate information on the full atmospheric column is usually not available. A generic in-flight calibration linked to a model based atmospheric correction method taking into account varying signal path lengths and viewing geometry would allow to counter above issues. Introduction of an in-flight calibration module will bolster residual atmospheric effects and sensor deviations not adequately modelled so far.

In summary, a generic pre-processing chain for thermal data in compliance with accepted standards already employed for reflective data by various operators (see [Habermeyer \*et al.\* \(2003, 2012\)](#) and EUFAR ([Beekhuizen \*et al.\*, 2009a](#))) can be implemented. Reproducible and quality controlled emissivity datasets can therefore be generated, irrespective of the selected sensor and the prevailing atmospheric conditions. This is a prerequisite for quantitative analyses and the integration of reflective and thermal products.

The in-flight calibration will be based on homogeneous targets of known emissivity and image-based temperatures retrieved from thermal bands least affected by the atmosphere. The latter circumvents difficulties to measure a representative surface kinetic temperature for the ground targets during the exact time of sensor overflight. The in-flight calibration module takes advantage of the ATCOR-4 atmospheric look-up-table and radiative transfer routines to ensure a seamless integration and independence of the sensor's spatial and spectral resolution.

The following section presents implications of varying atmospheric temperature and water vapour content on the thermal domain radiative transfer and thus atmospheric correction. This leads to a thermal in-flight calibration approach which is detailed subsequently.

### 4.2.1. Atmospheric Modelling

Recalling the radiative transfer in the thermal domain (Equation 2.3) given by [Dash \*et al.\* \(2002\)](#)

$$L_{as}^e(\lambda) = L_{path}^e(\lambda) + \tau^e(\lambda) \cdot L_{ag}^e(\lambda) + \tau^e(\lambda) \cdot [1 - \epsilon(\lambda)] \cdot \frac{F^e(\lambda)}{\pi} \quad (4.1)$$

the thermal path radiance  $L_{path}^e(\lambda)$  and transmittance  $\tau^e(\lambda)$  affect the sensor received signal  $L_{as}^e(\lambda)$ . Influence of the (reflected) downwelling thermal sky flux  $F^e(\lambda)$  is small for natural targets given their high absorptance. The atmosphere's vertical temperature and moisture profile as well as the signal's path length govern both  $L_{path}^e(\lambda)$  and  $\tau^e(\lambda)$ .

Signal propagation experiments based on the MODTRAN-4 atmospheric radiative transfer model (Berk *et al.*, 2006) were undertaken in collaboration with the Remote Sensing Laboratory of the Spanish Instituto Nacional de Técnica Aeroespacial (INTA). Computations were based on the predefined standard atmospheres “mid-latitude summer” (45 ° north latitude) and “sub-arctic summer” (60 ° north latitude). Former represents a temperate climate daytime summer atmosphere with a water vapour total column of 3.0 cm and a ground temperature of 21 °C at m.s.l. while latter cooler atmosphere is characterized by a 2.1 cm water column and 14 °C m.s.l. temperature. Radiative transfer was modelled for typical nadir and slant viewing geometry of AHS in Sokolov (see Table 4.1). The water vapour profile was scaled to fixed total amounts.

Table 4.1.: Summary of parameters used for atmospheric modelling in MODTRAN.

parameter	value	note
sensor height	2,400 m	typical airborne survey setup
ground elevation	600 m	for central European location
latitude / longitude	49.43 ° / 11.90 °	(North-West Czech, GMT+1)
time	22.75 GMT	night, decimal time
day of year	203	summer, 22. July
zenith angle	180 ° 135 °	nadir look case 45 ° slant look case
water vapour profile	scaled to fixed value	0.4, 1.0, 2.0 and 2.9 cm total
ground temperature	air temp. profile	scaled temperature of atmospheric model
ground emissivity	1.0	blackbody emitter

Path radiance and transmittance showed a strong viewing geometry dependence (see Figure 4.1). Between 1.2 to 1.4 times higher thermal path radiance values were reached in wavelengths 9,000-13,000 nm for the 45 ° slant experiment compared to the nadir case. Atmospheric transmittance reduced to 95-80% of its nadir value and dropped dramatically towards the 8,000 and 14,000 nm boundaries of the TIR window. Atmospheric correction of imagery of airborne sensors with their typical wide field of view must take the view angle dependence into account.

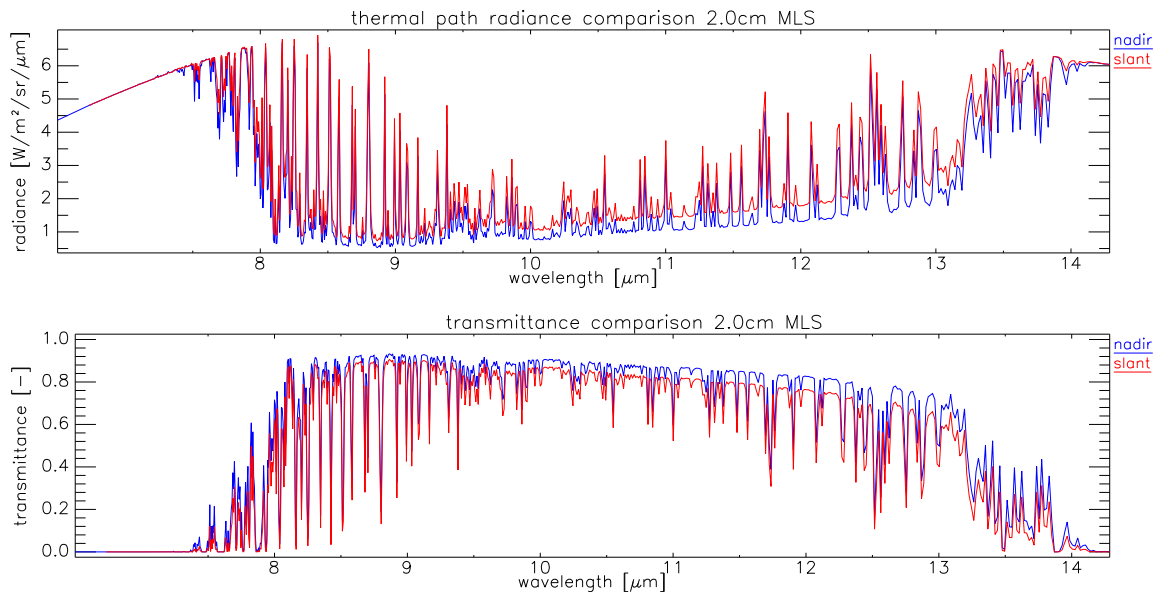


Figure 4.1.: Thermal path radiance (top) and transmittance (bottom) for the 2.0 cm water vapour mid-latitude summer model for nadir (blue) and slant (red) viewing geometry.

Several effects became apparent when comparing the influence of atmospheric temperature and total water vapour (see Figure 4.2): Between 9,000 - 12,000 nm thermal path radiance was relatively low and transmittance relatively high. Outside this spectral window both increased/decreased rapidly. Total atmospheric water vapour content had a major effect on both as expected. An increase from 0.4 to 2.0 cm resulted in a decrease of transmittance to 94% of its original value at 9,000 nm which gradually fell to 82% at 12,000 nm irrespective of the temperature. Path radiance increased 4 to 12 fold over that spectral range. A temperature dependence of path radiance was observed between both of the 2.0 cm and of the 0.4 cm cases where path radiance increased 1.3 to 1.2 fold from subarctic to mid-latitude summer model. Thus humid atmospheres with different temperature profiles show a considerable influence on the at-sensor radiance, which has to be taken into account during atmospheric correction.

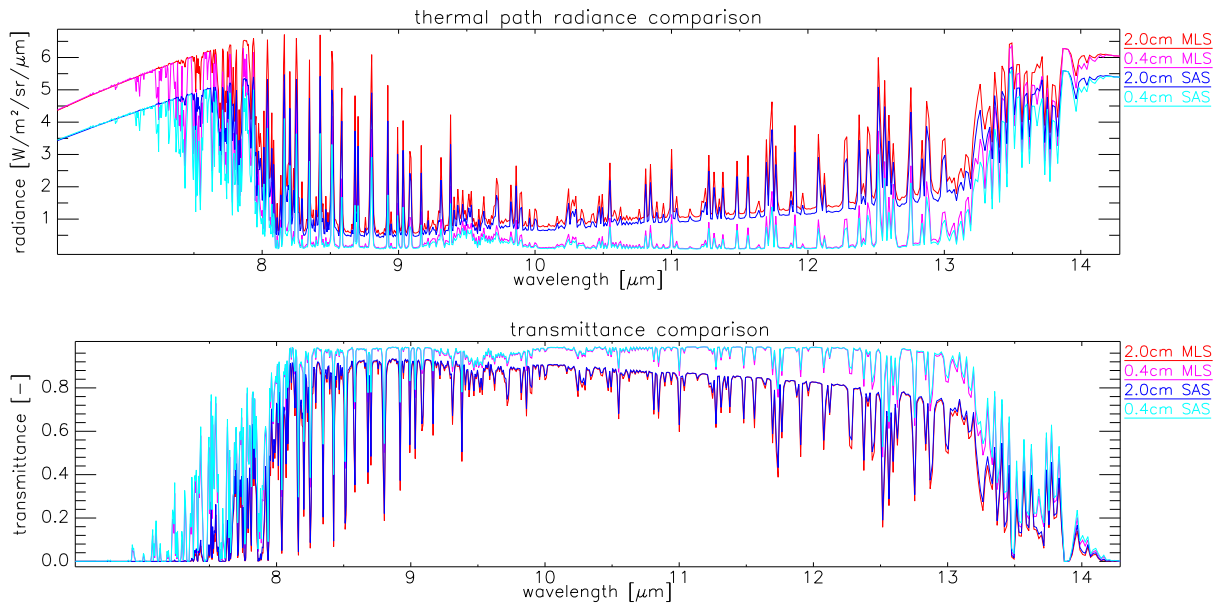


Figure 4.2.: Thermal path radiance (top) and transmittance (bottom) for the mid-latitude summer models of 2.0 cm (red) and 0.4 cm water vapour (magenta) as well as for the subarctic summer models of 2.0 cm (blue) and 0.4 cm water vapour (cyan).

Next to the strong dependence of atmospheric radiative transfer on the total water vapour column between sensor and ground, it was shown that the atmospheric temperature profile also plays an important role. An accurate characterisation of the atmosphere hence is needed.

Often the required profile data, e.g. by radiosonde ascent, is not available for the time of sensor overpass and in proximity of the study area. Usage of model atmospheres which can be scaled to the local conditions by few readily available input data or estimation of the atmospheric parameters directly from the imagery are solutions. Referring to the thermal atmospheric correction methods introduced in Section 4.1 - *State-of-the-Art & Open Issues*, it can be concluded that those methods neglecting the view angle dependent signal path length are not appropriate for wide field-of-view sensors such as AHS. Other methods requiring narrow-band hyperspectral thermal data obviously are not applicable to multispectral thermal sensors currently prevailing.

For the work at hand, the look-up-table based atmospheric correction approach of ATCOR-4 (Richter, 2010b) was chosen. Its atmospheric model in form of a pre-computed look-up-table is scalable to a wide range of sensor configurations and handles view angle dependence as well



as ground surface topography. The applied physical-based corrections resemble the radiative transfer equation components as seen in Equation 2.3. However, there remains a discrepancy between scaled model atmosphere and its complex true state. When neglecting the reflected downwelling thermal sky flux the radiative transfer Equation 2.3 simplifies to

$$\hat{L}_{as}^e(\lambda) = L_{path}^e(\lambda) + \tau^e(\lambda) \cdot L_{ag}^e(\lambda) \quad (4.2)$$

which resembles the ISAC (Young *et al.*, 2002) or empirical line (Roberts *et al.*, 1985; DiStasio & Resmini, 2010) approach. This linear approximation allows to estimate both atmospheric parameters from at least two calibration targets with known at-sensor and at-ground radiance. The general linear sensor calibration formula

$$L_{as}^{cal}(\lambda) = offset(\lambda) + gain(\lambda) \cdot L_{as}(\lambda) \quad (4.3)$$

has an identical pattern as Equation 4.2. Application of Equation 4.3 as an in-flight calibration step on top of the atmospheric correction therefore can eliminate remaining atmospheric effects as well as other constant offsets and linear shifts, e.g. sensor dark current and sensor gain adjustment respectively. Latter calibration values may slightly be changed by installing the sensor in the aircraft and operating it at variable air pressure and temperature regimes. Therefore the in-flight calibration is an additional fine-adjustment of the laboratory calibration.

#### 4.2.2. In-Flight Calibration Module

The in-flight calibration, commonly applied in the reflective domain (Habermeyer *et al.*, 2003; Beekhuizen *et al.*, 2009a; Habermeyer *et al.*, 2012), is complicated in the thermal domain by several peculiarities. The more complex thermal measuring equipment and its time-consuming setup limit the ground crew's mobility and the number of mobilised units. Even calibration targets of known emissivity need permanent measurement of their surface kinetic temperatures and upscaling to a spatial average. Surface cooling due to wind gusts or transiting cloud shadows complicate the task. Temperature of turbulent river water may be accurately measured whereas for stagnant lake water the internal temperature gradient, warm currents along the shores and the flat, mirror-like surface impede a reliable measurement. As a consequence relatively few quality measurements of emissivity and temperature or ground-emitted radiance are available to calibration.

In the following an in-flight calibration method is introduced that closely links to the atmospheric correction in order to estimate the calibration target's representative surface temperature directly from the thermal imagery. Hence temperature measurements are required solely for validation. Calibration can include targets not reachable for ground crews but of known emissivity. The emissivity can be assumed stable with respect to temperature. This circumvents the need of extensive radiance or temperature and emissivity measurements exactly at the time of overflight.

For each sensor band the in-flight calibration transforms the L1 at-sensor signal to a calibrated at-sensor signal by applying an additive offset and a multiplicative gain (see Equation 4.3). To

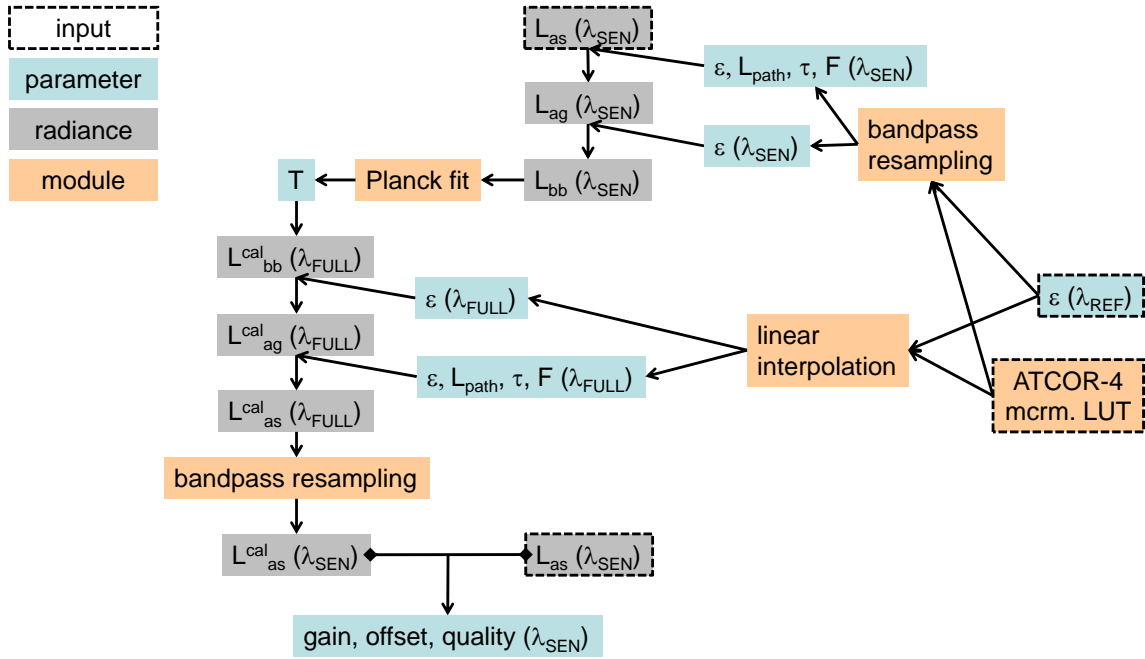


Figure 4.3.: Generic in-flight calibration procedure established for thermal sensors. Calibration target temperature is estimated from atmospheric corrected L1 imagery.

estimate both calibration coefficients at least two calibration targets are necessary that should bracket the expected radiance values in the image, e.g. a dark and a bright target. Favourably more calibration targets are used which are located within different flight lines allowing to reject outliers, to check the sensor stability and thus yield overall more robust calibration parameters. The sensor response is assumed linear and the calibration coefficients are calculated using a least squares fit. A module was implemented which accesses the ATCOR-4 monochromatic look-up-tables and ATCOR-4 sensor definitions while using identical bandpass-resampling routines and basic radiative transfer equations. This ensures compliance with the reflective case and the following thermal atmospheric correction step in the pre-processing work flow. Figure 4.3 presents the in-flight calibration module, in the following explained verbally:

- An average total radiance at sensor spectrum  $L_{as}(\lambda_{SEN})$  is computed for each calibration target from the L1 radiance imagery. It is transformed to at-ground emitted radiance  $L_{ag}(\lambda_{SEN})$  according to the radiative transfer Equation 2.3 requiring surface emissivity, path radiance, atmospheric transmittance and downwelling thermal flux. Latter three atmospheric parameters are taken from the high-resolution ATCOR-4 monochromatic look-up-table ( $d\lambda = 1 \text{ cm}^{-1}$  for TIR). The emissivity is taken from library spectra or averaged ground reference measurements. All four inputs are bandpass-resampled to the respective sensor's band response functions. Nadir viewing geometry with a fixed sensor-to-ground path length, constant visibility and constant water vapour content are assumed.
- With the known target emissivity spectrum the blackbody radiance  $L_{bb}(\lambda_{SEN}, T)$  is isolated from the at-ground emitted radiance  $L_{ag}(\lambda_{SEN})$  according to Equation 2.8.
- The calibration target's kinetic temperature  $T$  is estimated from a Planck function fitted to the blackbody radiance  $L_{bb}(\lambda_{SEN}, T)$ .

- From the temperature retrieval the whole process is reversed. Emissivity and atmospheric parameters are linearly interpolated to a full resolution defined by joining the wavelength sampling points of the ATCOR-4 look-up-table and the reference emissivity. The calibration blackbody radiance  $L_{bb}^{cal}(\lambda_{FULL}, T)$  is attained by evaluating Equation 2.1 at the full resolution sampling points with the retrieved temperature. The conversion to calibrated ground emitted radiance  $L_{ag}^{cal}(\lambda_{FULL})$  and the atmospheric transfer to  $L_{as}^{cal}(\lambda_{FULL})$  is performed using target emissivity and atmospheric parameters at full resolution.
- The expected total radiance at sensor  $L_{as}^{cal}(\lambda_{FULL})$  is bandpass-resampled,  $L_{as}^{cal}(\lambda_{SEN})$ , and compared to the initial at-sensor radiance  $L_{as}(\lambda_{SEN})$  from the L1 imagery. The gain and offset calibration parameters are calculated for each sensor band using a least squares fit which quality is given as Pearson's correlation coefficient.

Errors in emissivity and temperature of the calibration targets cause erroneous offset and gain values which therefore are quality controlled and validated against other flights or other flight lines. Also the estimated surface temperatures are compared to field measurements. Finally the per-band calibration parameters are applied to the L1 imagery just before atmospheric correction and temperature-emissivity-separation. Figure 4.4 gives an example of the fitting and calibration procedure for a vegetation target. The influence of the in-flight calibration on water, forest and soil spectra is exemplified in Figure 4.5.

### 4.2.3. Calibration Target Selection

Following the recommendation of Hook *et al.* (1992), water and dense vegetation targets are well suited for TIR sensor calibration. Referring back to Section 2.3.5 - *Spectral Characteristics of relevant Minerals*, for both landcover types the emissivity spectra are well known, of little variability and spectrally flat, thus virtually approaching that of a graybody. Salisbury & Daria (1992) demonstrate negligible effect (< 1% reflectance) of suspended quartz on water emissivity spectra nearly diminishing for suspended clay. By contrast, windblown dust floating due to the water's surface tension may produce a 1-3 % reflectance increase in TIR reststrahlen band for clay and pure quartz, respectively. Around nadir, dense green and especially coniferous vegetation can be considered a graybody due to the complex canopy promoting multiple scattering (Salisbury & Daria, 1992).

At night time water bodies appear as warm targets due to their high thermal inertia whereas vegetation is relatively cool at about the ambient air temperature. Hence the temperature range for natural objects in the imagery to be calibrated is bracketed. Under day time conditions water and vegetation both appear as cool targets. Asphalt and concrete may be suitable warm substitutes. Bare soils showing distinct emissivity features should be avoided as the features can get superimposed onto the image spectra.

Selection of calibration targets close to nadir avoids view angle effects. Targets distributed over multiple flight lines allow to validate the sensor's stability.

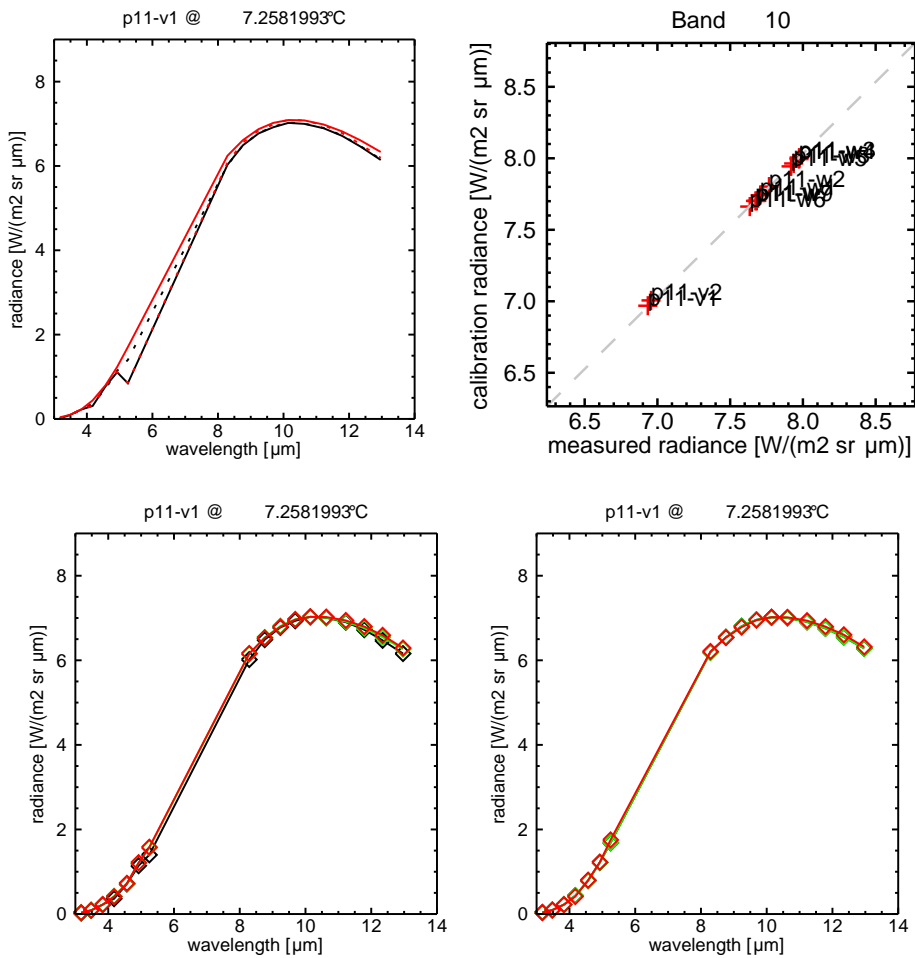


Figure 4.4.: In-flight calibration procedure for a cool vegetation target for AHS MIR and TIR bands. Planck function fitting to target (upper left) displaying  $L_{as}$  (dotted black),  $L_{ag}$  (solid black),  $L_{bb}$  (dotted red) and fitted Planck function  $L_{bb}^{cal}$  (solid red) yielding temperature  $T$  of 7.26 °C. Least squares linear fit (upper right) determining calibration parameters in AHS band 10 from 11 calibration targets. Application of calibration parameters (lower left) showing measured at-sensor radiance  $L_{as}$  (black), theoretical at-sensor radiance  $L_{as}^{cal}$  (green) and actual calibrated at-sensor radiance (red) after applying per-band gain  $\mathcal{E}$  and offset. Difference between theoretical at-ground radiance  $L_{ag}^{cal}$  (green) and actual calibrated at-ground radiance (red) (lower right).

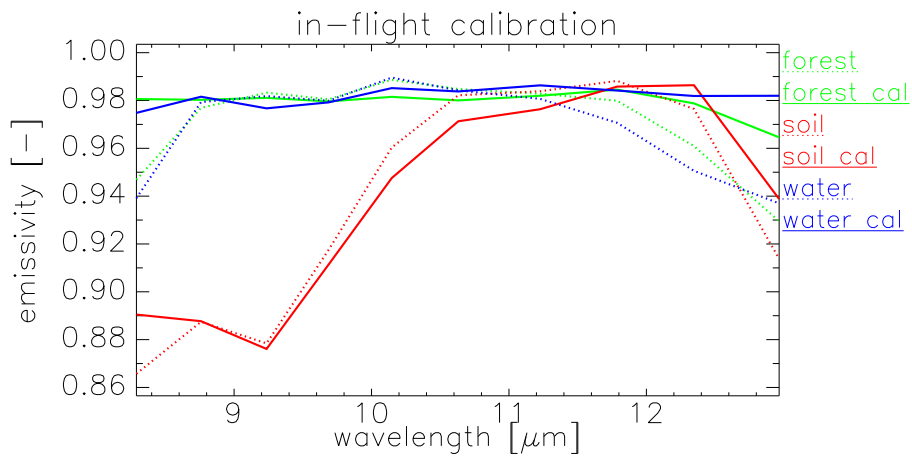


Figure 4.5.: Examples of the influence of the in-flight calibration on ATCOR NEM mean emissivity spectra of water, forest and soil targets leading to improvement of the spectral shape towards the TIR window margins  $<9,000$  nm and  $>11,500$  nm. Emissivity values in the other bands are nearly unchanged (soil emissivity reduction attributes to NEM peculiarity). For water and vegetation the expected graybody behaviour is attained while soil is little affected.

#### 4.2.4. Temperature Estimation

The calibration target's kinetic temperature estimate is the critical part of this approach. Based on the known emissivity and an atmospheric correction the calibration target's blackbody radiance is calculated. A Planck curve fitting approach as proposed by [Hook & Kahle \(1996\)](#) and [Salisbury \(1998\)](#) is used to retrieve the temperature. The fitting was implemented based on a Pascal code (pers. communication Dr. Peter Gege, DLR) using the simplex algorithm of [Caceci & Cacheris \(1984\)](#). Instead of using a single band, it uses the full spectral range of the sensor.

In place of the true surface kinetic temperature the fit retrieves an estimate which is influenced by the fit parameters, residual atmospheric effects, noisy bands and sensor performance. Usage of a Planck fit dampens the noise influence. Bands in regions of low atmospheric transmittance can be excluded. The absolute value of the at-sensor radiance is governed by the sensor's laboratory calibration. Temperature retrieval from the imagery hence leads to a "relative calibration" of the radiance's spectral shape.

#### 4.2.5. At-Ground Calibration

Comparing L2 imagery extracted radiance spectra of the calibration targets, i.e. after in-flight calibration and atmospheric correction, a discrepancy to the expected radiance calculated for these targets and used for the in-flight calibration can be found. This attributes to approximation errors introduced by the full resolution forward modelling of the radiative transfer to the sensor but using coarser, sensor band-averaged atmospheric parameters for the inverse process when atmospherically correcting the imagery. [Richter & Coll \(2002\)](#) documented such bandpass-resampling errors for simulated surface emissivity retrievals of narrowband (150 nm FWHM) and broadband (1,000-2,000 nm FWHM) sensors. For a constant surface emissivity the errors can reach up to 0.8% for the narrowband and 3% for the broadband sensors. A spectrally varying emissivity adds another 3-4% and 2% of respective retrieval error in bands with strong emissivity changes (0.05-0.10). Bands in regions of lower atmospheric transmittance (8,000-8,500 nm and 12,500-14,000 nm) are affected the most.

Above in-flight calibration procedure was consequently given the option to apply the sensor bandpass-resampling before the atmospheric transfer from  $L_{ag}^{cal}$  to  $L_{as}^{cal}$ . Effectively this changes the calibration from a sensor-centred to a ground-centred viewpoint ensuring matching radiance spectra for the calibration targets after atmospheric correction.

## 5. Workflow & Product Validation

“[...] the accuracy of a remotely sensed data product is equally important as the information presented in the product. Without a known accuracy, the product can not be used reliably and, therefore, has limited applicability” (Liang, 2004).

The work at hand turns calibrated at-sensor radiance (L1) through at-ground radiance (L2) into higher level mineral classification and quantification products (L3) to support an assessment of the AMD problem in Sokolov. Not only data from a single sensor in the reflective domain with an established pre-processing chain is used, but a second sensor operating in the thermal domain is adjoined. This offers to cover additional key minerals but necessitates the development of an equivalent pre-processing chain. In addition are novel analysis methods proposed and applied in the reflective domain. This leads to a long and complex processing workflow that combines several sensors, calibration methods, analysis approaches and modelling steps (see Figure 6.2). The many unknowns inevitably need to be addressed by a comprehensive validation concept spanning the whole workflow.

Optimally, the validation should cover the complete product generation process from L1 to L3. This would allow to describe the accumulated uncertainties inherent in the final product and to break down the error budget to the individual sources.

### 5.1. State-of-the-Art & Open Issues

The Guide to the expression of Uncertainty in Measurement (GUM, JCGM100:2008 (2008)) illustrates the importance of quantitative quality indicators to assess the reliability of results. While need of quantitative quality measures is generally accepted, the complexity of the various involved processing steps coupled with an enormous number of differing analysis paths prevent a “readily implemented, easily understood, and generally accepted procedure for characterizing the quality of a result” as requested by the GUM.

Therefore, validation is typically split into several subdomains that treat the processing levels L1, L2 and L3 independently. They can be categorised into instrument driven approaches for the L1 data and product driven approaches for L2 and above. Focusing on airborne sensors, the primary postlaunch instrument validation means is vicarious calibration. This is the simultaneous radiometric measurement of ground targets by an independent and calibrated instrument. The major product driven validation means are direct correlative measurements, i.e. collection of representative in-situ data, and intercomparisons of algorithms/products. The latter encompasses comparison to the same product from different sensors or algorithms as well as to related products (Liang, 2004).

Chander *et al.* (2013) list a number of national and international entities that have put considerable efforts in establishing standards for validation: the NASA EOS Validation Program, the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation, National Institute of Standards and Technology (NIST), the Quality Assurance Framework for Earth Observation (QA4EO) or the envisioned Global Space-based Inter-Calibration System (GSICS).

While routinely implemented for operational spaceborne missions with standardised product output, e.g. MODIS (Liang *et al.*, 2002) or upcoming EnMAP (Storch *et al.*, 2008; Gerighausen *et al.*, 2008), the airborne community joined harmonisation efforts in the European Facility For Airborne Research (EUFAR). Developments geared towards quality assurance of hyperspectral L2 imagery and derived products (HYQUAPRO) with the establishment of standardised quality measures (Bachmann *et al.*, 2007) and metadata models (Holzwarth *et al.*, 2011). Also within the EUFAR framework, European sensor operators agreed on and implemented an error propagation concept for the pre-processing of airborne hyperspectral imagery (Beekhuizen *et al.*, 2009a,b). An analytic modelling is impossible due to the number and complexity of the error influences. Yet the favoured Monte Carlo statistical analysis faces limitations due to the exponential computational costs with increasing degree of freedom. Practical concepts applicable for higher-level products therefore still need to be elaborated.

For L3 products descriptive and analytical statistical techniques such as confusion matrices, Kappa analysis and difference images are typically applied (Lunetta & Lyon, 2004). This requires reference data (ground truth) from in-situ measurements, manual image interpretation or assumed correct outputs from other sensors/models. Congalton (2001) proposes five progressive steps for evaluating L3 product accuracy: visual inspection, non-site specific analysis, difference image creation, error budgeting and quantitative accuracy assessment. The only step that links back to L1 or L2 is error budgeting, yet the impact of individual inaccuracies onto the final product are rarely known as described earlier.

As a consequence do standard classification and quantification studies limit their validation and accuracy assessment to the primary input data, the established regression/classification model and the final result. The intermediate sources of uncertainty inherent to the concerted processing steps are little discussed, hence leading to questionable confidence into the study's findings.

## 5.2. Validation Concept Developments

**Hypothesis and Approach** Validation of the inputs and outputs of the individual steps throughout the whole processing workflow will allow the empirical evaluation of (aggregated) uncertainties in the pre-processing, modelling and analysis. Eventually the major uncertainty sources can be identified. This will back the transferability and repeatability of the results. Furthermore the accuracy and stability of the chosen methods and models can possibly be judged.

The validation efforts will make use of instrument and product driven approaches which may be categorised into direct and indirect validation. The direct methods encompass the vicarious calibration and correlative measurements both relying on independent reference data collected in time and condition representative for the sensor overflight. For most quantitative studies the following data can be expected to be available:

- in-situ spectroradiometric measurements
- auxiliary measurements defining environmental conditions
- targeted geochemical properties of collected samples (mineral fractions)

The indirect validation encompasses product intercomparisons and plausibility checks. Aiming at mineral classification and quantification the availability of a number of additional non in-situ reference data and generated products can be expected:

- laboratory spectroradiometric measurements of collected samples
- spectral libraries of mineral reference samples
- quantitative regression models
- quantitative mineral maps
- mineral classification maps

Above data and products allow a number of straightforward validations as detailed in the following Section 5.2.1 - *Empirical Validation Options*.

The imagery of two airborne sensors, covering the reflective and thermal domain, is available to this study. The processing workflow therefore is two-tier, producing results from each sensor that are complementing or overlapping. This can be used for additional intercomparisons on product level to evaluate the plausibility and value of the individual results. Section 5.2.2 - *Two-Tier Workflow Model & Validation Concept* introduces a conceptual model of such a two-tier processing workflow. The data driven validation options are integrated such that a continuous quality control from primary inputs to the final products exists. The cross link between the two tiers is established for the final products.

In summary this is a small step towards the envisioned ambition to establish a standardised quality & accuracy assessment of the whole product generating workflow (L1-L3), such that quantitative uncertainty measures can be given and the influences of individual errors on the result can be described.

### 5.2.1. Empirical Validation Options

While generic error budgets on a theoretical/statistical basis are very hard to determine, an empirical estimate of the aggregated error can be derived from the data, models or products relatively easy. The standard methods are widely applied and comprise of vicarious calibration, cross-validation, intercomparisons and plausibility checks. To a typical mineral classification and quantification study using data from two airborne sensors the following non-exhaustive validation options are generally available:

VICARIOUS CALIBRATION of the L1 airborne sensor products with in-situ spectroradiometric measurements taken during the overflight to evaluate the postlaunch performance of the sensors and the need for an additional in-flight calibration.

REFERENCE SPECTRA INTERCOMPARISON between the pre-processed in-situ/field and laboratory spectroradiometric measurements of the samples to evaluate representativeness and effects due to environmental/illumination conditions and scaling.



IMAGE SPECTRA INTERCOMPARISON between the pre-processed in-situ spectroradiometric measurements and respective L2 image spectra to evaluate performance of the in-flight calibration and atmospheric/illumination correction.

STATISTICAL CROSS-VALIDATION of the quantitative regression models to evaluate their stability and prediction error.

REGRESSION MODEL PLAUSIBILITY CHECKS by applying the calibrated models to in-situ, library and image spectra in order to evaluate transferability and stability.

CLASSIFICATION MODEL PLAUSIBILITY CHECKS by investigating the feature space to check for ambiguity and separability of the classes as well as their representativeness.

ABUNDANCES INTERCOMPARISON of quantitative maps for overlapping minerals mapped by both airborne sensors to evaluate the mineral distribution patterns, model shortcomings (false positives/negatives) and bias in the derived quantitative information.

ABUNDANCES PLAUSIBILITY CHECKS by verifying the physical 0...100 wt-% constraint for individual mineral maps and their summation (individual sensor and cross sensor) as well as verifying congruent spatial coverage of quantification and classification results or their non overlap due to their mutually exclusive mineralogy.

### 5.2.2. Two-Tier Workflow Model & Validation Concept

A classical quantitative study may be divided into four stages: data acquisition, pre-processing, analysis & modelling, and evaluation. Figure 5.1 shows a scheme for a two-tier workflow combining imagery from two airborne imaging spectrometers A and B covering different spectral domains. Next to the raw image data, also field spectrometers C and D are used to collect in-situ spectroscopic measurements. Spectrometers E and F are used for laboratory measurements of the collected samples. Additional auxiliary data are collected, e.g. geochemical analysis. In the pre-processing stage several correction and quality control measures are applied to ensure the reliability and reproducibility of the data such that it is suitable for the following analysis & modelling stage. Here qualitative and quantitative models are developed based on the reference data which are eventually applied to the imagery. The resulting quantitative mineral maps and mineral classifications are evaluated in the last stage.

The empirical validation tasks are interlaced between the four stages. The raw reference (ground truth) measurements are used to calibrate and validate all further processing steps. Being the starting point, the raw reference data themselves are hardly validated which is a concern. They essentially need to be of high quality, acquired with calibrated and well characterised instruments, and pre-processed to correct for systematic and random errors. Targets characterised multiple times in successive flight campaigns may be used to overcome this source of uncertainty.

From the pre-processed field reference spectra validation links can be established to the raw imagery (vicarious calibration), the pre-processed imagery (intercomparison) and the laboratory spectra of the collected samples (intercomparison) through corresponding spectra. Validated imagery and field/lab spectra are therefore provided as inputs for the analysis & modelling stage.

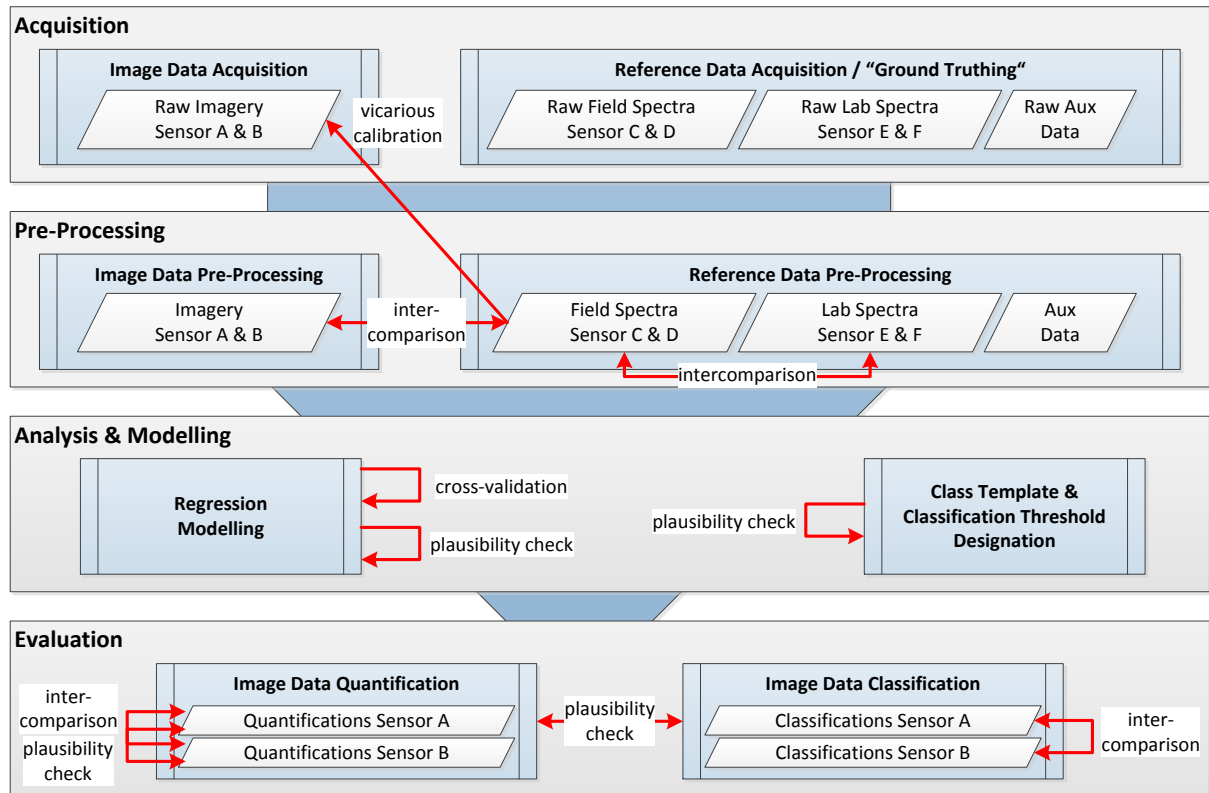


Figure 5.1.: Workflow model for two airborne imaging spectrometers A & B showing four processing stages (gray boxes) with sub-processes (blue boxes) and data products (parallelograms). The generic validation concept is indicated by red arrows.

The regression models, typically established on laboratory spectra, can be integrated into the validation link by statistical cross-validation methods and plausibility checks when applying them to corresponding field and image spectra. Likewise can the trained classification model be validated with the image/field/laboratory reference spectra.

The evaluation of the final data products closes the validation chain. Plausibility checks empirically validate the products derived from individual sensors while intercomparisons of related products establish a validation link between the two sensors.

In summary a validation chain is established which links the primary inputs to the final products. The individual validation tasks are based on data products which allow a straightforward but empirical estimate of the aggregated uncertainties inherent in the processing workflow. The major uncertainty sources can be identified. Naturally, these estimates however will be scene dependent and governed by the representativeness and amount of reference data. The accuracy, including possible systematic errors, of the presumed "true" raw reference data will not be revealed. The quantitative influence of individual sources of error onto the results (error budgeting) also can not be broken down.

## 6. Application Scenario Mining Environment

This chapter brings together the methods developed in Chapters 3 through 5 for a practical mineral classification and quantification study at the Sokolov demonstration site. Section 2.1.1 - *Geologic Setting* introduced its geology and the environmental issues arising from extensive coal mining leading to the formation of AMD. In Section 2.2 - *Application Objective & Mineral Mapping Candidates* the objectives for this practical application were given and the relevant minerals were reasoned.

In brief, the wavelet based methods presented in Sections 3.3.2 - *Wavelet Coherence Mapper* and 3.3.3 - *Wavelet Transform Feature Analysis* will be used to derive qualitative distributions of siderite, goethite and clay minerals as well as quantitative distributions of goethite and kaolinite from the hyperspectral reflective data respectively. The improved pre-processing according to Section 4.2.2 - *In-Flight Calibration Module* will bring the multispectral thermal data to a similar pre-processing quality as the reflective data. Therefore it will be possible to use it to derive quantitative distributions of quartz and kaolinite. This will complement the information derived from the reflective data and will allow for a plausibility check between the results from both spectral domains. The latter is part of the empirical validation concept from Section 5.2.2 - *Two-Tier Workflow Model & Validation Concept* which will be applied to evaluate the individual processing stages based on the generated (intermediate) data products. The processing workflow is developed in the following section.



Figure 6.1.: Impressions of an active coal mine (top) and of an AMD affected ancient mine site (bottom) in Sokolov. Photographed during field visits 2010 and 2011, © EO-MINERS project.

## 6.1. Overall Workflow Concept

Starting from the available data sets that will be detailed in Section 6.2 - *Data Sets and Measurement Protocols*, the following processing outline describes the major processing steps that will be accomplished. They follow the stages postulated in Section 5.2.2 - *Two-Tier Workflow Model & Validation Concept*. Scheme 6.2 displays the outline in graphical form.

1. Pre-processing of field and laboratory spectroscopic measurements following established standards and protocols for reflective data and establishing a compliant workflow for thermal data (Section 6.3). This includes quality control and validation of the field and library spectra against each other (Section 6.7.1.1).
2. Pre-processing of reflective and thermal airborne imagery according to DLR's standardised workflow (Section 6.3) including in-flight calibration, atmospheric correction, geocoding, data quality control and validation against various reference data sources (DLR standard calibration targets, field reference measurements from point 1, JHU library spectra (Salisbury *et al.*, 1991)) (Section 6.7.1.2). Inclusion of the thermal in-flight calibration module and updating ATCOR's atmospheric look-up-table (based on MODTRAN simulations and collected climate data) to attain a compliant quality level of the pre-processing also for the thermal imagery.
3. Normative modelling of mineral abundances based on elemental compositions determined by laboratory wet chemistry of the collected field samples (Section 6.4). Plausibility check of the derived mineral abundances against qualitative XRD analysis and quantitative elemental composition and in addition cross-validation between two models normative models, NORMA and SEDNORM (Section 6.7.1.3).
4. Wavelet-based classification of siderite, goethite, kaolinite, illite and montmorillonite using the Wavelet Coherence Mapper (Section 6.5) and class templates from the USGS spectral library. Comparison and cross-validation against Spectral Angle Mapper results.
5. Wavelet-based quantification of goethite, Fe<sub>2</sub>O<sub>3</sub> and kaolinite using the Wavelet Transform Feature Analysis (Section 6.6.1). The regression models are build on the sample's laboratory spectra and normative mineral abundances (from points 1 and 3). Validation of the models against a test set (statistical cross-validation), field spectrometric measurements (from point 1) and USGS library spectra (Clark *et al.*, 1993) (Section 6.7.1.4).
6. Band-ratio-based quantification of quartz and kaolinite in the thermal domain imagery (Section 6.6.2). Establishment of the regression models on image spectra and corresponding sample's normative mineral abundances (from points 2 and 3). Validation of the models against a test set (statistical cross-validation), field spectrometric measurements (from point 1) and ASU library spectra (Christensen *et al.*, 2000) (Section 6.7.1.5).
7. Evaluation of the individual results, their inherent accuracies and plausibility of the derived abundances and spatial patterns when combining results from reflective and thermal domain, classification and quantification (Section 6.7.2). Identification of the main sources of uncertainty in the workflow (Section 6.7.1.6).

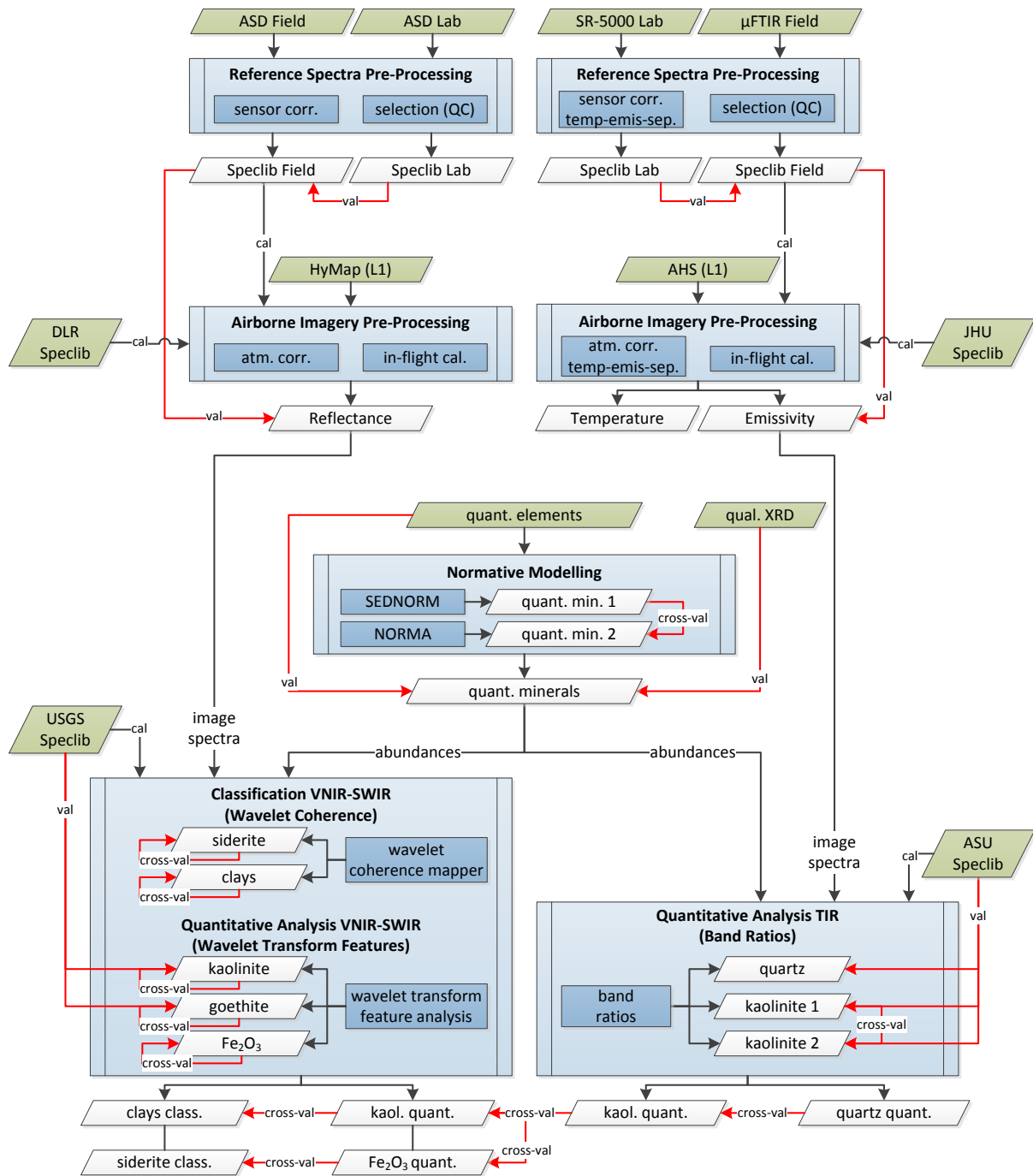


Figure 6.2.: Concept of the data processing workflow showing data layers (parallelograms) and linkage of the processing steps (light blue boxes) to form higher-level products. The methods used for the processing are depicted as enclosed blue boxes. Original input data layers given as green, intermediate outputs and final products as white parallelograms. The validation steps are depicted as red arrows.

## 6.2. Data Sets and Measurement Protocols

This study uses airborne hyperspectral reflective and multispectral thermal imagery plus ground and laboratory reference data collected during two field campaigns. The first campaign was conducted in August 2010 parallel to a HyMap (VNIR-SWIR) flight. The collected data includes:

- 7 daytime hyperspectral flight lines (VNIR-SWIR) covering approximately 300 km<sup>2</sup> of the Sokolov demonstration site at 4 m ground sampling distance,
- 37 spectroscopic (VNIR-SWIR) field reference measurements for calibration and validation using a portable ASD FieldSpec Pro 2 spectrometer,
- 50 mineral soil samples sent for laboratory geochemical analyses and also spectrally measured with an ASD FieldSpec Pro 2 under laboratory conditions.

The second campaign was conducted in July 2011 parallel to an AHS airborne hyperspectral (VNIR-SWIR-MIR-TIR) flight. The data collected includes:

- 10 nighttime multispectral flight lines (TIR) covering approximately 500 km<sup>2</sup> of the Sokolov demonstration site at 5 m ground sampling distance,
- 44 spectroscopic (TIR) field reference measurements for calibration and validation using a portable D&P FT-IR Model 102 spectrometer,
- 30 mineral soil samples sent for laboratory geochemical analyses.

### Airborne Imagery

Hyperspectral reflective imagery of the Sokolov demonstration site were acquired with the HyMap instrument, provided by HyVista Corporation and operated by German Aerospace Center (DLR), on 21st August 2010 around solar noon time during almost cloud free conditions. The sensor covered the wavelength range 450 to 2,500 nm with 128 spectral bands at an average ground spatial resolution of 4 m. On 22nd July 2011 the multispectral thermal imagery were acquired with the AHS instrument, provided and operated by Instituto Nacional de Técnica Aeroespacial (INTA). Cloud free conditions were found during the acquisition time around midnight. The sensor's TIR port covered the wavelength range 8,000 to 13,000 nm with 10 spectral bands at an average ground spatial resolution of 5 m. Detailed specifications of the two sensors and flight parameters can be found in Appendix C - *Sensor Descriptions*.

### Sampling and Geochemical Analysis

During both campaigns mineral soil samples were collected to representatively cover the mineralogy found on the demonstration site. Eighty samples representing the uppermost 1 cm of vegetation-free targets favourably big enough to be identifiable in the airborne imagery that span the major geologic units were selected. These samples underwent chemical and mineralogical laboratory analysis conducted by the ČSN ISO/IEC 17025 accredited Prague-Barrandov laboratory of the Czech Geological Survey. Analysis comprises of whole rock elemental analysis, qualitative X-ray diffraction (XRD) mineralogical identification and complementary analyses (see Appendix B.3 - *Sample Mineralogy*). The 2011 samples could not be XRD analysed.

## Reference Spectra

Parallel to the airborne surveys ground reference spectra were measured with portable instruments. This allowed to establish calibration and validation targets for the airborne imagery as well as reference spectral libraries for the materials and mixtures found at the Sokolov site. The community accepted best-practise guidelines of [NERC-FSF-Homepage \(2013\)](#), [Hook & Kahle \(1996\)](#) and [Salisbury \(1998\)](#) were adhered to.

ASD FieldSpec Pro 2 spectrometers were used to acquire point reflectance spectra of spatially extend quasi-homogenous targets for calibration and validation purposes during the HyMap flight. Field reflectance spectra of the undisturbed sites were taken at nadir view by measuring 5 to 10 spots for each target with a co-adding of 40 individual measurements and a footprint of 20 cm diameter.

Laboratory reflectance measurements were conducted under the controlled environment of the DLR spectroscopic laboratory equipped with absorptive materials shielding the samples. The air-dried sample material was evenly distributed on a tray with no prior treatment to resemble the field situation as closely as possible. For each sample five successive measurements were done at nadir view for each tray rotation of 0°, 90°, 180° and 270° with a co-adding of 80 individual measurements of a footprint of 10 cm diameter.

Sensor calibration on a white reference panel, dark current and reference panel measurements preceded each target/sample measurement block. Another reference panel measurement followed. In case an illumination change, sensor drift, sample inhomogeneity or strong BRDF was detected the measurement was completely redone.

Thermal radiance calibration and validation spectra were collected with a D&P FT-IR Model 102 spectrometer over the course of 3 days during the AHS campaign. The instrument was calibrated with four blackbody measurements bracketing the target's kinetic temperature. For each target 3 radiance measurements at differing positions were conducted with a footprint of 10 cm diameter, at nadir view and a co-adding of 20 individual measurements. Simultaneously the soil surface temperature was measured with a contact thermometer and a handheld pyrometer. The 3 target measurements were bracketed by two gold reference panel measurements of the down-welling radiance. The gold panel's temperature was measured using a contact thermometer.

## Auxiliary Data

In order to support subsequent processing and validation tasks numerous additional information were collected. For this study the following are of relevance:

- climate data from national and on-site weather stations, AERONET network and MODIS water vapour products used for the atmospheric correction,
- water temperature logs for validation of the temperature-emissivity-separation,
- thermal radiance spectra of 79 soil samples measured by the Remote Sensing and GIS Laboratory, Tel-Aviv University using a CI Systems SR-5000 spectrometer used for validation of the D&P FT-IR and AHS derived emissivities.

### 6.3. Spectroscopic Data Pre-Processing

Pre-processing of the spectral measurements and imagery corrects for systematic and random perturbations in the data. A calibrated data set of physical meaning is established. This allows the data to be compared to temporally and spatially offset measurements, e.g. from another site, library spectra or from a previous campaign. Coupled with a validation concept the pre-processed data can be assured reproducible and only then can be used for quantitative analysis.

This study defines spectral data pre-processing as the transformations and corrections that are applied to convert the raw sensor data (L0 product) over the system-corrected at-sensor radiance (L1 product) to the acquisition condition independent at-ground reflectance/emissivity (L2 product). Obviously the used models of sensor, atmosphere, illumination and viewing geometry all make model assumptions or use vicarious gross corrections that always can be only approximations of the real-world processes. Here the validation comes into play (see Section 6.7 - *Evaluation of Processing Workflow & Results*).

#### 6.3.1. Field & Laboratory Measurement Corrections

##### Reflective Domain

Pre-processing of the laboratory and field reference spectra followed the recommendations of EUFAR (Holzwarth *et al.*, 2010) and NERC (NERC-FSF-Homepage, 2013) using a toolbox developed by the Team Applied Spectroscopy (German Remote Sensing Data Center, DLR). The ASD FieldSpec Pro instrument handles L0 to L2 reflectance conversion internally by repeated detector gain optimisation, calibration onto a white reference panel, dark current measurements and internal co-adding. Additional corrections were necessary to handle reflectance value inconsistencies at the two detector transitions (jump correction) and to account for the non-whitebody behaviour of the Spectralon<sup>®</sup> reference panel (calibration on a traceable standard). Finally mean reflectance spectra for each measurement target were calculated by averaging the individual co-added and corrected measurements. Table 6.1 details the parameter values used.

Table 6.1.: Summary of parameters used for reference spectra pre-processing in the reflective domain.

parameter	value	note
<b>detector jump correction</b>		
between detectors	1-2 (#651 / #652) 2-3 (#1481 / #1482)	both jumps for laboratory measurements only first jump for field measurements
reference detector	2	provides reflectance master value
method	constant additive offset	absorption depths not altered
<b>reference panel correction</b>		
reference	DLR internal standard	traceable to PTB
method	bandwise multiplicative gain	reciprocal of reference panel reflectivity



## Thermal Domain

Pre-processing of the D&P FT-IR measurements comprises of hot/cold blackbody instrument calibration, downwelling radiation correction and temperature-emissivity-separation. For latter the simultaneously acquired temperatures using contact thermometer and pyrometer proved too unreliable for an absolute calculation. Instead a Planck curve fitting approach as proposed by [Hook & Kahle \(1996\)](#) and [Salisbury \(1998\)](#) was implemented based on a Pascal code (pers. communication Dr. Peter Gege, DLR) using the simplex algorithm of [Caceci & Cacheris \(1984\)](#). It counters noise by allowing a certain “sinking into” the measured radiance curve and limiting the fit to certain wavelength intervals in which common natural materials approach maximum emissivity. In a strict sense, the derived emissivities are relative emissivities since the true kinetic temperature and maximum emissivity are unknown. For each measurement target the mean emissivity spectrum was calculated. Table 6.2 lists the parameter values used.

Table 6.2.: Summary of parameters used for reference spectra pre-processing in the thermal domain.

parameter	value	note
<b>instrument calibration</b>		
blackbody measurements	2 warm, 2 cold	bracketing target surface temperature
calibration gain & offset	bandwise mean	linear regression of blackbody measurements versus Planck radiance
<b>Planck curve fitting</b>		
fit intervals	7,700-7,900 nm 11,800-14,000 nm 14,500-15,000 nm	exclude noisy spectral regions (simultaneous fit over all intervals)
max. emissivity	1.0	maximum emissivity allowed in fit interval

### 6.3.2. Airborne Imagery Corrections

The individual steps of the imagery pre-processing are detailed in Figure 6.3. For the reflective domain an accepted standard procedure exists. Starting from L1 imagery, the in-flight calibration checks the sensor’s laboratory calibration and makes corrections if necessary. It is followed by the correction of atmospheric influences and of illumination conditions. Remaining artefacts are removed by filtering and band selection. Eventually geometric distortions introduced by movements of the airborne platform and ground topography are corrected for. For the thermal domain the pre-processing procedure is similar. The in-flight calibration step however was missing. The generic method developed in Chapter 4 - *Thermal Imagery Pre-Processing* was applied. Due to the cool and humid climatic conditions, the nighttime acquisition and the noise characteristics of the AHS sensor the atmospheric model needed revision and a specialised destriping method needed to be developed as outlined below.

## Reflective Domain

The daytime reflective hyperspectral data set was pre-processed following the standardised and ISO 9001:2008 certified DLR OpAiRS processing chain. It consists of several (semi-) automated modules for system correction, in-flight calibration, atmospheric correction, orthorectification

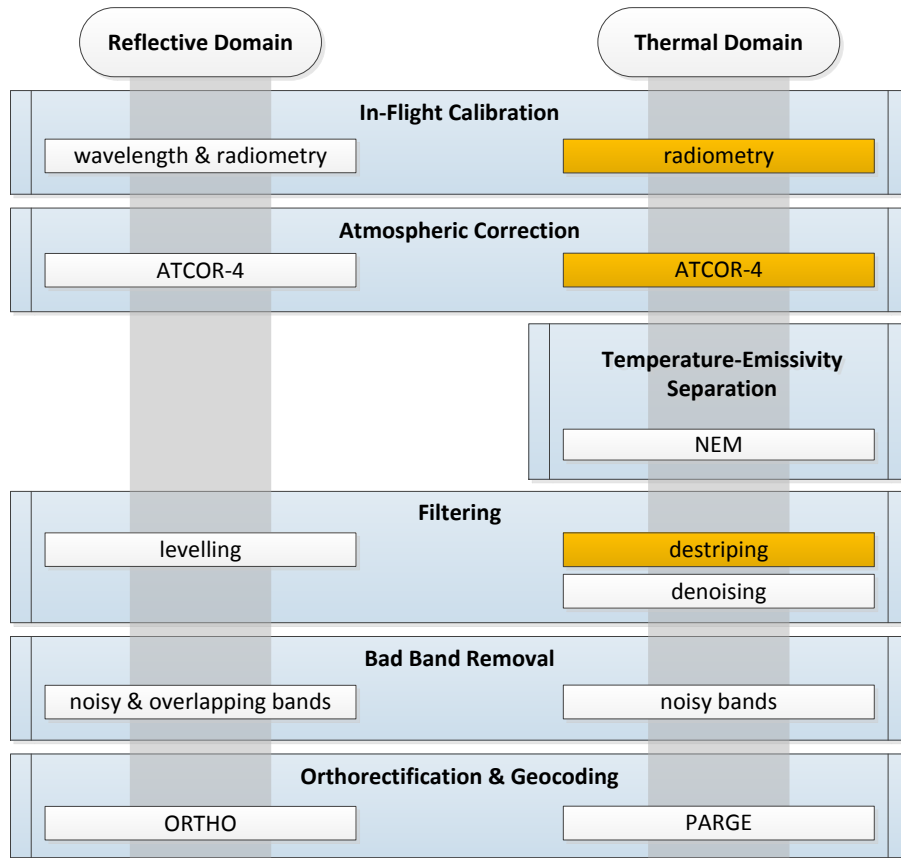


Figure 6.3.: Comparison of reflective and thermal imagery pre-processing steps. Enclosed boxes depict used methods/software. Revised existing or devised new methods/software marked orange.

and geocoding as documented in [Habermeier et al. \(2003, 2012\)](#). Additionally, a levelling filter was applied ([Rogge et al., 2012](#)). Pre-processing started at L1 system corrected at-sensor radiance as the HyMap sensor provider HyVista undertook the laboratory calibration. The relevant parameters are listed in [Table 6.3](#) at the end of this section.

**In-Flight Calibration** The in-flight calibration is an additional fine adjustment of the laboratory calibration checking for wavelength shifts and sensor gain deviations. Spectral shifts up to 1.1 nm were found for the Sokolov campaign which magnitude is in accordance with earlier HyMap experiences and compares well to HyMap’s spectral bandwidth of 15-20 nm FWHM. The radiometric calibration was based on five well-characterized reference targets that are measured and controlled by DLR every year yielding a sensor gain and offset correction function (see [Figure 6.4](#)). The offset values were very small and are known to be stable for HyMap. Therefore the offset correction was neglected. Usage of the Sokolov calibration field measurements had little effect on the gain and offset. The Sokolov field measurements were used as validation targets since the controlled standard reference target measurements were regarded more reliable.

**Atmospheric Correction** The correction for atmospheric effects and illumination conditions was performed with ATCOR-4 ([Richter, 2010b](#)) using the standard MODTRAN mid-latitude summer based look-up-table (LUT). Visibility and per-pixel water vapour content were determined automatically from the daytime hyperspectral imagery. Flat terrain was assumed since the terrain elevation and aspect correction did not show benefits for the hilly Sokolov site.

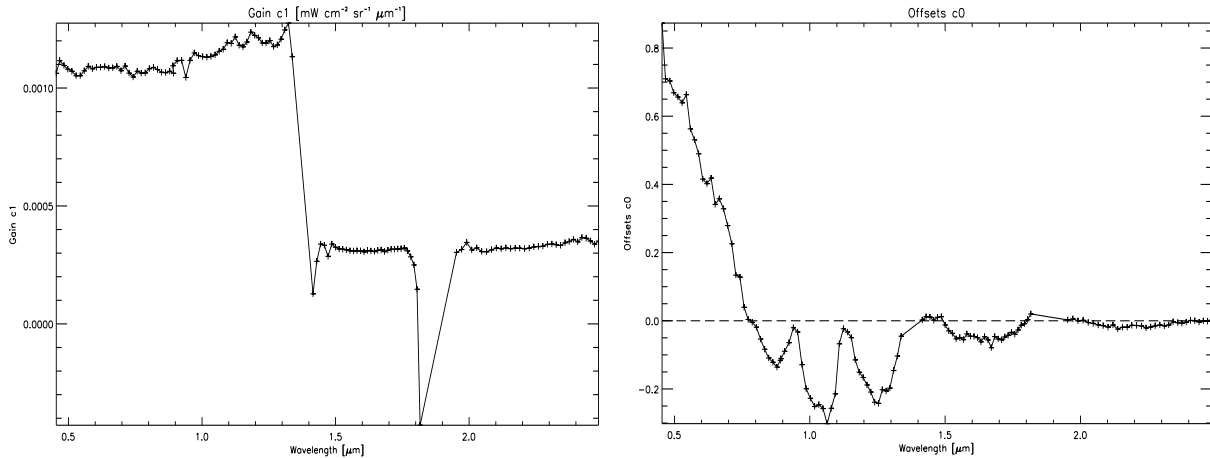


Figure 6.4.: Sensor gain [ $mW/(cm^2 sr \mu m DN)$ ] (left) and offset [ $mW/(cm^2 sr \mu m)$ ] (right) computed from 5 calibration targets. Note that ports 1 & 2 have different scaling than 3 & 4. The laboratory calibration standard coefficients are 0.00100/0.00025 for gain and 0 for offset.

**Filtering** The north-east/south-west alignment of the flight lines pronounced line-to-line radiometric inconsistencies due to cover dependent BRDF and strong cross-track illumination effects. Rogge *et al.* (2012) presented an empirical two-fold scattering correction and flight-line levelling approach which was applied to the HyMap imagery. It does not rely on a master line but treats all flight lines equally. A seamless mosaic resulted. No further filtering was applied to preserve as much of the original spectral characteristics as possible for the subsequent quantitative analysis. The wavelet transform's ability to separate noise into the smallest scale component together with the overall high signal-to-noise ratio allowed to omit noise reduction filtering and spectral polishing.

**Bad Band Removal** Continuous wavelet transforms are susceptible to border effects when the cone of influence (COI, see Section 2.4.3 - *Cone Of Influence*) exceeds the spectral coverage margins, i.e. overlaps with the start or the end of the spectrum. With increasing scale the COI grows and thus the transform values close to the coverage margins depend more and more on the padded data. The chosen mirror padding translated spectral distortions of noisy HyMap bands onto the wavelet transform. This impeded the application of spectral similarity measures since the transforms of otherwise similar spectra showed very different shape within the coverage margin induced COI. This effect could be greatly reduced by removing the first two and last two HyMap bands from the further processing, thus allowing to interrogate the spectrum also with higher scales close to the spectral coverage margins.

Two bands occurred with reversed ordering of the centre wavelength at HyMap's 890 nm VIS/NIR detector transition. The last VIS detector band was excluded to resolve this.

**Orthorectification and Geocoding** The airborne images were geo-referenced using the parametric orthorectification/geocoding software ORTHO (Müller *et al.*, 2005). The image pixels were projected based on a SRTM digital elevation model of the demonstration site together with position and attitude information recorded by the airborne platform. Nearest neighbour resampling was chosen to not alter the spectral signatures.

Table 6.3.: Summary of parameters used for image data pre-processing in the reflective domain.

parameter	value	note
<b>atmospheric correction</b>		
visibility	80 km	image derived
aerosol type	rural	selected preset
water vapour	variable, 2.2 cm	image derived
atmosphere	mid-latitude summer	LUT from MODTRAN standard atmosphere
ground elevation	flat, 0.5 km	Sokolov mean, little terrain effect
sensor altitude	2.4 km	as reported in flight reports
<b>in-flight calibration</b>		
calibration types	wavelength & radiometry	standard approach for hyperspectral sensors
target count	5, near-nadir	standard targets
target type	asphalt & concrete	spectrally flat, well-characterised
<b>levelling</b>		
dark mask	< 5 % mean reflectance	empirical, derived from image histograms, exclusion of dark pixels from levelling stats
<b>bad band removal</b>		
excluded bands	#1, #2, #124, #125 #30	noisy detector overlap

### Thermal Domain

Pre-processing of the nighttime thermal multispectral data set generally followed the same steps as for the reflective case (Fernández-Renau *et al.*, 2005). Peculiarities of the TIR data and AHS sensor necessitated additionally a temperature-emissivity-separation, destriping and denoising of the data. For the in-flight calibration the generic method presented in Chapter 4 - *Thermal Imagery Pre-Processing* was applied. A levelling of the emissivity product was not necessary. The Spanish sensor operator INTA accomplished the laboratory calibration of the AHS sensor and provided L1 system corrected at-sensor radiance. The relevant parameters are listed in Table 6.4 at the end of this section.

**Atmospheric Correction** The atmospheric parameters necessary for atmospheric correction could not be directly estimated from the imagery because the typically used reflective bands were not available due to the night time acquisition and the broad thermal bands of AHS prevented an approach like AAC. The two most important parameters governing the thermal radiative transfer are atmospheric water vapour content and the temperature profile (Liang, 2004). On-site weather station records, radiosonde data, AERONET network readings and the MODIS water vapour product have been used to fix both. Due to the night flight's particular cool and humid conditions, two issues arose when performing the thermal atmospheric correction with ATCOR-4 using the standard MODTRAN mid-latitude summer atmosphere model and not including a thermal in-flight calibration:

- The retrieved surface temperatures generally appeared too low. Measured in-situ air temperatures of 8-10 °C contradicted with brightness temperatures of 4-6 °C for dense Norway

Spruce vegetation and temperatures of freezing point or less for meadows. Brightness temperatures close to the ambient air temperature would be expected for this near black-body targets. Variation of the water vapour content had little effect. Generally surface temperatures decreased with increased water vapour modelled.

- An evident decrease of at-sensor radiance towards the short and long wavelength margins of the TIR atmospheric window occurred. This drop exceeded well beyond the known target spectral emissivity response (water, dense vegetation) and worsened the more water vapour was modelled (see Figure 6.5).

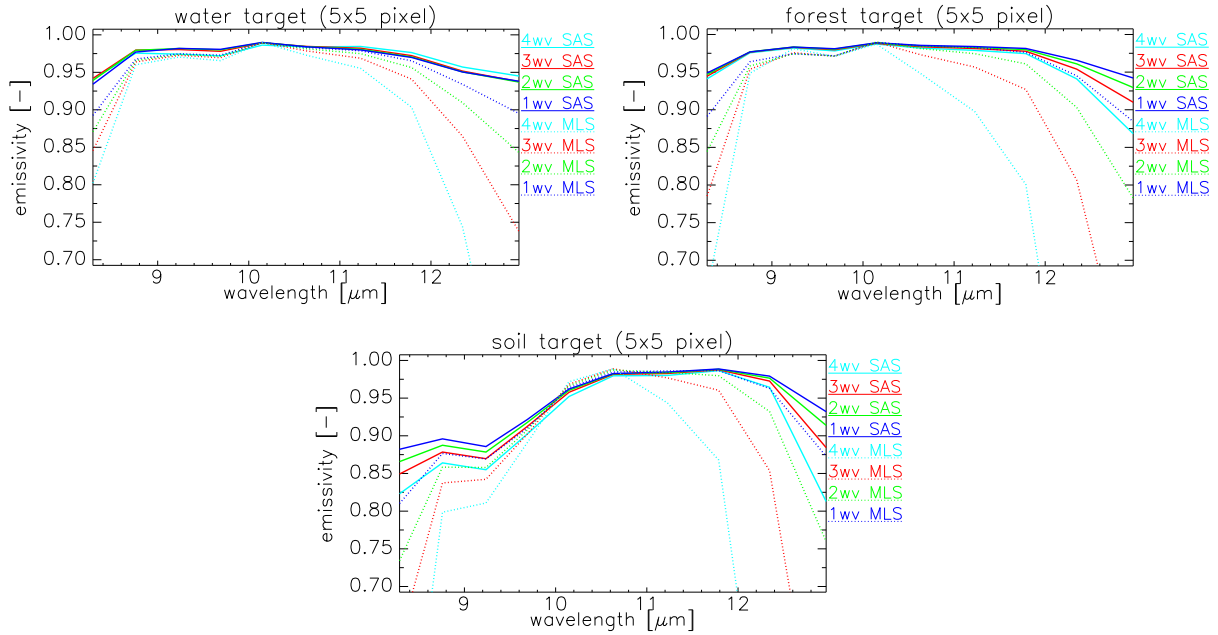


Figure 6.5.: ATCOR NEM mean emissivity spectra of water, forest and soil targets comparing different water vapour settings [cm] (colors) of the mid-latitude summer (MLS, dotted) and subarctic summer (SAS, solid) atmospheres.

MODTRAN simulations (see Section 4.2.1 - *Atmospheric Modelling*) revealed a reversal of the atmospheric influence on the total at-sensor signal compared to the mid-latitude summer atmosphere when choosing the cooler subarctic summer model. Latter temperature profile starts with 14 °C (instead of 21 °C) which scales to 11 °C at the Sokolov average elevation. This compares well to in-situ measurements. A higher atmospheric absorption and lower atmospheric self emission results. Instead of  $L_{as}^e > L_{ag}^e$  an inverse relation  $L_{as}^e < L_{ag}^e$  was observed for warm targets. Increasing water vapour content had little effect on at-ground radiance of cool targets (e.g. vegetation) but a reverse effect for warm targets (e.g. water). Due to the increasing influence of water vapour absorptions towards the margins of the TIR window this effect compensated the observed emissivity drop (see Figure 6.5). Mean temperature increased by 2 K for the 2.9 cm water vapour case but is cover dependent. Warm water surfaces increased the least (1.0-1.5 K) and cool surfaces such as forests, meadows and soils the most (2.0-3.5 K). In summary it could be demonstrated that under cool and humid atmospheric conditions an updated atmospheric model is necessitated to adequately model the TIR radiative transfer. An updated LUT was generated for ATCOR-4 using the MODTRAN predefined subarctic summer atmosphere as a best match.

**In-Flight Calibration** The updated LUT could compensate the temperature discrepancies and the severe drop of radiance/emissivity towards the TIR window margins. However latter effect could not be fully eliminated. This might be due to further deviations of the temperature and water content of modelled to real atmosphere or due to sensor calibration. Instead of fine tuning the atmospheric parameters leading to a highly customised LUT it is proposed to use a realistic set of atmospheric parameters backed by auxiliary climate measurements and apply the in-flight calibration presented in Section 4.2.2 - *In-Flight Calibration Module*. In doing so 74 extensive near-nadir water and dense spruce forest targets were identified within the 10 flight lines. For both surface types the emissivity spectra are known, nearly featureless and approaching that of a blackbody. Hence introduction of artifacts due to spectral shape, wrong temperature estimate and spectral mixing are minimised. The target's thermal properties led to a stretching of the regression line between the warm water and cool vegetation thus increasing stability and limiting extrapolation errors. The bandwise scatter plots showed excellent linear behaviour manifesting in high correlation coefficients  $> 0.999$  for most bands and all flight lines. The offset was forced to zero to prevent mathematically optimised but physically impossible values of the band regressions. This resulted in a stable sensor gain correction function of a mean  $< 1.01$  reaching a maximum of 1.02 for the sensor bands towards the TIR window margins. Figure 6.6 displays the improved emissivity spectra after applying the in-flight calibration to the L1 radiance data. The emissivity drop could be successfully corrected for.

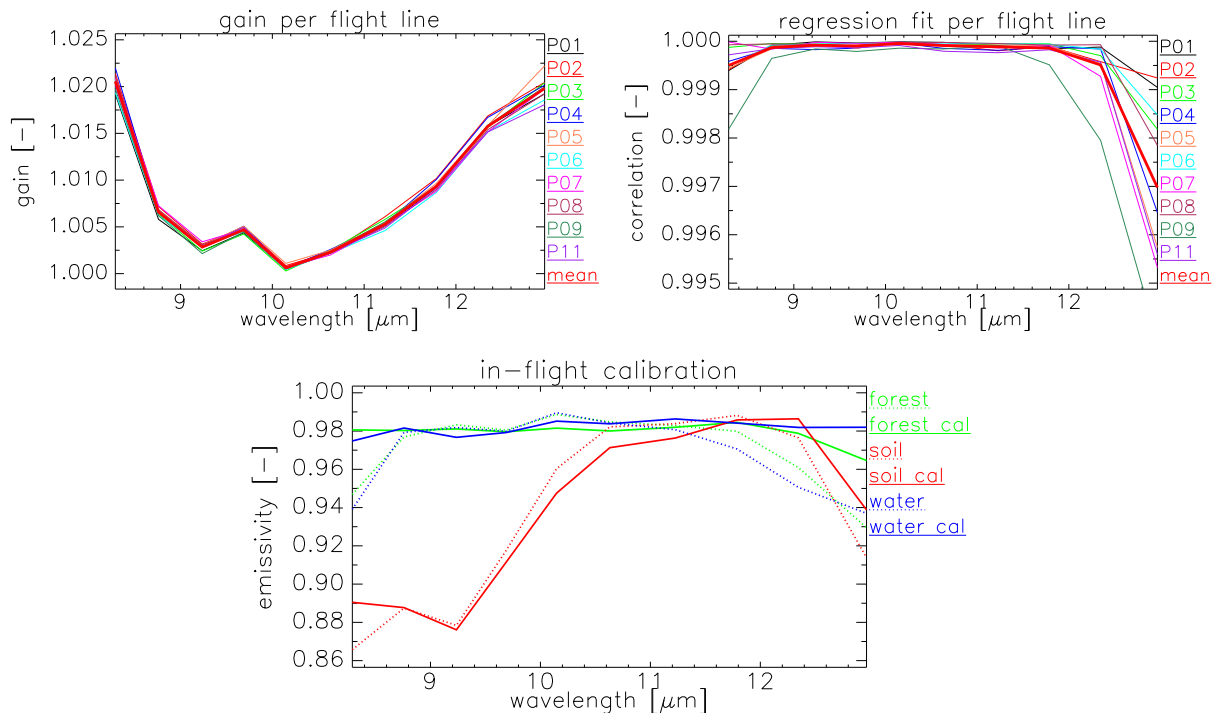


Figure 6.6.: Top shows sensor gain correction and goodness-of-fit of linear regression model computed for each TIR flight line. Bottom shows influence of the calibration on ATCOR NEM mean emissivity spectra of water, forest and soil targets.

**Temperature-Emissivity-Separation (TES)** The ATCOR-4 normalised emissivity method (NEM) implementation was chosen because it handles retrieval of multispectral emissivity somewhat similar to the Planck curve fitting approach chosen for the thermal field measurements. For each image pixel NEM selects the sensor band exhibiting the highest brightness temper-

```

1 // ITERATIVEIMAGEDESTRIPING() //
2 for each iteration
3   for each band
4     for each scan line (= row in the image)
5       calculate median scan line brightness (= emissivity)
6       calculate mean brightness of current & neighboring scan lines
7       for each pixel in current scan line
8         normalise by dividing its value by median scan line brightness
9         multiply back mean neighbourhood brightness
10      next scan line in band
11    next band
12  next iteration

```

Listing 6.1: Algorithm for the iterative image destriping in pseudo-code.

ature as reference channel and assigns a fixed expert-defined emissivity  $\epsilon_{max}$ . This resolves the underdetermination of the set of thermal band equations and emissivities for all other bands can be calculated according to Planck's law. Relative emissivity spectra result.  $\epsilon_{max} = 0.985$  was chosen because it suits water, vegetation and the longer wavelength behaviour of mineral assemblages expected for Sokolov containing high fractions of quartz, clay and carbonate (see Sections 2.1.1 - *Geologic Setting* and 2.3.5 - *Spectral Characteristics of relevant Minerals*). NEM has the drawback of translating the noise in the reference channel to all other bands. Since the maximum brightness temperature typically appears in bands more affected by atmospheric water absorption and hence of reduced signal-to-noise ratio this imposes a problem.

**Filtering** The L2 emissivity product exhibited severe cross-track striping in combination with random noise. Several sources contribute: TES removes the temperature governed signal continuum subsequently also noise becomes more apparent. The night time acquisition results in a lower signal-to-noise ratio due to lower surface temperatures and a lower thermal contrast at supposedly constant sensor noise level. For a line scanning instrument such as AHS random noise is likely to be introduced by the electronics whereas systematic noise (striping) probably originates from errors/instabilities in the gain and offset calibration parameters which are recalculated for each scan line from the onboard cold and hot calibration blackbodies.

The AHS noise is spectrally and spatially correlated making it difficult to correct for. The striping is only partially correlated over the sensor bands and thus the striping is not simply linked to errors in the calibration blackbody temperature. The sensor operator assessed the noise characteristics by numerous means however no satisfactory correction scheme was found so far (pers. communication Dr. Eduardo de Miguel Llanes, INTA). Jiménez-Munoz *et al.* (2011) applied maximum/minimum noise fraction (MNF) transform. Noise was significantly reduced feeding only the first two or three MNF components into the inverse transform but residual striping was apparent. The authors suggested to use the transform on the L1 data also noting the use of that few MNF components is critical.

For the work at hand a straightforward approach operating on the L2 emissivity maps was preferred resulting in a three-fold filtering process to suppress striping and random noise. First a destriping is applied followed by two diffusion filtering steps.

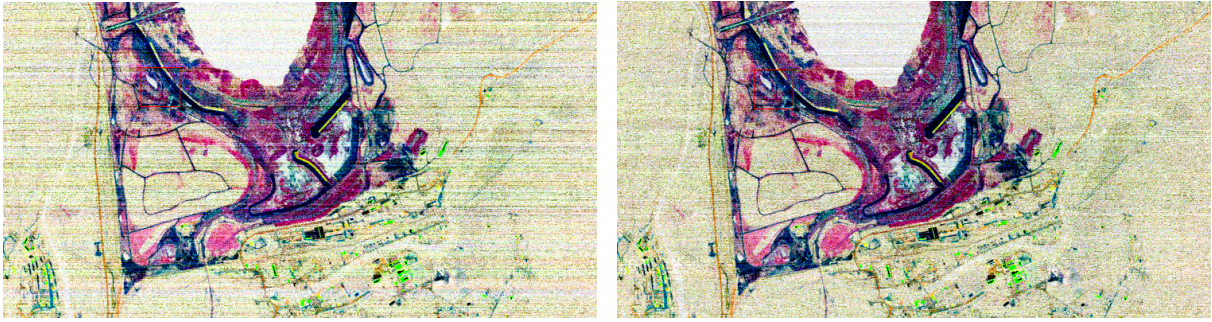


Figure 6.7.: Cross-track striping as it appears in the AHS nighttime emissivity product (left) and destriped product (right) using 200 iterations and 4 neighbour scan lines. False color composite from AHS bands #72, #74 & #76 (RGB) of flight line 11.

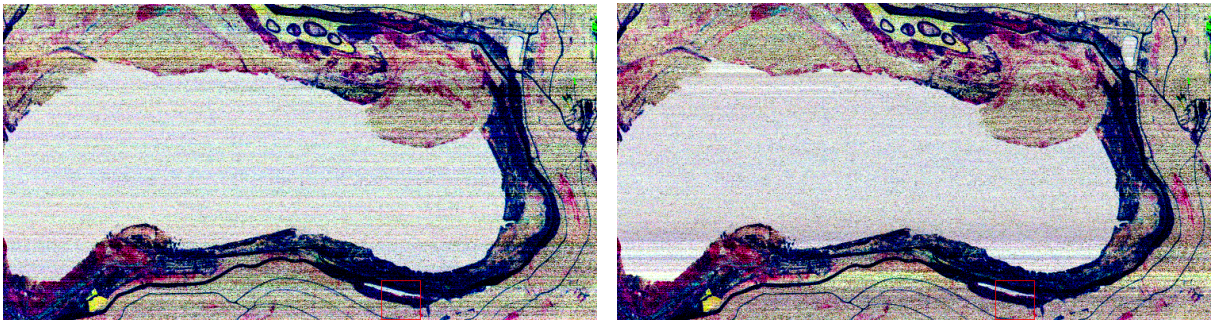


Figure 6.8.: Stripe artifacts in destriping result of Sokolov night line 3 introduced by dark shoreline materials aligned cross-track parallel. Left original image vs. destriped image on the right.

**DESTIPING** A basic destriping algorithm producing acceptable results for the Sokolov scenes was realised (see Figure 6.7). The filter assumes a spectrally and spatially independent striping and that true brightness variations on the ground can be approximated by the mean brightness of the local neighbourhood. It was implemented as iterative filtering algorithm given by Listing 6.1. Using the median brightness for normalisation of the scan line makes the filter less susceptible to pixel outliers of very high/low brightness. The scan line brightness variation distribution is assumed zero-centred. The filter does not alter the brightness variation within a scan line. Only its average brightness is adjusted to that of the neighboring scan lines. This changes the shape of the individual pixel spectra as every band is treated independently. Globally the mean emissivity spectrum is not changed. An issue arose for cross-track parallel contrasting structures that made up a considerable part of the scan line, e.g. roads or housing areas. Such structures were treated as false positive stripes and were therefore brightened/darkened while introducing artifacts on the remaining scan line parts. For the Sokolov data such artefacts were found in three flight lines attributing to a motorway, a road and a shoreline (see Figure 6.8 for latter). Application of this destriping filter to not striped imagery introduced many striping artefacts. Also clouds led to problems. In summary the destriping approach is geared towards cloud-free rural scenes exhibiting a severe cross-track striping.

**DENOISING** In order to remove random noise and to reduce the apparent spectral intra-class variability a set of diffusion filters was applied after removing the striping. The filters are an improved version of Rogge & Rivard (2010), (pers. communication Dr. Derek Rogge, DLR), which sustain spatial features. Pixels within the filter window are assigned a



weight depending on the spectral similarity to the current pixel. A zero weight is assigned to pixels that are deemed too different as defined by a similarity threshold. For each iteration the threshold is lowered in an exponential fashion. First random noise is removed by applying a root-mean-square (RMS) diffusion filter. Neighbouring pixels with similar absolute spectral shape are smoothed together. In a second step brightness differences are removed by applying a spectral-angle (SA) diffusion filter. Neighbouring pixels with similar spectral shape but overall brightness difference are smoothed together. The filters reduced random noise and intra-class variability dramatically while keeping the inter-class variability (see Figure 6.9). Spectrally similar classes, e.g. water and vegetation, were now clearly separated from each other. The filters kept the spatial pattern to a maximum while pulling out much more spectral detail which was drowned in the noise beforehand. The filters however lead to a smoothing of gradual spectral changes and loss of spectral variations below the set threshold and therefore impact the results of quantitative studies. The thresholds were empirically adjusted defining the tradeoff between noise reduction and spectral smoothing acceptable.

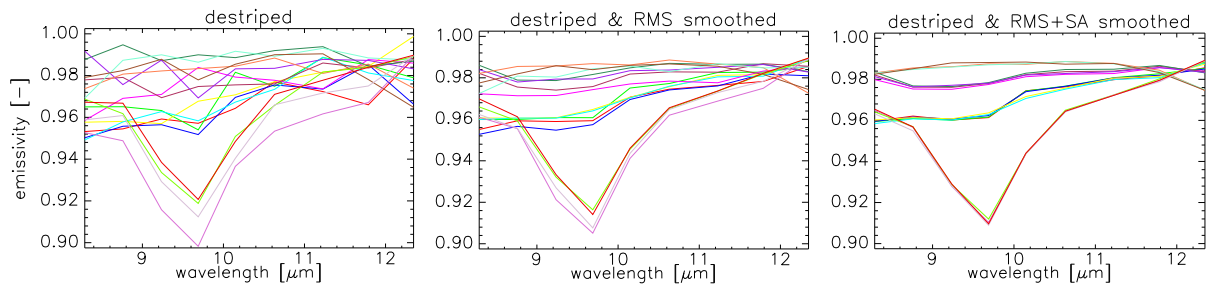


Figure 6.9.: Pixel emissivity spectra of destriped image (left), same after additional RMS diffusion filtering (middle) and after additional SA diffusion filtering (right). Noisy AHS band #80 cropped.

**Bad Band Removal** An inspection of the operator provided signal-to-noise ratio (SNR) and noise-equivalent delta temperature ( $NE\Delta T$ ) plots for the AHS sensor showed a degraded performance in bands #71 and #77 to #80. Especially band #80 stood out negatively with  $SNR < 8$  and  $NE\Delta T > 0.75$ . Visual inspection of the L1 at-sensor radiance of latter band showed virtually no land surface features besides water and metal roofs embedded in heavy noise. The in-flight calibration and filtering could not satisfactorily stabilise the erratic radiometric behaviour. Eventually band #80 was excluded from further processing as it did not contain meaningful information.

**Orthorectification and Geocoding** INTA provided geo-referencing information calculated with the PARGE parametric orthorectification/geocoding software (Schläpfer & Richter, 2002) based on the same digital elevation model as was used for the reflective data. Nearest neighbour resampling was used to preserve the original spectral signatures.

Table 6.4.: Summary of parameters used for image data pre-processing in the thermal domain.

parameter	value	note
<b>atmospheric correction</b>		
visibility	25 km	not relevant for TIR
aerosol type	rural	not relevant for TIR
water vapour	2.9 cm	scales to 1.6 cm for Sokolov mean elevation
atmosphere	subarctic summer	LUT from MODTRAN standard atmosphere
ground elevation	flat, 0.5 km	Sokolov mean, little terrain effect
sensor altitude	2.4 km	as reported in flight reports
<b>in-flight calibration</b>		
calibration types	radiometry	multispectral sensor
target count	74, near-nadir	ca. 7 per flight line
target type	water & dense forest	graybody like, spectrally flat, known emissivity
<b>temperature-emissivity-separation</b>		
temp. channel	10,600 nm	band #66 least atmospheric affected
TES method	NEM, $\epsilon_{max} = 0.985$	brightest band used as reference band
<b>iterative destriping</b>		
iterations	200	empirical, allows diffusion over whole image
neighbour count	4	empirical compromise
<b>diffusion filtering</b>		
RMS threshold	1.5 mean scene RMS	empirical compromise
RMS iterations	5	empirical, exponential decreasing effect
SA threshold	1.5 mean scene SA	empirical compromise
SA iterations	5	empirical, exponential decreasing effect
<b>bad band removal</b>		
excluded bands	#80	noisy

## 6.4. Normative Modelling

Eighty selected mineral soil samples of the 2010 and 2011 campaigns underwent chemical and mineralogical laboratory analysis. The analysis was conducted by the ČSN ISO/IEC 17025 accredited Prague-Barrandov laboratory of the Czech Geological Survey. It comprises of whole rock elemental analysis, qualitative X-ray diffraction (XRD) mineralogical identification and several complementary analyses as summarised by Table B.2. For the thirty 2011 samples no XRD analysis was conducted.

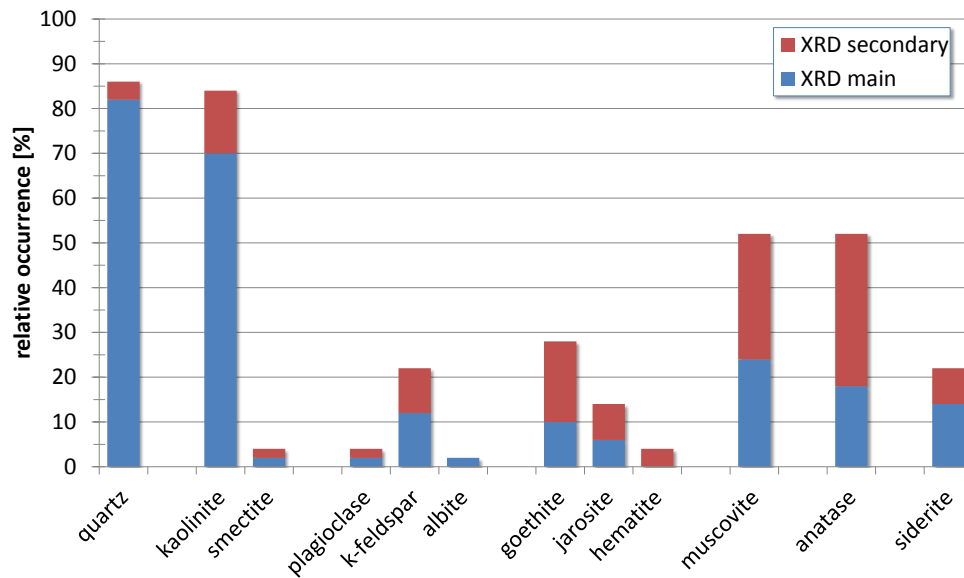


Figure 6.10.: Aggregated mineral occurrences in the fifty Sokolov 2010 samples according to XRD analysis.

Qualitative XRD analysis reports minerals into three classes - main, secondary and accessory - based on XRD peak intensities and matching quality to library patterns. Minerals identified into former two classes can be assumed certain to occur in the sample's mineral assemblage whereas the accessory class may contain several false positives. The qualitative XRD analysis data hence only allows for a rough estimation of major minerals.

Figure 6.10 shows the relative mineral distribution for the fifty Sokolov 2010 samples based on the occurrence count reported in the XRD main and secondary classes. As expected quartz and kaolinite are the two major minerals in the Sokolov mineral soils. A second clay mineral, smectite, is reported for only two samples. The feldspar family is represented by low-temperature variants rich in sodium (albite) and primarily potassium (orthoclase or microcline). The AMD minerals geared sample selection lead to a high number of Fe-rich minerals being reported. Muscovite is reported for 26 of the samples whereas no biotite was found. Interestingly anatase, a titan dioxide, is present in about 50% of the samples. The fast neutralising iron carbonate siderite is present in 20%.

Since this study aims at quantitative abundance mapping of selected AMD relevant minerals, it is a necessity to derive quantitative information from the laboratory references in order to calibrate and validate the mapping algorithms. To do so, the XRD intensity plots were examined but found to have been normalised, i.e. the tallest peak rescaled to 100, and thus not usable for inter-sample quantitative abundance estimation based on apparent peak areas. However, relative

abundances of minerals within one sample might be derived from ratios of the respective XRD intensity peaks. Due to inconsistencies in the XRD reports and the complexity of possible peak interferences this approach was dropped. Alternatively, quantitative information was given by the whole rock analysis by means of total elemental concentrations given as oxides.

However, occurrence and location of spectral features in the SWIR and TIR originate from molecular vibrational processes which are modified by lattice structures and atomic bonding of the respective minerals. Therefore it is not sufficient to correlate spectral features to pure chemistry but the breakdown into minerals is necessary (Salisbury & Daria, 1992). As an example, the amount of total silica must be corrected for the fraction bound by clays and feldspars in order to correlate the remaining free silica, supposedly forming quartz, with the distinct emissivity feature seen for quartz in the TIR. The same applies to alumina of which only a fraction is used to form kaolinite with its characteristic SWIR absorption doublet. It is hence necessary to resolve the given total elemental concentrations into mineral abundances.

Mineralogical normative calculations were seen a promising approach. Hereby a theoretical quantitative mineralogy is calculated for a rock/soil under the assumption of certain minerals present and idealised formation processes based on given element concentrations. The following section reviews popular normative approaches and their application to the Sokolov samples.

Several normative back-calculation models were considered to reconstruct the mineralogy from the given total elemental concentrations. All considered models successively distribute the elements to a sequence of minerals, constrained by the depletion of the initial element concentrations. The mineral sequence is based on petrographic relationships observed for the rocks/soils under investigation. Element depletion is governed by simplified a priori stoichiometric mineral compositions valid for the studied mineralogic domain. Both the mineral sequence and stoichiometry particularly influence the model's result.

The CIPW norm (Best, 2002) was developed for intrusive rocks in 1931. It focuses on primary minerals and is widely used in studies of igneous rocks. It does not cover micas and secondary minerals typically found in soils. Cohen & Ward (1991) devised SEDNORM which specifically focuses on clastic sedimentary rocks and coal ash taking the peculiarities of mineral mixtures during sedimentation into account. It includes muscovite mica and secondary minerals, i.e. clays and gibbsite. The UPPSALA model (Holmqvist *et al.*, 2002) was developed for Swedish forest soils in 1992. It includes minerals of relevance in soil science, i.e. the hydrosilicates muscovite and chlorite but does not resolve individual clay minerals. Also it lacks common Fe- and Ca-bearing minerals. The NORMA method (Räisänen *et al.*, 1995) is an advancement of the UPPSALA model with strong parallels to the CIPW norm. It was originally developed for podzolic soils found on glacial tills in Northern Europe and was successfully applied to various soil types on differing parent materials (Salminen *et al.* (2008) and references therein). NORMA focuses on easily weathered minerals, secondary minerals and silicates.

Following above remarks SEDNORM and especially NORMA appeared most suitable for normative calculations in the Sokolov setting. The exposed mineral soils there appeared geochemically roughly similar to podzolic soils characterised by a high content of quartz and K-feldspars, little organic carbon, low soil pH and depleted in nutrients (Scheffer *et al.*, 1984; Kuntze *et al.*,

1988). Both models allow for some flexibility in the mineralogy and available chemical analysis data. Appendix B.2 - *Adaptation of Normative Models* covers both models in detail and gives a reasoning for the adaptations made for the Sokolov data.

Appendix B.3 - *Sample Mineralogy* summarises the whole rock geochemistry, SEDNORM and NORMA calculation results for all samples. A good overall correspondence between XRD analysis and both normative models was observed as confirmed by the validations in Section 6.7.1.3 - *Normative Modelling — Model Cross-Validation & Mineral Validation*.

Backed by the overall small difference between NORMA and SEDNORM for the abundant minerals, the NORMA results were favoured for all further investigations. NORMA allowed the differentiation between ferrous and ferric ion as well as a higher flexibility in adjusting the normative mineral sequence to the geologic setting found at the demonstration site Sokolov.

It must be noted that, despite the agreement with the XRD results, the normative modelling can only be an approximation due to the simplified mineralogic dependencies and species covered. Hence a normative kaolinite abundance may include abundances of similar clay minerals, e.g. illite or montmorillonite. Additional minerals, other than the modelled ones, are likely to occur and to influence the derived mineral abundances. However their fraction should be small given the NORMA coverage of all major minerals identified by the XRD analysis. Finally, normative modelling is an accepted approach to derive quantitative mineral information and serves as a complement/alternative to quantitative XRD analysis or thin section point counting as demonstrated by Melkerud *et al.* (2000).

## 6.5. Mineral Classification

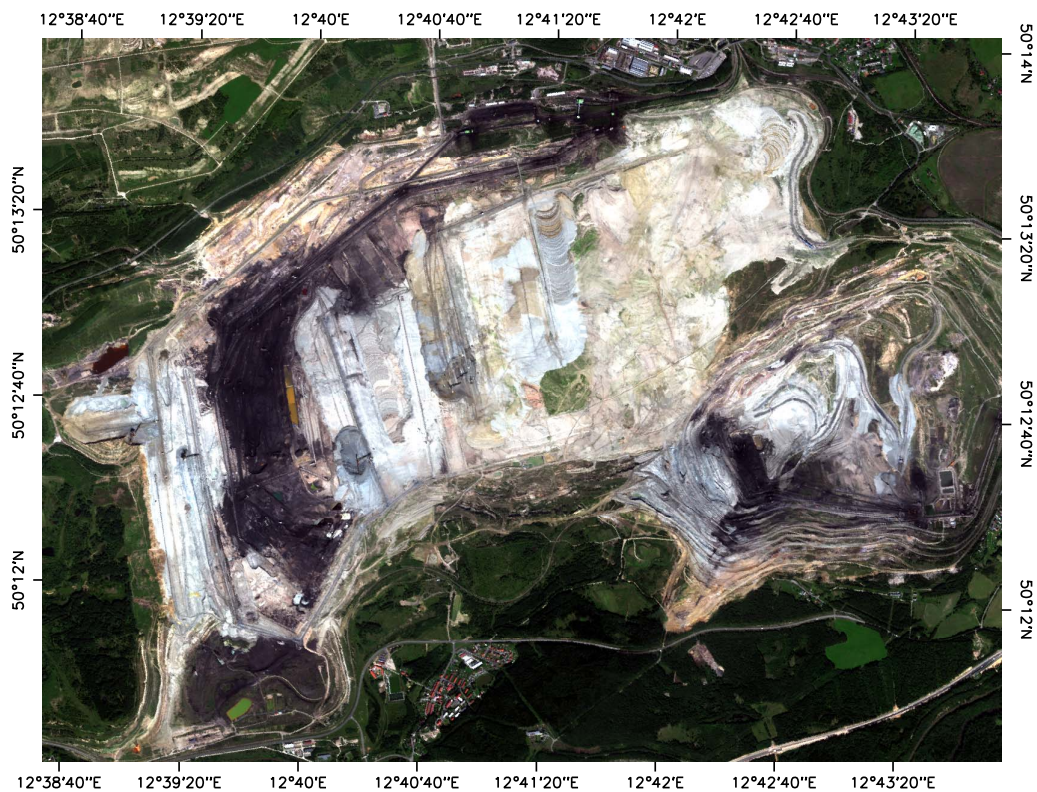


Figure 6.11.: True color composite of the HyMap reflectance imagery (R:640 nm, G:540 nm, B:480 nm). Active mines subset used for showcasing the classification and quantification results.

Minerals of interest for the evaluation of AMD were reasoned in Appendix A - *Acid Mine Drainage Supplements* and Section 2.2 - *Application Objective & Mineral Mapping Candidates*. The carbonate siderite is the major sink for acidity in Sokolov. Spectral resolution of the AHS sensor however proved too coarse for the thermal carbonate feature. Ultimately, the rare occurrence within the mineral soil samples restricted siderite to mineral identification based on the HyMap imagery. Further classification candidates are kaolinite and goethite. They describe slow acidity buffering areas and possibly AMD affected sites respectively. In Section 2.3.5 - *Spectral Characteristics of relevant Minerals* the features of the three minerals are introduced.

In the following two sections the above minerals are classified with the Wavelet Coherence Mapper applied to the pre-processed hyperspectral reflective imagery. A comparison to Spectral Angle Mapper is made. A critical review of methodical and practical advantages and disadvantages is given in Chapter 7 - *Discussion & Summary*. Figure 6.11 displays the spatial subset which was chosen for visualisation and comparison purposes. It encompasses of two lignite open casts.

### 6.5.1. Wavelet based Classification of Siderite, Goethite & Kaolinite

In a first classification experiment the Wavelet Coherence Mapper was used to classify siderite, goethite and kaolinite. Their relevance for assessing AMD dispersion patterns in Sokolov was outlined in Section 2.2 - *Application Objective & Mineral Mapping Candidates*. From a methodological perspective, this experiment was chosen to demonstrate i) the influence of the wavelet

transform scale on detectable features, ii) the characteristics of the yielded spectral similarity measure over the wavelength, and iii) the limitations due to border effects inside the COI area at the margins of the spectral coverage.

The three minerals expose characteristics in different scales. The rather narrow Al–OH doublet in kaolinite can be expected to be picked up at small scales. The characteristic  $\text{Fe}^{3+}$  absorptions in goethite are spanning the VNIR region and should dominate the small to medium scales. Siderite is characterised by the combination of the very broad  $\text{Fe}^{2+}$  and the narrower  $\text{CO}_3^{2-}$  absorptions which should show up at small and high scales.

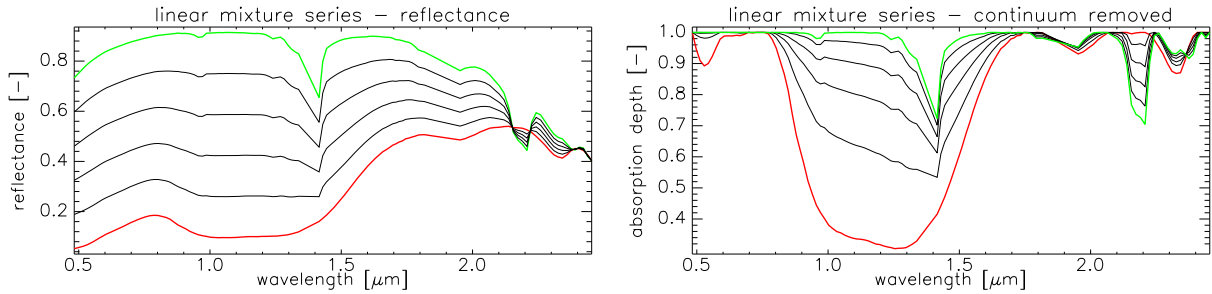


Figure 6.12.: Synthetic linear mixture series of the resampled kaolinite (green) and siderite (red) class templates in 20% increments (black).

**Data** The collected samples did not contain pure minerals. The NORMA modelled abundances for kaolinite and goethite reached up to 60 wt-% which was in accordance with localised enrichments found during field visits. For siderite no such accumulations were found. This was reflected in normative abundances reaching 45 wt-% for one sample but yielding an average of only 5 wt-% for the 23 samples that contained siderite according to XRD. Inspection of the spectra of 7 samples with siderite as a XRD main mineral showed little evidence of the expected  $\text{CO}_3^{2-}$  feature. Likewise, the  $\text{Fe}^{2+}$  absorption was not well developed. Both were likely degraded by admixed mineral fractions. Synthetic linear mixtures of the kaolinite and siderite class templates confirmed this observation (see Figure 6.12). In the reflectance spectra the  $\text{Fe}^{2+}$  and  $\text{CO}_3^{2-}$  features became apparent for approximately  $> 50\%$  siderite in the mixture. The continuum removed spectra showed a subtle curvature reversal in the carbonate feature at 2,350 nm for  $> 40\%$  siderite.

Therefore USGS library spectra (Clark *et al.*, 1993) were used as class templates. The variants “kaolini6.spc Kaolinite CM3”, “goethit4.spc Goethite WS220” and “siderite.spc Siderite HS271.3B” were chosen based on the absorption characteristics and high spectral contrast (see Figures 2.6, 2.3 and 2.4). Final input to the Wavelet Coherence Mapper were the pre-processed reflectance imagery and the three resampled mineral template spectra.

**Method** The wavelet coherence between all pixel spectra and the three individual class templates was calculated over the full wavelength range 483–2,454 nm for the dyadic scales 2, 4, 8 and 16 (see Scheme 3.10). The coherence values inside the COI area of the spectral coverage margins were set to “no data” to exclude border effects. Practically, this cropped spectral information from the beginning and end of the spectral signatures. Table 6.5 lists the remaining effectively usable spectral subset for the individual scales.

Table 6.5.: Size of the COI area for the coherence spectra at various scales and resulting effective spectral subset left from the 120 bands HyMap data covering 483 - 2,454 nm.

scale	coherence COI area	effective spectral subset
1	2 bands	513 - 2,422 nm
2	3 bands	528 - 2,407 nm
4	6 bands	574 - 2,359 nm
8	12 bands	666 - 2,242 nm
16	23 bands	833 - 2,066 nm

The average coherence over the effective spectral subset was calculated. This aggregated value holds information on the “mean similarity” of the curvature of the unknown image spectrum compared to the template spectrum at a spectral level of detail governed by the selected scale. One rule image for each scale and class resulted.

The classification was based on a minimum coherence threshold, conducted on a per-pixel basis and for each class independently. Table 6.6 shows thresholds of the histograms of the rule images. For each scale the highest coherence value and selected quantiles of the coherence are given that may serve as classification thresholds. By nature, this quantile thresholds are scene dependent. The here chosen 97%-quantile for example classifies the 3% most similar image pixels, regardless of their actual similarity score. Reasons for this choice were i) to consider the variable “fit” of the class templates to the mixtures found in the scene, ii) to compensate for the effectively used spectral subset changing with scale, and iii) to intentionally allow for some false positives.

Table 6.6.: Statistics for Wavelet Coherence Mapper results of kaolinite, goethite & siderite classification. Figures are the maximum spectrum averaged coherence and its selected quantiles for each scale based on the rule images.

mineral	scale	quantiles of average coherence			
		max.	99.99%	99.00%	97.00%
kaolinite	2	0.709	0.672	0.612	0.596
	4	0.859	0.829	0.678	0.636
	8	0.928	0.816	0.597	0.556
	16	0.995	0.983	0.894	0.854
goethite	2	0.763	0.668	0.590	0.556
	4	0.828	0.750	0.690	0.660
	8	0.972	0.930	0.796	0.744
	16	0.983	0.945	0.864	0.833
siderite	2	0.756	0.666	0.590	0.579
	4	0.910	0.805	0.688	0.670
	8	0.989	0.962	0.839	0.805
	16	0.996	0.965	0.832	0.797

**Results** Looking at Table 6.6, in general the coherence values increased with scale because focus shifted from detailed similarity investigation of the small spectral features to more general shape similarity of the smoothed spectra. Also the increasing COI excluded possibly discriminative spectral features leading to an apparently higher similarity. The maximum coherence values around 0.7 in scale 2 indicate a considerable difference in spectral detail between the best-matching image pixels and the class templates. The decrease of the coherence over the quantiles is governed by the number of possibly matching pixels. Kaolinite, one of the main minerals



found in Sokolov, would be expected to be matched in great numbers. The contradicting drop in coherence already for the 99% quantile gives an indication that the kaolinite class template was actually not a good representation, especially at the medium level of spectral detail (scales 4 and 8). At the spectral continuum level (scale 16) apparently a number of better matches existed. The higher coherence however must be interpreted with caution due to the limited effective spectral subset the similarity measure is based on.

Figure 6.13 plots the kaolinite class template and the best-matching image pixel scoring the highest average coherence in scale 2. Hence the match was focused on spectral detail rather than the spectral continuum. Looking towards the curvature of the continuum, i.e. scale 16, a general good match is indicated by the high wavelet coherence. The scale 16 CWT plot shows a concave up curvature (negative CWT) for the template at 1,300 nm while for the match no curvature to concave down (zero to positive CWT) is indicated. The resulting drop in coherence from 1,100-1,500 nm localises this deviating shape behaviour. At medium level of detail, in scales 8 and 4, curvature differences around the 1,400 nm and 1,900 nm water absorptions are detected with increased penalty at scale 4 due to less inherent spectral smoothing. The concave up behaviour in the match at 1,000-1,150 nm compared to the slight concave down of the template causes a distinct drop in coherence for both scales. The deviations in scale 8 beyond 2,100 nm appear to be an artefact introduced by border effects. Together it illustrates the high penalty of differences in curvature direction on the coherence while deviations in curvature strength have less effect. The scale 2 coherence quantifies the similarity in spectral curvature at a high level of detail. Strong deviations at the margins of the linear interpolation over the water absorptions at 1,350 and 1,850/2,000 nm are detected while a good match around the 2,200 nm Al–OH feature is reported. The CWT of scale 1 exactly localises the narrow absorption features.

Based on above observations, a kaolinite classification based on the average coherence over the full spectral range is problematic. The characteristic kaolinite Al–OH absorption is best represented in a small scale, however averaging over the full spectral range includes a lot of irrelevant information where no characteristic features are present. The detailed interrogation of this spectral regions would yield low coherence values. This would force the classification threshold to be lowered and hence would lead to many false positives. On the other hand was a significant shape difference between class template and the image spectra revealed in the 500-2,000 nm wavelength range. A classification based on high scale coherence (continuum shape) can therefore not produce the envisaged results. Instead medium scale 4 was used as a compromise that can represent the Al–OH absorption and remaining featureless behaviour while largely eliminating an influence of the spectral continuum. Figure 6.14 shows the spatial kaolinite distribution patterns based on the 574-2,359 nm averaged coherence in scale 4. Areal enrichments were identified at the northern flanks of the two lignite mines for the 97%-quantile (magenta). Additional areas in the overburden and backfill were classified when lowering the threshold to the 80%-quantile. Basing the classification on scale 4 excludes mixtures of kaolinite with vegetation or iron minerals due to the superimposed medium scale features. Mixtures with in that scale featureless coal has less impact. In summary relatively bright and, besides a strong clay absorption, featureless image pixels were classified.

The spectral character of the coherence allows to easily exclude irrelevant spectral information

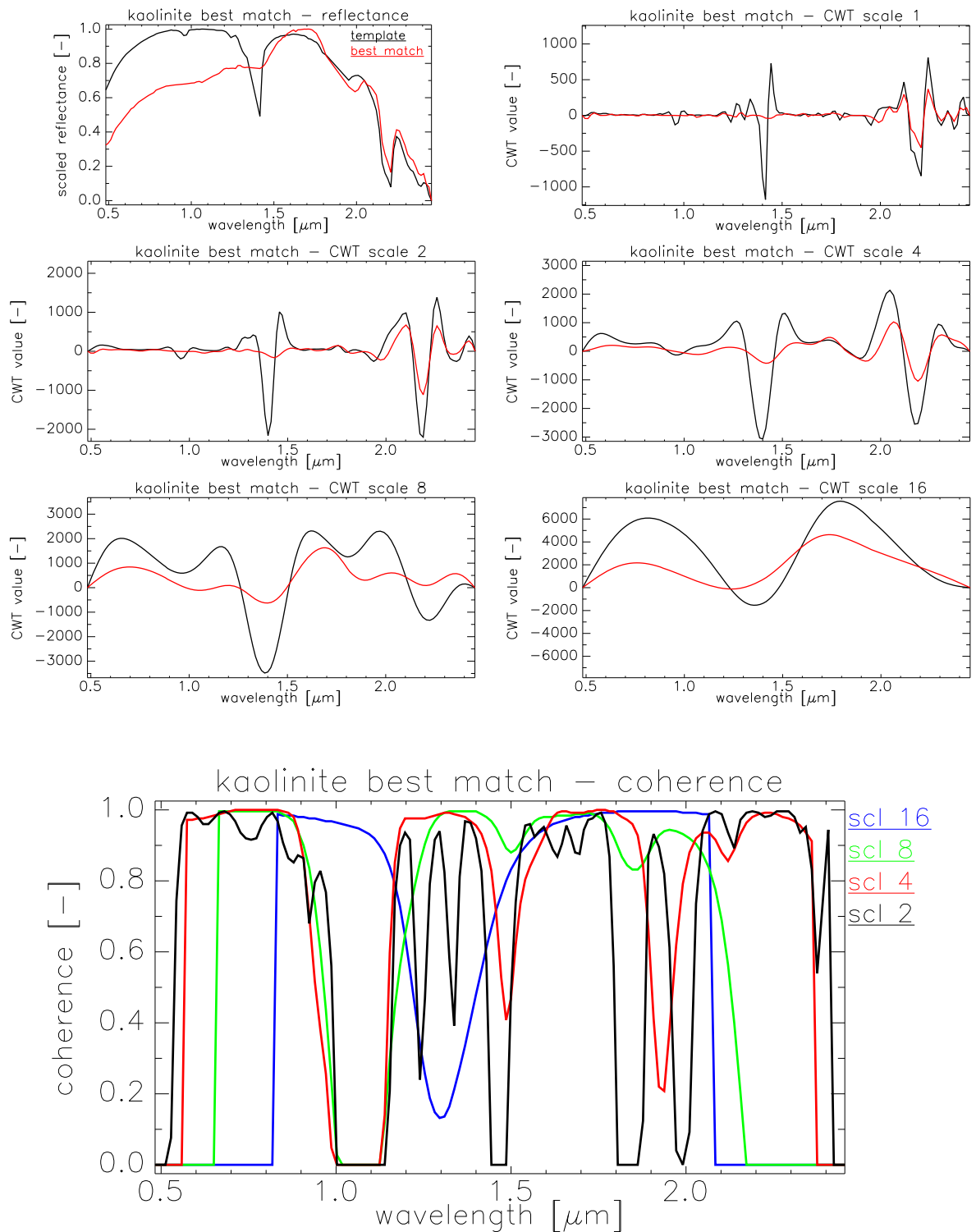


Figure 6.13.: Comparison between kaolinite class template and best-matching image pixel in respect to average coherence at scale 2. Reflectance spectra (top left) and respective continuous wavelet transforms for dyadic scales 1 through 16 (top and middle). The resulting wavelet coherence for scales 2, 4, 8 & 16 (bottom) showing match quality over wavelength.

The drop in coherence in scale 16 around  $1.3 \mu\text{m}$  is caused by the mismatching sign and gain of the CWT of both spectra in the vicinity of  $1.3 \mu\text{m}$  in that scale 16.

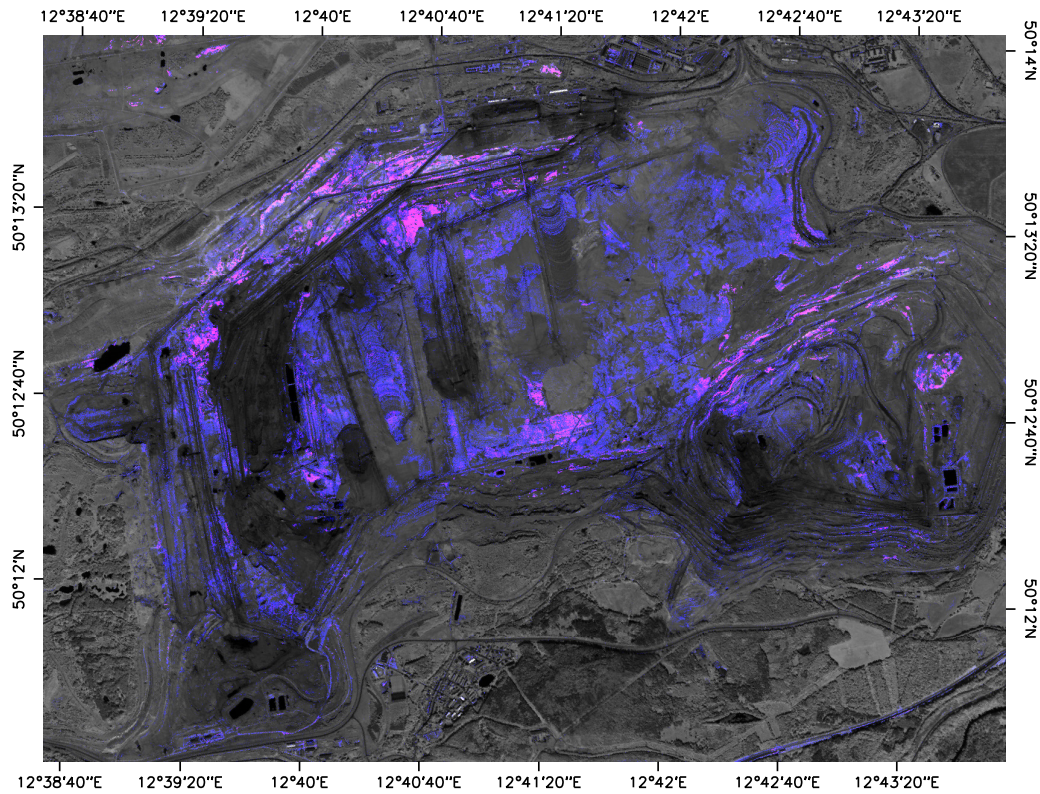


Figure 6.14.: Wavelet coherence for kaolinite showing distributions for scale 4 using a 0.80-quantile (blue) and 0.97-quantile (magenta) threshold (const. lin. color stretch [quantile ... max]).

and to explicitly focus on characteristic features. The clay mineral classification conducted in Section 6.5.2 - *Wavelet based Classification of Clays* demonstrates the advantages of such a selective analysis.

For goethite Figure 6.15 plots the class template and best-match. The small to medium scales indicate a comparably good match for the iron features until 1,300 nm, followed by a considerable mismatch until the beginning of the waver vapour interpolation at 1,850 nm. Additional deviations are located at both flanks of the 2,200 nm reflectance bulge in the SWIR. A mismatch till 1,200 nm in scale 16 appears to be caused by the strong inherent smoothing in combination with border effects. A concave up template spectrum but rather concave down pixel spectrum result in the coherence to drop to zero. A likewise behaviour is apparent for the spectral region beyond 1,800 nm.

The goethite classification was based on scale 2 given the rather high spectral detail of the  $\text{Fe}^{3+}$  absorptions and their localisation close to the short-wavelength margin of the spectral coverage. Interference with the COI was therefore prevented. Figure 6.16 shows the spatial goethite distribution patterns. Areal enrichments were identified at the northern and western flanks of the lignite mines. These occurrences in the uppermost layer of the Cypris overburden occur naturally in the Sokolov country rocks. A small hotspot appeared close to a pool of water (1) where shallow orange water changes over to goethite precipitation. Restricted access to the mine prevented the field verification of this probably AMD affected site. The imagery however gives strong evidence (see Figure 6.36 on page 126 at the discussion of the goethite quantification).

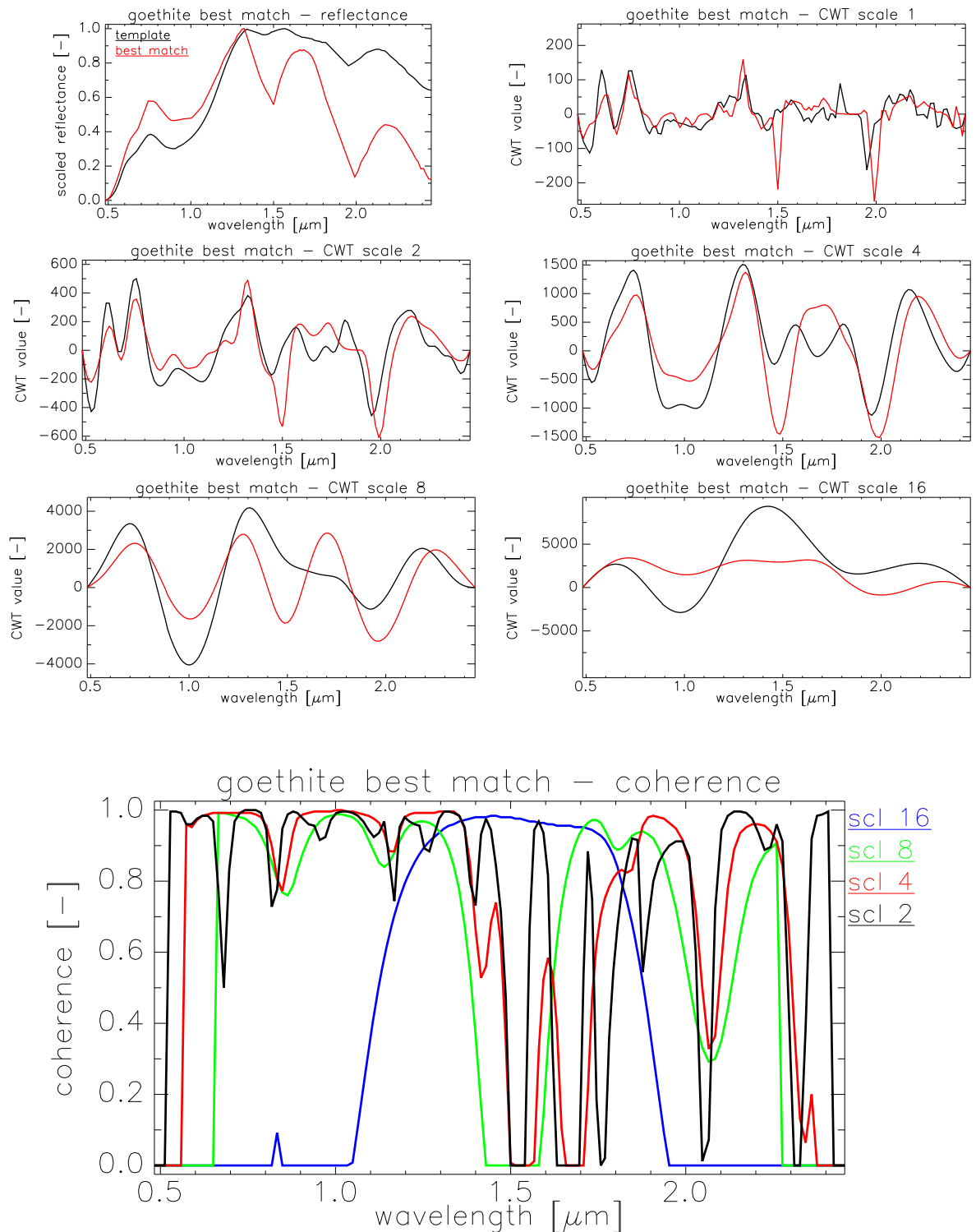


Figure 6.15.: Comparison between goethite class template and best-matching image pixel in respect to average coherence at scale 2. Reflectance spectra (top left) and respective continuous wavelet transforms for dyadic scales 1 through 16 (top and middle). The resulting wavelet coherence for scales 2, 4, 8 & 16 (bottom) showing match quality over wavelength.

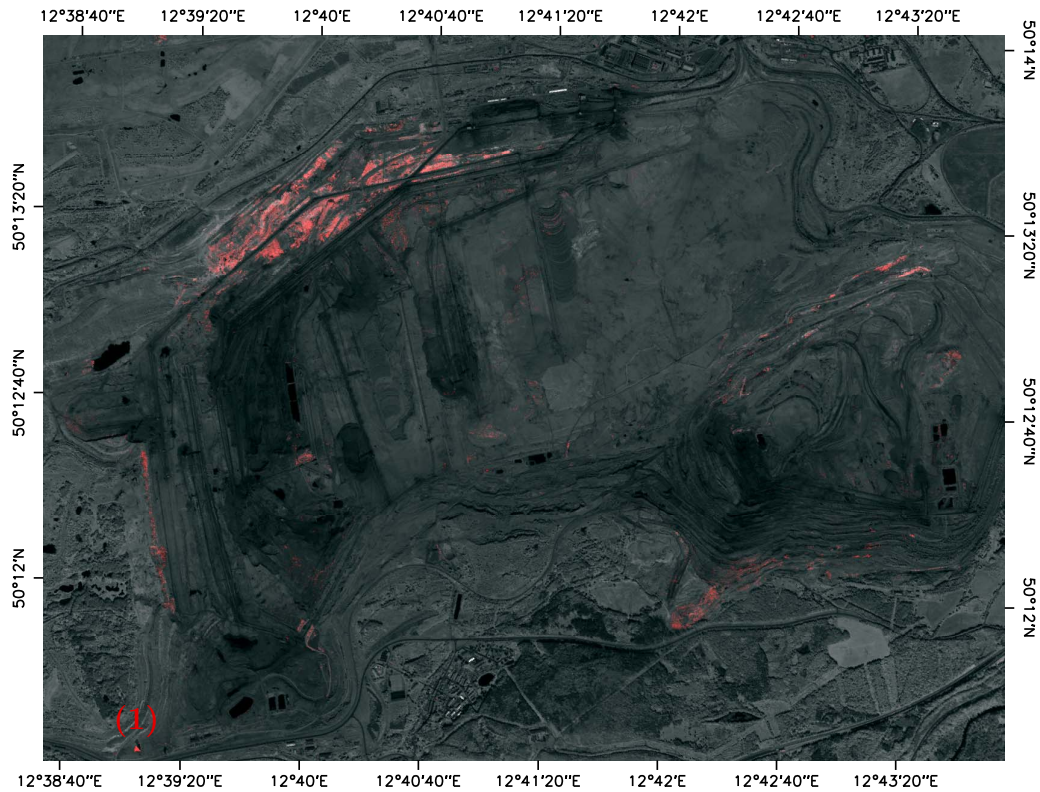


Figure 6.16.: Wavelet coherence for goethite showing distributions for scale 2 in red (const. lin. color stretch [0.97-quantile ... max]). Possibly AMD affected area around a water pool (1).

The siderite class template and best-match are displayed in Figure 6.17. The higher scale wavelet transforms and coherence indicate a good match in spectral shape. At scales 16 and 8 the differing Al–OH and  $\text{CO}_3^{2-}$  absorptions are not visible due to the strong inherent smoothing. They are picked up by scale 4 and smaller where the differences in spectral curvature manifest in the coherence drop after 2,100 nm. Further smaller deviations are detected at 850 and 1850 nm, the former originating from the differences in the short-wavelength shoulder of the broad  $\text{Fe}^{2+}$  feature. At scale 2 the subtle curvature differences within the  $\text{Fe}^{2+}$  trough lead to a coherence drop around 1,200 nm. Scales 8 or 16 instead are suitable to evaluate the gross shape of this broad absorption irrespective of small scale differences, however the limitations due to the COI must be respected. The scale 1 CWT allows to specifically focus onto the discussed subtle 2,350 nm curvature differences found there for kaolinite and siderite.

The siderite classification was conducted in the medium to high scales 4, 8 and 16 over the respective full effective spectral subset. With increasing scale the wide  $\text{Fe}^{2+}$  trough gained higher influence onto the classification while the influence of the smaller scale deviations diminished. The characteristic  $\text{CO}_3^{2-}$  absorption was cropped due to the COI area for scales above 4. Figure 6.18 depicts the derived spatial siderite distributions, color coded for the three scales in a linear stretch from the respective 97%-quantile to the maximum value reached. The threshold intentionally allowed for some false positives discussed below. Scale 4, coded in red, classified spatially limited distributions of pixels to the north, south and east of the smaller lignite mine. Scale 8, given in green, additionally classified larger areas in the backfill and northern flank of the bigger lignite mine. Scale 16, coded in blue, had no overlap with the scale 8 classification in the backfill but instead classified areas in the Cypris clay overburden of the bigger mine. It

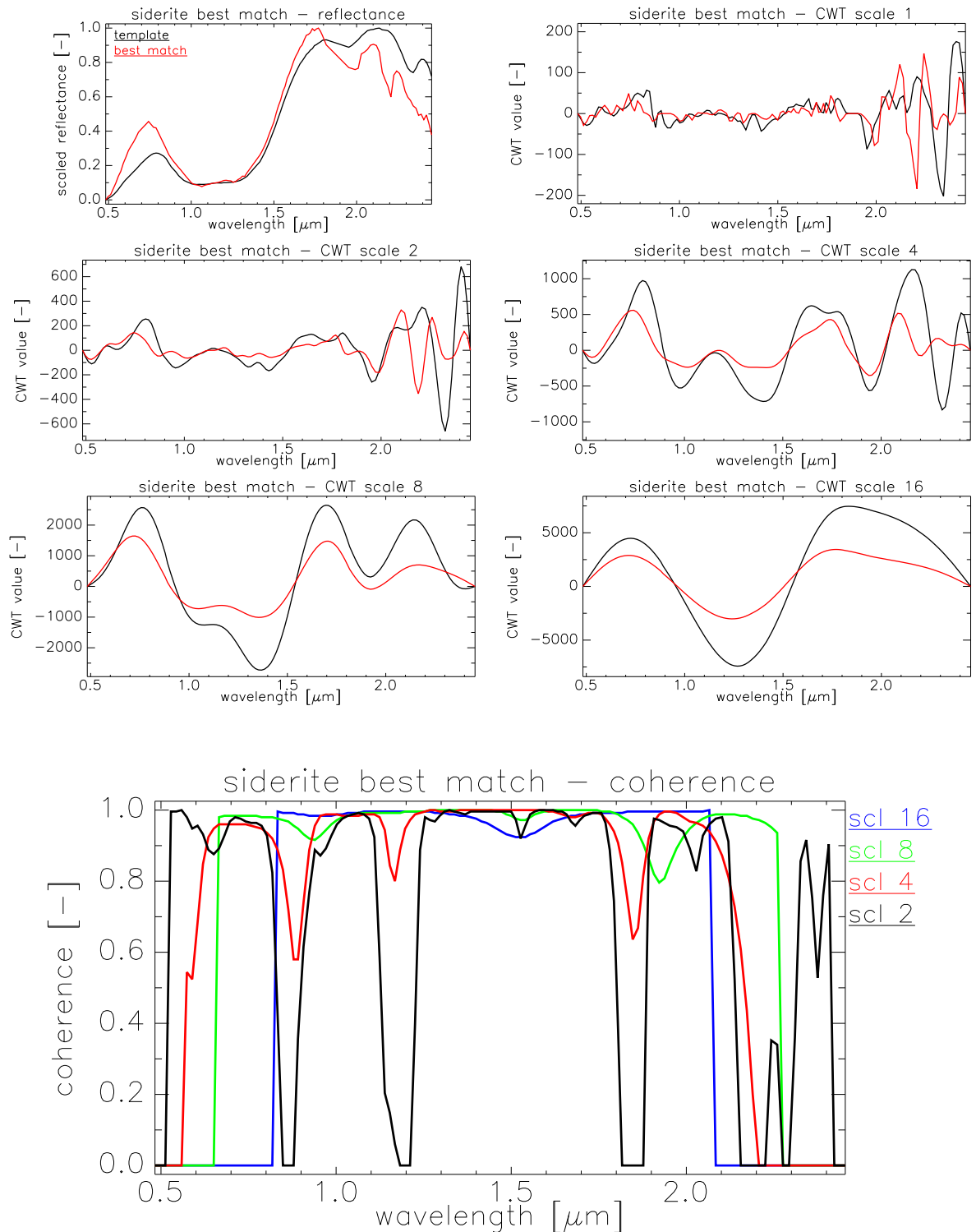


Figure 6.17.: Comparison between siderite class template and best-matching image pixel in respect to average coherence at scale 2. Reflectance spectra (top left) and respective continuous wavelet transforms for dyadic scales 1 through 16 (top and middle). The resulting wavelet coherence for scales 2, 4, 8 & 16 (bottom) showing match quality over wavelength.

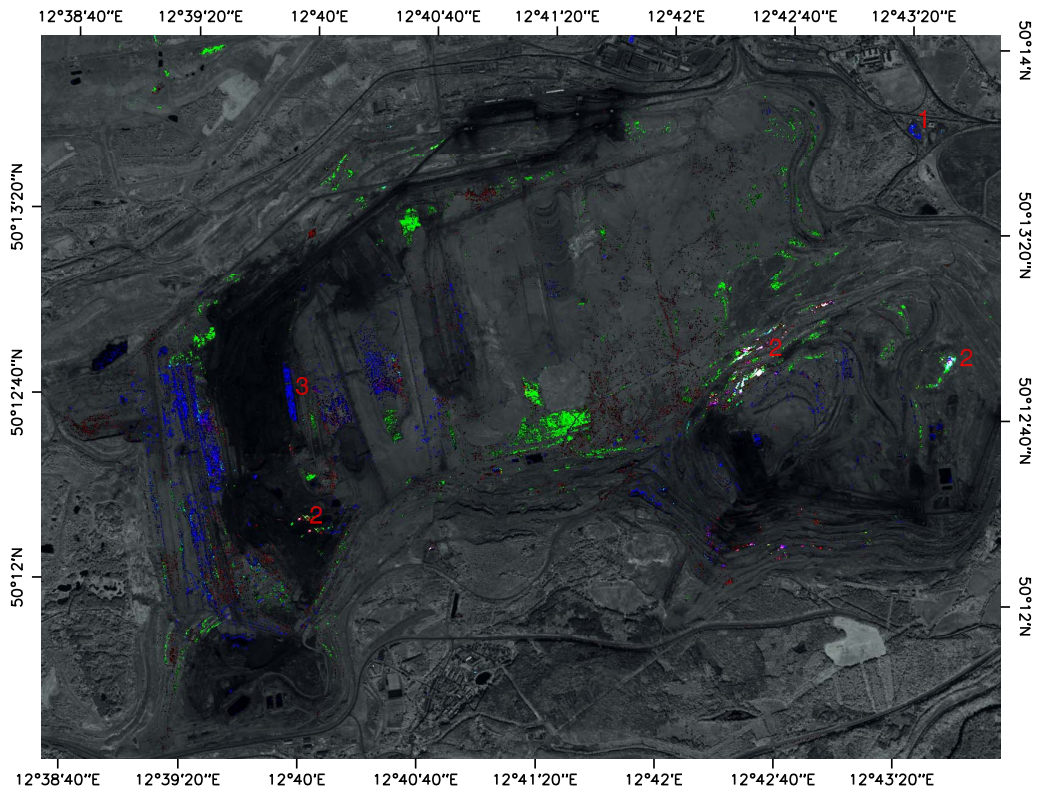


Figure 6.18.: Wavelet coherence for siderite showing distributions for scale 4 (R), 8 (G) and 16 (B) (const. lin. color stretch [0.97-quantile ... max]). Areas of high confidence (2) detected in all three scales versus false positives in scale 16 (1 and 3) as described in text.

however showed apparent false positives in water (3) and a green glass dump at a recycling site (1). The areas marked by (2), appearing white in Figure 6.18, yielded a high coherence in all three scales.

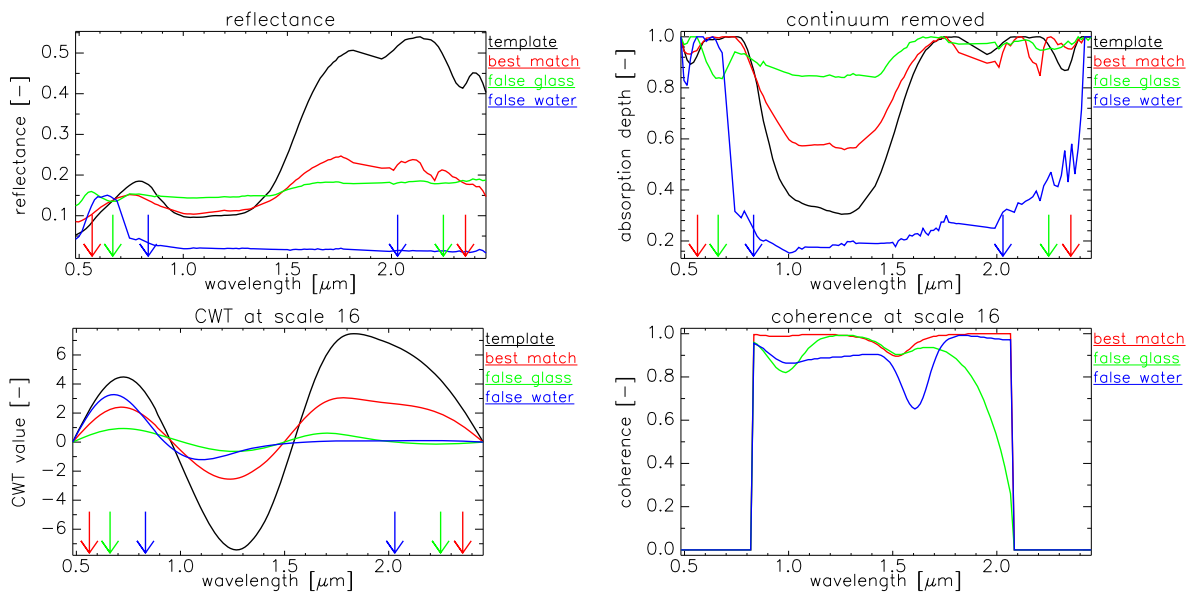


Figure 6.19.: Comparison of siderite class template spectrum (black) and image spectra of a classification match (red), glass false positive (green) and water false positive (blue). Reflectance spectra (top left) and continuum removed absorption depths (top right). The scale 16 CWT amplitude spectra (bottom left) and the resulting coherence spectra (bottom right). Arrows on the wavelength axis depict size of the COI area for scales 4/8/16 (red/green/blue).

Representative spectra of the observed false positives in scale 16 (1, 3) can be seen in Figure 6.19. They were caused by a combination of factors. The rather low coherence corresponding to the 97%-quantile threshold allowed for a significant deviation of the image spectra from the class template. In addition, the effective spectral range shrinks with increasing scale. The arrows on the wavelength axis in Figure 6.19 mark the extent of the COI for the respective scales. Within the COI the coherence was set to zero. For scale 16 this obviously excluded many of the discriminative continuum curvatures. The combination of wavelet transform and coherence omitted additive constants, multiplicative constants and linear bias. Altogether this led to the cropped water and glass spectra having similar curvature characteristics as the class template. This can be roughly reproduced by a continuum removal (see Figure 6.19).

Highly correlating curvatures can be seen when comparing the scale 16 CWT of the class template and match. This manifests in a high average coherence of 0.97 (spatially averaged). The glass false positives also show a well correlating curvature until 1,850 nm beyond which the glass spectra show a concave up behaviour while the template maintains concave down curvature. This leads to the drop in coherence which however is masked by the COI after 2,066 nm. The spatially averaged coherence was 0.88 for the glass. The water false positives had a distinct offset of the CWT minimum and maximum compared to the template. However the direction of the curvature, i.e. sign of the CWT value, generally agreed. This is reflected by a reduced coherence from 1,000-1,700 nm beyond which the curvatures match and high coherence is reached. In total the spatially averaged coherence for water was 0.85 which may appear comparably high. Reason for this behaviour is the high sensitivity of the coherence measure to differences in curvature direction, i.e. the sign of the CWT, as seen for the glass beyond 1,900 nm. Differences in curvature strength, i.e. the CWT value, cause less penalty as seen in the water from 1,000-1,400 nm.

The small scale curvature differences around 2,350 nm found for synthetic kaolinite/siderite mixtures were investigated with the scale 1 coherence. Retrieving the coherence value at 2,342 nm, 2,391 nm or averaged over 2,242-2,422 nm to take advantage of an absorption doublet found there for kaolinite versus the broader featureless  $\text{CO}_3^{2-}$  absorption for siderite did not produce clear results. It was found that especially lignin and holocellulose in vegetation (Elvidge, 1990) but also hydrocarbon in coal (Lyder *et al.*, 2010) had similar shaped featureless absorptions close to that particular wavelength. Noise in the image spectra did also seem to frequently mimic matching curvature properties when relying on scale 1 coherence. In consequence the spectral details around 2,350 nm did not bear characteristic information usable for a complementing classification of siderite from the  $\text{CO}_3^{2-}$  absorption. For the siderite occurrences overlapping in 3 scales, marked with (2), it was possible to detect supporting detail features at 2,350 nm. The spatial patterns however were strongly influenced by noise and not fully consistent. The latter may attribute to the high siderite fraction  $> 40\%$  necessary for this features to develop.

For a confident siderite classification one would choose a tighter threshold (e.g. 99.99% quantile) and use the intersection of the scale 4 and scale 8 class assignments. This was done for Section 6.7.2 - *Validation of Quantification Results* leaving the overlap areas (2) from above and excluding the false positives. The siderite classification hence relied on the uniqueness of the



broad iron absorption for this mineral. The carbonate feature was not developed sufficiently and not characteristic enough at the HyMap resolution to support a multi-scale classification.

In summary, this experiment demonstrated the high flexibility in interrogating the spectral signature at selectable level of detail. The decomposition of the signal into various scales allows to focus on narrow discriminatory spectral features or on characteristic curvatures of the continuum. Information from a combination of scales complementing each other can be exploited for a well adapted target detection or classification.

The transform values at higher scales must be interpreted with caution close to the margins of the spectral coverage due to the possibility of distortions by border effects. The COI describes the jeopardised spectral regions. This limitation in interrogating spectral features close to the margins of the spectral coverage degraded the classification performance for higher scales.

The coherence gives a spectral similarity representation evolving over wavelength. It was shown to be very sensitive to differences in curvature direction, i.e. the sign of the underlying wavelet transform amplitude spectra. Mismatches in curvature strength, i.e. the absolute value of the underlying transform spectra, cause less penalty on the coherence value. The coherence measure is therefore well suited for spectral regions that exhibit characteristic curvature properties while featureless regions should be excluded. The latter can be readily achieved by focusing on specific wavelengths or wavelength ranges (demonstrated by the following experiment). Possible noise and ambiguity in the very small scale features must be considered when interpreting coherence at scale 1.

### 6.5.2. Wavelet based Classification of Clays

In a second classification experiment the Wavelet Coherence Mapper was used for a more differentiated clay mineral classification. The Cypris lacustrine clay formation overlaying the main coal seam in the two active mines was selected as test case. In Section 2.1.1 - *Geologic Setting* the gradual change from kaolinitic clays at the Cypris formation's bottom through illite-kaolinite claystones into kaolinite-montmorillonite-illite claystones at the top is introduced. This intimate mixture of three clay minerals together with the variable admixture of coal from the underlying seam and up to 16 wt-% siderite was seen a challenging case-study.

This experiment was chosen i) to demonstrate focusing analysis on a limited wavelength range, ii) to compare to Spectral Angle Mapper classification results, and iii) to test sensitivity and class assignment of the Wavelet Coherence Mapper in order to demonstrate the benefits and limitations of using the coherence spectral similarity measure.

**Data** Lacking pure clay endmembers from field/laboratory spectroscopic measurements, the classifications were based on USGS library spectra (Clark *et al.*, 1993) chosen as class templates. The variants “kaolini6.spc Kaolinite CM3”, “illite4.spc Illite IL101 (2M2)” and “montmor8.spc Montmorillonite STx-1” were selected based on the absorption characteristics and high spectral contrast. The class template spectra were resampled to the HyMap sensor's spectral resolution (see Figure 6.20). This resulted in the degeneration of the characteristic Al–OH absorption around 2,200 nm compared to the full resolution spectra (Figure 2.6). The most discriminative spectral features were the 2,000-2,150 nm slope, the localisation of the short-wavelength flank of the 2,200 nm absorption, the shape of the 2,200 nm absorption and the 2,300-2,400 nm features.

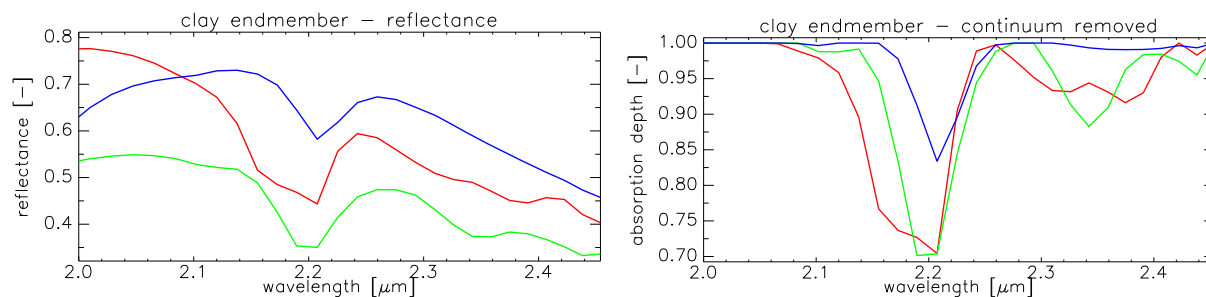


Figure 6.20.: Class template spectra from the USGS library (Clark *et al.*, 1993) used for clay classification. Reflectance spectra (left) and respective continuum removed absorption depths (right) for kaolinite (red), illite (green) and montmorillonite (blue).

**Method** Given the variable admixtures of coal and siderite, the classifications were based on two spectral subsets to exclude irrelevant and possibly misleading spectral information. The used 27 bands spectral subset 2,000-2,500 nm represents a wider spectral range limited by the strong atmospheric water vapour absorption at 1,900 nm and the margin of the SWIR atmospheric window and HyMap spectral coverage at about 2,500 nm. The narrower 9 bands spectral subset 2,102-2,242 nm explicitly focuses on the Al–OH absorption, approximately from its short to long-wavelength shoulder. By using this spectral subsets the necessity for several mixture endmembers was mitigated.

SAM was simply calculated on the two subsets separately. The wavelet coherence does not require prior selection of a spectral subset, hence it was calculated once for the full spectral range. The Wavelet Coherence Mapper classification was then based on the average coherence in the 2,000-2,500 nm spectral subset and on the specific coherence value at 2,170 nm which is located in the centre of the Al–OH absorption. Scale 2 was chosen such that the wavelet’s wavelength support (2-COI) was similar to the width of the Al–OH absorption. The smoothing during coherence calculation then led to a dilated support of about 9 wavelength sampling intervals (see Section 3.2.1 - *Signal Similarity Measure*). Hence the coherence at 2,170 nm actually contains (weighted) spectral similarity information from the wavelength range 2,102-2,242 nm. This range equals the subset used by SAM. This allowed comparison of the results of the two methods in respect of the spectral information used.

For directly comparing the Wavelet Coherence Mapper and SAM in respect of the similarity measure, SAM was also applied on the scale 2 wavelet transformed reflectance image. The combinations of SAM, coherence, reflectance spectra and their wavelet transforms led to six classification approaches (CA1 - CA6) as indicated in Table 6.7. All approaches use an empirically chosen 75%-quantile classification threshold, i.e. the 25% most similar image pixels, regardless of their actual similarity score, are assigned the individual classes. In case of multiple class assignments of a pixel, it is assigned the class with the highest similarity score. This quantile thresholds are scene dependent but allow for a spatial coverage of the classification inside the mining area and, most important, facilitate a comparison between the six approaches.

Table 6.7.: 75%-quantile thresholds for maximum SAM angle [radians] and minimum coherence [0...1] for the three clay classes used by the six classification approaches CA1 - CA6.

method	2,000-2,500 nm			2,102-2,242 nm		
	kao.	ill.	mont.	kao.	ill.	mont.
SAM on reflectance	CA1 : 0.153	0.109	0.052	CA2 : 0.115	0.106	0.047
SAM on CWT scale 2	CA3 : 0.862	0.736	0.792	CA4 : 0.773	0.514	0.719
WCM on CWT scale 2	CA5 : 0.524	0.596	0.579	CA6 : 0.628	0.956	0.824

**Results** The resulting classification thresholds (75% quantile of rule images) are given in Table 6.7. It must be noted that the thresholds can not be compared between the approaches. For SAM on reflectance spectra, the thresholds generally comply typical values found in the literature that range from 0.050 to 0.100 radians, e.g. Vaughan *et al.* (2005); Debba (2006); Bachmann (2007); Alfaro (2009). It appears that montmorillonite best matches the image pixels. SAM thresholds on CWT scale 2 are much higher because of the continuum removal and resulting oscillation of the transform around zero. The magnitude of the thresholds converge due to the exclusion of continuum differences between the three class templates. Illite now appears the best match. In summary it can be observed that the SAM thresholds are relatively stable for both subsets. The coherence thresholds instead show a strong dependence. Low thresholds resulted from the 2,000-2,500 nm subset whereas focus on the Al–OH feature resulted in the admissibility of high thresholds. The 2,102-2,242 nm threshold is considerably low for kaolinite, the most abundant clay mineral in Sokolov, indicating that the average image spectra matched kaolinite least and illite best. Each classification approach is discussed in the following.

**CA1 - SAM ON REFLECTANCE SPECTRA, 2,000-2,500 NM** The standard approach led to very different maximum angle thresholds 0.153/0.109/0.052 for the three clay minerals kaolinite/illite/montmorillonite. Figure 6.23 shows the classification result. In general the amount of unclassified pixels was rather high (A) but also false positives (e.g. roofs) occurred (B). Rather a relaxation nor tightening of the thresholds would therefore be beneficial. Most pixels were classified montmorillonite contradicting the expectations from the site's geology and sample analysis (see Sections 2.1.1 - *Geologic Setting* & 6.4 - *Normative Modelling*). Figure 6.21 compares image spectra around a sharp classification boundary showing their difference in slope but identity in absorption feature characteristic. False positives and false negatives appeared to be caused by SAM's bias towards the overall continuum. In the 2,000-2,500 nm range the montmorillonite template spectrum best matched the image pixel's continuum.

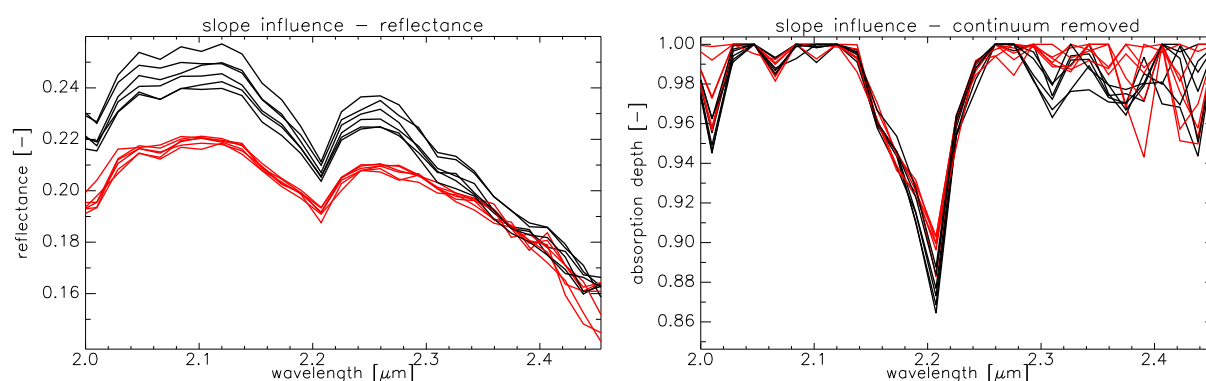


Figure 6.21.: Slope influence on SAM clay classification. Reflectance spectra of mixtures (left) and continuum removed absorption depths (right) for image pixels classified (black) versus not (red).

**CA2 - SAM ON REFLECTANCE SPECTRA, 2,102-2,242 NM** The further narrowing of the spectral range resulted in the classification explicitly focusing on the spectral shape of the clay's characteristic Al–OH absorption. The thresholds reduced to 0.115/0.106/0.047 radians, indicating a significant better match for kaolinite compared to CA1. Still, the image spectra were most similar to the montmorillonite class template. Comparing the CA2 classification depicted by Figure 6.24 to the results of CA1 some classification gaps were assigned montmorillonite (A) while in the north and east areas changed from illite/montmorillonite to kaolinite class (C). The number of false positives reduced (B). The differences originated from focusing on the Al–OH absorption hence limiting the influence of the overall spectral slope.

**CA3 - SAM ON WAVELET TRANSFORM SCALE, 2,000-2,500 NM** Application of SAM on the wavelet transform scale 2 produced interesting results. It showed the expected and area-filling distribution patterns of clay minerals inside the mines (see Figure 6.25). Clearly, the transition from kaolinite around the coal strata towards admixtures of illite and finally of montmorillonite in the overburden was differentiated. SAM on scale 1 (not shown) provided a classification with similar spatial coverage for montmorillonite but largely deviating distribution of illite versus kaolinite, the latter taking over nearly the complete coverage. SAM therefore correctly picked up the subtle curvature changes within the 2,200 nm Al–OH feature emphasised at scale 1, indicating that indeed kaolinite is the

major clay mineral present. Admixtures of illite or montmorillonite slightly change the clay feature's slope and left shoulder position which was picked up in scale 2 leading to the observed distribution patterns. The SAM thresholds of 0.862/0.736/0.792 radians for kao./ill./mont. in scale 2 approached each other but were very high. The latter is caused by the wavelet transform oscillating around zero which results in large spectral angles due to the vector interpretation used by SAM.

**CA4 - SAM ON WAVELET TRANSFORM SCALE, 2,102-2,242 NM** Similar in methodology to CA3, the narrowing of the spectral range limited the classification onto the Al–OH absorption. This narrowing was reflected in the smaller spectral angles of 0.773/0.514/0.719 radians. Little difference in the kaolinite distribution but considerable transition of montmorillonite to illite in the backfill and in the overburden (A) was observed (see Figure 6.26). This appeared to be caused by the exclusion of the 2,000-2,150 nm slope differences in the clay template spectra (see Figure 6.20). False positives appeared for industrial roofs and a green glass dump however (B).

Table 6.8.: Confusion matrix for clay minerals classified with Wavelet Coherence Mapper (scale 2, 2,000-2,500 nm) versus wavelet-based SAM classification (scale 2, 2,000-2,500 nm).

clays confusion matrix [%]					
SAM:2,000-2,500 nm vs. WCM:2,000-2,500 nm					
	SAM none	SAM kaol	SAM illi	SAM mont	total
WCM none	61.68	0.69	0.55	0.23	63.14
WCM kaol	1.27	3.58	0.38	0.00	5.23
WCM illi	2.22	1.75	10.58	1.37	15.91
WCM mont	2.53	0.20	5.44	7.56	15.72
total	67.69	6.22	16.94	9.15	100.00
overall accuracy = 83%, Kappa coefficient = 0.69					

**CA5 - AVERAGE COHERENCE IN SCALE 2, 2,000-2,500 NM** This classification based on the average coherence in scale 2 over the 2,000-2,500 nm spectral range resulted in the clay mineral distribution depicted in Figure 6.27. The coherence 75% quantiles 0.524/0.596/0.579 were similar in magnitude but of rather low value indicating that the class templates were not the best match over the complete 2,000-2,500 nm spectral range. Compared to the respective SAM results (CA3) the classification appeared “noisy” due to frequent swapping between illite and montmorillonite class assignment. The kaolinite distribution patterns were in general very similar, but SAM assigned illite was frequently classified montmorillonite by the coherence method. The confusion matrix given in Table 6.8 quantifies aforementioned deviations. About 11% of SAM kaolinite matches were not classified. Another 28% SAM kaolinite were assigned illite. For about 32% of SAM illite the Wavelet Coherence Mapper assigned montmorillonite. The reverse case occurred for 15% of the pixels. Figure 6.22 compares the mean spectra of the pixels assigned to the three clay classes by SAM and Wavelet Coherence Mapper. SAM mean spectra were very similar in overall shape and amplitude. Wavelet Coherence Mapper mean spectra instead showed a higher spread in amplitude and shape. The continuum removed spectral differences were subtle, especially in respect to the class templates (see Figure 6.20). The CWT plot indi-

cates that curvature direction differences at 2,020 / 2,275 / 2,300 & 2,425 nm, especially in montmorillonite, likely triggered the different class assignments.

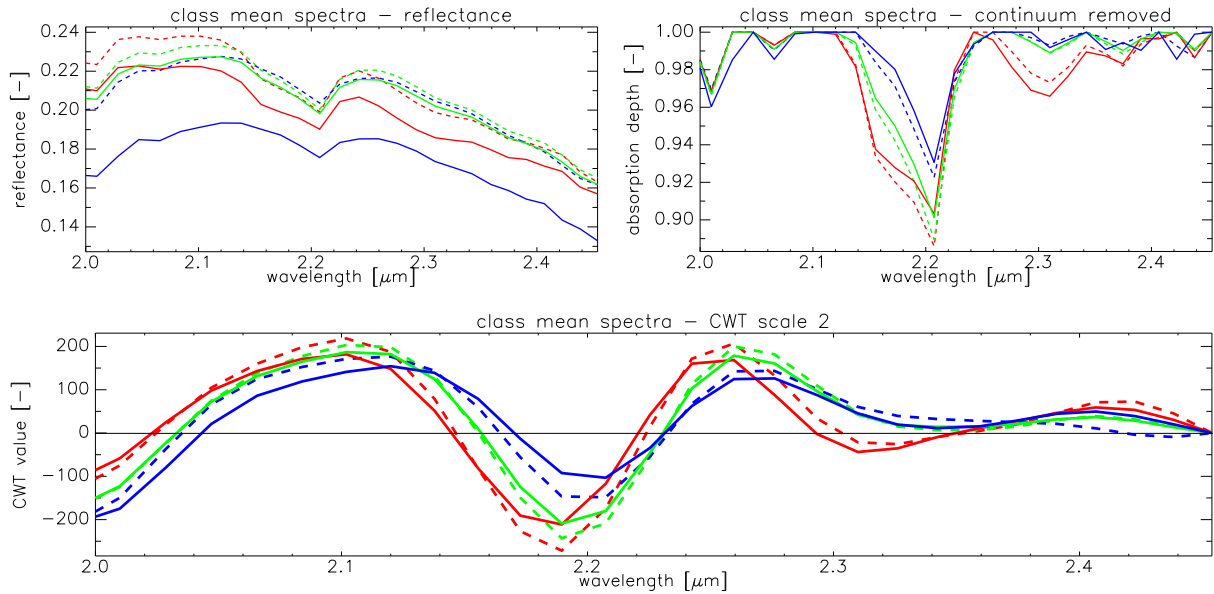


Figure 6.22.: Spectral comparison of mean class spectra for SAM (dashed) and wavelet coherence (solid). Reflectance spectra (left) of clay classes (kaolinite (R), illite (G), montmorillonite (B)) and their continuum removed absorption depths (right). Below the respective CWT of scale 2.

**CA6 - COHERENCE AT FIXED WAVELENGTH 2,170 NM** This classification approach relied on a single coherence value taken from the middle of the characteristic clay absorption. Due to the wavelength support of scale 2, this coherence value integrates the spectral similarity in the 2,102-2,242 nm neighbourhood, weighted by the wavelet function. The thresholds 0.628/0.956/0.824 for kao./ill./mont. showed a higher spread than compared to CA5. Their values were considerably higher, indicating a better match within the narrow spectral range. The illite class template was clearly favoured. The classification results showed the expected and area-filling clay distribution patterns (see Figure 6.28). The spatial patterns were very similar in comparison to the SAM results (CA4). The confusion matrix (Table 6.9) shows a good agreement for illite. However, 30% of SAM kaolinite was assigned illite. Another 33% of SAM montmorillonite was assigned no class. On the other hand did the Wavelet Coherence Mapper classify additional pixels montmorillonite (40% previously SAM unclassified) and kaolinite (30% previously SAM unclassified). The additional kaolinite assignments occurred primarily in the coal-clay mixing areas (A). The apparent false positives (roofs, glass dump) were not mitigated (B).

In summary the clay classification was a challenging experiment. The image spectra did not show the clear differentiation as visible for the class templates. Instead, mixing considerably affected the 2,000-2,150 nm spectral slope and shape of the 2,200 nm absorption. Image spectra of coal-clay mixtures exhibited the best developed kaolinite absorption characteristics, despite kaolinite being the dominant clay species in Sokolov. In this analysis the clay discriminating spectral feature appeared to be the position of the short-wavelength Al–OH absorption shoulder.

Spectral Angle Mapper, sensitive to the spectral shape, naturally had problems to focus on such a subtle difference. Its bias towards the overall continuum resulted in most pixels mapped as

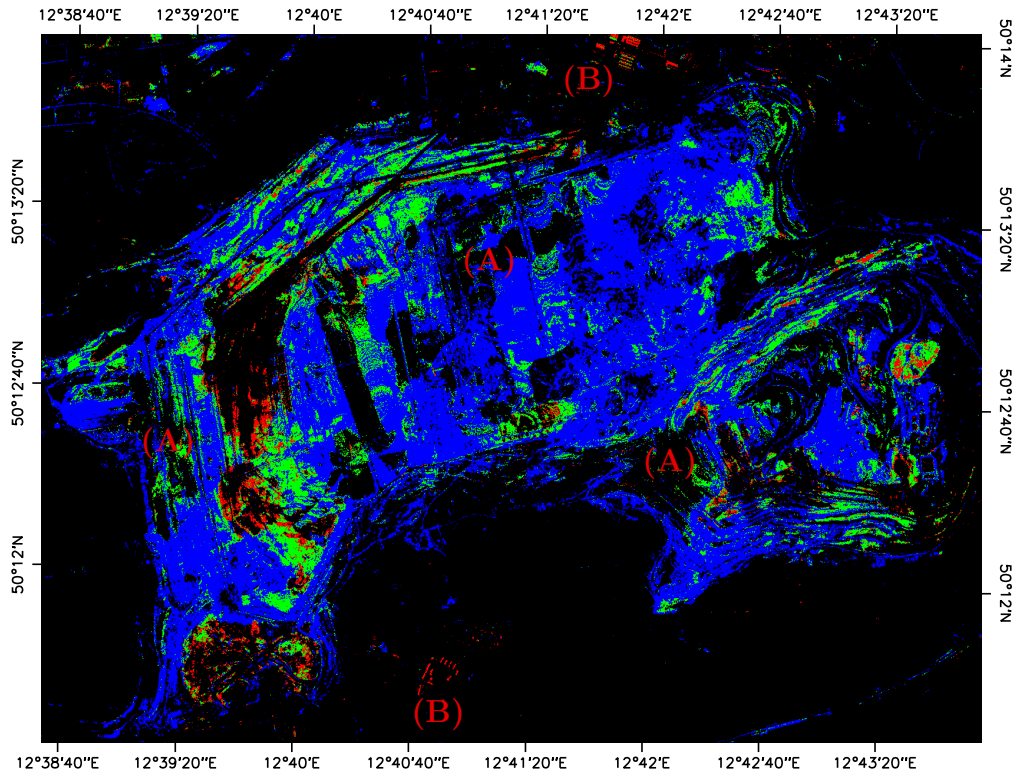


Figure 6.23.: **CA1** 2,000-2,500 nm SAM classification of reflectance image showing kaolinite (R), illite (G) and montmorillonite (B) (max. angle 0.153/0.109/0.052 respectively, class map). Unclassified Cypris clays (A) and roof false positives (B).

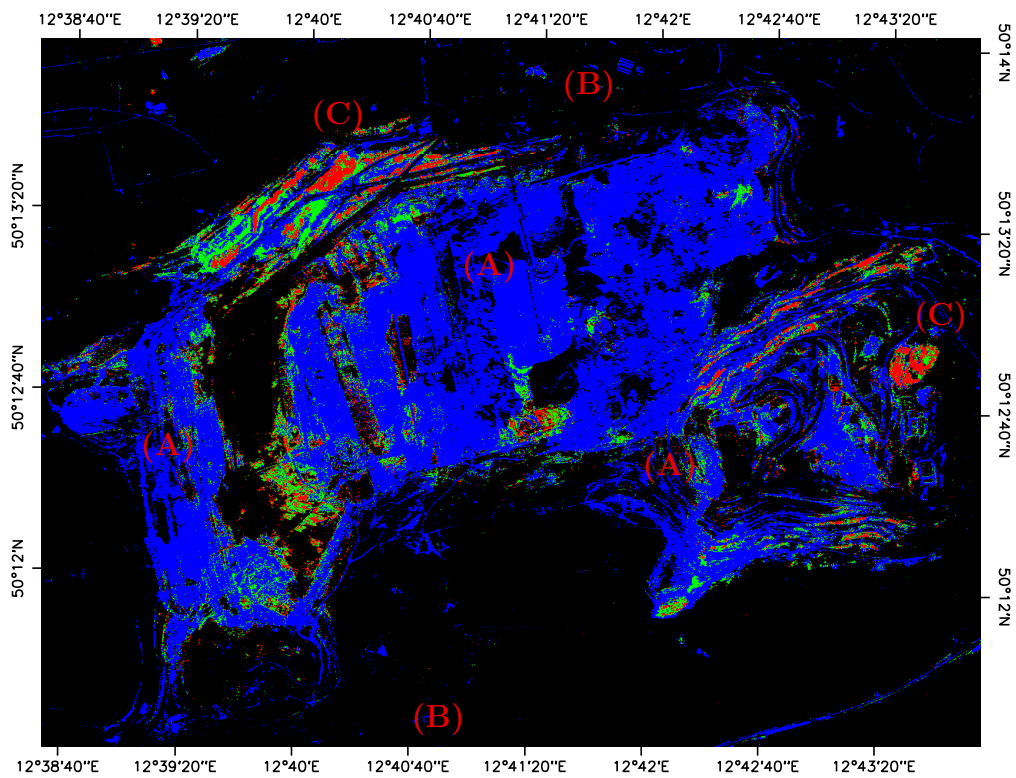


Figure 6.24.: **CA2** 2,102-2,242 nm limited SAM classification of reflectance image showing kaolinite (R), illite (G) and montmorillonite (B) (max. angle 0.115/0.106/0.047 respectively, class map). Classified Cypris clays (A), less false positives (B) & areas changed to kaolinite (C).

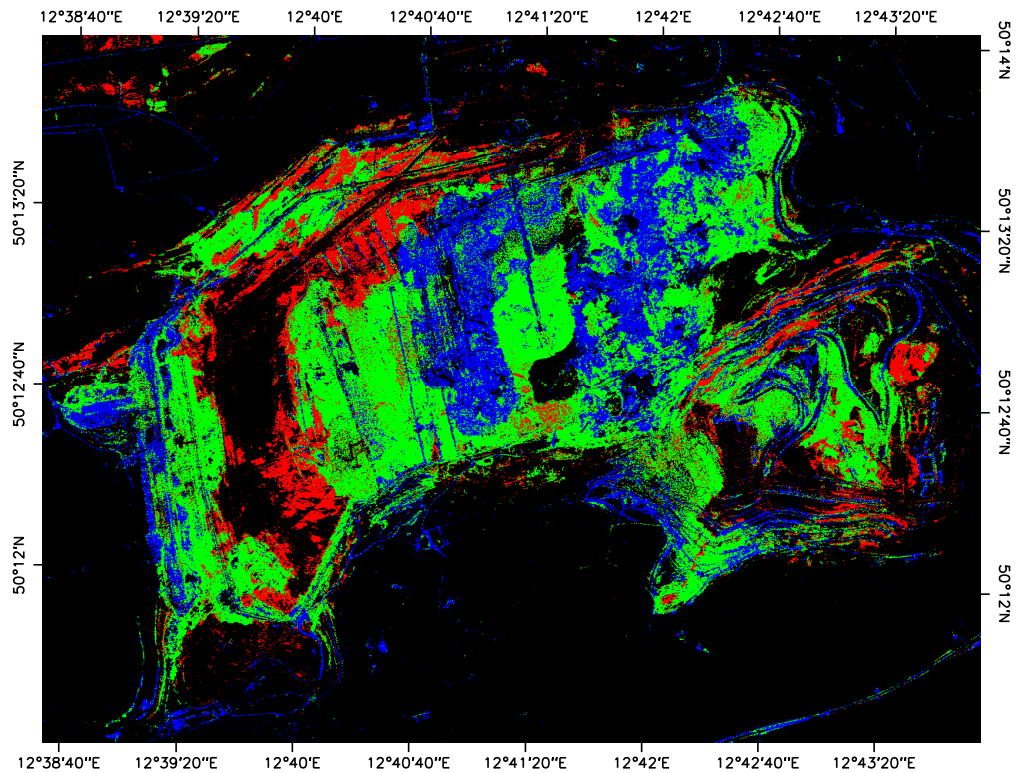


Figure 6.25.: CA3) 2,000-2,500 nm limited SAM classification on scale 2 of wavelet transformed image showing kao. (R), ill. (G) and mont. (B) (max. angle 0.862/0.736/0.793 respectively, class map).

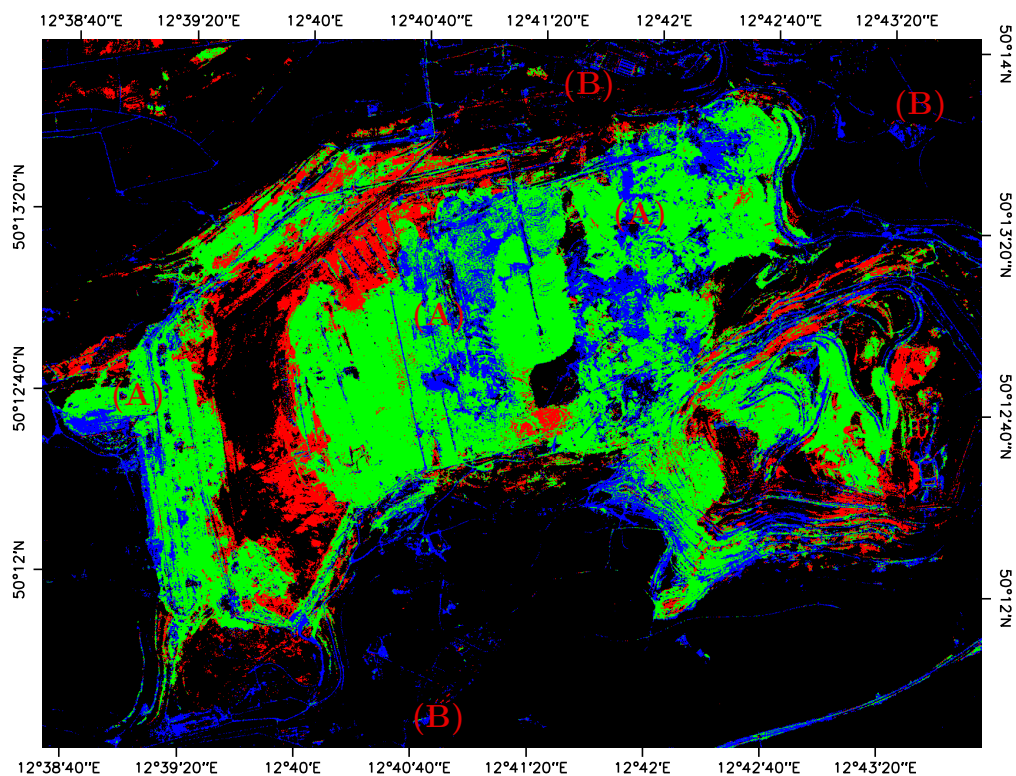


Figure 6.26.: CA4) 2,102-2,242 nm limited SAM classification on scale 2 of wavelet transformed image showing kao. (R), ill. (G) and mont. (B) (max. angle 0.773/0.514/0.719 respectively, class map). Transitions from montmorillonite to illite (A) & roof/glass false positives (B).



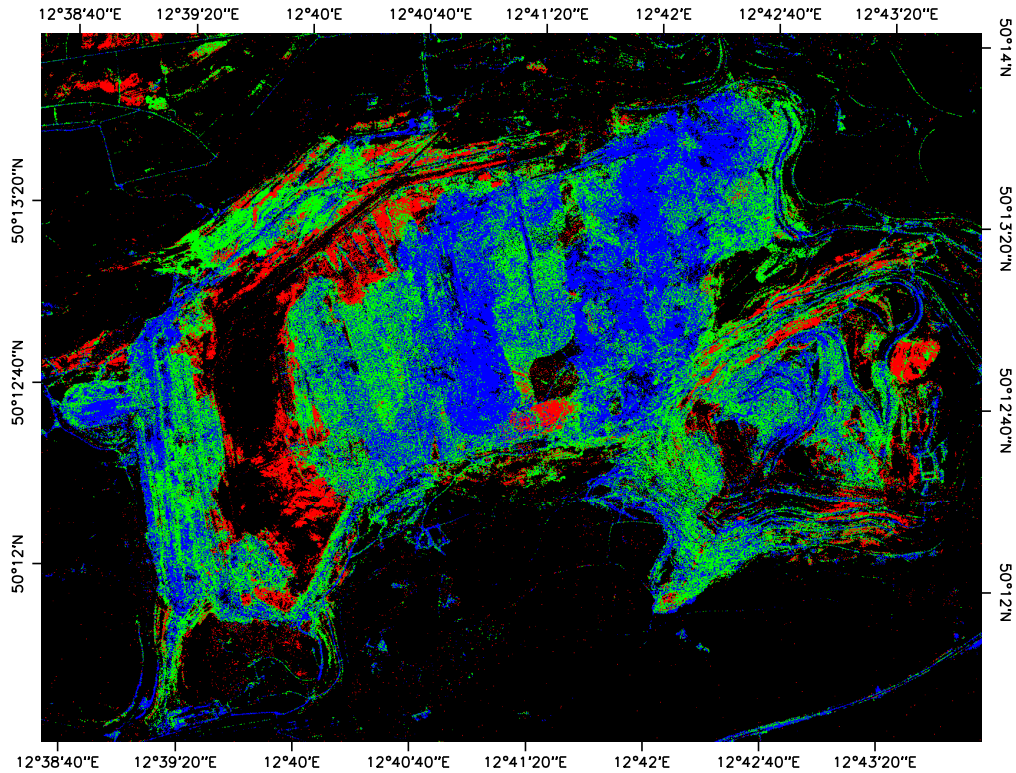


Figure 6.27.: **CA5**) 2,000-2,500 nm limited average coherence classification on scale 2 of image showing exposed kao. (R), ill. (G) and mont. (B) (threshold 0.524/0.596/0.579 respectively, class map).

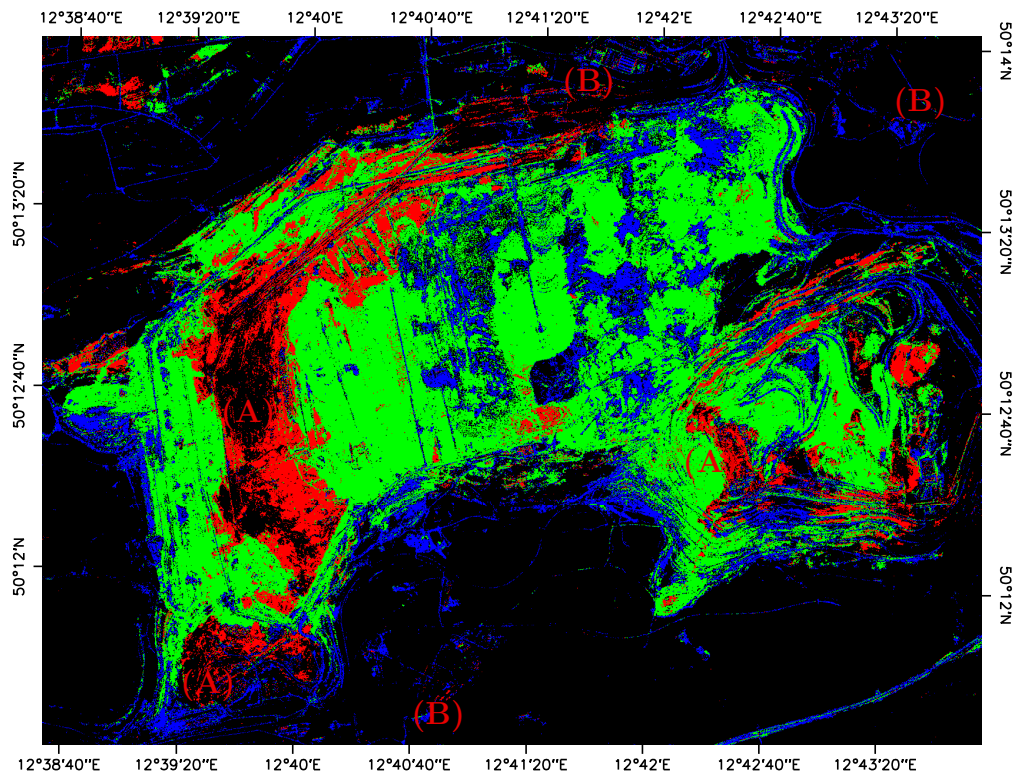


Figure 6.28.: **CA6**) 2,173 nm coherence classification on scale 2 of image showing exposed kao. (R), ill. (G) and mont. (B) (threshold 0.628/0.956/0.824 respectively, class map). Additional kaolinite assignments in coal (A) & roof/glass false positives (B).

Table 6.9.: Confusion matrix for kaolinite, illite and montmorillonite classified with Wavelet Coherence Mapper (scale 2, 2,173 nm) versus wavelet-based SAM classification (scale 2, 2,102-2,242 nm).

clays confusion matrix [%]					
SAM:2,102-2,242 nm vs. WCM:2,173 nm					
	SAM none	SAM kaol	SAM illi	SAM mont	total
WCM none	55.08	0.34	1.26	3.08	59.76
WCM kaol	2.04	4.62	0.14	0.00	6.80
WCM illi	0.67	2.08	20.40	0.71	23.86
WCM mont	3.87	0.00	0.00	5.71	9.59
total	61.66	7.04	21.79	9.50	100.00
overall accuracy = 86%, Kappa coefficient = 0.75					

montmorillonite for both spectral subsets. This strong weighting of the continuum was disadvantageous for the clay mixture classification in Sokolov.

Application of the  $\text{DOG}_{m=2}$  CWT allowed to suppress the continuum information (spectral slope) and to emphasise smaller scale features. The SAM classifications on scale 2 revealed the expected clay mineral distribution patterns for the Cypris formation overburden. Very similar clay mineral distributions resulted from the 2,000-2,500 nm and 2,102-2,242 nm subsets. This highlights SAM's bias towards strong features, i.e. the Al–OH absorption.

The Wavelet Coherence Mapper is biased towards curvature direction differences instead. Its wavelength localisation leads to a localised normalisation and hence equal weighting of features over the wavelength range. Therefore classification results differed for both spectral subsets. The 2,173 nm coherence classification compared well to the CWT scale 2 SAM approaches. All three mainly focused on the shape of the Al–OH absorption. The coherence appeared to have a considerably higher sensitivity for subtle clay admixtures inside the coal. In the 2,000-2,500 nm subset coherence reacted differently. It appeared to be less focused on the Al–OH absorption and to put weight also on the more subtle 2,000-2,100 nm and 2,250-2,500 nm features. The better match of montmorillonite template and image spectra found more consideration.

Comparing this spectral range limited kaolinite classifications to the full range experiment in Section 6.5.1 - *Wavelet based Classification of Siderite, Goethite & Kaolinite* the importance of a similarity measure localised in wavelength becomes apparent. Here the focus on the characteristic spectral subset allowed an area-filling classification relatively independent from spectral influences of other admixed minerals. Wavelet coherence allows to define/determine the relevant spectral subset after the calculation of the similarity measure. This eases investigative analysis and avoids iterative recalculation when the optimum is not known from the beginning.

Concluding, the use of coherence is proposed when the evolution of the similarity measure over wavelength is of benefit, when spectral curvature changes are indicative for the presence of a material, or when features of different amplitude in a spectral signature are equally characteristic (e.g. a strong deep feature paired with a shallow one at same transform scale). The latter equal weighting/sensitivity necessitates to limit investigations to characteristic wavelength regions. The adverse effect of included irrelevant spectral information is more pronounced than on SAM.

## 6.6. Mineral Quantification

Appendix A - *Acid Mine Drainage Supplements* lists minerals of interest for the evaluation of AMD affected areas. The geologic and climatic conditions of the Sokolov demonstration site in combination with the available data led to three minerals exposed in such abundances and spatial extent on the topographic surface that they were accessible to this quantitative study: kaolinite, goethite and quartz (see Section 2.2 - *Application Objective & Mineral Mapping Candidates*). In Section 2.3.5 - *Spectral Characteristics of relevant Minerals* their spectral features are introduced. The improved and validated pre-processing of the imagery, the 80 mineral soil samples taken in the field, their spectroscopic measurement in the laboratory and the normative modelling of the mineral fractions allowed for a reference data backed quantification of the three minerals.

Quartz does not contribute to the neutralisation of AMD. Hence quartz dominated soils can be an indicator of relatively low acidity buffering capacity. Its spectral characteristics allowed its quantification with the thermal AHS data. Quartz is not resolvable from the reflective data since no characteristic feature exists. The clay mineral kaolinite is the second most abundant mineral in the Sokolov samples. Its slow dissolution in acid environments provides some neutralisation potential. The characteristic kaolinite spectral features allow a quantification in both the reflective thermal imagery. Accumulations of the weathering-stable goethite are indicators for possibly AMD affected areas. It has characteristic features in the VIS and NIR that can be exploited for a quantification. For the TIR at AHS spectral resolution no characteristics exist.

In the following two sections the wavelet based quantification of goethite and kaolinite with the Wavelet Transform Feature Analysis is conducted on the hyperspectral reflective HyMap imagery and the band-ratio based quantification of quartz and kaolinite is performed on the multispectral thermal AHS imagery. A critical review of methodical and practical advantages and disadvantages is given in Chapter 7 - *Discussion & Summary*. Figure 6.11 displays the spatial subset which was chosen for visualisation and comparison purposes. It encompasses the two lignite open casts as were chosen for the classification.

### 6.6.1. Wavelet based Quantification of Goethite & Kaolinite

The hyperspectral reflective data allows the use of spectral analysis techniques taking advantage of the quasi-continuous signal. Rather than treating each spectral band independently, the Wavelet Transform Feature Analysis uses the signal's correlation over bands and wavelength.

This quantification experiment of the clay mineral kaolinite and the iron oxyhydroxide goethite give insight into the spatial distribution patterns of a slow neutralising acidity sink and a possible indicator for AMD affected areas. Methodically a practical application of the Wavelet Transform Feature Analysis is performed. Starting from the extraction of feature vectors, their sieving, clustering and parameterization, eventually a multivariate regression model for each mineral is established. The models are finally applied to the imagery using the common feature patterns as a filter. It demonstrates i) the choice of sieving and clustering parameters, ii) the impact of limiting to the spectral resolution of the airborne sensor and exclusion of atmospheric water absorption bands, and iii) the effect of feature pattern filtering.

**Data** Reflectance spectra derived from the HyMap imagery, from field reference measurements and from the laboratory measurements of the samples were available for developing the quantitative models. Validation in Sections 6.7.1.1 and 6.7.1.2 showed the problematics of transiting spatial scales and disturbance of the measured surface. Practically, weather conditions limited the number of available field measurements such that no meaningful model development was possible using field spectra. Model development on image derived spectra was refrained from due to the scale transitions from image pixel (16 m<sup>2</sup>) to collected field sample (500 g) and final geochemical analysis (5 g of homogenised sample). The laboratory spectra therefore were used in view of maximising the representativeness between spectral signatures and laboratory analysis. Hence, reflective quantitative models were developed on the 50 mineral soil samples that underwent laboratory geochemical analyses plus the corresponding 50 laboratory spectral measurements. The pre-processed mean reflectance spectra (2,150 bands) were resampled to the HyMap sensor resolution (120 bands) using the band response function parameters FWHM and central wavelength known from the HyMap laboratory calibration. The NORMA modelling results provided the estimated mineral abundances for the respective 50 samples.

**Correlation Screening** All normative minerals were tested for associations with the samples' spectral properties by a simple correlation screening technique. Here, the so called correlation scaleograms as introduced in Lyder *et al.* (2010) were used. The correlation scaleograms are formed by calculating the CWT for each input spectrum and stacking the resulting wavelet transform plots into a cube. An individual wavelet transform plot is a two-dimensional scale-wavelength representation of the CWT of one spectrum. Each profile through the cube lists the CWT values of all input spectra for the respective scale-wavelength combination. The dependence between the CWT values and the targeted abundances can be assessed with the standard correlation measures yielding a value for each scale-wavelength point which, repeated for all possible combinations, gives a two-dimensional representation of the correlation over scale and wavelength - the correlation scaleogram.

Elevated correlations were found for the three minerals kaolinite, goethite and siderite. They are relevant for studying AMD processes in Sokolov, occur in sufficient spatial extend for an airborne mapping and have an adequate abundance variability covered by the 50 samples. Figure 6.29 displays the correlation scaleograms for kaolinite and goethite. Table 6.10 lists Pearson's  $r$  and Spearman's  $\rho$  for scale-wavelength combinations yielding the highest positive and highest negative correlations to the NORMA abundances. The respective areas in the kaolinite and goethite correlation scaleograms are coloured red and green.

For both kaolinite ( $\rho^2 \approx 0.6$ ) and goethite ( $\rho^2 \approx 0.5$ ) promising correlation coefficients were reached and the respective scale-wavelength combinations were located in proximity to expected wavelengths that correspond to iron or clay features. Kaolinite in particular showed a stable localisation of highest correlations in scale and wavelength for both Pearson's  $r$  and Spearman's  $\rho$  indicating robust spectral features. Goethite and siderite were notably less stable especially in scale localisation. With scale 4 or above, most correlations were rather focused on the spectral continuum than on narrow absorption bands. This caused concern for the 2,300 nm/scale 8 indicated kaolinite correlation due to border effects.

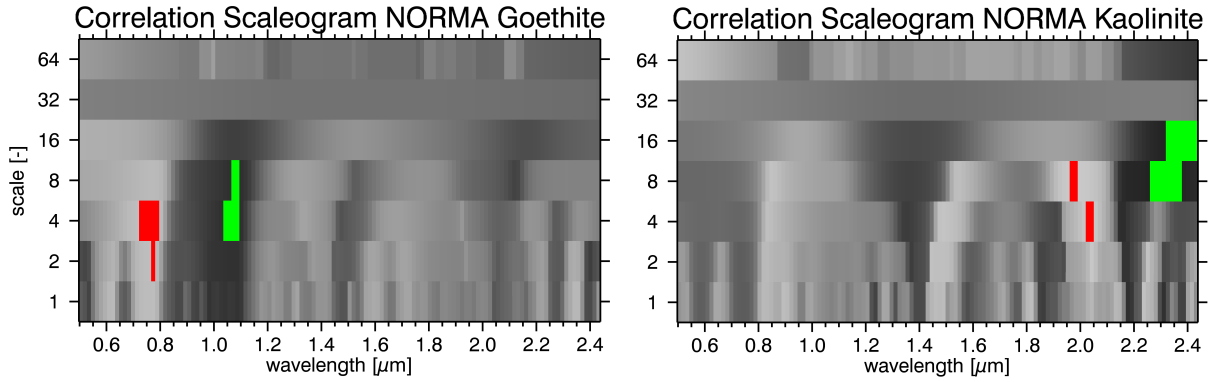


Figure 6.29.: Correlation scaleograms for NORMA goethite and kaolinite based on 50 HyMap resampled laboratory spectra. Grayscale ramp gives Pearson's  $r$  from  $-1$  (black) to  $+1$  (white). Scale-wavelength regions of highest positive (red) and negative (green) correlation are marked.

Table 6.10.: Dependence between NORMA main minerals and 50 ASD laboratory spectra for selected scale-wavelength combinations measured as Pearson's  $r$  and Spearman's  $\rho$ .

mineral	full ASD resolution (2150 bands)						HyMap resampled (120 bands)					
	$r$	wvl	scl	$\rho$	wvl	scl	$r$	wvl	scl	$\rho$	wvl	scl
Goethite	0.87	1.302	16	0.70	0.708	32	0.52	0.772	4	0.63	0.712	2
	-0.91	1.100	16	-0.73	2.398	32	-0.66	1.079	4	-0.68	1.094	8
Kaolinite	0.67	2.219	8	0.76	1.094	32	0.60	1.991	4	0.73	1.991	4
	-0.79	1.381	16	-0.82	1.382	16	-0.76	2.326	8	-0.79	2.310	8
Siderite	0.60	0.797	2	0.62	1.921	4	0.35	0.833	8	0.54	0.891	1
	-0.61	1.220	256	-0.67	1.315	16	-0.61	1.211	16	-0.65	1.197	8

**Wavelet Transform Feature Analysis** Based on above correlation screening and a manual inspection of the abundance versus CWT value scatter plots, kaolinite and goethite were chosen for setting up quantitative models. Figure 6.30 shows the distribution of NORMA abundances for kaolinite and goethite indicating two samples of comparably high goethite abundance  $> 50$  wt-%. XRD labeled both samples (MIX\_LO\_01, LO10\_03) as primary goethite with secondary fractions of jarosite. As they led to erroneous high correlation indications both samples were excluded from the model development. During validation however the final goethite regression model was tested for conformity to the outliers. Virtually no correlation between both minerals with  $r^2 = 0.02$  was present. Inter-correlations of goethite and kaolinite to NORMA quartz are given as reference showing medium inverse relation due to the physical sum-to-100 wt-% constraint with  $r^2 = 0.12$  and  $r^2 = 0.13$  respectively. In order to bridge the gap between the low and (excluded) two high wt-% goethite samples also  $\text{Fe}_2\text{O}_3$  was used for development of an abundance model. Latter represents a sum of NORMA modelled goethite and jarosite abundances. The additional samples of  $\geq 20$  wt-%  $\text{Fe}_2\text{O}_3$  allowed to establish a non-extrapolating model towards higher abundances in contrast to goethite which was limited to  $\leq 15$  wt-%. Inter-correlations of  $\text{Fe}_2\text{O}_3$  to goethite/kaolinite/quartz were  $r^2=0.71/0.09/0.16$  respectively.

Using the method described in Section 3.3.3 - *Wavelet Transform Feature Analysis* an overall workflow as depicted by Scheme 3.11 was followed. First each sample's mean reflectance spectrum was wavelet transformed, the transform features extracted at each scale and chained into feature vectors tracing maxima, minima, upward and downward zero crossings. Spurious feature vectors attributing to noise were filtered by requiring the vectors to persist over at least 2 scales.

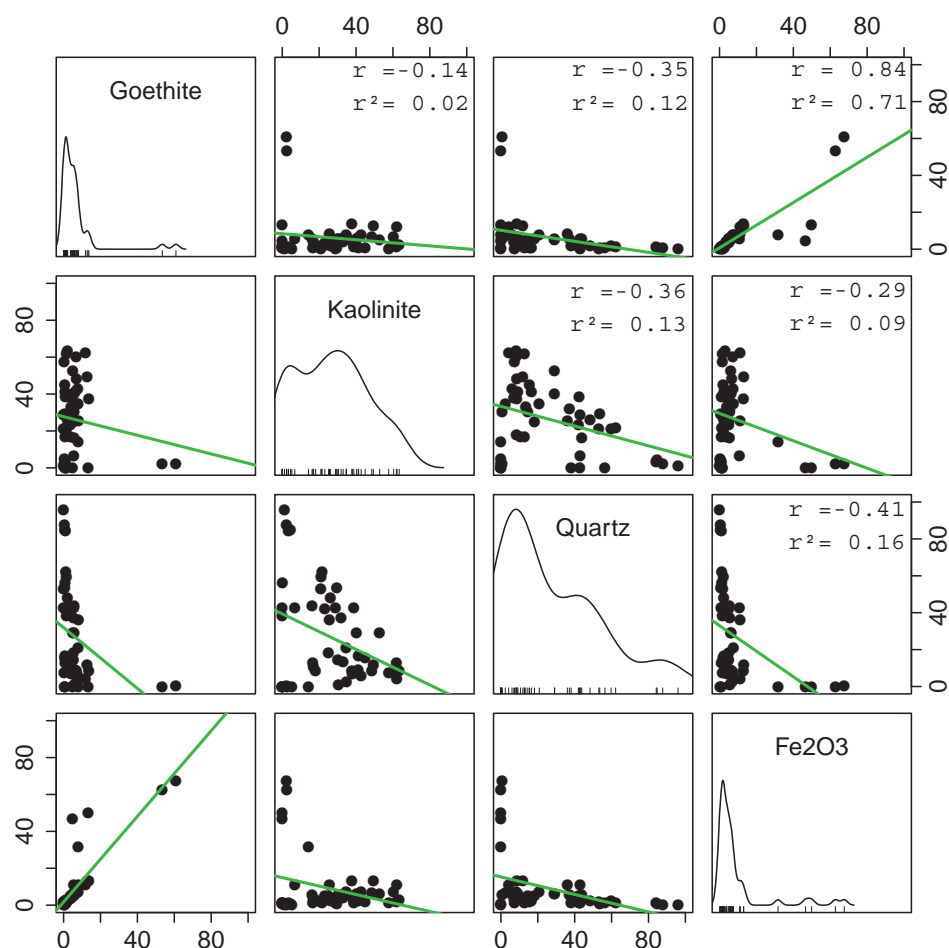


Figure 6.30.: Scatter plot matrix showing inter-correlations and value distributions [wt-%] of selected NORMA mineral / elemental abundances for all 50 samples that underwent lab spectroscopy.

The top section of Figure 6.31 gives an example of the extracted transform maxima for one spectrum and the overlain maxima vectors of all 50 samples.

The ensuing semi-automatic clustering of the feature vectors into common vectors avoided the need for manual definition of wavelength ranges for feature searching and possible bias introduced by such an approach. Treating the four feature vector types separately, the GDBSCAN clustering algorithm was generally successful operating on a two-dimensional feature space only. The feature space is spanned by the feature vector's wavelength localisation at scale 1 ( $wvl$ ) and the spectrum id ( $sid$ ) it belonged to (see lower left of Figure 6.31). As expected, the wavelength neighbourhood described by parameter  $\epsilon_{wvl}$  was most distinctive followed by  $\epsilon_{sid}$  and  $minCard$  describing the completeness of the set.  $\epsilon_{val}$  was chosen sufficiently large to not impose any constraint because no clusters formed in the transform value ( $val$ ) - wavelength ( $wvl$ ) feature space (see lower right of Figure 6.31).

Nevertheless, the clustering parameters needed fine tuning beyond the methods presented in Section 3.3.3 - *Wavelet Transform Feature Analysis* and a manual merge of splinter clusters as summarised by Table 6.11. Reasons for the very relaxed specifications of  $\epsilon_{sid}$  and  $minCard$  was the overall good wavelength grouping of the feature vectors separated by distinct void. In some cases multiple features of the same type in one spectrum assigned the same cluster arose from the manual merge or from diffuse clusters spreading over an extended wavelength range. These conflicts were resolved by keeping the feature closest to the cluster's mean wavelength.

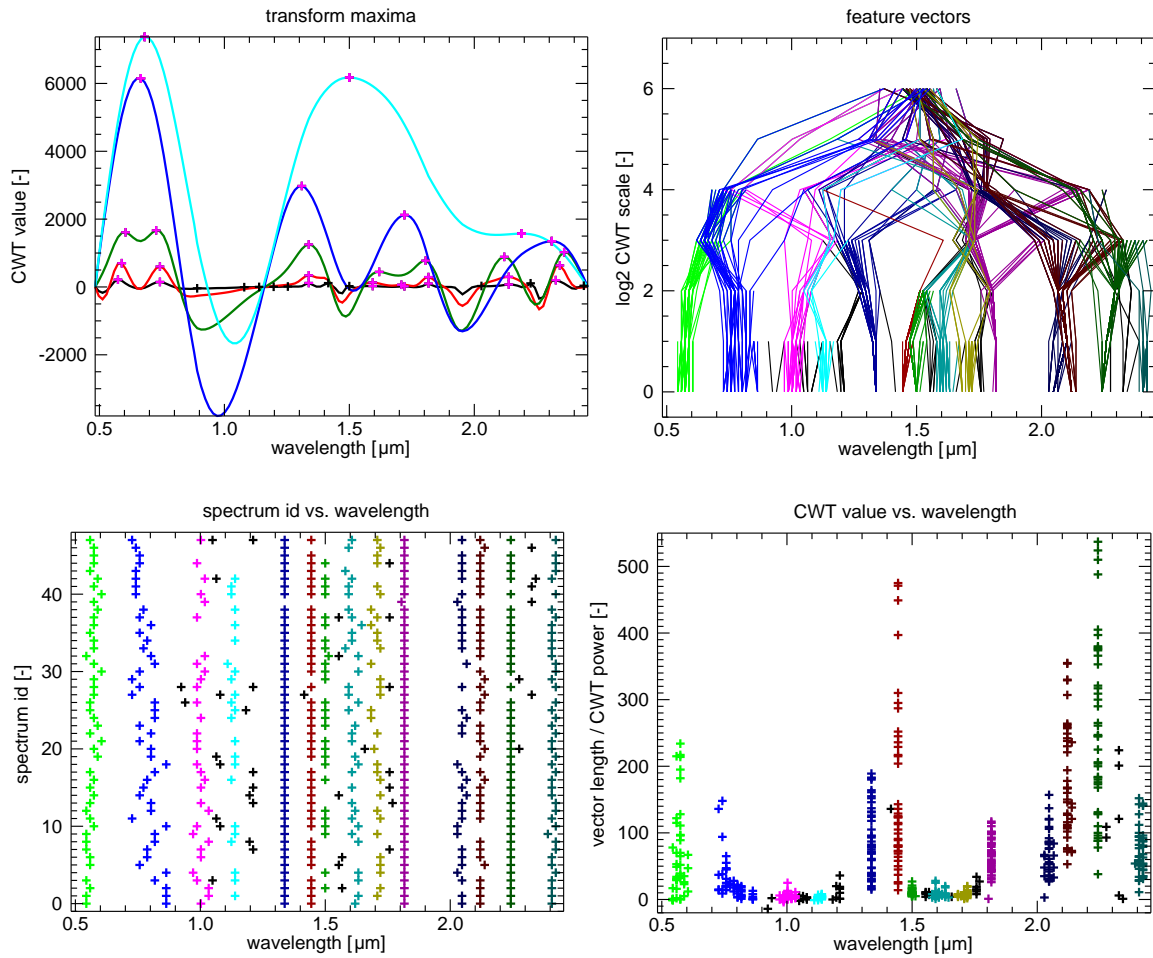


Figure 6.31.: Extracted transform maxima for multiple scales of one spectrum (upper left). Maxima belonging to a vector fulfilling the length threshold marked as magenta cross or black cross otherwise. Clustering result of maxima vectors extracted from all 50 spectra showing common maxima vectors (upper right) in same color together with sid/wvl (lower left) and val/wvl (lower right) feature spaces. Noise & not clustered features coloured black.

Clustering of the 50 Sokolov mineral soil samples resulted in 16 common features for upward zero crossings, 14 for downward zero crossings, 16 for maxima and 17 for minima, hence a total of 63. Figure 6.31 gives an example of the extracted transform maxima for one spectrum and the clustering result of the maxima vectors of all 50 samples.

Each of the 63 common features represents a distinct spectral feature which is present in the majority of the 50 input spectra. Hence, each common feature contains up to 50 individual feature vectors. Recalling, the feature vectors were extracted from the individual spectra and each represents one maximum, minimum, downward or upward zero crossing. Depending on the

Table 6.11.: Parameters used for the GDBSCAN clustering of the transform feature vectors into common features and mean wavelengths of manually merged clusters.

transform feature type	clustering parameters			
	$minCard$	$\epsilon_{sid}$	$\epsilon_{wvl}$	manual cluster merge
upward zero crossings	5	20	1	1,116/1,188/1,247 nm
downward zero crossings	5	20	1	1,673/1,719 nm
maxima	10	20	1	774/863 nm
minima	10	20	1	892/956 nm

type of transform feature, the following parameters were calculated from each feature vector: feature wavelength at scale, feature power at scale, feature wavelength drift at scale, feature width, feature amplitude, feature area and feature significance (see Section 3.3.3 - *Wavelet Transform Feature Analysis*).

In a next step following the parameterization, linear regression models were fitted that linked the parameter values to the NORMA modelled mineral abundances. This was repeated for all 63 common features. Parameter values of spectra not showing the required transform feature defaulted to 0, except for the feature wavelength parameter. The latter defaulted to the median wavelength of all other vectors at the respective scale. A maximum of 6 spectra not showing the transform feature were allowed. The threshold of 6 was motivated by 6 samples with a NORMA modelled abundance  $< 0.5$  wt-% for kaolinite (5 for goethite). Another 2 ( $\approx 5\%$ ) samples were tolerated that contained the required transform feature but which did not persist up to the required scale. Note that a minimum persistence over 2 scales was imposed in the initial noise filtering. Their parameter values were filled with above defaults. These constraints together with the 7 scales resulted in a total of 400 transform feature vector parameter combinations that were tested for correlation with the NORMA modelled mineral abundances.

Combinations that yielded both a Pearson's product-moment correlation and Spearman's rank correlation above a certain threshold were regarded candidates for setting up a Multiple Linear Regression (MLR) model for abundance estimation. Simply speaking, Pearson's  $r$  assesses the linear relationship between two variables assuming both are distributed normally. By contrast, Spearman's  $\rho$  assesses their monotonic relationship not requiring a certain distribution and treating the data as ranked, i.e. converting from interval to ordinal scale, and thus reducing the influence of outliers. Demanding  $r$  and  $\rho$  to both satisfy the threshold, a linear and monotonic relationship between feature parameter and abundance can be assured to a certain degree.

The correlation threshold was empirically chosen 0.5 for kaolinite and 0.4 for goethite resulting in 13 transform feature parameters for each mineral. They pooled the highest indicated dependence and are candidates for building the quantitative MLR model on. The straightforward link between wavelet transform features and spectral features allows to easily interpret the retrieved 13 parameters (cf. Section 2.3.5 - *Spectral Characteristics of relevant Minerals*).

In case of kaolinite, the 13 parameters of highest correlation belonged to six transform features explained in the following. The upward zero crossing around 1,992 nm corresponds to the long wavelength inflexion point on the 1,900 nm water absorption flank overlain by the weaker 1,960 nm carbonate absorption. The 1,366 nm downward zero crossing relates to the short wavelength inflexion point of the 1,400 nm hydroxyl and water absorption. The maximum transform feature at 1,338 nm corresponds to the short wavelength shoulder of the 1,400 nm absorption. Three transform minima features around 914 nm, 1,953 nm and 2,206 nm relate to the ferric ion ( $\text{Fe}^{3+}$ ) absorption, 1,900 nm water/1,960 nm carbonate absorption and 2,200 nm Al–OH absorption, respectively.

The 13 parameters for goethite belonged to five transform features. The transform maximum around 568 nm represents the long wavelength shoulder of the wide  $\text{Fe}^{3+}$  absorption in the ultra-violet. The transform maximum around 787 nm corresponds to the short wavelength shoulder of



overlying ferric ion (860-900 nm) and ferrous ion (1,100-1,300 nm) absorptions. The 2,413 nm transform maximum appears related to the long wavelength shoulder of the 2,300-2,350 nm carbonate absorption. The 917 nm transform minimum relates to the ferric ion absorption while the 1,756 nm minimum may relate to an Al–OH absorption shown in Clark (1999) for alunite.

Obviously, many of the aforementioned transform features do not link directly to the targeted mineral or can not be used on airborne imagery. Features likely masked or strongly affected by the broad atmospheric water absorption bands at 1,400 and 1,900 nm had to be dropped. The inter-correlation tests (see Figure 6.30) between goethite/Fe<sub>2</sub>O<sub>3</sub> and kaolinite indicated no association. Features not directly linked to the targeted mineral (proxies) were therefore dropped too. Validation of the laboratory spectra against corresponding image spectra revealed mismatching spectral ranges that also had to be excluded (see Section 6.7.1.1). Note that the feature based approach of Wavelet Transform Feature Analysis allows to easily implement this selective omission of inadequate features. This is an important benefit over statistical methods like Partial Least Squares, Support Vector Machines or Linear Discriminant Analysis.

The necessary omissions well illustrate issues arising from transferring a laboratory spectra calibrated model onto image data. Here, kaolinite with its hydroxyl bonds was especially affected. The compelled exclusion of well correlating feature vector parameters and admittance of inferior ones (correlation threshold lowered to 0.3) led to a decreased kaolinite abundance model performance. Tables 6.12 and 6.13 list the surviving best correlating feature parameters, grouped by feature type and identified by their mean wavelength, scale and parameter type. Given are their correlation to the targeted abundance and their inter-correlations amongst each other.

The remaining vector parameters were investigated for inter-correlations to avoid collinearity of the predictor variables which may cause instabilities and redundancy in the MLR model. Dependent vector parameters with  $r$  and  $\rho \geq 0.8$  were resolved following two rules:

1. High correlations between different transform feature types (MAX, MIN, DZC, UZC) or between different common features (differing wavelength) were ignored. Hence, the respective vector parameters were kept.
2. High correlations within the same feature were resolved by keeping only the one parameter yielding the highest correlation. In case of a close match between multiple parameters the simpler one was kept, thus dropping composite parameters such as area or amplitude.

Finally, a statistical cross-validation was conducted on the surviving individual vector parameters (see Section 6.7.1.4 - *Reflective Domain Quantification — Model Cross-Validation & Plausibility Checks*). The parameters related to the identified instable linear regressions were excluded. In summary, merely five vector parameters were retained for the kaolinite quantification and only four parameters passed the tests for goethite (see Table 6.15).

While the kaolinite 1,992 nm upward zero crossing and the 787 nm goethite maximum corresponded to the findings of the correlation scaleogram analysis (correlation screening), the indicated scales were lower and the major vector parameter was wavelength. This highlights an important extension brought by the Wavelet Transform Feature Analysis allowing the retrieval of additional absorption feature describing parameters. The wavelength localisation of spec-

Table 6.12.: 2,000-2,500 nm transform feature parameters yielding correlations  $|r| \wedge |\rho| \geq 0.3$  to NORMA kaolinite abundance. Parameter inter-correlations  $\geq 0.8$  marked yellow. Feature types are upward and downward zero crossings (UZC and DZC), maxima (MAX) and minima (MIN). Parameter types are wavelength (WV), power (PO), area (AR), amplitude (AP), width (WT), drift (WD) and significance (SI). Scale  $s=X$  is given as  $sX$ .

	abundance	DZC								MIN								
		2144nm-s1-WV		2144nm-SI		2454nm-SI		2206nm-AR		2206nm-s2-PO		2206nm-s2-WD		2206nm-s4-PO		2206nm-SI		
		r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	
DZC	2144nm-s1-WV	-0.37	-0.44															
	2144nm-SI	0.46	0.63	-0.43	-0.37													
	2454nm-SI	0.33	0.46	-0.41	-0.50	0.69	0.69											
MIN	2206nm-AR	0.40	0.44	-0.46	-0.49	0.70	0.77	0.46	0.58									
	2206nm-s2-PO	-0.37	-0.42	0.43	0.46	-0.59	-0.75	-0.44	-0.58	-0.95	-0.98							
	2206nm-s2-WD	-0.35	-0.50	0.28	0.33	-0.19	-0.49	-0.36	-0.55	-0.26	-0.53	0.38	0.57					
	2206nm-s4-PO	-0.42	-0.51	0.52	0.56	-0.69	-0.78	-0.46	-0.59	-0.98	-0.98	0.96	0.97	0.29	0.57			
	2206nm-SI	0.59	0.60	-0.35	-0.45	0.53	0.69	0.41	0.57	0.72	0.86	-0.77	-0.86	-0.42	-0.62	-0.75	-0.88	
	2206nm-WD	0.56	0.50	-0.33	-0.42	0.61	0.79	0.46	0.63	0.79	0.76	-0.70	-0.67	-0.14	-0.43	-0.73	-0.74	0.69

Table 6.13.: 400-1,300 nm transform feature parameters yielding  $|r| \wedge |\rho| \geq 0.4$  to NORMA goethite abundance. See description of Table 6.12.

	abundance	MAX								MIN										
		568nm-WT		568nm-s2-WV		787nm-s1-PO		787nm-s1-WV		787nm-s2-AR		787nm-s2-PO		787nm-s2-WV		917nm-s1-WV		917nm-s2-WV		
		r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	r	$\rho$	
MAX	568nm-WT	0.49	0.57																	
	568nm-s2-WV	0.44	0.58	0.27	0.31															
	787nm-s1-PO	0.43	0.43	0.31	0.52	0.17	0.27													
	787nm-s1-WV	-0.59	-0.62	-0.57	-0.52	-0.50	-0.48	-0.54	-0.68											
	787nm-s2-AR	0.41	0.41	0.35	0.45	0.17	0.29	0.97	0.95	-0.59	-0.68									
	787nm-s2-PO	0.45	0.45	0.35	0.53	0.19	0.29	1.00	0.99	-0.59	-0.70	0.98	0.96							
	787nm-s2-WV	-0.60	-0.65	-0.63	-0.57	-0.54	-0.52	-0.61	-0.76	0.94	0.95	-0.66	-0.75	-0.66	-0.79					
MIN	917nm-s1-WV	0.47	0.60	0.44	0.50	0.20	0.32	0.01	0.41	-0.25	-0.24	0.04	0.34	0.06	0.42	-0.28	-0.30			
	917nm-s2-WV	0.40	0.52	0.41	0.44	0.16	0.22	-0.11	0.28	-0.20	-0.20	-0.06	0.26	-0.06	0.31	-0.23	-0.23	0.94	0.88	
	1756nm-s1-PO	0.49	0.51	0.29	0.34	0.19	0.27	0.02	0.09	-0.33	-0.32	0.03	0.08	0.05	0.11	-0.30	-0.28	0.16	0.23	0.20

Table 6.14.: 400-1,300 nm transform feature parameters yielding  $|r| \wedge |\rho| \geq 0.3$  to  $Fe_2O_3$  abundance. See description of Table 6.12.

	abundance	MAX				MIN		
		568nm-s1-PO		568nm-s2-PO		914nm-AP		
		r	$\rho$	r	$\rho$	r	$\rho$	
MAX	568nm-s1-PO	0.55	0.35					
	568nm-s2-PO	0.50	0.35	1.00	0.99			
MIN	914nm-AP	0.62	0.38	0.60	0.59	0.58	0.59	
	914nm-s2-PO	-0.76	-0.38	-0.86	-0.86	-0.84	-0.86	-0.75

Table 6.15.: Final feature parameters used for kaolinite, goethite and  $\text{Fe}_2\text{O}_3$  MLR models. For each parameter a short interpretation is given, followed by the individual linear regression model's training and cross-validation error. Localisation of absorptions & absorption shoulders defined by wavelength of maximum curvature. See Table 6.12 for parameter naming convention.

mineral	feature parameter	interpretation	RMSE [wt-%]	
			train.	vali.
kaolinite	DZC-2144nm-s1-WV	wavelength of the inflection point on the short-wavelength flank of the Al–OH absorption	18.35	19.17
	MIN-2206nm-s2-PO	similarity of the absorption to a second derivative of Gaussian of width $\sigma = 2$ ( $\approx 4$ sampling distances)	18.31	18.74
	MIN-2206nm-s2-WD	asymmetry of the absorption (scale 2 excludes the spectral surrounding $\gtrsim 3$ sampling distances (COI))	18.45	20.31
	MIN-2206nm-SI	significance of the absorption, i.e. contrast / how well it is developed	15.92	16.19
	MIN-2206nm-WT	width of the Al–OH absorption	16.35	16.63
goethite	MAX-568nm-s2-WV	wavelength of the long-wvl shoulder of the smoothed UV $\text{Fe}^{3+}$ absorption	3.30	3.33
	MAX-787nm-s2-PO	similarity of the short-wavelength shoulder of the NIR $\text{Fe}^{2+}/\text{Fe}^{3+}$ absorption to a second derivative of Gaussian of width $\sigma = 2$	3.37	3.50
	MAX-787nm-s2-WV	wavelength of the smoothed NIR $\text{Fe}^{2+}/\text{Fe}^{3+}$ absorption	3.01	3.06
	MIN-914nm-s1-WV	wavelength of the $\text{Fe}_3^+$ absorption	3.33	3.40
$\text{Fe}_2\text{O}_3$	MAX-568nm-s1-PO	similarity of the long-wvl shoulder of the UV $\text{Fe}^{3+}$ absorption to a 2nd derivative of Gaussian of width $\sigma = 1$	12.94	13.81
	MIN-914nm-AP	depth of the NIR $\text{Fe}^{3+}$ absorption	12.15	13.00
	MIN-914nm-s2-PO	similarity of the $\text{Fe}^{3+}$ absorption to a 2nd deriv. of Gaussian of width $\sigma = 2$	10.02	10.46

tral curvature features, i.e. information beyond CWT power at a fixed scale and wavelength, apparently was of relevance to kaolinite and especially goethite quantification.

Finally, two quantitative models were set up by a multiple linear regression (MLR) using the five kaolinite, respectively four goethite, individual vector parameters as predictor variables. The full 50/48 samples training set was used. Figure 6.34 depicts the overall model quality, yielding  $r^2 = 0.52$  over the abundance range  $[0 \dots 65]$  wt-% with a training error of about 14 wt-% RMSE for kaolinite and  $r^2 = 0.50$  over the abundance range  $[0 \dots 15]$  wt-% with a training error of 3 wt-% RMSE for goethite. The prediction accuracy was estimated by model cross-validation to 16 wt-% and 3 wt-% RMSE respectively (see Section 6.7.1.4).

The scatter plot of NORMA modelled versus calculated kaolinite abundance (Figure 6.34, top left) showed overestimation of small abundances, while high abundances appeared to be underestimated. The goethite scatter plot (Figure 6.34, top right) can be interpreted similarly but needs more samples of high and low abundance for a more definite answer. Several samples of negative calculated abundance are apparent for kaolinite and goethite which however lie in-

side the model's training accuracy. The goodness of prediction of the MLR models ( $r^2 \approx 0.5$ ,  $RPD \approx 1.4$ ) ranks at the lower end of medium quality models. Chang *et al.* (2001) rate soil constituent models with a  $RPD \in [1.4 \dots 2.0]$  and  $r^2 \in [0.5 \dots 0.8]$  as medium quality category.

Validation of the goethite model against the two previously excluded high abundance samples and library endmember spectra revealed a severe underestimation of high goethite fractions by the model (see Section 6.7.1.4 - *Reflective Domain Quantification — Model Cross-Validation & Plausibility Checks*). It rendered not applicable for abundances beyond the  $[0 \dots 15]$  wt-% range.

To overcome this limitation, a  $\text{Fe}_2\text{O}_3$  abundance model was established following the same procedures as applied for kaolinite and goethite. The full 50 samples training set and wet chemistry elemental concentrations were used. Three individual feature parameters (see Tables 6.14 and 6.15) resulted for the  $\text{Fe}_2\text{O}_3$  MLR model, which belonged to the 568 nm transform maximum and 914 nm minimum just like the goethite model. The 787 nm maximum however did show no correlation. The bottom plot of Figure 6.34 depicts the overall model quality, yielding  $r^2 = 0.63$  over the abundance range  $[0 \dots 70]$  wt-% with a training error of about 10 wt-% RMSE. The prediction accuracy was estimated 11 wt-% RMSE. Several samples of negative calculated abundance are apparent. They lie within the model's training accuracy. Notably, the two high NORMA goethite abundance samples MIX\_LO\_01 and LO10.03 appeared again as outliers, while three high NORMA jarosite abundance samples LO10.05A, LO10.05B and PVS10.01 plotted near the 1-to-1 line. The goodness of prediction measures  $r^2 \approx 0.6$  and  $RPD \approx 1.6$  ranked the  $\text{Fe}_2\text{O}_3$  model of medium quality.

Application of the trained kaolinite and  $\text{Fe}_2\text{O}_3$  MLR abundance models to the HyMap imagery made use of the common feature pattern matching described in Section 3.3.3. Scale-wavelength patterns for the two common features at 2,144 nm (downward zero crossing) and 2,206 nm (minimum) that were used by the kaolinite MLR model were determined from the training set. The same was done for the two common features at 568 nm (maximum) and 914 nm (minimum) used by the  $\text{Fe}_2\text{O}_3$  MLR model. Figure 6.35 displays the four common feature patterns in the scale-wavelength space. The pattern's left and right wavelength boundaries were defined by the maximum/minimum wavelength span of the 50 contained feature vectors (training set), relaxed by a margin of  $\pm 0.5$  times the standard deviation. The lower and upper scale boundaries were defined by the shortest and longest feature vector in the training set.

Iterating over all image pixels, the individual pixel spectra were wavelet transformed, the transform features extracted, chained into vectors and compared to the respective two patterns for kaolinite and  $\text{Fe}_2\text{O}_3$ . Image pixels not showing the required two feature vectors or their vectors not conforming to the scale-wavelength constraints of the patterns were rejected and flagged with a background value. The feature vectors of the pixels passing the tests were parameterized and quantified according to the respective MLR model. Some water pixels were quantified despite the pattern matching. This occurred where random noise in the water spectra mimicked the targeted absorption features. A simple water mask using a reflectance threshold of HyMap band #84 (1,708 nm)  $\geq 0.1$  was applied. Figures 6.32 and 6.33 show the quantification results for the active mines subset.

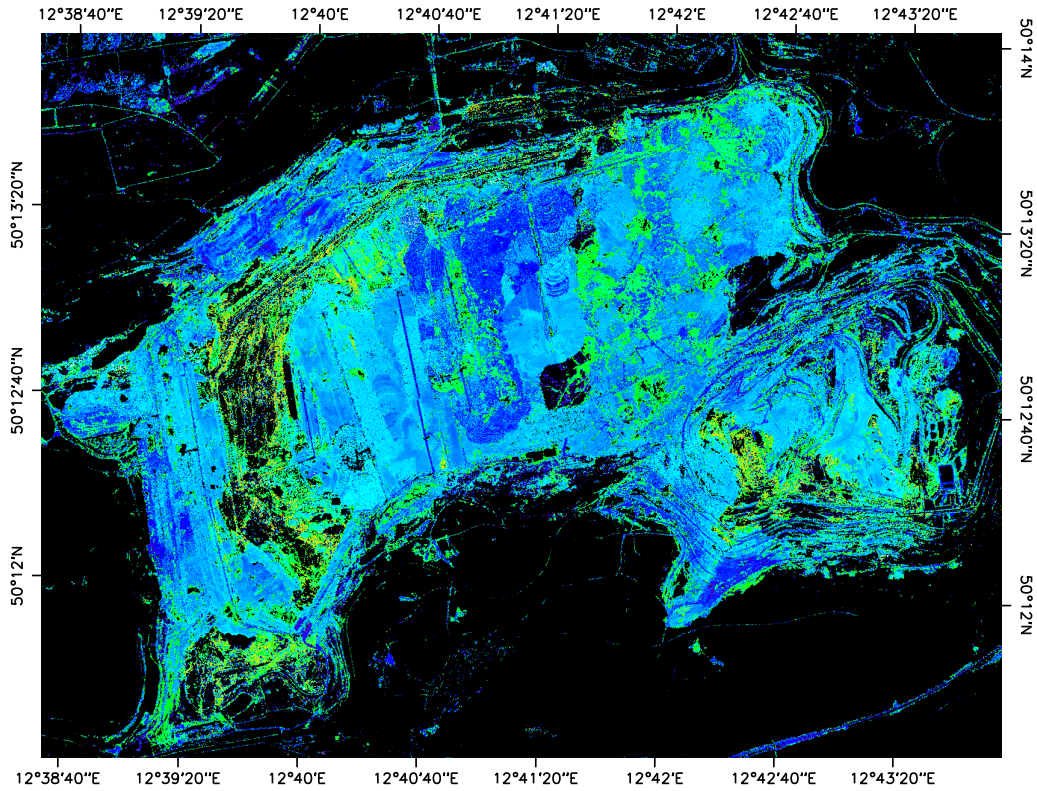


Figure 6.32.: Quantitative kaolinite map based on HyMap Wavelet Transform Feature Analysis derived MLR model (rainbow color ramp, 0-75 wt-%).

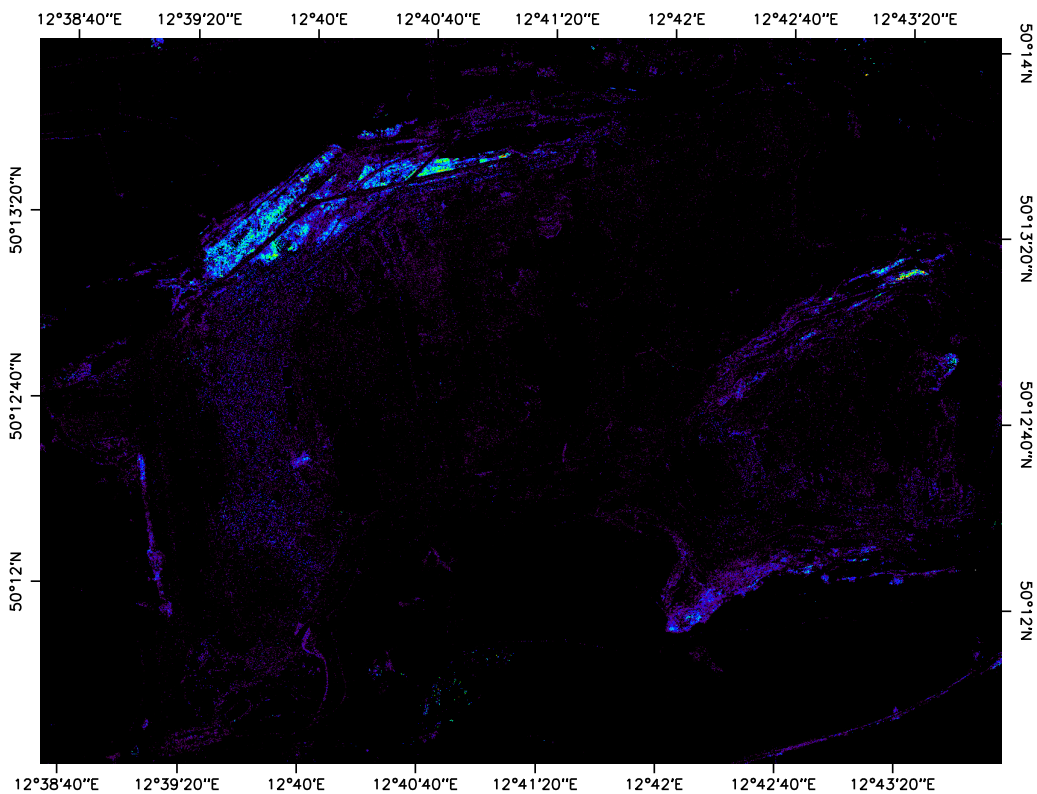


Figure 6.33.: Quantitative  $Fe_2O_3$  map based on HyMap Wavelet Transform Feature Analysis derived MLR model (rainbow color ramp, 0-40 wt-%).

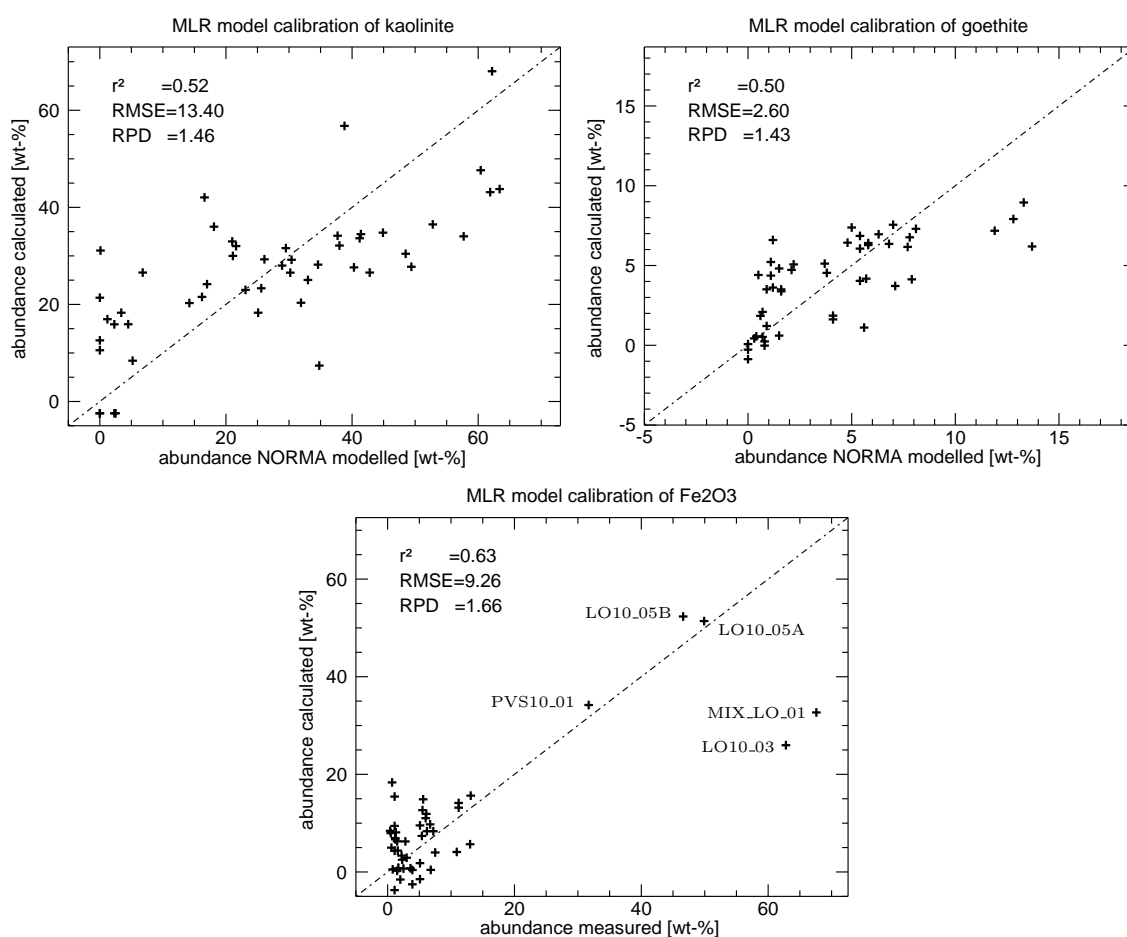


Figure 6.34.: MLR model quality for kaolinite (top left), goethite (top right) and  $\text{Fe}_2\text{O}_3$  (bottom). Shown are calculated versus NORMA modelled mineral abundances or measured element fraction for the 50/48/50 sample sets.

The estimated kaolinite abundance of the Cypris formation overburden and backfill was typically in the range 15-30 wt-% with a mean of 27 wt-%. High kaolinite abundances of locally up to 40-50 wt-% were indicated in the coal seams. Elevated coal abundances were also present for other areas appearing dark in the visible region, e.g. coal-clay mixtures along the coal transportation paths, but not for coal stockpiles.

The  $\text{Fe}_2\text{O}_3$  map showed elevated abundances at the northern and southern flanks of the two active mines (10 wt-% mean, locally up to 25-40 wt-%). A weak accumulation (4 wt-% mean) was apparent in the topmost brown-yellowish layer of the Cypris formation. The deeper blue-brownish layers of the Cypris formation overburden and backfill showed practically no  $\text{Fe}_2\text{O}_3$  (mean < 1 wt-%). The coal seams appeared with a mean  $\text{Fe}_2\text{O}_3$  content of 3 wt-%, locally ranging up to 10 wt-%. These are naturally occurring iron enrichments known from stratigraphy and field surveys.

The small goethite hotspot close to a pool of water that was identified in Section 6.5.1 - *Wavelet based Classification of Siderite, Goethite & Kaolinite* (marked (1) in Figure 6.16) was not quantified. The hotspot's image spectra did not conform the required  $\text{Fe}_2\text{O}_3$  common feature patterns. Figure 6.37 compares average image spectra of a goethite rich sampling site that was used in the training set hence spanning the patterns, iron rich soil spectra that conformed the patterns, and the goethite precipitate at the pool which did not conform to the patterns. It can be observed

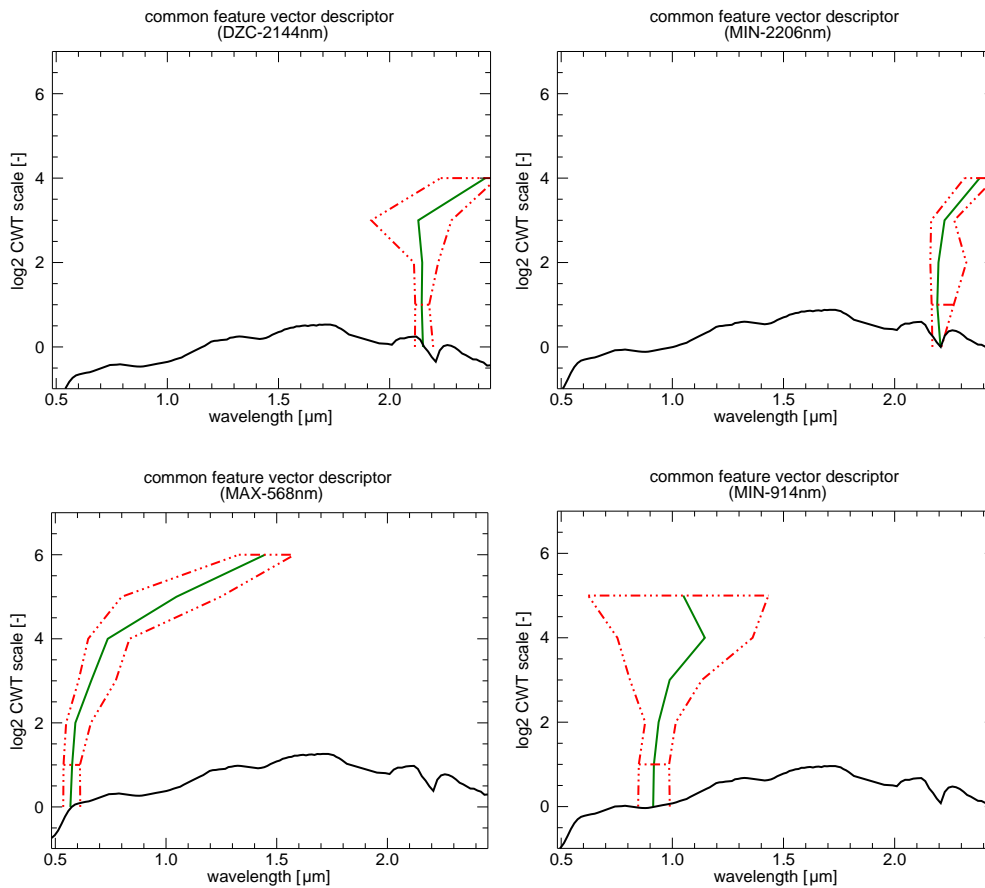


Figure 6.35.: Common feature patterns for kaolinite (top row) and  $Fe_2O_3$  (bottom row) used by the image pixel matching. Shown are a sample spectrum (black) with the mean (green) and maximum/minimum wavelength (red) of the respective common feature over the transform scales.

that a maximum curvature in the latter precipitate spectra occurs at 620 nm. In comparison to the sampling site's and soil's maxima located at a wavelength of about 568 nm, this is considerably shifted and therefore not meeting the required pattern. As a consequence was the area of the sampling site successfully quantified while the precipitate at the hotspot was not (see Figure 6.36). The pattern matching prevented application of the  $Fe_2O_3$  model to a material considerably deviating from the training data.

In summary, standardized pre-processing of the hyperspectral reflective imagery and laboratory spectrometric measurements of the collected samples together with the normative modelling allowed for a successful quantification of  $Fe_2O_3$  and kaolinite. The multivariate quantification models were based on common transform features in the scale-wavelength space which could be directly related back to spectral features such as absorptions, inflection points or absorption shoulders. The Wavelet Transform Feature Analysis made use of the multi-scale decomposition of the spectra. This allowed to test a large number of properties for correlation with the training abundances, while keeping their interpretation straightforward. Limiting the regressions to few best correlating parameters avoided model overfit. The final features and parameters used for the quantification matched spectral properties known from literature to indicate presence of kaolinite and  $Fe_2O_3$ .

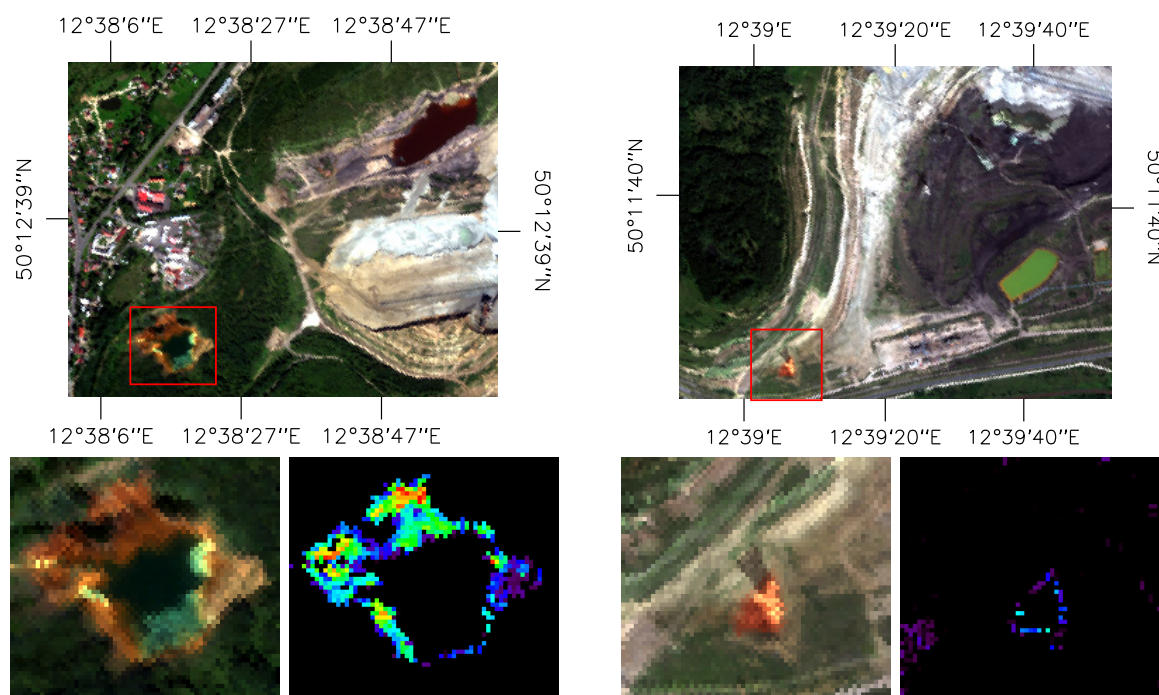


Figure 6.36.: Impact of  $\text{Fe}_2\text{O}_3$  common feature pattern matching on quantitative mapping result. True color image subsets showing acidic pond framed by goethite rich soil (left) and presumed AMD hotspot of goethite rich precipitate close to a pool of orange water (right). Samples from the former site were used for  $\text{Fe}_2\text{O}_3$  model development. Below true color image zoom and respective zoom from  $\text{Fe}_2\text{O}_3$  quantitative map (rainbow color ramp, 0-40 wt-%). Note the omission of the hotspot due to pattern mismatch.

Spectral resolution reduction from laboratory to image spectra impaired the model's performance considerably. Diagnostic wavelength ranges had to be excluded. Thus both quantitative models were ranked of medium prediction quality. The difficulties in applying a laboratory data calibrated model to imagery was exemplified by the apparent kaolinite enrichment in coal. Therefore, the abundances must be interpreted with caution. Nevertheless, comparing all individual pixel spectra to common feature patterns rigorously prevented application of the quantitative models to spectral features not represented in the training data. Hence an extrapolation of the quantitative model, possibly leading to erroneous abundances, was avoided.

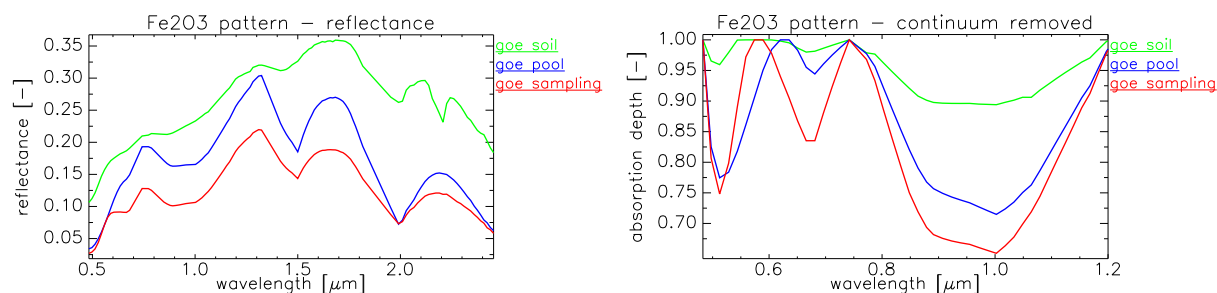


Figure 6.37.: Match of image spectra to the common feature patterns of  $\text{Fe}_2\text{O}_3$ . Average spectrum for goethite rich sampling position (red) in the training set. Average spectrum of iron rich soil (green) matching pattern and goethite rich precipitate close to pool of orange water (blue) not matching the pattern. Reflectance spectra (left) versus continuum removed absorption depths (right). Note the distinct shift of the 620 nm maximum for the pool target.



### 6.6.2. Band-Ratio based Quantification of Quartz & Kaolinite

The ten broad thermal AHS bands allowed investigation of only the most discriminative materials and were obviously insufficient for hyperspectral techniques as presented for the reflective domain. For deriving quantitative information from the emissivity data, the standard methods of multispectral data analysis were favoured. Single band ratios were chosen due to the ease of interpretation, the relatively small risk of model overfitting and their robustness against absolute calibration inaccuracies.

This quantification experiment covers the two most abundant minerals found in Sokolov. Quartz and kaolinite exhibit distinct spectral features in the thermal domain. Quartz is spectrally featureless in the reflective domain and hence almost indiscernible there. Kaolinite also shows characteristic features in the reflective domain that allowed for its quantification from the HyMap data (see Section 6.6.1 - *Wavelet based Quantification of Goethite & Kaolinite*). From the perspective of potential AMD related processes is kaolinite of interest due to its slow acidity neutralising reaction. The inert quartz does not counter soil acidification. Knowledge of the soil dominating mineral therefore can be an indicator for its acidity buffering capacity. Their quantification also allows i) to cross check the kaolinite distribution patterns between thermal and reflective, ii) to check the plausibility of the per-pixel sum of mineral fractions, and iii) to evaluate the applicability of the thermal data pre-processing.

**Data** Emissivity spectra derived from the AHS imagery, from field reference measurements and from the laboratory measurements of the samples were available for developing the quantitative models. Validation in Sections 6.7.1.1 and 6.7.1.2 showed the problematics of transiting between spatial scales and different sensors. Practically, the number of field measurements backed by a sample which underwent geochemical analysis was too small for a meaningful model development. The thermal laboratory measurements showed issues with spectral contrast and linear bias. Consequently, the thermal quantitative models were developed directly on image spectra. Out of the 80 samples with laboratory reference data, 64 corresponded to targets identifiable within the AHS imagery. Based on expert knowledge this set was further reduced to 35 targets of high confidence by eliminating targets possibly influenced by vegetation, of high spatial variability or of small spatial extent. For all 64 targets mean emissivity spectra were extracted from the AHS imagery using expert derived regions of interest in case the ground target extent was known or otherwise circular regions of 5 pixel diameter. The 5 pixel diameter was chosen to accommodate an estimated geolocation error of 1-2 pixel of the AHS data.

**Correlation Screening** The correlations between band values/ratios and normative calculation results were screened using Pearson's product-moment correlation coefficient and Spearman's rank correlation coefficient. Using the confident targets subset good correlations were found for the main minerals quartz, kaolinite and, to a lesser degree, feldspar. For other minerals no such correlation could be found that bridges the differing mineral mixtures and geologies of the various sampling sites. For an investigation of iron oxides or iron sulphates the thermal reference data and abundance variation over the samples was rejected as too sparse.

The left plot of Figure 6.38 shows the relation of the samples' NORMA kaolinite abundances in respect to their quartz abundances. Little difference is seen for the full 64 samples set (black+red) and the 35 samples confident subset (red). For reference, also the scatter plot between  $\text{Fe}_2\text{O}_3$  (covered by the reflective quantification) and quartz is given (right plot of Figure 6.38). Kaolinite showed an inverse tendency to quartz with  $r^2 = 0.33$ . This is expected for the two most abundant minerals found in Sokolov from the physical sum-to-100 wt-% constraint. Several samples combining low quartz and low kaolinite fractions existed which attributed to elevated coal, siderite or goethite fractions. No dependency ( $r^2 = 0.16$ ) between  $\text{Fe}_2\text{O}_3$  and quartz was apparent that could possibly indicate a link between AMD sites and non-buffering quartz rich soils.

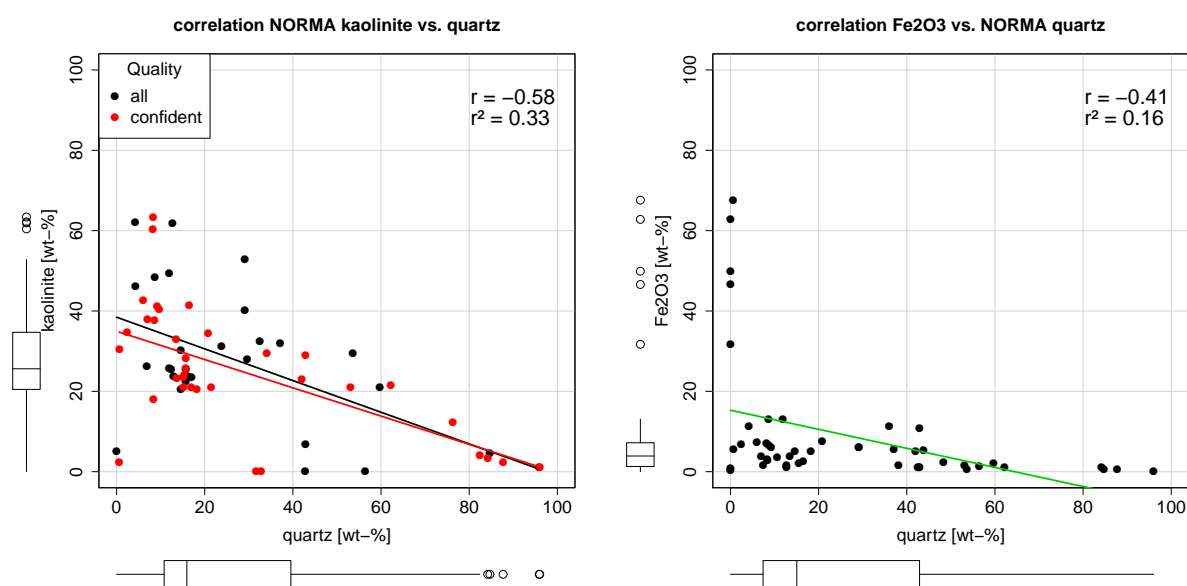


Figure 6.38.: Scatter plot of NORMA kaolinite vs. quartz abundances (left) for all 64 samples (black) and confident subset (red). For comparison scatter plot of  $\text{Fe}_2\text{O}_3$  vs. NORMA quartz (right).

Table 6.16.: Dependence between NORMA main minerals and selected AHS band combinations measured as Pearson's  $r$  and Spearman's  $\rho$  for the high confidence targets sampled in 2010 and 2011.

mineral	band	wvl [nm]	2010		2011		combined	
			$r$	$\rho$	$r$	$\rho$	$r$	$\rho$
Quartz	#72	8,770	-0.97	-0.76	-0.93	-0.58	-0.97	-0.74
	#72/#78	8,770/11,796	-0.97	-0.74	-0.94	-0.63	-0.97	-0.74
Kaolinite	#72/#74	8,770/ 9,680	0.75	0.84	0.63	0.56	0.67	0.74
	#76/#77	10,624/11,230	0.75	0.78	0.78	0.49	0.75	0.65
Feldspar	#72/#73	8,770/ 9,237	-0.56	-0.33	-0.07	-0.44	-0.48	-0.47
	#74/#75	9,680/10,143	0.39	0.22	0.52	0.56	0.46	0.49

Table 6.16 summarises the measures of dependency for the most promising band combinations to derive quartz, kaolinite and feldspar abundances based on NORMA model results. The rank correlation coefficient reduces the influence of outliers. Latter is the reason for the generally lower  $\rho$  values observed for quartz, since the strong (enhancing) influence of large normative quartz abundance values on  $r$  was mitigated. Generally, a strong association between quartz abundance and AHS band #72 (8,770 nm) emissivity value was observed. It occurred for both sample sets taken in 2010 and 2011. A similarly strong association existed for the AHS band ratio #72/#78

(8,770/11,796 nm). For kaolinite two band ratios #72/#74 (8,770/9,680 nm) and #76/#77 (10,624/11,230 nm) with a moderate association to the normative abundance existed. The drop in  $r$  and  $\rho$  for the 2011 samples may partially be attributed to a reduced abundance variation compared to the 2010 samples. For feldspar, two band ratios #72/#73 (8,770/9,237 nm) and #74/#75 (9,680/10,143 nm) stood out from all other possibilities. However, the association was only minor, primarily basing on the 2010 samples and not supported by the 2011 ones in respect of  $r$ . Rank correlation  $\rho$  showed a more stable behaviour indicating a nonlinear relation. Further analyses focused on AHS band ratio #72/#78 (8,770/11,796 nm) for quartz because a ratio is invariant towards absolute emissivity shifts. It relates the quartz emissivity minimum around 8,500 nm to the long-wavelength “background” emissivity. For kaolinite, both band ratios #72/#74 (8,770/9,680 nm) and #76/#77 (10,624/11,230 nm) were investigated. They aim at independent spectral features, respectively the 9,500 nm main emissivity minimum and the 11,000 nm secondary minimum. Feldspar was excluded from further analysis due to the inferior (overall) dependency measures and the observation of contradicting (individual) correlations for the three feldspar endmembers orthoclase, albite and anorthite. The feldspar endmembers exhibit differing emissivity characteristics but could not be appropriately differentiated at AHS resolution nor with the limited set of samples. In summary, it should be noted that the band combinations identified by the correlation screening correspond well to band combinations expected by the analysis of library spectra of pure minerals (see Figure 6.40).

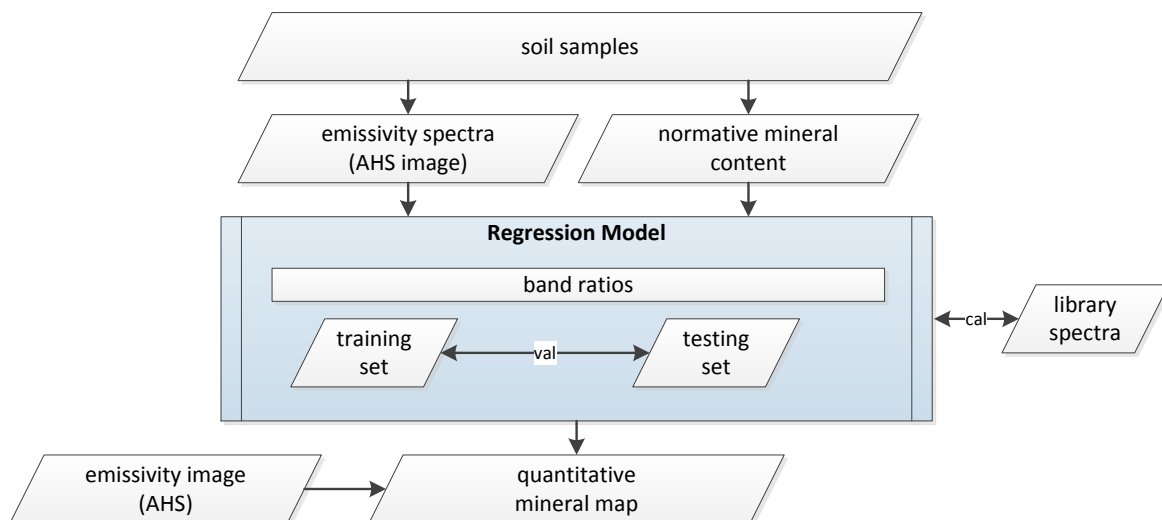


Figure 6.39.: Processing workflow for the application of band ratios in TIR quantification. Data layers depicted as white parallelograms, processing steps as blue rectangles.

**Regression Models** Regression models for quartz and kaolinite were established following Scheme 6.39. A training set formed from the 35 confident targets was used to calibrate the models. Theoretical ratio values for 0 wt-% and 100 wt-% derived from ASU library spectra of pure minerals (Christensen *et al.*, 2000) were included into the training set to support the regression models at the very extreme abundances otherwise poorly supported by the data.

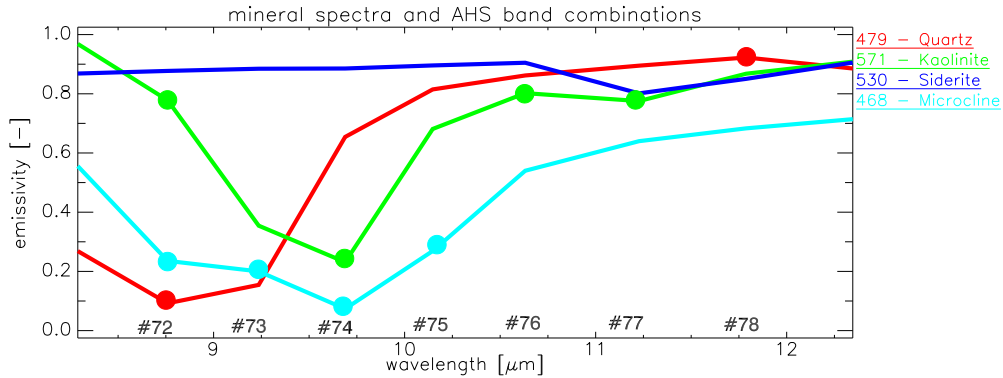


Figure 6.40.: Visualisation of the candidate AHS band combinations for TIR domain quantification displayed as dots on common minerals from ASU library (Christensen et al., 2000). Quartz (red), kaolinite (green), K-feldspar microcline (cyan), carbonate siderite (blue) for reference.

An exponential regression model could be established for quartz (see Figure 6.43). The training error was 9 wt-% RMSE and the prediction accuracy was estimated 15 wt-% RMSE based on the complement testing set. For kaolinite both #72/#74 and #76/#77 ratio scatter plots showed similar distributions (see Figures 6.41 and 6.42). Besides an exponential trend, a cluster of samples with a band ratio near 1 and low normative kaolinite abundance (< 10 wt-%) was identified. The emissivity ratio #76/#77 slightly mitigated this effect. These samples resembled mainly targets for which a considerable weight fraction was not explained by quartz and kaolinite. A cutoff threshold of maximum 60 wt-% quartz & kaolinite unexplained fraction was set to exclude respective samples from the training set. These 5 targets (DR10\_05, LO10\_05A, MIX\_LO\_01, 2LI11M1 and 2LI11M5) were moved into the testing set.

For the #72/#74 ratio a training error of 9 wt-% RMSE and a prediction accuracy of 15 wt-% RMSE were achieved while for the #76/#77 ratio this increased to a training error of 29 wt-% RMSE and a prediction accuracy of 21 wt-% RMSE. Latter training RMSE reduced to 12 wt-% when excluding the 100 wt-% kaolinite support point which was extremely overestimated by the low intercepting regression function. The #76/#77 ratio model basing on the secondary kaolinite feature near 11,000 nm with a comparably low value range of [0.98 . . . 1.01] hence proved inferior to the #72/#74 ratio model. The value range [0.65 . . . 1.15] of the latter was comparably high as was based on the stronger primary kaolinite feature around 9,500 nm (see Section 2.3.5 - *Spectral Characteristics of relevant Minerals*). A ratio value of 1 occurred at about 20 wt-% where the slope reversed. Hence spectrally flat materials like water and vegetation needed to be masked, otherwise yielding a mistakenly uniform kaolinite abundance in these areas. A simple emissivity threshold was imposed on three AHS bands #72 (8,760 nm)  $\vee$  #73 (9,236 nm)  $\vee$  #74 (9,689 nm)  $\leq$  0.95 to mask such materials in the imagery.

Application of above three exponential regression models to the emissivity imagery produced the aspired quantitative mineral maps. Figure 6.44 displays the result for quartz. In the coal seams no quartz was found. The overall quartz fraction in the overburden and backfill was rather low with a mean of about 17 wt-%. Higher quartz fractions were found along the northern flank of the large mine and the south-western flank of the smaller mine (mean 40 wt-%, locally up to 80 wt-%). The kaolinite quantification based on the #72/#74 ratio is shown in Figure 6.45. Overburden and backfill yield a mean kaolinite abundance of 38 wt-%, the northern and southern

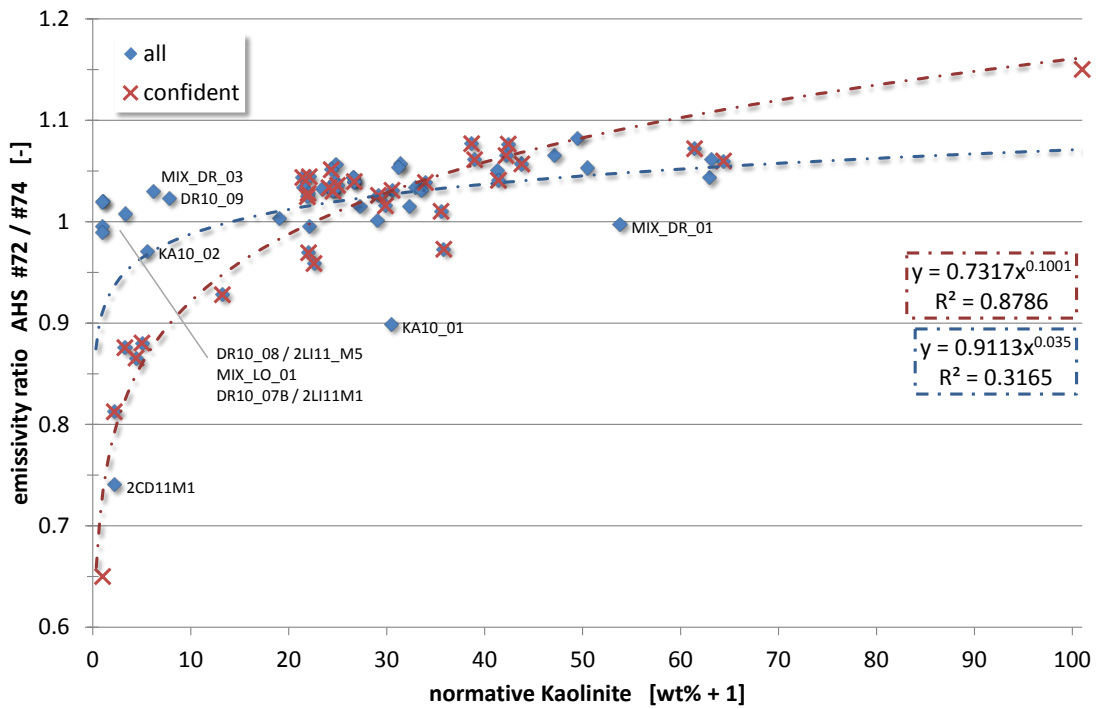


Figure 6.41.: Scatter plot of AHS band ratio #72/#74 (8,770/9,680 nm) versus normative kaolinite fraction calculated with NORMA. Shown are the complete 64 sample set and the 32 sample filtered confident subset with their respective exponential regressions.

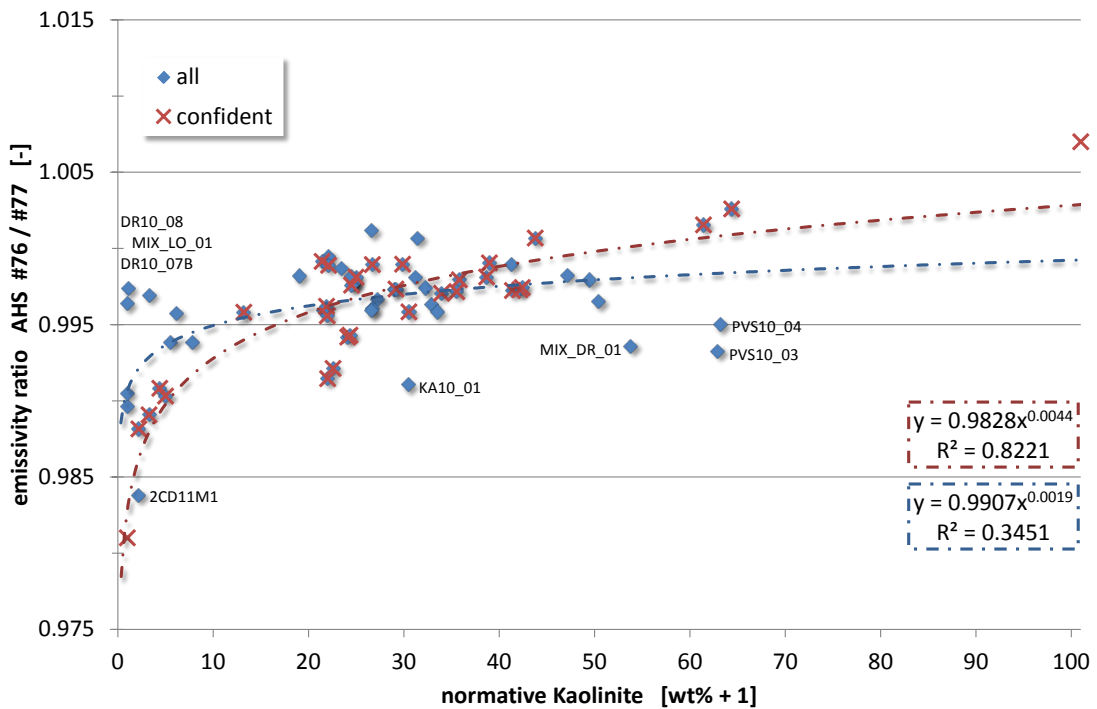


Figure 6.42.: Scatter plot of AHS band ratio #76/#77 (10,624/11,230 nm) versus normative kaolinite fraction calculated with NORMA. Shown are the complete 64 sample set and the 32 sample filtered confident subset with their respective exponential regressions.

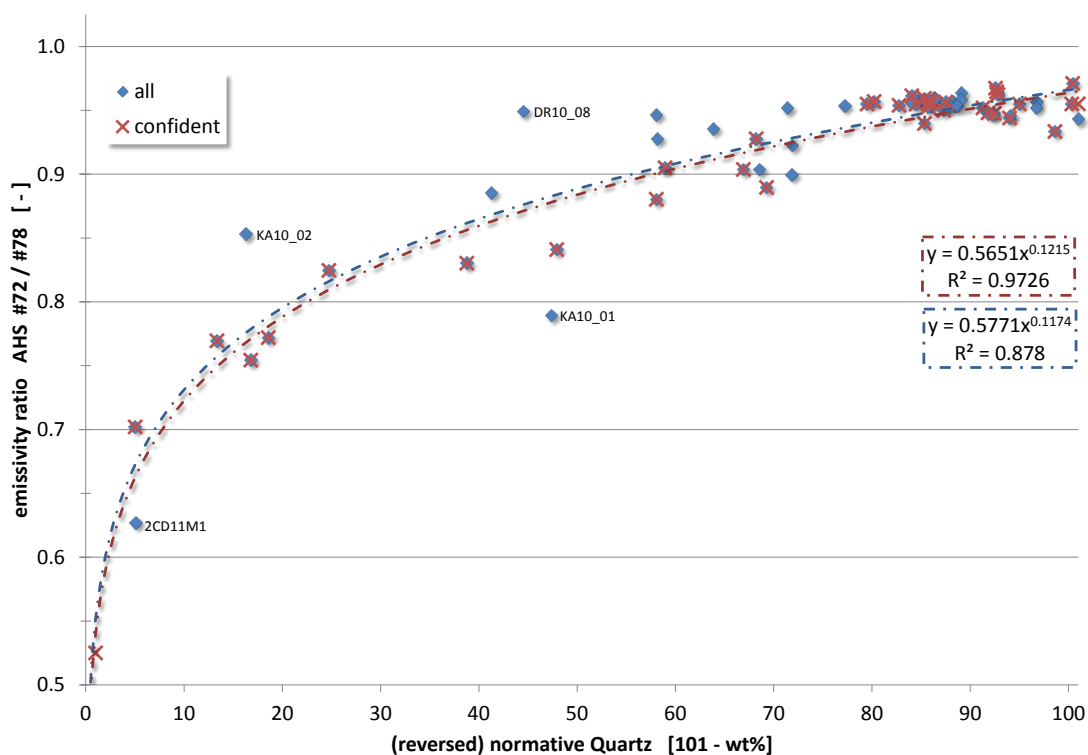


Figure 6.43.: Scatter plot of AHS band ratio #72/#78 (8,770/11,796 nm) versus reversed normative quartz fraction calculated with NORMA. Shown are the complete 64 sample set and the 35 sample confident subset with their respective exponential regressions.

flanks of the two mines yield a mean of 25 wt-%. A kaolinite enrichment of average 52 wt-% and locally up to 80 wt-% is visible for freshly dumped material in the centre of the larger mine. Kaolinite quantification from the #76/#77 ratio is not shown here because the low data range of the ratio made it prone to noise in the emissivity data which led to a poor result.

In summary, pre-processing of the thermal imagery with the in-flight calibration module and adapted model atmosphere together with the collected samples and normative modelling allowed for a successful quantification of quartz and kaolinite from the multispectral thermal imagery. Robustness of the models against absolute emissivity offsets was achieved by the use of band ratios. The bands themselves corresponded to wavelength regions known to be related to the targeted minerals' absorption features. The quartz model fit was excellent and stable while for kaolinite a medium quality fit and stability could be reached by including two additional support points (0 and 100 wt-%) from library spectra. The derived spatial distribution patterns of both minerals followed the expectations.

It must be noted that the multispectral character of the data did not allow to differentiate between spectrally similar minerals, e.g. kaolinite, illite and feldspars, that exhibit their main emissivity minima at close wavelengths (see Section 2.3.5 - *Spectral Characteristics of relevant Minerals*). The support of the kaolinite quantification with a secondary feature at 11,000 nm was hampered by noise (latter feature would also be influenced by carbonates, e.g. siderite). Therefore, the abundances must be interpreted with caution.

Reflective and thermal data derived mineral maps are compared in the validation section next.

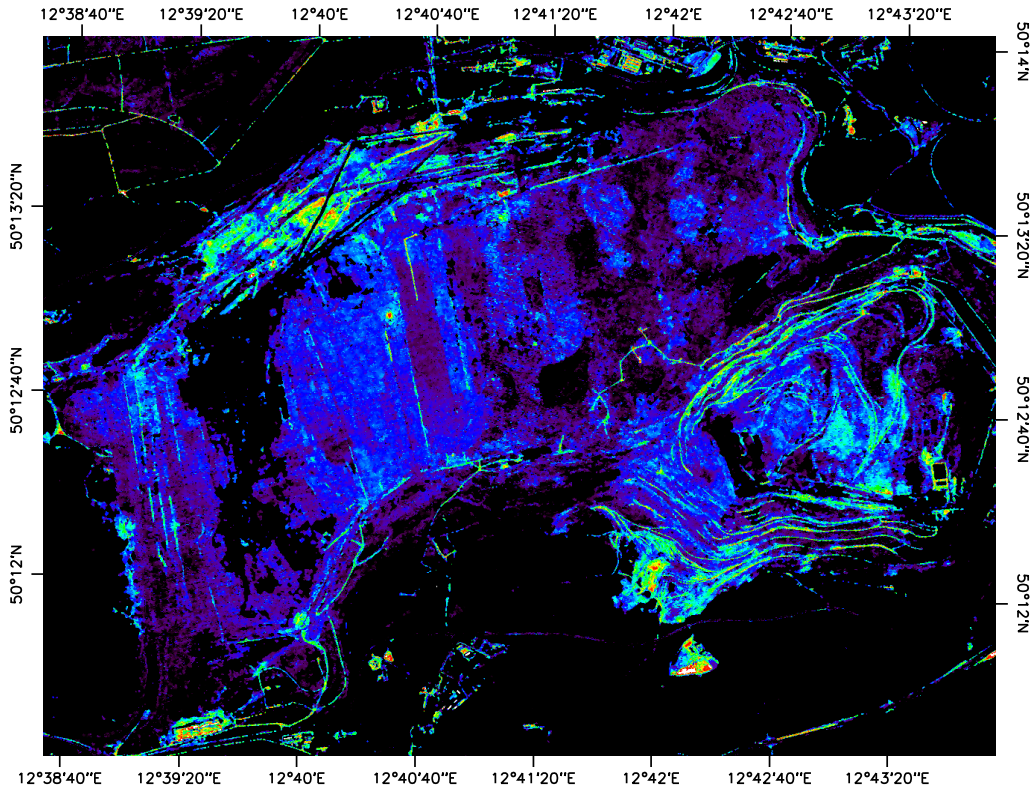


Figure 6.44.: Quantitative quartz map based on AHS band ratio regression model (rainbow color ramp, 0-75 wt-%).

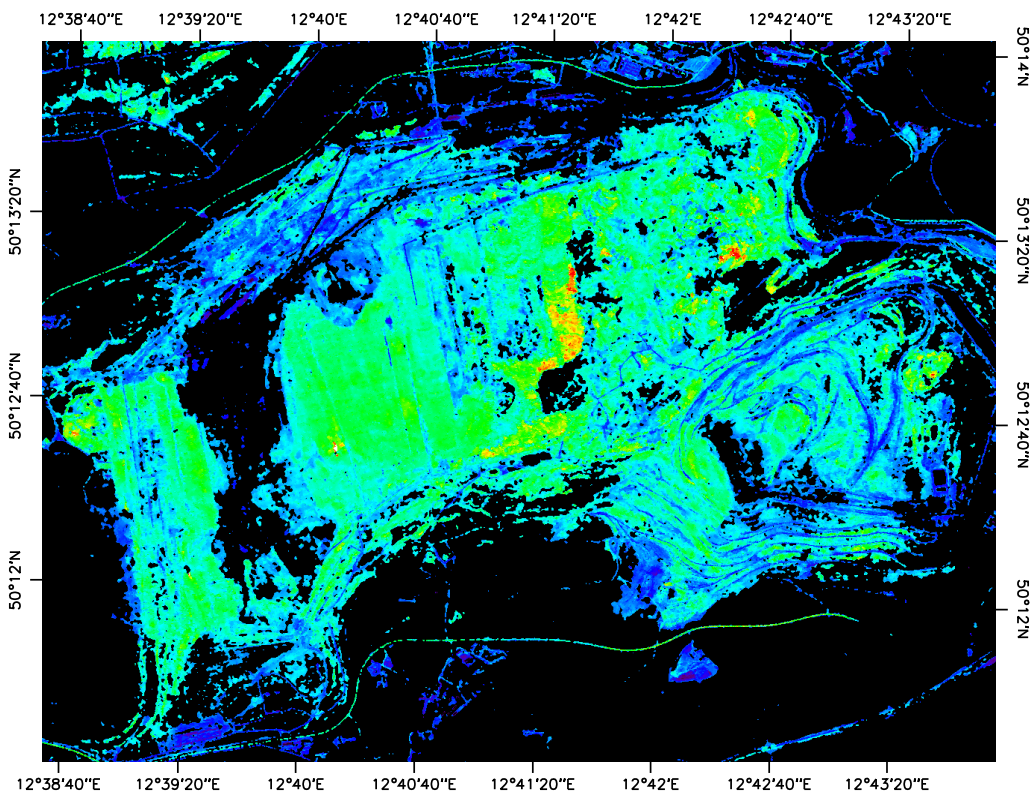


Figure 6.45.: Quantitative kaolinite map based on AHS band ratio regression model (rainbow color ramp, 0-75 wt-%).

## 6.7. Evaluation of Processing Workflow & Results

### 6.7.1. Validation of Processing Steps

In the following, the main steps from spectral data pre-processing to the development of quantification models are empirically validated. The validation bases on comparison of intermediate results of the individual processing steps assessing the result's plausibility and accuracy. Influencing factors and error budgets of sub-processes in one step can not be resolved.

In case of the pre-processed reference spectra and airborne imagery, a validation against independent data was possible. The normative modelling results mark the transition from validation against independent data (qualitative XRD) to intercomparisons and plausibility checks of the different models/methods among each other (NORMA vs. SEDNORM). The quantitative models can be validated only indirectly. They depend on the normative results. Statistical model cross-validation and plausibility checks of the quantitative results for library spectra were used.

#### 6.7.1.1. Lab & Field Reference Spectra Pre-Processing — Reference Spectra Validation

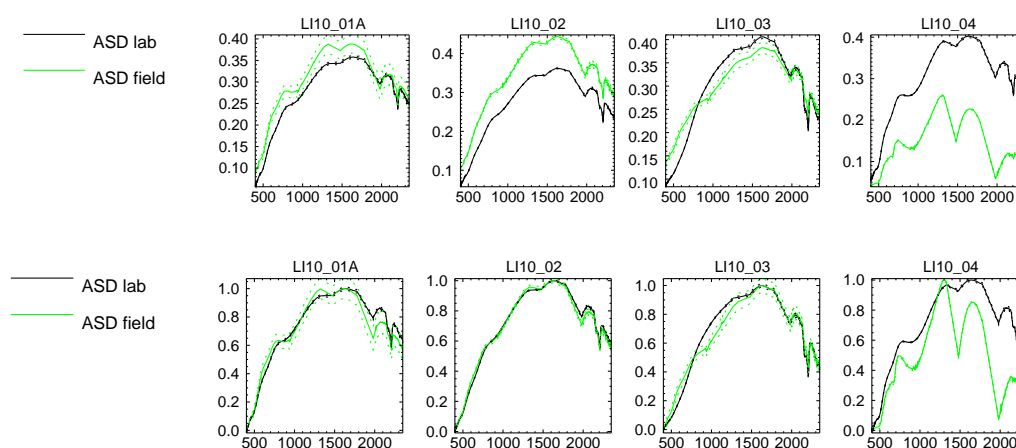


Figure 6.46.: Validation of field (green) and laboratory (black) reflectance measurements exemplified for 4 corresponding targets. Reflectance [-] (top) and normalised spectra [0..1] (bottom) over wavelength [nm] with standard deviations (dotted). Plot of all 19 targets in Figure E.2.

The pre-processed reflectance spectra derived from the ASD laboratory measurements were validated against the pre-processed field reference measurements. A total of 19 corresponding targets existed that were measured in the laboratory and in the field. The field spectra were taken one day after the HyMap sensor overflight during deteriorating atmospheric conditions. At 14:30 o'clock the formation of cirrus and cumulus clouds forced abortion of the measurements. For each measurement target in the field, typically 5 to 10 spectroscopic measurements were taken to capture the natural variability of the surface. One sample of the top 1 cm of the mineral soil was taken for spectroscopic measurements under laboratory conditions. These samples were air-dried for a month before measurement. Section 6.2 - *Data Sets and Measurement Protocols* gives the details. The spectral data pre-processing gave a mean spectrum and the variability for each target, as exemplified for four targets in Figure 6.46. The full plot of all 19 corresponding targets can be found in Figure E.2 of Appendix E - *Validation Supplements*. It was apparent that differences in spectral intensity and also spectral shape existed. This has multiple reasons.



Firstly, it is difficult to represent the target's natural variability by laboratory measurements. This manifested in their typically low variability compared to the wider spreading standard deviations seen for the field reflectance measurements. Exceptions may be found for field targets where individual measurements did not meet an illumination stability requirement and hence were excluded leading to an apparently reduced variability.

Secondly, transportation and drying of the samples changed surface structure, grain size distribution and water content compared to the natural surface. This homogenised the samples leading to a reduced intra-sample variability but also to deviating spectral amplitude and spectral contrast compared to the field measurements. Moreover, particles of the different fractions may cover each other, e.g. clay on iron aggregates, leading to nonlinear effects that can selectively alter the spectral signature (Gupta, 1991; Ben-Dor *et al.*, 1999).

Thirdly, the representativeness of the collected sample filling a sample tray of 20 cm diameter in relation to measurements of a possibly heterogeneous natural surface of several square meter size can hardly be ensured. To meet this, apparently homogeneous field targets were selected and the samples were taken from different points within.

The lower part of Figure 6.46 shows the normalised reflectance spectra, resulting from shifting the individual spectrum's minimum to 0 and scaling its maximum to 1. The plots therefore show field and laboratory reference measurements after removal of constant additive and multiplicative differences. The spectral shape and spectral contrast of most normalised reference targets were matching, especially when considering the field measurement's variability. Apparently sample LI10.04 was not representative.

As expected, deviations existed around the 1,400 and 1,900 nm water absorptions. Several targets showed a considerable change in curvature between 800-1,000 nm, others between 500-900 nm. The cause could not be determined, however the systematic reduction of the 860 nm ferric ion absorption may be linked to the air-drying procedure. An influence of adjacent vegetation can be ruled out at this wavelengths according to Richter (2010a). The 2,200 nm Al–OH absorption depths and shapes were generally matching, except for samples PVS10.13 through PVS10.15 and PVS10.18. Also here the cause could not be determined.

Previously mentioned spectral shape and absorption depth differences in the 2,100-2,300 nm and 500-1,000 nm wavelength ranges could decrease the accuracy of quantitative models for kaolinite and goethite/Fe<sub>2</sub>O<sub>3</sub> since the deviations may directly propagate into the models. Spectral normalisation techniques like mean centring, continuum removal or derivatives are typically used to lessen the above mentioned influences. The continuous wavelet transform based classification and quantification methods have a similar effect as presented in Section 3.3 - *Application to Hyperspectral Signatures*.

The emissivity spectra derived from the D&P FT-IR field measurements were validated against open air laboratory measurements of the corresponding samples using a CI Systems SR-5000 spectrometer (pers. communication Dr. Gila Natesco, Remote Sensing and GIS Laboratory, Tel-Aviv University). The FT-IR emissivity spectra were downsampled to the bands of the SR-5000 instrument (see Appendix C - *Sensor Descriptions* ).

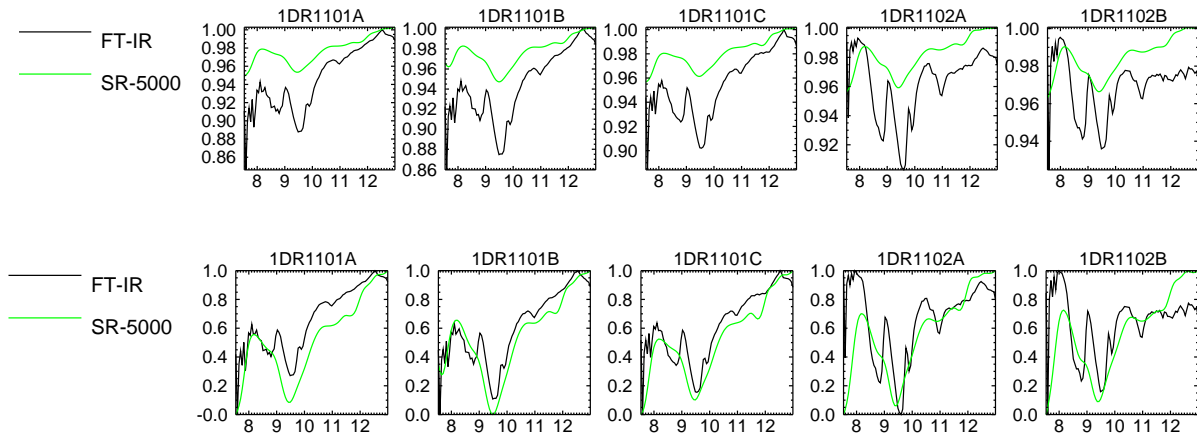


Figure 6.47.: Validation of field (black) and laboratory (green) emissivity spectra exemplified for 5 corresponding targets. Emissivity [-] (top) and normalised spectra [0..1] (bottom) over wavelength [ $\mu\text{m}$ ] together with standard deviations (dotted). Plot of all 26 targets in Figure E.3.

Figure 6.47 exemplifies the comparison for 5 out of the 26 corresponding samples. An overall agreement in spectral shape and wavelength position of the main features was observed. The spectral resolution and spectral contrast however were largely different. Besides sensor characteristics, a decreased spectral contrast may be caused by a lower thermal contrast between the sample and its surroundings, increased downwelling radiance, finer grain size or increased pore space. Latter two likely occurred due to the destruction of the compacted natural surface in the process of taking the sample and placing it onto the sample tray for measurement.

A comparison of the normalised emissivity spectra (bottom of Figure 6.47) confirmed the overall agreement in spectral shape and relative absorption depths from 8,000-11,500 nm. Several targets showed a linear bias (1DR1101B through 1MD1101E) also observed by other authors (Hook & Kahle, 1996; Vaughan *et al.*, 2003) which was likely caused by a differently estimated sample temperature and assumed blackbody-like spectral region during the Planck curve fitting. Field and laboratory measurement of 1MD1103C mismatch. In samples 1MD1102A through 1MD1103B the location of the minimum emissivity found in the primary quartz reststrahlen band (8,000-9,000 nm) was reported differently. This could not be explained and may relate to differences in the clay/quartz fraction visible to the field and laboratory spectrometers due to particle redistribution.

Comparison of both VNIR-SWIR and TIR reference measurements demonstrated the difficulties in obtaining reliable, repeatable and comparable spectra (absolute and relative) when crossing scales from field to laboratory or changing from one instrument to another. Many differences could be approximated by a per-target linear normalisation model (illumination, moisture, grain size, temperature) however nonlinearities persisted, which in case of the iron and clay features as well as the emissivity bias may influence the transferability of models. The deviations were not consistent over the targets hence no global linear correction model was found that would match the spectra between the instruments. The quantitative model validation in Section 6.7.1.4 - *Reflective Domain Quantification — Model Cross-Validation & Plausibility Checks* evaluates the impact of above effects by comparing the derived abundances from the 19 corresponding field and laboratory spectra.

### 6.7.1.2. Airborne Imagery Pre-Processing — Image Spectra Validation

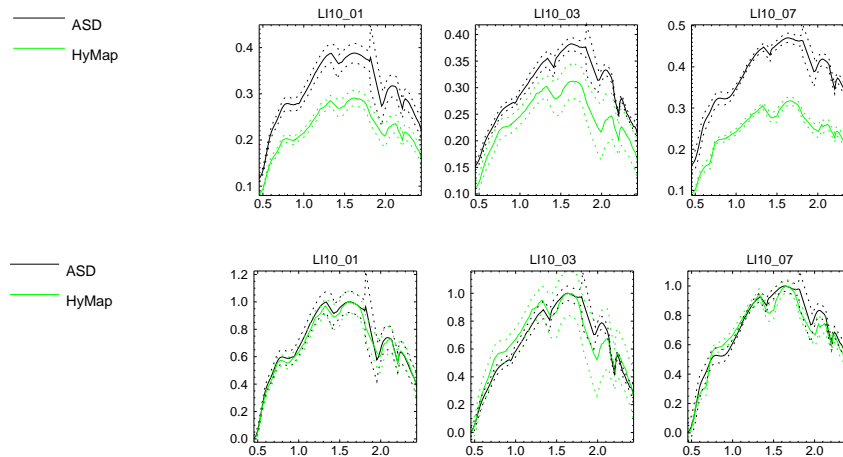


Figure 6.48.: Validation of field (black) and image (green) reflectance spectra exemplified for 3 corresponding targets. Reflectance [-] (top) and normalised spectra [0...1] (bottom) over wavelength [ $\mu\text{m}$ ] with standard deviations (dotted). Plot of all 11 targets in Figure E.4.

A comparison of the pre-processed reflectance spectra derived from the ASD field measurements and HyMap image spectra is exemplified in Figure 6.48. A total of 11 field reference targets were identifiable in the airborne imagery. The CCal1 targets were measured during the HyMap sensor overflight in order to provide in-flight calibration and validation information. The LI10 field spectra were taken one day after the HyMap sensor overflight during deteriorating atmospheric conditions. The corresponding image spectra (mean and variability) were extracted from regions of interest covering the apparent spatial extent of the targets.

The normalised spectra showed general good agreement. Spectral shape and relative absorption depths were matching, especially when taking the spectral variability and the scale difference between field (0.2 m GIFOV) and image pixel (4 m GIFOV) into account. Two dark asphalt calibration targets (CCali1a, CCal1b) were affected by spectral blooming. This was caused by adjacency radiance, where “reflected radiation from the neighborhood [is] scattered by the air volume into the [sensor’s] current instantaneous [field of view]” (Richter, 2010b). The bright red edge and NIR plateau of adjacent vegetation was superimposed onto the dark asphalt’s spectral signature in spite of the car park’s spatial extent. The Sokolov calibration measurements hence were used as validation targets. DLR’s well-characterized standard reference targets were regarded more reliable for in-flight calibration (see Section 6.3.2 - *Airborne Imagery Corrections*).

Validation of 18 emissivity targets derived from the pre-processed AHS imagery and corresponding FT-IR field measurements showed a very good match for the per-target normalised emissivity spectra (exemplified in Figure 6.49). The image spectra were extracted using regions of interest in case the ground target extent was known. Otherwise circular regions of 5 pixel diameter were used. The spatial scale of the field instrument (0.1 m GIFOV) compares to 5 m GIFOV (pixel size) of the AHS imagery.

The per-target linear normalisation model covered the differences in spectral contrast and absolute emissivity offsets. Targets 1DR1102F and, to a lesser degree, 1DR1103B, 1DR1102G and 1DR1103A show the linear bias already observed with the field/laboratory measurements probably introduced by mismatches in the temperature estimates.

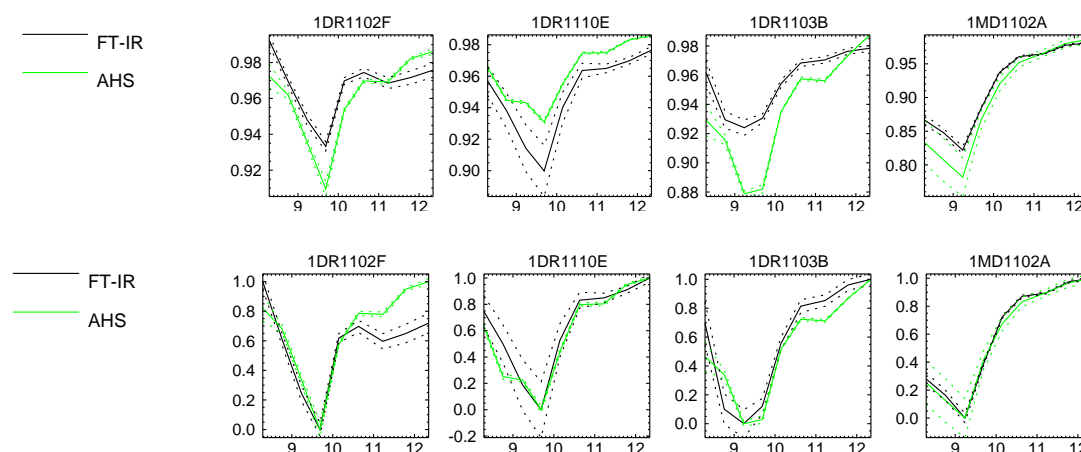


Figure 6.49.: Validation of field (black) and image (green) emissivity spectra exemplified for 4 corresponding targets. Emissivity [-] (top) and normalised spectra [0...1] (bottom) over wavelength [ $\mu\text{m}$ ] with standard deviations (dotted). Plot of all 18 targets in Figure E.5.

In summary, the image derived spectral signatures were in agreement with the field measurements. Deviations were largely explained by constant offsets and gain factors, however not consistent over the various targets. Thus, no global linear correction model was found that would match image and field spectra to each other. This indicated highly individual influences such as moisture, mixing, spatial scale and illumination as source of the differences rather than a systematic sensor deviation. It appears that spectral normalisation techniques could largely reduce these influences and hence would be beneficial for classification and quantification of minerals aiming at relative shape differences in the spectra. For the hyperspectral reflective data such an approach could be realised with the Wavelet Coherence Mapper and partially the Wavelet Transform Feature Analysis eliminating constant offsets, linear bias and, at small scale, reducing the influence of gain differences. The 10 multispectral thermal bands did not allow for such an approach. Instead, the quantitative models were developed on image spectra using band ratios. Hence the problematics of different sensors was circumvented at the cost of representativeness of the geochemical analysis in respect to a  $25 \text{ m}^2$  pixel. The band ratios eliminated constant emissivity offsets, leaving possible linear bias and gain differences. The latter is of little concern for quantitative models developed on and applied to the same sensor. A possible per-pixel linear bias is practically uncorrectable given the few spectral bands and NEM temperature-emissivity-separation method.

### 6.7.1.3. Normative Modelling — Model Cross-Validation & Mineral Validation

A good overall correspondence between qualitative XRD analysis and the quantitative results of both SEDNORM and NORMA was observed (see Section 6.4 - *Normative Modelling* and Appendix B.3 - *Sample Mineralogy*). The calculated mineral abundances matched the XRD class assignment, were plausible to the rock analysis elemental concentrations and were in accordance between the two normative models. More specific, for the abundant minerals quartz, feldspar, muscovite and kaolinite, NORMA and SEDNORM yielded very similar abundances expressed by a standard deviation of only 0.9, 1.0, 1.3 and 3.0 wt-% respectively. This range also held for the other minerals in common (see Figure E.1 in Appendix E - *Validation Supplements*).

For sample PVS10.01 a notably high deviation was observed. SEDNORM underestimated kaolinite by 14 wt% whereas feldspar and muscovite were overestimated 5 wt% and 9 wt% compared to NORMA. This behaviour could be explained by the high abundance of jarosite which SEDNORM did not account for. The excess of potassium led to higher abundances of normative feldspar and muscovite which depleted silica to the point that normative kaolinite formation was limited. The remaining iron and aluminium were reported as remarkably high amounts of normative haematite and gibbsite. This characteristic was not repeated for the other jarosite rich samples LO10.05A and LO10.05B because of their negligible  $\text{SiO}_2$  and  $\text{Al}_2\text{O}_3$  concentrations. For seventeen Druzba (DR) mine and PVS dump samples of the 2011 campaign, the wet chemistry showed elevated  $\text{CO}_2$  concentrations compared to the other samples of that campaign and compared to samples collected the year before. Since no corresponding CaO was available for the formation of normative calcite, SEDNORM assigned it to magnesite. Thus the MgO concentration was depleted and also less  $\text{CO}_2$  for normative siderite was left. By contrast NORMA was configured to not include magnesite but chlorite. Latter was assigned the MgO and therefore led to an  $\text{Al}_2\text{O}_3$  deficit compared to SEDNORM. This finally expressed as a difference in normative kaolinite of 5 wt%.

Based on the SEDNORM user option whether to impose a water limit on the formation of normative minerals, it was verified that for all samples the structural water reported as loss on ignition ( $\text{H}_2\text{O}^+$ ) was sufficient or nearly sufficient to allow for an unconstrained normative mineral formation. The only outlier was sample VR10.01 taken from a coal power plant fly ash dump. It showed discrepancies of 13 wt% kaolinite and 8 wt% quartz compared to the unconstrained formation due to the water deficit.

#### 6.7.1.4. Reflective Domain Quantification — Model Cross-Validation & Plausibility Checks

Candidate feature parameters and the final MLR models in the reflective domain were individually validated. K-fold cross-validation with five  $K_i \in [2 \dots N]$ , with  $N$  the number of samples, was used. This spans the special cases holdout method, dividing the samples 50:50 in training and testing set, until leave-one-out cross-validation. To cover all possible combinations in case of the holdout method with only 50 samples already a practically infeasible  $\binom{50}{25}$  runs would be required. Instead, repeated random assignment of the samples to the folds was chosen here. With 1000 repetitions the probability  $p = \frac{1}{2}^{1000}$  of a sample never chosen into the testing set was marginally low.

On the level of individual feature parameters, a convergence of the mean prediction accuracy for all five  $K_i$  was observed for the linear regression model. The variance of the RMSE was highly asymmetric towards higher mean errors and declined with increasing  $K_i$  as expected. Nevertheless, a notable model instability occurred for the ‘significance’ parameter of the 2,144 nm downward crossing in the kaolinite model which was eliminated from the MLR model in consequence (see Figure 6.50). Table 6.15 on page 121 gives the individual feature parameter’s prediction accuracy in form of the RMSE reported by the leave-one-out cross-validation. The individual prediction accuracies were 16-21 wt-% for kaolinite, 3.0-3.5 wt-% for goethite (low weight fractions only) and 11-14 wt-% for  $\text{Fe}_2\text{O}_3$ .

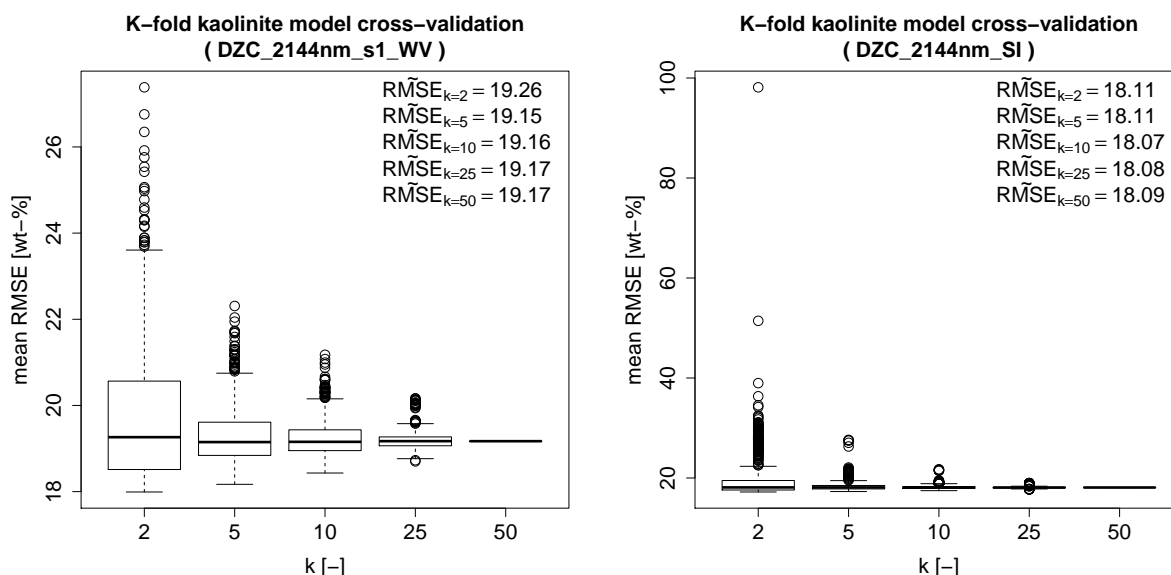


Figure 6.50.: K-fold kaolinite model cross-validation result showing the well-performing individual parameter ‘wavelength’ (left) and the worst-performing parameter ‘significance’ (right) of the 2,144 nm downward crossing transform feature.

The final MLR models for kaolinite, goethite and  $\text{Fe}_2\text{O}_3$  in the reflective domain underwent an identical cross-validation procedure. Figure 6.51 depicts the results. After 1,000 and also 10,000 iterations only a partial convergence of the median prediction error over all  $K_i$  was apparent. The respective 12%, 5% and 11% relative error decrease (from holdout method to leave-one-out cross-validation) showed some MLR model instabilities, i.e. the model’s dependence on the full sample data set. The error variance approached symmetry, its decline persisted. Overall, a kaolinite prediction accuracy of 16 wt-% RMSE was estimated, which compares to 4 wt-% RMSE for weight fraction limited goethite and 11 wt-% RMSE for  $\text{Fe}_2\text{O}_3$ .

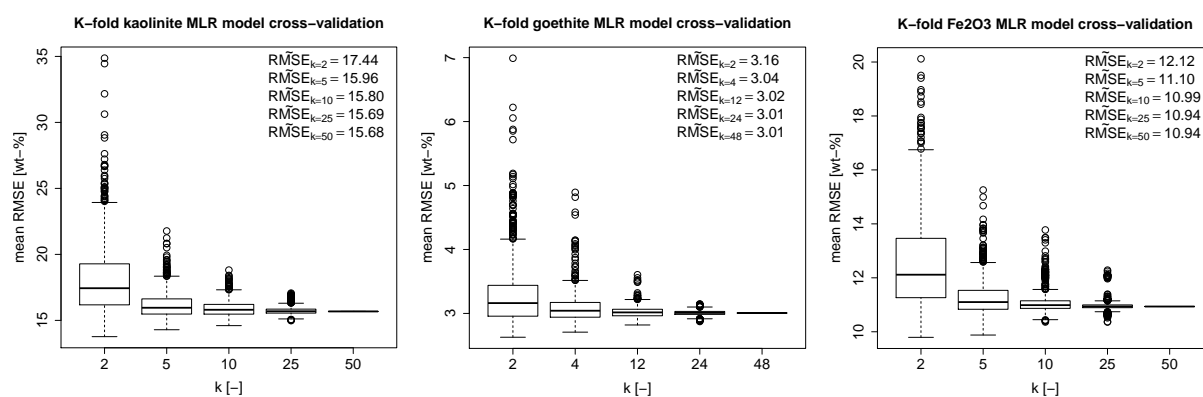


Figure 6.51.: K-fold cross-validation result for VNIR-SWIR MLR models of kaolinite (left), weight fraction limited goethite (middle) and  $\text{Fe}_2\text{O}_3$  (right).

The influence of the differences between corresponding field and laboratory reflectance spectra (see Section 6.7.1.1 - Lab & Field Reference Spectra Pre-Processing — Reference Spectra Validation) on the quantification results needed to be evaluated. Hence, the laboratory spectra trained kaolinite and  $\text{Fe}_2\text{O}_3$  MLR models were run on 19 corresponding field spectra. Figure 6.52 plots the laboratory spectra derived abundances against the field spectra derived ones. Both models yielded a RMSE of 2 wt-%, however the point distributions in the scatter plots revealed

a more differentiated model performance.

The 19 kaolinite abundances had little spread. The majority fell in the range 20-30 wt-%, thus limiting the validity of conclusion drawn. A group of outliers was apparent. It encompassed samples PVS10\_15, PVS10\_16 and PVS10\_18 which earlier stood out due to their reduced intensity and smaller relative clay absorption depth in the field spectra. In contrast, their field spectra derived abundances were about 1.6 times as high as the lab spectra derived values. This may seem contradicting but results from the multivariate nature of the MLR model. It depends not solely on an absorption depth but combines five feature vector parameters in a concurring manner. The model parameter MIN-2206nm-s2-PO (transform feature power) appeared to propagate relative absorption depth into the abundances. Exclusion of this parameter led to a significant decrease of the prediction accuracy of the model. It was therefore retained. In conclusion it must be noted that the kaolinite model was developed, and hence optimised, to the laboratory spectra. The majority of the laboratory spectra showed a pronounced 2,200 nm Al–OH absorption. An image or field spectrum, exhibiting a well developed but instead contrast reduced absorption, lays therefore outside of the model’s calibration range. An erroneous abundance estimate may result from this extrapolation.

The  $\text{Fe}_2\text{O}_3$  abundances instead were well spread between 0-20 wt-% and followed a linear trend. Negative abundances for samples PVS10\_14 and PVS10\_21 lay within the model’s training accuracy. An increasing overestimation of the field spectra derived values was apparent for higher abundances. This was well explained by the more pronounced iron absorption features in the field spectra compared to the laboratory ones.

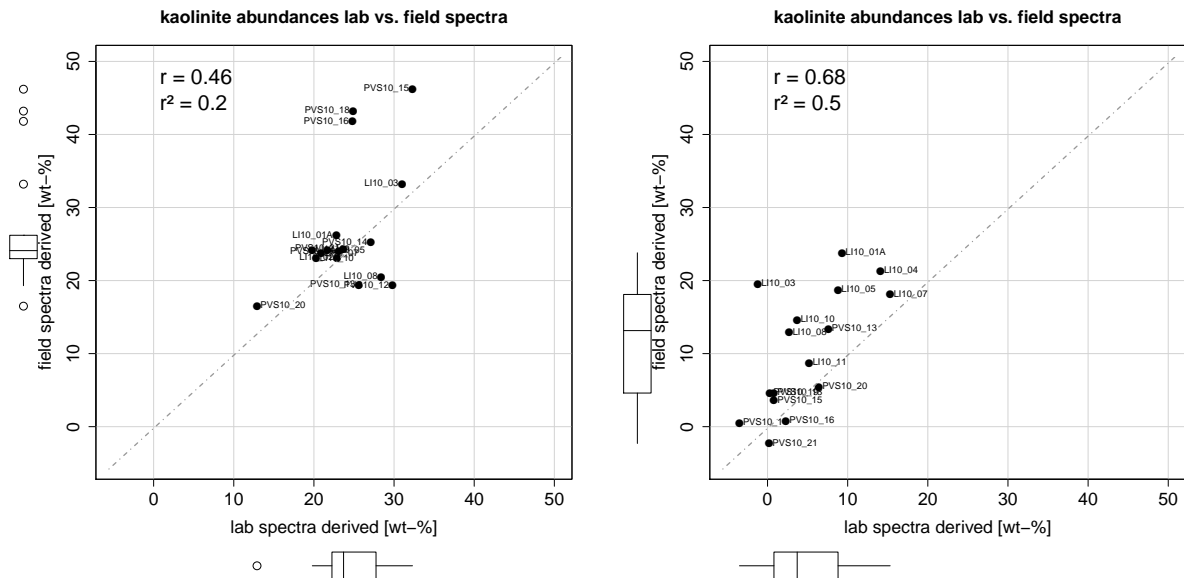


Figure 6.52.: MLR model quality for kaolinite (left) and  $\text{Fe}_2\text{O}_3$  (right). Shown are laboratory versus field spectra derived mineral abundances / element fractions for the 19 corresponding targets.

In a final validation step, the kaolinite, goethite and  $\text{Fe}_2\text{O}_3$  MLR models were applied to HyMap resampled USGS library spectra and the previously excluded MIX\_LO\_01 and LO10\_03 samples rich in goethite. Table 6.17 summarises the abundances calculated by the kaolinite MLR model. Spectrum “kaolini1.spc Kaolinite CM9” was not quantified as its short-wavelength shoulder of the Al–OH absorption had a considerably different shape than the other library and training

spectra. In general the computed kaolinite fractions ranging 19-31 wt-% appeared rather low. The probable cause for this apparent underestimation was a largely deviating spectral characteristic of the Al–OH absorption between laboratory and library spectra. The latter showed a considerably higher intensity and spectral contrast, but also a different absorption shape with a shifted short-wavelength flank. As expected, this once more underlined the limited transferability of the in-situ data trained model. Table 6.18 summarises the abundances calculated by the kaolinite and Fe<sub>2</sub>O<sub>3</sub> MLR models. The goethite abundances ranging 2-9 wt-% were dramatically underestimated. The two high goethite abundance samples MIX\_LO\_01 and LO10\_03, which were excluded during the MLR model development, should have yielded values close to the NORMA modelled 60 wt-%, but instead were estimated only about 6 wt-%. This demonstrated that extrapolation of the goethite model trained on the abundance range 0...15 wt-% to higher abundances fails. As a consequence was the goethite model not applied to the imagery since compliance of the imaged surface with the training data abundance range of 0...15 wt-% can not be warranted. The Fe<sub>2</sub>O<sub>3</sub> abundances were plausible. The low 3 wt-% value was caused by the nearly featureless library spectrum “goethit2.spc Goethite HS36.3”, clearly out of the scope of the MLR model. High NORMA goethite samples MIX\_LO\_01 and LO10\_03 appeared as outliers already during model training (cf. Figure 6.34 on page 124).

Table 6.17.: Validation of Sokolov trained kaolinite MLR model on USGS library spectra of kaolinite (Clark et al., 1993). Given are spectra names, kaolinite fractions after USGS description (NORMA modelled if composition provided) and MLR calculated abundances.

USGS library kaolinite spectra		MLR modelled abundance [wt-%]
name	description	
kaolini1.spc Kaolinite CM9	87 wt-% (NORMA)	pattern mismatch
kaolini2.spc Kaolinite KGa-1 (wxy1)	89 wt-% (NORMA)	21
kaolini3.spc Kaolinite KGa-2 (pxyl)	87 wt-% (NORMA)	19
kaolini4.spc Kaolinite KL502 (pxyl)	90 vol-% (microscope)	31
kaolini5.spc Kaolinite GDS11 <63um	kaolinite & anatase (XRD)	23
kaolini6.spc Kaolinite CM3	87 wt-% (NORMA)	21
kaolini7.spc Kaolinite CM5	85 wt-% (NORMA)	19
kaolini8.spc Kaolinite CM7	83 wt-% (NORMA)	25

Table 6.18.: Validation of Sokolov trained goethite & Fe<sub>2</sub>O<sub>3</sub> MLR models on USGS library spectra of goethite (Clark et al., 1993) and previously excluded samples. Given are spectra names, goethite fractions after USGS description (NORMA modelled for Sokolov samples) and MLR calculated abundances.

USGS library goethite spectra & excluded		MLR modelled abundance [wt-%]	
name	description	goethite	Fe <sub>2</sub> O <sub>3</sub>
goethit1.spc Goethite WS222	no visible impurities	6	36
goethit2.spc Goethite HS36.3	90 vol-% (microscope)	2	3
goethit4.spc Goethite WS220	40 vol-% (microscope)	9	52
LO10_03	54 wt-% (NORMA)	5	26
MIX_LO_1	61 wt-% (NORMA)	6	33



### 6.7.1.5. Thermal Domain Quantification — Model Cross-Validation & Plausibility Checks

TIR model cross-validation, in contrast to the reflective case, was based on fixed expert-derived training and testing sets due to the particularity of using image spectra rather than laboratory spectra. The testing set of 29 targets was formed from the complement of the training set in the complete target set. A second extended testing set comprised of all 64 targets. Obvious outliers, whether explainable or not, were not eliminated from both testing sets.

The quartz regression model's training error of 9 wt-% RMSE compares to a validation error of 15 wt-% RMSE based on the complement testing set and 12 wt-% RMSE based on the extended testing set covering all 64 targets. Noteworthy was the excellent match between quartz regression models set up with only the confident samples (including 0/100 wt-% support points) and all samples (excluding 0/100 wt-% support points) as shown in Figure 6.43 on page 132. This did not apply to kaolinite where the described cluster of low wt-% outliers and the sample sparsity towards higher abundances led to a considerable rotation of the regression function (see Figures 6.41 and 6.42 on page 131). For kaolinite #72/#74 (8,770/9,680 nm) the validation error was 15 wt-% RMSE (complement testing set) and 13 wt-% RMSE (extended testing set) while for #76/#77 (10,624/11,230 nm) the validation error was 21 wt-% RMSE (complement testing set) and 17 wt-% RMSE (extended testing set).

Though appearing better correlated in respect of  $r$  value, the kaolinite #76/#77 regression evaluated inferior compared to the #72/#74 one. The exponential trend was disturbed by sample points scattering comparably strong in respect to the low span of the ratio values. Latter proved to be heavily influenced by sensor noise. Hence the #76/#77 (10,624/11,230 nm) ratio had to be disregarded.

Given the scatter plot distributions, first- or second-order polynomial models would have had difficulties in adapting the strong drop observed for  $> 80$  wt-% quartz and  $< 20$  wt% kaolinite. A severe quartz overestimation ( $\gg 100$  wt-%) or kaolinite underestimation ( $\ll 0$  wt-%) from emissivity ratios would have resulted for such surfaces. The inclusion of theoretical ratio values derived from pure library spectra emphasised the non-linear trend observed with the sample points. It must be noted that the pure quartz and pure kaolinite library spectra could not be verified against appropriate image spectra, thus the derived ratios remain somewhat uncertain. Due to the sparsity of the kaolinite samples towards higher abundances ( $> 50$  wt-%) a verification of the regression function was not possible there. However, for quartz, a good match was observed and the apparent exponential behavior fitted well to expectations, e.g. covering of coarse quartz grains with finer kaolinite particles and weathering of feldspars forming a sericite coating mimicking clay features.

Comparison of the NORMA with the SEDNORM #72/#74 kaolinite model yielded an abundance difference of 2 wt-% RMSE over the full range. Given the stable regression for quartz and the three times lower abundance differences seen for the two normative models, a negligible abundance difference caused by differing regressions can be expected.

### 6.7.1.6. Summary Validation of Processing Steps

Table 6.19 lists aggregated statistical figures for the workflow-integrated validation steps. These values strongly depend on the used instruments, the available data and the intended application.

Summing up, the normalised field, laboratory and image spectra were very similar in shape and absorption feature contrast. This showed that the spectral data acquisition and pre-processing was accomplished successfully. Individually, large differences in spectral intensity were found. An average intensity increase of spectral reflectance of 30% (stddev 8%) from image to field and another 1% (stddev 7%) from field to laboratory was identified. The emissivity spectra showed less differences and variability with 3% (stddev 2%) from field to laboratory and 1% (stddev 1%) from image to field. Given the humid weather conditions and sampling/measurement protocols (see Section 6.2 - *Data Sets and Measurement Protocols*), it is assumed that the main factors for this discrepancy were differences in moisture, spatial scale effects, and perturbation of the natural surface by the sampling.

The normative modelling was an indispensable step towards the aspired mineral quantification. The model cross-validation between SEDNORM and NORMA showed promisingly small differences in the normative abundances with a RMSE <1 wt-% (stddev 1.2 wt-%). However, no error estimate against an independent quantitative measure was possible. Instead, plausibility checks against the qualitative XRD results and quantitative elemental analysis were conducted by an expert (pers. communication Dr. Derek Rogge, DLR). Clearly, the quantitative models built on the normative mineral fractions and the derived mineral abundance maps can only be as good as the normative modelling result itself.

Validation of the quantitative models in the thermal domain proved an excellent stability of the quartz model. The kaolinite model instead was unstable and strongly dependent on the training data. The addition of two supporting mineral endmembers from a spectral library overcame this limitation. The regression line was stabilised and the model's plausibility was enhanced. Both models settled at a prediction error of 15 wt-% RMSE.

Cross-validation in the reflective domain showed ample stability of the individual predictors and final MLR models. This indicates a sufficient distribution of the training data. The models settled at prediction errors comparable to the thermal domain with 16 wt-% RMSE for kaolinite, 4 wt-% RMSE for goethite (0-15 wt-%) and 11 wt-% RMSE for  $\text{Fe}_2\text{O}_3$ .

Change from laboratory to image spectra led to the loss of well correlating spectral features. The model's accuracy significantly decreased. The observed differences in spectral intensity and feature contrast could be mitigated, but not eliminated, by the use of CWT based methods. Comparison of the field and laboratory spectra quantification results gave an average transfer error of 2 wt-% RMSE for kaolinite and  $\text{Fe}_2\text{O}_3$ . The use of multivariate statistics leads to a strong model adaptation to the data. The quantification results depend on how well and how balanced the training data can represent the different mineral assemblages found on site. The strong kaolinite abundance overestimation (60%) for three validation samples (PVS10\_15, PVS10\_16, PVS10\_18) revealed the difficulty to find generic features for quantification and to balance the contributions of differing mineral assemblages in the training set. A systematic error must be expected when leaving the mineral assemblages or geochemical characteristics covered by the majority of the training set members.

Table 6.19.: Summary of the validation methods and results for the individual processing steps.

Pre-processed Lab & Field Spectral Measurements		
variable	validation method	error estimate
reflectance [0...1]	pre-processed field versus laboratory reflectance spectra ⇒ per band linear regression, offset forced to 0 ⇒ sample size = 19	average gain correction ⇒ 1.01, stddev=0.07
emissivity [0...1]	pre-processed field versus laboratory emissivity spectra ⇒ per band linear regression, offset forced to 0 ⇒ sample size = 26	average gain correction ⇒ 0.97, stddev=0.02

Pre-processed Airborne Imagery		
variable	validation method	error estimate
reflectance [0...1]	field versus image reflectance spectra ⇒ per band linear regression ⇒ sample size = 11	average gain correction ⇒ 1.31, stddev=0.08 average offset correction ⇒ -0.01, stddev=0.02
emissivity [0...1]	field versus image emissivity spectra ⇒ per band linear regression, offset forced to 0 ⇒ sample size = 18	average gain correction ⇒ 0.99, stddev=0.01

Normative Modelling		
variable	validation method	error estimate
abundance [0...100 wt-%]	NORMA versus SEDNORM derived mineral abundances ⇒ statistics for 10 minerals in common ⇒ sample size = 80	average median difference ⇒ -0.3, stddev=1.2 wt-% average error ⇒ RMSE=0.7 wt-%

Quantitative Models Reflective Domain		
variable	validation method	error estimate
kaolinite [0...65 wt-%]	MLR model based on NORMA modelled abundances and laboratory spectra derived transform vector parameters <sup>(1)</sup> ⇒ K-fold cross-validation <sup>(2)</sup> ⇒ laboratory versus field spectra derived abundances <sup>(3)</sup> ⇒ sample size = 50 <sup>(1)</sup> / 50 <sup>(2)</sup> / 19 <sup>(3)</sup>	training error <sup>(1)</sup> ⇒ r <sup>2</sup> =0.52, RMSE=14 wt-% validation error <sup>(2)</sup> ⇒ RMSE=16 wt-% transfer error <sup>(3)</sup> ⇒ RMSE=2 wt-%
goethite [0...15 wt-%]	MLR model based on NORMA modelled abundances and laboratory spectra derived transform vector parameters <sup>(1)</sup> ⇒ K-fold cross-validation <sup>(2)</sup> ⇒ sample size = 48 <sup>(1)</sup> / 48 <sup>(2)</sup>	training error <sup>(1)</sup> ⇒ r <sup>2</sup> =0.63, RMSE=3 wt-% validation error <sup>(2)</sup> ⇒ RMSE=4 wt-%
Fe <sub>2</sub> O <sub>3</sub> [0...70 wt-%]	MLR model based on element fractions and laboratory spectra derived transform vector parameters <sup>(1)</sup> ⇒ K-fold cross-validation <sup>(2)</sup> ⇒ laboratory versus field spectra derived abundances <sup>(3)</sup> ⇒ sample size = 50 <sup>(1)</sup> / 50 <sup>(2)</sup> / 19 <sup>(3)</sup>	training error <sup>(1)</sup> ⇒ r <sup>2</sup> =0.50, RMSE=10 wt-% validation error <sup>(2)</sup> ⇒ RMSE=11 wt-% transfer error <sup>(3)</sup> ⇒ RMSE=2 wt-%

Quantitative Models Thermal Domain		
variable	validation method	error estimate
quartz [0...95 wt-%]	exponential model based on NORMA modelled abundances and image spectra band ratio <sup>(1)</sup> ⇒ confident vs. non-confident cross-validation <sup>(2)</sup> ⇒ all spectra derived abundances <sup>(3)</sup> ⇒ sample size = 37 <sup>(1)</sup> / 64 <sup>(2)</sup> / 64 <sup>(3)</sup>	training error <sup>(1)</sup> ⇒ r <sup>2</sup> =0.97, RMSE=9 wt-% validation error <sup>(2)</sup> ⇒ RMSE=15 wt-% transfer error <sup>(3)</sup> ⇒ r <sup>2</sup> =0.88, RMSE=12 wt-%
kaolinite [0...65 wt-%]	exponential model based on NORMA modelled abundances and image spectra band ratio <sup>(1)</sup> ⇒ confident vs. non-confident cross-validation <sup>(2)</sup> ⇒ all spectra derived abundances <sup>(3)</sup> ⇒ sample size = 37 <sup>(1)</sup> / 64 <sup>(2)</sup> / 64 <sup>(3)</sup>	training error <sup>(1)</sup> ⇒ r <sup>2</sup> =0.83, RMSE=9 wt-% validation error <sup>(2)</sup> ⇒ RMSE=15 wt-% transfer error <sup>(3)</sup> ⇒ r <sup>2</sup> =0.32, RMSE=13 wt-%

## 6.7.2. Validation of Quantification Results

Complementary to the validation of the individual processing steps covered by the previous section, this section evaluates the quantitative mineral mapping results themselves.

### 6.7.2.1. Mineral Abundances — Sum in Range 0...100 Validation

In the reflective domain, the kaolinite and  $\text{Fe}_2\text{O}_3$  abundance maps were tested individually for the admissible 0...100 wt-% abundance range. 230 pixels with a calculated  $\text{Fe}_2\text{O}_3$  fraction below the model's error margin of -11 wt-% RMSE, i.e. 0.01% of all pixels in the active mines subset, existed. The most extreme outliers in the range -15 to -20 wt-% (totalling 30 pixels) were caused by water surfaces. The kaolinite abundances met the admissible range  $0...100 \pm 16$  wt-% everywhere. Derived from the same imagery, the kaolinite and  $\text{Fe}_2\text{O}_3$  abundance maps could be readily summed on a per-pixel basis. The physically possible abundance range 0...100 wt-% was met for practically the whole active mines image subset (31 km<sup>2</sup>). Outliers of up to 113 wt-% and down to -9 wt-% existed that however were within the individual model's prediction errors of 16 and 11 wt-% RMSE. The latter summation is a weak test as the average sum of NORMA kaolinite and  $\text{Fe}_2\text{O}_3$  in the 50 samples was  $35 \pm 21$  wt-%. Both abundances must be compared with care as one is a mineral and the other an elemental oxide standing for goethite ( $2\text{Fe}_2\text{O}_3 + 1\text{H}_2\text{O}$ ) and jarosite ( $1\text{K}_2\text{O} + 3\text{Fe}_2\text{O}_3 + 4\text{SO}_3 + 6\text{H}_2\text{O}$ ) in the normative modelling. They have greatly varying molar mass ratios of  $\frac{M_{\text{goethite}}}{2M_{\text{Fe}_2\text{O}_3}} = 1.06$  and  $\frac{M_{\text{jarosite}}}{3M_{\text{Fe}_2\text{O}_3}} = 2.09$  in relation to the bound  $\text{Fe}_2\text{O}_3$ . Since jarosite is rarely exposed in Sokolov above comparison holds.

Likewise in the thermal domain, the kaolinite and quartz abundance maps could be readily summed due to their identical spatial map information. Quartz and kaolinite are the major minerals found in Sokolov. The sum of both minerals' NORMA abundances over the 50 samples gave an average of  $53 \pm 26$  wt-% which increased to  $62 \pm 18$  wt-% when excluding the 8 samples with a normative quartz fraction below 1 wt-%. The summed range 0...100 wt-% was met for practically the whole image coverage (500 km<sup>2</sup>). Negative outliers occurred for metal roofs yielding erroneous quartz abundances down to -190 wt-%. This was caused by a slope reversal of the AHS bands #72/#78. Positive outliers up to +220 wt-% appeared in the kaolinite abundance map for a group of 40 soil pixels. An exceptional deep emissivity drop in AHS band #74 led to this overestimation by the #72/#74 ratio model.

Superposition of the  $\text{Fe}_2\text{O}_3$  map with the AHS derived quartz and kaolinite maps was done by nearest neighbour resampling latter two onto the required higher spatial resolution (see Figure 6.53). The sum of all three abundances conformed the physically possible range 0...100 wt-% for the active mines subset except for a small region showing both high  $\text{Fe}_2\text{O}_3$  and quartz abundances (marked by A in Figure 6.53). The abundance sum maximum of 113 wt-% reached in this 50 pixel area lies within the individual model's tolerance of 11-15 wt-% RMSE.

### 6.7.2.2. Ferric Oxide Abundances — Overlap with Siderite Classification Plausibility Check

Figure 6.54 displays the  $\text{Fe}_2\text{O}_3$  abundance distribution in comparison to high confidence siderite occurrences according to the Wavelet Coherence Mapper investigation of Section 6.5.1. A pixel was regarded a confident siderite occurrence when the coherence in wavelet scale 4 or 8 reached a value  $\geq$  the 99.99%-quantile. For the 443 pixels classified as siderite, an average  $\text{Fe}_2\text{O}_3$

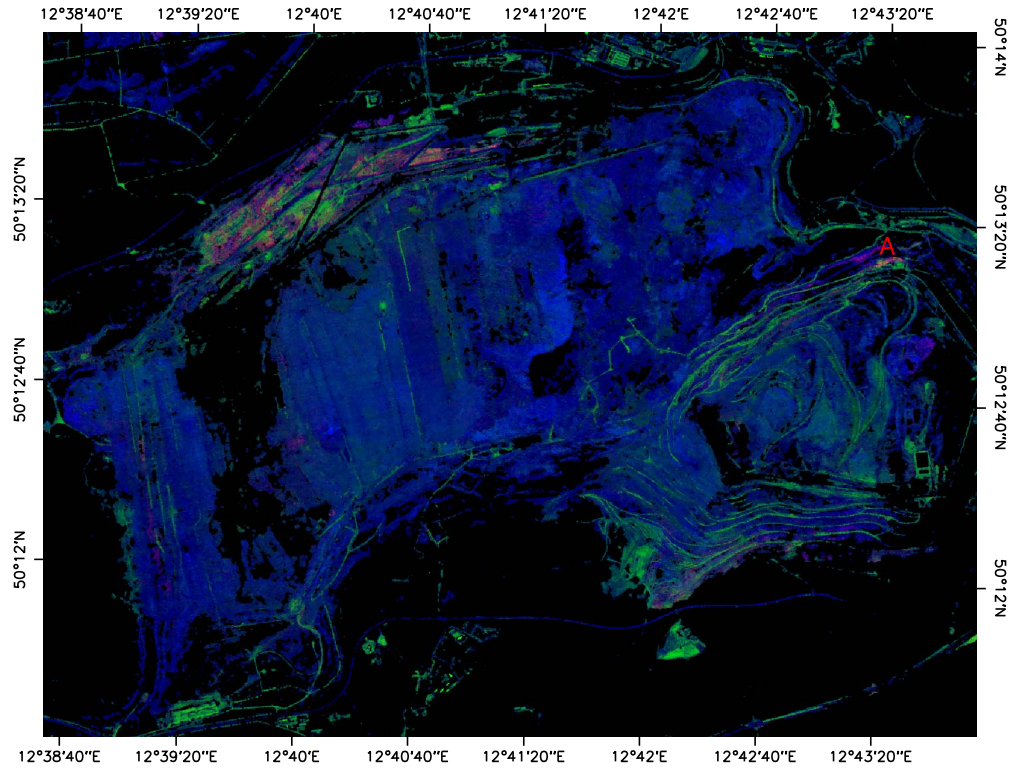


Figure 6.53.: Color composite showing abundance sum constraint conformity of the reflective domain  $\text{Fe}_2\text{O}_3$  (R, 0-40 wt-%) together with the thermal domain quartz (G, 0-40 wt-%) and kaolinite (B, 0-80 wt-%) abundances (linear stretch). Area exceeding 100 wt-% margin marked by A.

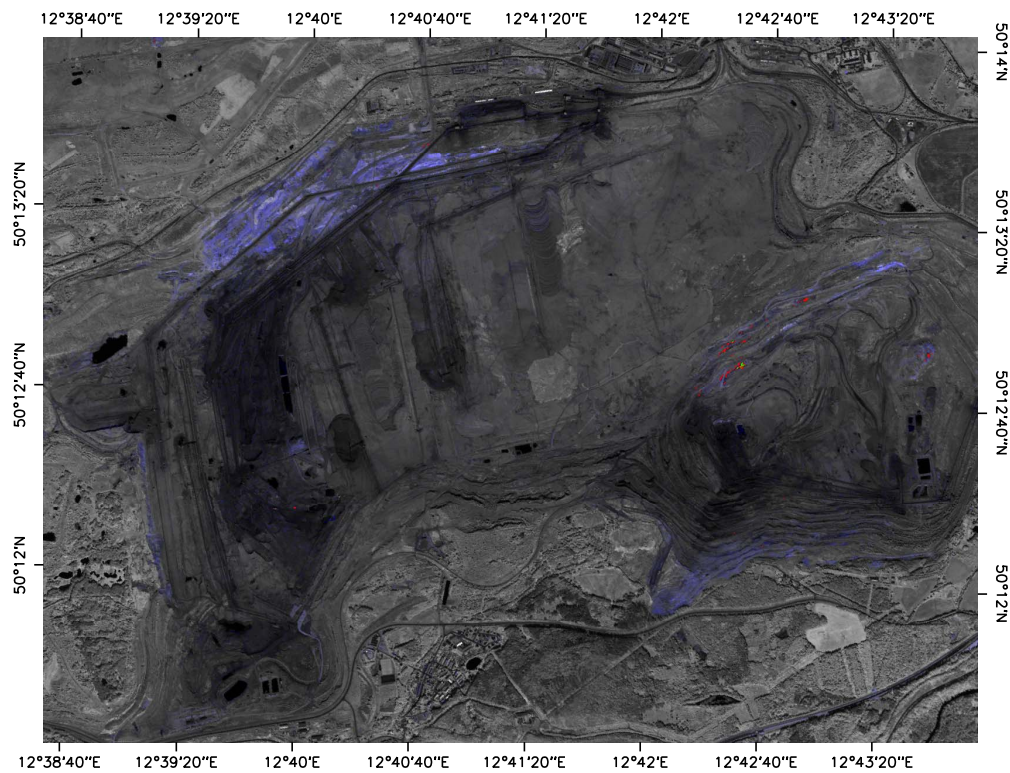


Figure 6.54.: Comparison of siderite versus  $\text{Fe}_2\text{O}_3$  distributions.  $\text{Fe}_2\text{O}_3$  abundances given in blue hues (blue-white color ramp, 0-80 wt-%). Confident abundances of siderite according to wavelet coherence  $\geq 99,99\%$ -quantile marked in red (scales 4 and 8, intersection marked green).

abundance of only 3 wt-% with a standard deviation of 4 wt-% was reported. Hence practically no overlap existed in the active mines subset. This fits well to the assumption that siderite occurs as primary mineral in the Sokolov country rocks, and not as secondary precipitate which could indicate AMD. By contrast,  $\text{Fe}_2\text{O}_3$  accumulations are related to goethite or jarosite that both may indicate AMD affected areas in which fast neutralising minerals such as siderite would have been depleted. No overlap between siderite and ferric oxide abundances are hence expected.

### 6.7.2.3. Kaolinite Abundances — Overlap with Clay Classification Plausibility Check

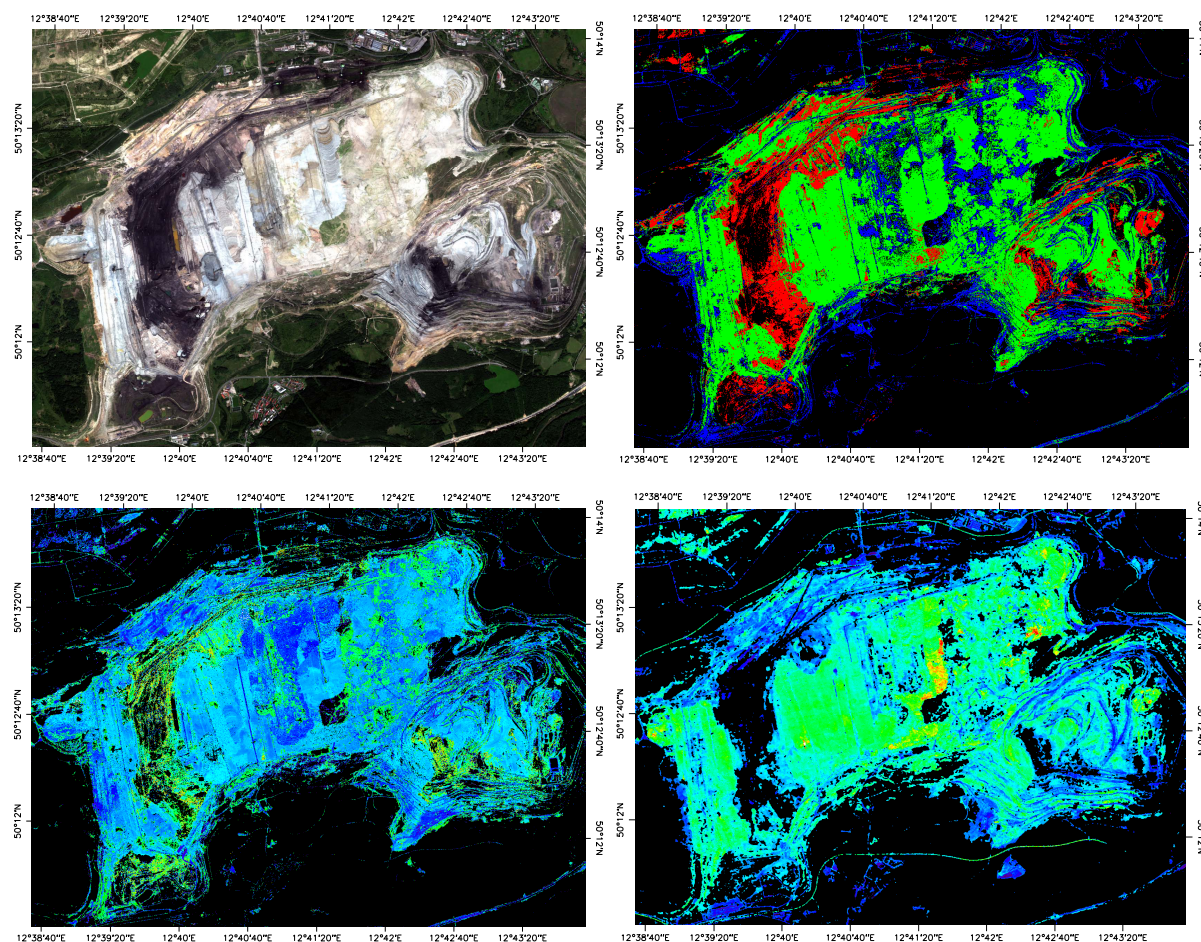


Figure 6.55.: Comparison of the kaolinite abundance result in the reflective domain (bottom left) & thermal domain (bottom right) versus the clay classification (top right). Top left shows true color composite for reference. See the original Figures 6.11, 6.28, 6.32 and 6.45 for explanations and color keys, respectively.

The kaolinite abundance patterns in both reflective and thermal domain followed in general the clay classification results (see Figure 6.55). In the reflective domain, high kaolinite abundances (green-yellow) appeared in coal-clay mixtures around the coal seams and coal transportation paths. It matches the kaolinite classification (red). This indicates that kaolinite quantification relied on the shape of the Al–OH, best developed in coal-clay mixtures as confirmed by the classification. This is an obvious but not surprising shortcoming of the model, which was not developed on coal-clay mixtures. The backfill areas show an apparent link between elevated kaolinite abundances and montmorillonite class assignment (blue). This was caused by the wavelength shift of the Al–OH absorption’s left shoulder. Whether this relation is real could not be determined as no samples have been collected in the respective areas.

In the thermal domain, no link between coal admixture and high kaolinite abundances exists, but the opposite relation is apparent. Coal-clay mixtures around the coal seams and coal transportation paths appear with a very low kaolinite abundance as would be expected. A negative relation of montmorillonite class assignment with the abundances may be interpreted. This should be exercised with caution due to the time gap between the HyMap and AHS flights.

#### 6.7.2.4. Kaolinite Abundances — Reflective versus Thermal Domain Cross-Validation

The direct comparison of the kaolinite abundances derived from the reflective and thermal datasets showed largely deviating abundance patterns for the active mines subset (see Figures 6.32 and 6.45 or Figure 6.55 for a direct comparison). Looking back on the abundance overestimation found when applying the kaolinite MLR model to field spectra (see Section 6.7.1.4 - *Reflective Domain Quantification — Model Cross-Validation & Plausibility Checks*), the observed high kaolinite abundances in the coal seams appeared to be caused by a shortcoming of the MLR model. Model training was conducted on a sample set that contained only few coal mixtures ( $8 \times C_{tot} > 10$  wt-%,  $3 \times C_{tot} > 20$  wt-%). The clay classification also showed that the kaolinite Al–OH absorption feature was best developed in the coal-mixing areas. In combination, this led to the observed overestimation of the kaolinite abundance. The reflective domain MLR kaolinite model hence must be ranked insufficient for mixtures of kaolinite and coal. Another probable cause for differences is the one year time offset between reflective and thermal flight campaigns during which large scale material relocations can be seen to have taken place when comparing the active mining fonts in Figure 6.55.

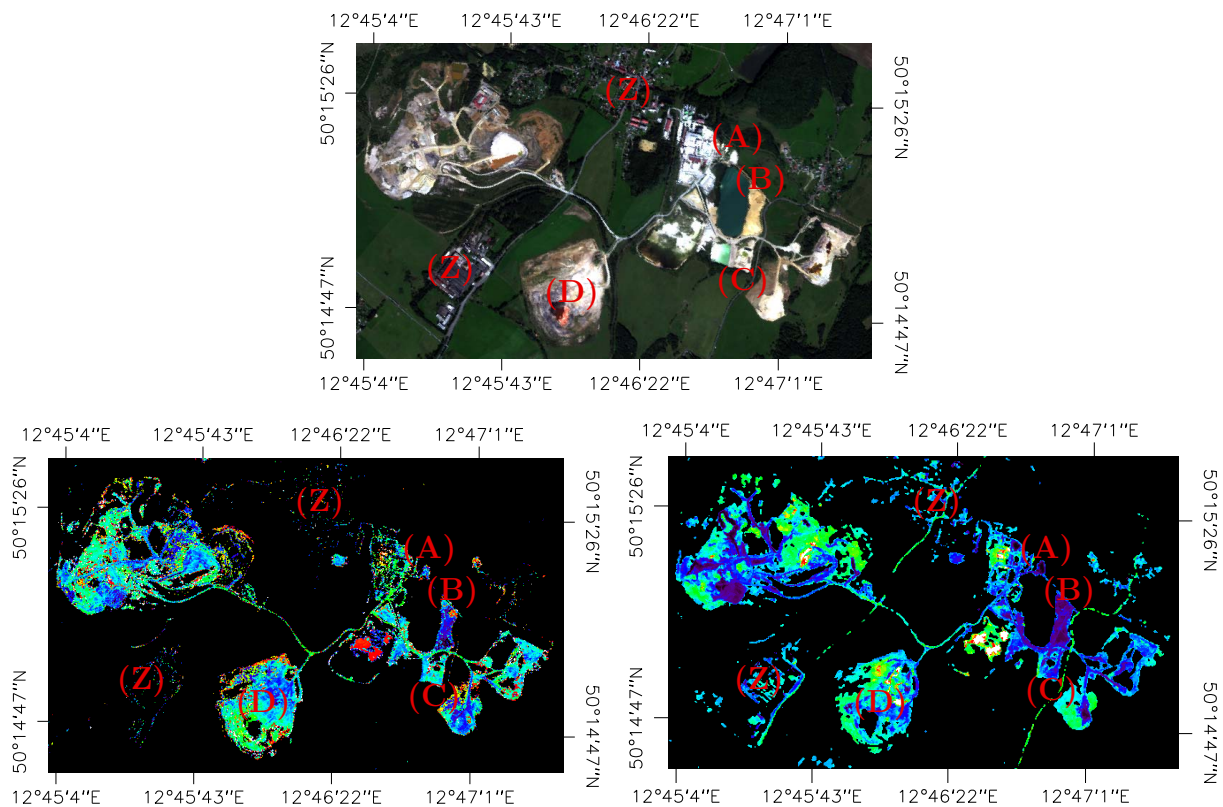


Figure 6.56.: Comparison of kaolinite quantification for a kaolinite extraction site. True color composite of the reflectance image subset (top). Kaolinite abundance maps based on HyMap Wavelet Transform Feature Analysis derived MLR model (left) and based on AHS band ratio exponential regression model (right) both linearly stretched 0-60 wt-% with rainbow color ramp.

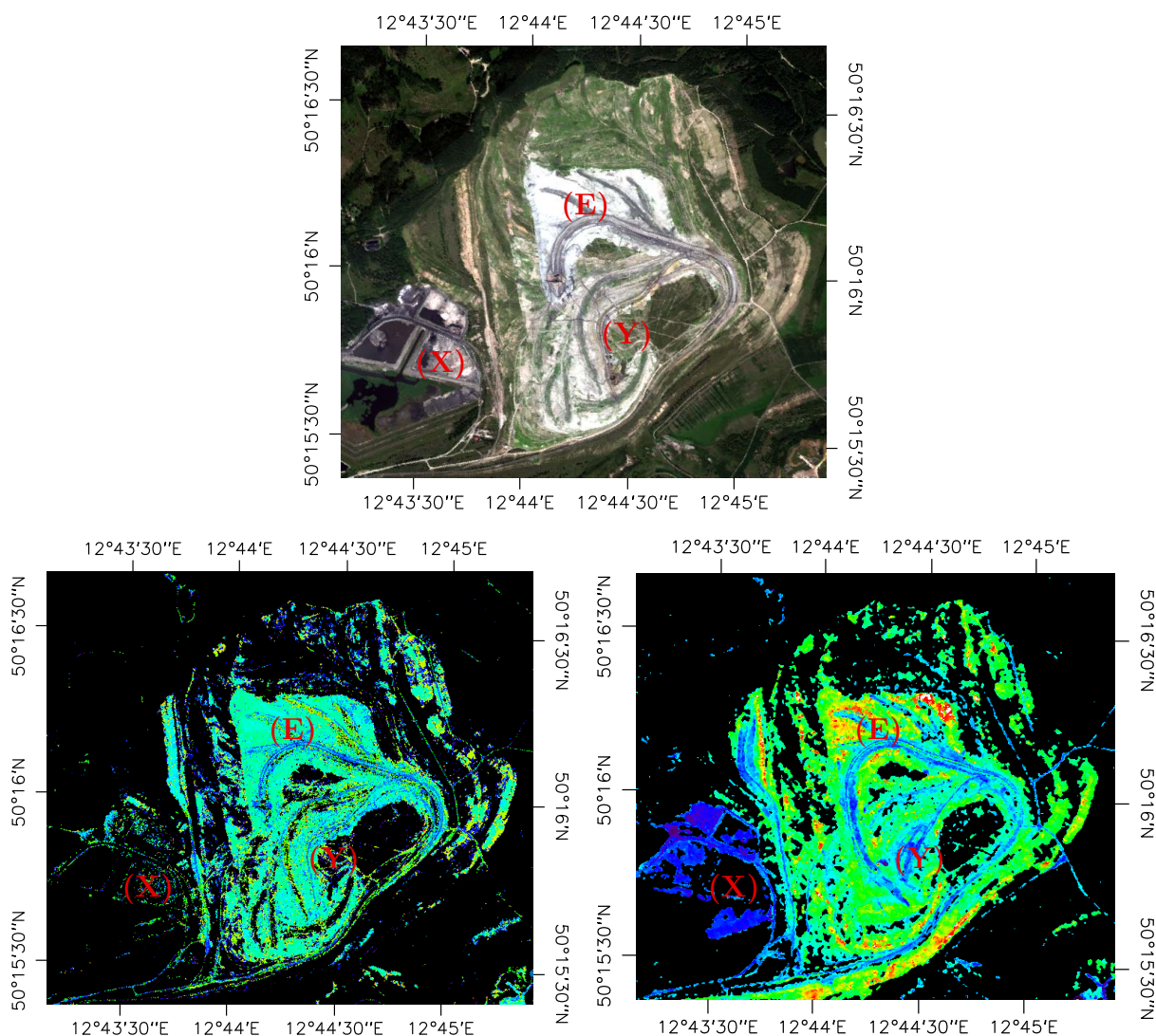


Figure 6.57.: Comparison of kaolinite quantification for an overburden dump site. True color composite of the reflectance image subset (top). Kaolinite abundance maps based on HyMap Wavelet Transform Feature Analysis derived MLR model (left) and based on AHS band ratio exponential regression model (right) both linearly stretched 0-60 wt-% with rainbow color ramp.

Two areas that appeared less reformed during this time span, that were better represented by the taken field samples and that showed a large span between highest and lowest kaolinite abundance were chosen for the result evaluation. Figures 6.56 and 6.57 depict an kaolinite extraction site and an overburden dump site, respectively.

The kaolinite distribution patterns derived from the reflective and thermal domain data agreed well in these areas. Generally, qualitative differences in the abundances showed similar patterns albeit their absolute values were offset. The labels (X), (Y) and (Z) in both figures demonstrate an important difference between the simple band ratio model and the wavelet transform based MLR model with common feature pattern matching. The band ratio is calculated for all image pixels that passed the water and vegetation masking, irrespective whether the model was trained onto such a surface type or not. The Wavelet Transform Feature Analysis instead prevents application of the quantitative model to pixel spectra not meeting the feature pattern established during the model training procedure. The dry-fallen slurry ponds (label X) to the west of the dump not exhibiting the typical Al–OH absorption feature hence were omitted.



Fractional vegetation cover led to omissions in the reflective domain, while in the thermal domain vegetation (having a uniform emissivity contrast reducing influence) frequently passed the masking (see label Y). Also roads and car parks in settlements were typically omitted in the reflective domain (label Z).

Table 6.20.: Deviations of kaolinite quantitative maps in the reflective & thermal domain. Given are the average abundance over selected image regions of interest and the standard deviation together with NORMA modelled kaolinite abundances of samples taken in the area.

label & description	area [m <sup>2</sup> ]	sample NORMA [wt-%]	reflective domain [wt-%]	thermal domain [wt-%]
(A) fresh, white quartz sand pile left from kaolinite separation process	2,300	1.4	20 ± 4	4 ± 4
(B) oxidised, yellow quartz sand pile left from kaolinite separation process	12,000	4.4/2.3/3.4	17 ± 6	7 ± 3
(C) weathered, gray-brownish approach to kaolinite extraction site	15,000	21.6/38.8	26 ± 7	15 ± 9
(D) weathered, dark gray-brownish former kaolinite extraction site	151,000	N/A	25 ± 8	20 ± 14
(E) fresh, white-grayish overburden material on top of dump	320,000	N/A	27 ± 7	32 ± 12

The supposedly invariant areas (A) to (E) were used to estimate average kaolinite abundance differences between reflective and thermal domain as well as the NORMA reference. Table 6.20 lists the deviations. The kaolinite MLR model in the reflective domain appeared to overestimate small kaolinite fractions, as was expected from the respective scatter plot (Figure 6.34) discussed in the concluding remarks of Section 6.6.1 - *Wavelet based Quantification of Goethite & Kaolinite*. The thermal domain abundances seemed to follow the general trends given by the NORMA kaolinite fractions of the samples and the expectations from the field site descriptions. The average deviation of 9 wt-% between both domains was within the model's performance margin of 16 wt-% RMSE.

### 6.7.2.5. Summary Validation of Quantification Results

Summing up, the validation of the quantitative results showed that the data and model quality was sufficient to derive physically sound quantitative mineral distribution patterns for the Sokolov demonstration site. It was demonstrated that the individual quantitative layers as well as the summed abundances, over reflective and thermal domain, conformed the 0...100 wt-% constraint. Violations of this physically possible abundance range were sparse and typically within the model's estimated prediction error margins. Exceptions were found for the thermal domain band ratio models that had a systematic problem with metal roofs and a locally constrained area of atypically increased thermal clay feature contrast. Both resulted from the models not being trained on such surfaces and the simple band ratio models not rejecting image pixels that differ from the training set characteristics.

The cross-validation of reflective and thermal domain kaolinite abundances yielded a mixed result. It was apparent that the wavelet transform feature based MLR model systematically overestimated the kaolinite abundance for admixtures of coal. This shortcoming could be re-

lated to such mixtures being underrepresented in the model's training data. The thermal domain band ratio model instead seemed less affected by the admixture of coal.

When focus was set on areas dominated by kaolinite-quartz mixtures, the quantification results in the reflective and thermal domain matched in spatial coverage and agreed in the general abundance trends. However, an overestimation of small kaolinite fractions and an underestimation of high fractions by the MLR model were found. This confirmed the observed model quality in the previous model cross-validation section. The thermal band ratio model well followed the trends seen in field samples.

A comparison of the siderite classification results with the  $\text{Fe}_2\text{O}_3$  abundance map showed no overlap. This was expected, as acidity neutralising siderite should not occur at locations where AMD precipitates such as jarosite or goethite could potentially exist and would be indicated by high  $\text{Fe}_2\text{O}_3$  abundances. Natural co-occurrence of siderite and  $\text{Fe}_2\text{O}_3$  enrichments as well as siderite as AMD precipitate were ruled out due to the Sokolov stratigraphy and field visits. This demonstrates the possibility to differentiate ferrous from ferric iron on the basis of spectral information.

In conclusion, the undertaken pre-processing steps and developed classification/quantification methods facilitated the generation of validated results that were plausible, individually and also in combination over the two distinct spectral domains VNIR-SWIR and TIR. Inconsistencies in coal-clay mixtures could be related to limited transferability of the kaolinite regression model caused by a training set which did not adequately resemble such mineral mixtures.

## 7. Discussion & Summary

The chapter examines the developed wavelet based analysis, thermal pre-processing and workflow validation concepts with regard to the fulfilment of the specific research objectives and to generic methodological aspects. The viewpoint is then changed to the developed concepts application on the Sokolov data, where limitations of the results are discussed. This leads to a review of the concepts practical performance which is concluded by recommendations for their usage.

### 7.1. Critical Discussion of Developments

The major development work of this research was focused on the three specific objectives detailed in Section 1.1 - *Specific Study Objectives*. The following of this section will summarise advantages and disadvantages of the developed methods and the implications for their practical application.

#### 7.1.1. Review of Specific Research Objectives

The first objective was to advance wavelet transform based classification and quantification methods and to demonstrate their practical applicability to hyperspectral imagery.

This goal was reached by the development of Wavelet Coherence Mapper and Wavelet Transform Feature Analysis. Wavelet Coherence Mapper is a classification method that operates on spectral curvature direction and provides similarity spectra. As such it complements shape based methods like Spectral Angle Mapper or Spectral Correlation Mapper which yield only an aggregated similarity score. Wavelet Transform Feature Analysis is a quantification method. It forms a generalisation of the classical Spectral Derivative Analysis and Spectral Feature Analysis while retaining ease of interpretation of a feature based quantification.

Practical applicability of the wavelet based methods on hyperspectral image mosaics was facilitated by limiting the wavelet transform to selected frequency scales in order to confine data volume. The utilisation of the scale number limited wavelet methods to a real-world mineral classification and quantification study was successfully demonstrated.

The second objective was to implement an in-flight calibration module which allows pre-processing of thermal data to a comparable quality level as reached in the reflective domain.

The aim was reached by development of a generic module that integrates ATCOR sensor definitions, radiative transfer calculations, MODTRAN derived atmospheric parameters, image derived temperature estimates and the field reference measurements. As such the thermal module is in compliance with best practises from the reflective domain and integrates into existing pre-processing chains. It can be easily adapted to any multi- or hyperspectral thermal sensor.

The third objective was to develop a concept for validation of a multi-sensor workflow such that the main sources of uncertainty can be identified.

This goal was tackled by incorporating empirical validation steps throughout the whole processing workflow such that a validation chain from input data to the final products was established. The major uncertainty sources can be identified, but their individual contribution to the error budget can not be allotted. It contributes to better understand the uncertainty sources, leading towards the ambition of a quality and accuracy assessment of the whole product generation.

The following Sections 7.1.2, 7.1.3 and 7.1.4 focus on generic methodological aspects of the individual study objectives. The discussion of practical aspects based on the Sokolov study is covered by Section 7.2 - *Critical Discussion of Developed Concepts' Application*.

### 7.1.2. Wavelet based Classification & Quantification Concepts Review

The continuous wavelet transform (CWT) was chosen as basis for the classification and quantification method development due to its favourable linear, multi-scale and complete decomposition of spectral signatures into frequency components. The wavelength localisation of spectral features and their shape information are preserved for subsequent analysis.

#### Theoretical Reflections

CWT represents a bandpass filtering operation. Spectral features of a width that corresponds to the current scale are emphasised, while narrower or wider spectral features are suppressed. Understanding CWT as multi-scale differential operator and cross correlation between spectral signature and wavelet, the choice of the  $\text{DOG}_{m=2}$  as mother wavelet is motivated (see Section 2.4.2 - *Wavelet Choice*). The transforms resemble the second derivative (curvature) of the smoothed spectral signature. The extent of Gaussian smoothing is proportional to the scale. High transform amplitudes localise spectral features matching the wavelet shape.

In summary CWT offers multi-scale decomposition, linear bias elimination, noise separation, genericity and robustness for the analysis of spectral signatures. The multi-scale decomposition effectively separates noise, absorption features and spectral continuum. The former are readily described by local curvature changes which are picked up by the CWT at small to medium scales. High frequency noise is separated into the smallest scale. The continuum contains broad absorptions plus influences of e.g. illumination, moisture and viewing geometry. The continuum curvatures are retrieved in high scales while undifferentiated other influences are largely remedied by the cancelling of linear components. The preservation of spectral shape information in several levels-of-detail is certain to improve feature based classification and quantification. Firstly, the multi-scale decomposition allows to focus on the relevant level-of-detail of spectral features, or a combination thereof. This is especially advantageous for feature superpositions in mixtures. Secondly, complex shapes are retained for similarity comparison and feature parameterization. The method is generic, i.e. no methodological differentiation between maxima, minima, absorptions or absorption shoulders and reflectance, emissivity or absorbance need to be made.

In opposition to these remarkable benefits are some disadvantages of CWT. The high redundancy results in an enormous data volume. The CWT of an image produces an equally sized image for each scale. This imposes a technical limitation depending on the hardware resources available. Despite no methodological limitation, this probably is the reason for a very small number of studies on imagery found in the literature. To avoid excess data volume and processing times, this study limited CWT computation to dyadic scales and the wavelength sampling intervals provided by the sensor. The transform is complete, the full wavelength resolution is retained, but only selected frequency components are available. It keeps fast filter bank implementation of CWT an option for future use.

A second important issue are possible border effects. This is of interest when spectral features close to the margins of the spectral coverage are of interest. An appropriate signal padding must be found to ensure the reliability of the transform values. Here, the first and last two noisy HyMap sensor bands were cropped and a “mirror padding” was used that avoided signal discontinuities at all scales. This influenced all transforms to have zero crossings at the borders, which may mask signal differences there. The possibly affected number of bands (COI) was estimated  $\sqrt{2}s$  which was a tradeoff to preserve valuable bands for analysis. A COI of  $3s$  would be a more cautious choice, but double the number of presumably affected bands.

Novel wavelet based classification and quantification methods exploiting the benefits of CWT were developed under the constraints of dyadic scales and mirror padding. It must be noted that the calibrated/trained models are not directly transferable to another sensor. The information highlighted in a specific scale depends on the sensor’s spectral resolution. Transferability however can be easily achieved by resampling to a common wavelength grid.

### Wavelet Coherence Mapper

The wavelet coherence brings together two highly desirable properties. Firstly, the linear decomposition of signals into multiple scales with the benefits listed above. Secondly, the localisation of the similarity measure in wavelength. This enables to compare two spectra at selectable spectral detail (scale) and at chosen wavelengths for matching spectral behaviour.  $\text{DOG}_{m=2}$  CWT removes local additive and linear bias, while the coherence is a normalised linear correlation measure and thus independent of multiplicative factors and amplitude differences. The coherence measure WTC used here can differentiate positive and negative correlation.

In contrast to the shape driven SAM and SCM does the WTC disregard the feature strength and signal amplitude. WTC works by curvature changes rather than by exact shape match. It accounts for deep and shallow signal features equally. WTC depends on the CWT such that a transform oscillating around zero is the input for the coherence calculation. Then its strength in localising phase locked behaviour can be put to optimal use. In conclusion,  $\text{DOG}_{m=2}$  CWT based wavelet coherence does localise wavelength ranges in which the changes in direction of the spectral curvature match. Deviations in the strength of the changes are causing much less penalty on the coherence than deviations in the direction of the curvature. WTC is hence ineffective when applied directly to reflectance/emissivity spectra or to very large scales.

The developed Wavelet Coherence Mapper makes use of above WTC properties for a pixel-based classification. For each image pixel the coherence is calculated once over all wavelengths and for the desired scales. Selection of scale, wavelength and averaging range to produce rule images as well as selection of a appropriate threshold are decoupled from the expense of calculation. This introduces new degrees of freedom to be mastered by the user. Threshold selection is simplified as i) the normalising character of coherence make it independent from intra- and inter-class amplitude differences, and ii) coherence has a statistical meaning (linear relation between the two signals). Nevertheless a dependence of the threshold on the scene, sensor, scale and selected wavelength range exists. The latter two however are intended and the benefit of the coherence’s localisation in scale and wavelength. Sensor dependence can be greatly reduced by resampling to a common wavelength grid. Scene dependence is reduced by data pre-processing.

For each scale and endmember a coherence image equal in size as the input image is computed. Limitation to dyadic scales and mapping of only the positive coherence values on byte range (negative values set to 0) relieved the data load to practical volume. The application to hyper-spectral image mosaics was feasible. The coherence in bands possibly affected by border effects was set to 0. This precaution limits the similarity interrogation close to the margins of the spectral coverage at higher scales. Possibly discriminative features may be excluded. A higher confidence in the appropriateness of the signal padding can supersede this in future.

### **Wavelet Transform Feature Analysis**

The features of a spectral signature leave distinct patterns in the scale-wavelength space, namely amplitude extrema and zero crossings. Classical spectral feature analysis, relying on simple geometric primitives and hence limited in parameterizing complex spectral signatures, was complemented by a more dynamic wavelet based parameterization approach which operates on these characteristic patterns. It was shown that the development of modulus maxima and zero-crossings over the transform scales convey information about the shape of the inducing spectral feature and its surroundings. A quantitative link to absorption depth, width, area and asymmetry was established.

The developed Wavelet Transform Feature Analysis exploits these scale-wavelength patterns for the development of quantitative models. The extended set of feature describing parameters, relatively insensitive to noise, might better suit to capture the inherent characteristics of the data. A generalised feature based quantification approach emerges, which brings together classical Spectral Derivative Analysis and Spectral Feature Analysis.

In summary, this feature based quantification approach exploits the multi-scale decomposition of the input signal with its benefits of selective accentuation of spectral components, suppression of noise, normalisation of the input signal and selective focus on varying levels of spectral detail and neighbourhood. Limiting the regression's predictor variables to parameters derived from spectral features, their interpretation is eased and a validation against literature sources is facilitated. The Wavelet Transform Feature Analysis requires no previous knowledge of shape and location of characteristic features and no methodological differentiation between reflectance, absorbance or emissivity nor absorption minima or maxima. Complex spectral feature shapes can be conveyed into the regression model as no geometric primitives are needed for the approximation of signal structures.

The extraction and clustering of the transform feature vectors are critical for the application of the Wavelet Transform Feature Analysis. Extraction and chaining into vectors was facilitated by making use of the COI as feature search area and treating maxima, minima, upward and downward zero crossing features separately. Clustering into common feature vectors used an algorithm invariant to the starting point and number of clusters. Several clustering parameters are required, which are dependent on the data. A careful adjustment of the parameters and a manual check of the clustering result are advised. Nevertheless, an educated guess for the most important parameters  $\epsilon_{wvl}$ ,  $\epsilon_{sid}$  and  $minCard$  can be made from the data.

The “mirror padding” was especially useful for feature vector retrieval. The forced zero crossings at the borders promoted the formation of local extrema close to the spectral margins. Subtle features localised there, which otherwise were prone to border effects, persisted for a number of scales and hence were accessible for the vector extraction and subsequent quantification.

Feature reconstruction from modulus maxima lines was analytically shown for Gaussian shaped absorptions. The feature vectors are derived from dyadic scales, hence are approximations of the modulus maxima lines. An exact calculation of the area, width and depth would require a quasi-continuous frequency resolution (scales). Also absorption features may deviate considerably from Gaussian shapes or overlay each other. The parameterization of spectral features from the CWT hence is a generalisation that does not retrieve an area, width and depth in classical geometric sense for all types of features. Nevertheless, the derived relations also hold for absorptions of differing shape. Proportionality factors exactly determining absorption parameters were therefore dropped. Since CWT is a linear decomposition the proportionality is retained. Scale resolution only affects the proportionality factors.

Common feature patterns were used to exclude pixel spectra which do not conform the training data. A model application to materials not trained on, or an extrapolation to parameter values not backed by the training data, is avoided. This requires a comprehensive training set that adequately describes the natural variability found in the imagery. An incomplete set, or differences between image and laboratory spectra, can lead to areas unexpectedly not being quantified.

### 7.1.3. Thermal Pre-Processing Concept Review

The developed thermal in-flight calibration module is a generic extension to the processing chains existing at INTA and DLR. It takes advantage of the ATCOR-4 atmospheric look-up-table (atmospheric parameters), its sensor model based spectral resampling routines and uses an equivalent radiative transfer model. This ensures its adaptation to arbitrary multi-band thermal sensors and a seamless integration. The calibration bases on homogeneous targets of known emissivity and image-derived temperatures. The image-based temperature retrieval circumvents difficulties to measure a representative surface kinetic temperature for the ground targets during the exact time of sensor overflight.

This in-flight calibration resembles an additional fine-adjustment of the sensor’s laboratory calibration in the form of an empirical line method. To estimate the per-band gain and offset calibration coefficients at least two calibration targets are necessary. They should bracket the expected radiance values in the image, e.g. a dark and a bright target. For the Sokolov night time flight relatively warm water bodies and cool dense vegetation stands were available as targets in abundance of which emissivity spectra a known and nearly uniform graybodies. Under day time conditions the then relatively cool water and vegetation could be complemented by warm asphalt and concrete targets. Calibration targets showing distinct emissivity features should be avoided. Artefacts resulting from inaccurate emissivity measurements may get superimposed onto the image spectra. Calibration targets are required to be located close to nadir to avoid view angle effects which have not been modelled here. Targets distributed over multiple flight lines allow to validate the sensor’s stability.

The estimate of the calibration target's kinetic temperature is the critical part of this approach. A Planck curve fitting approach on atmospherically corrected at-ground radiance spectra was used for temperature retrieval. This dampens noise influences and is less susceptible to errors in individual sensor bands. A proper correction for atmospheric influences is therefore essential. This was assured by using a radiative transfer model and accessing the ATCOR look-up-table of atmospheric model parameters created with MODTRAN. Temperature retrieval from the imagery leads to a relative calibration of the spectral shape of the radiance. Its amplitude is not modified/corrected. In view of the majority of temperature-emissivity-separation algorithms yielding only relative emissivities in respect to a reference band this is less of an issue.

Errors in emissivity and temperature of the calibration targets cause erroneous calibration coefficients. They should be validated against ground reference measurements. The resulting coefficients should also be quality controlled amongst individual flight lines and against other flight campaigns (in case sensor gain is not adjusted). This was done for Sokolov (cf. Section 6.3.2).

Approximation errors can be introduced by the full resolution forward modelling of the radiative transfer in the in-flight calibration but later use of coarser sensor band-averaged atmospheric parameters for the inverse process in the atmospheric correction of the imagery. To counter this effect, calibration was changed to a ground-centred viewpoint ensuring matching at-ground radiance spectra for the calibration targets after atmospheric correction. The calibration coefficients therefore contain compensating information. If calibrated at-sensor radiance is desired, an in-flight calibration with the provided sensor-centred calculation option should be performed.

#### **7.1.4. Workflow Validation Concept Review**

Currently, no solution is available for a full error propagation concept over a complex multi-sensor processing workflow. The empirical validation of inputs and outputs to the individual steps pursued here allows an evaluation of aggregated uncertainties. The major uncertainty sources can be identified. This helps in rating the transferability and repeatability of the results. based on the available reference data, the various sources of uncertainty of the different sensors and products can be better understood. The proposed empirical data driven evaluation however can not verify the general accuracy, stability, repeatability and inherent uncertainties of the algorithms or methods. Such assessments on theoretical or statistical basis are typically very expensive given the complexity of applied methods.

Empirical estimates of the aggregated errors instead can be relatively easily derived from data, models and products. These data driven validation options are integrated into the workflow. A continuous quality control from primary inputs to the final products exists. Availability of both reflective and thermal data allows to check the final products for plausibility and consistency between both spectral domains.

These estimates however will be scene dependent and obviously be governed by the representativeness and amount of reference data. The accuracy, including possible systematic errors, of the presumed "true" reference data can not be revealed.



## 7.2. Critical Discussion of Developed Concepts' Application

A critical review of the individual processing steps and final results is given here. Practical lessons learnt from applying the developed concepts to the Sokolov demonstration site are given.

### 7.2.1. Review of Processing Results

REFLECTIVE DOMAIN PRE-PROCESSING followed the established standards at DLR. The validation of corresponding spectra from the imagery, field and laboratory measurements showed the known scaling issues due to the transition between several orders of spatial averaging inherent to the spectra. In addition did differences in illumination, surface structure, grain size distribution and water content modify the spectral signatures. Normalisation of the spectra could largely eliminate the differences, leading to a good agreement between laboratory, field and image in spectral shape and spectral contrast. It is therefore recommended to use a spectral normalisation technique before classification and quantification. Here, the inherent normalisation of continuous wavelet transform and using its smaller scales (quantification) or spectral amplitude invariant coherence (classification) fulfilled that purpose. Critical were differences in spectral shape of the 2,100-2,300 nm and 500-1,000 nm ranges found between some laboratory and field measurements. They have the potential to decrease the accuracy of quantitative models for kaolinite and goethite/Fe<sub>2</sub>O<sub>3</sub>. Model development on field spectra resembling the natural setting also seen by the airborne sensor would have been optimal. This was not feasible due to the weather conditions.

THERMAL DOMAIN PRE-PROCESSING integrated the newly developed concepts into the INTA & DLR standards for imagery and field data correction. Validation of normalised corresponding spectra from the multispectral imagery and field measurements showed a very well agreement. Differences in spectral shape and localisation of the main absorption features between normalised field and laboratory measurements however could not be fully explained. Spectral contrast differences, additive shifts and occasional linear bias existed between the non-normalised imagery, field and laboratory emissivity spectra. This was expected and originates from differences in thermal contrast, temperature-emissivity separation, measurement setup, moisture and treatment of the samples. Band-ratios of image spectra were used for quantification to exclude additive shifts and to circumvent the problematics of different sensors and high information loss when normalising multispectral signatures. It came at the cost of representativeness of the geochemical analysis.

NORMATIVE MODELLING of the field sample's mineralogy allowed quantitative model development and subsequent spatial mineral mapping based on the exposed characteristic spectral features. The normative models were adapted to the mineralogy of the Sokolov site. Validation showed a good overall correspondence between qualitative XRD analysis and the results of both normative models. Validation against another quantitative mineral determination was not possible due to time and cost reasons. It must be noted that the normative models depend on the unambiguity of the normative mineral sequence and its accordance with the site's mineralogic setting. Hence, the modelling can only be an approximation due to the simplified mineralogic dependencies and limited species covered.

Similar minerals, e.g. clay species, can not be differentiated. Additional minerals, other than the modelled ones, are likely to occur and to have influenced the derived mineral abundances. Their fraction should be small, given the NORMA coverage of all major XRD identified minerals. The quantitative models were trained on the normative mineral results. The derived maps must be interpreted under the caveat of aggregation of similar minerals in respect to their mineral formulae and the applicability of the normative models to the local mineral assemblage. NORMA was hence carefully adjusted to the Sokolov mineralogy and was found well suited for the modelling of lignite mining sites.

REFLECTIVE DOMAIN QUANTIFICATION made use of the developed Wavelet Transform Feature Analysis feeding spectral feature parameters into MLR models. Validation demonstrated sufficient data and model quality to derive physically sound quantitative mineral distribution patterns for Sokolov that conformed the physical 0...100 wt-% constraint. The quantitative map's accuracy depends on the transferability of laboratory spectra to the imagery and the general applicability of the normative calculations as discussed earlier. The kaolinite abundance maps must be interpreted cautious. It was apparent that a systematic overestimation in clay-coal mixtures exist. This originated from such mixtures being underrepresented in the model's training set. The model showed better performance on areas dominated by kaolinite-quartz mixtures. Yet, an overestimation of small kaolinite fractions and an underestimation of high fractions were found. This observations indicate a limited transferability of the quantitative models as was expected from their training on samples taken from the Sokolov site exclusively.

THERMAL DOMAIN QUANTIFICATION made use of band-ratios in exponential regression models. Validation showed that the data and model quality was sufficient to derive physically sound quantitative mineral distribution patterns for the Sokolov demonstration site. Individually and in sum did the abundances conform the physical 0...100 wt-% constraint. Exceptions were found for metal roofs and a small area of atypically increased clay feature contrast. Both resulted from the models not being trained on such surfaces and the simple band ratio models not rejecting image pixels that differ from the training set characteristics. The quantitative map's accuracy depends on the representativeness of the geochemical analysis at pixel scale and the general applicability of the normative calculations as discussed earlier. It must be noted that intermixed mineral fractions of feldspars, micas and carbonates modify the kaolinite absorption in the thermal domain. The multispectral AHS data did not allow to separate these possible influences. Hence, the trained quantitative models may be of reduced accuracy and are likely of limited transferability.

REFLECTIVE DOMAIN CLASSIFICATION was conducted with the developed Wavelet Coherence Mapper. The derived mineral patterns were plausible and followed the expected distributions known from field visits and stratigraphy. Comparisons to the quantification results showed matching distributions. An independent validation against collected samples however was not possible as limited accessibility to the active mines prevented a systematic sampling in the targeted areas. A validation campaign after the establishment of the thematic and quantitative maps unfortunately could not be realised.

### 7.2.2. Major Uncertainty Sources

From the empirical validation of the processing steps in the workflow (Section 6.7 - *Evaluation of Processing Workflow & Results*) and above review of the processing results it is possible to determine the major uncertainty sources. Concluding, the normative modelling, the transfer of the quantitative models from laboratory to image spectra (reflective domain), the representativeness of the sample's laboratory analysis at image pixel scale (thermal domain) and the transferability of the quantitative models to mineral assemblages not adequately covered by the sampling procedure caused the highest uncertainties.

The latter two major uncertainties are universal transferability/scaling issues that strongly depend on the variability in the scene in relation to the used sensor's spatial/spectral resolution and the completeness of the reference data to describe this variability. As such, both uncertainties are impossible to determine from "within" the workflow. They can be reduced from "outside" by extending the reference data collection and limiting analysis to areas conforming to the training data. The thesis at hand implemented the latter by the scale-wavelength pattern matching for the reflective domain quantification.

Normative model discrepancies of RMSE 0.7 wt-% and transfer errors from laboratory to field spectra of 2 wt-% RMSE may appear low compared to the average model's prediction uncertainty of 10-15 wt-% RMSE. However, validity of these empirical uncertainties is limited. The first describes only the relative deviation between two normative models, their systematic bias compared to the "true" mineral fractions stays unknown. The second is valid only for the comparably sparse validation set of 19 corresponding spectra and does not cover the transition to image spectra. The availability of independent and a greater number of validation data in a future study would greatly benefit to judge the influences with a higher confidence.

### 7.2.3. Review of Developed Concepts' Performance

WAVELET COHERENCE MAPPER was used in two classification experiments using library spectra as class templates. Deviating class references and image spectra challenged the classification. Use of library spectra instead of mixed image spectra is a common approach.

Classification using the averaged coherence over the full spectral range geared towards finding exact matches of the class templates in the imagery. Localisation and normalisation of the coherence measure led to local reactions to mismatches, e.g. 1,400/1,900 nm water absorptions. Even strong artefacts in these wavelength regions had a limited impact on the average coherence (in small to medium scales) and needed no masking to avoid false negatives. Scale selection allowed to focus classification on characteristic spectral detail. This was especially beneficial when class template and image spectra deviated in their continuum. Handling of the COI area was problematic. Coherence set to 0 in the affected bands limited interrogation of spectral features close to the spectral coverage margins. This excluded spectral information for higher scales and degraded the classification performance there (false positives). Yet, border effects appeared even beyond the expected COI area. The transform values at higher scales therefore must be interpreted with caution close to the margins of the spectral coverage due to the possibility of distortions by border effects.

Classification limiting to a narrow spectral range took advantage of the coherence's wavelength localisation. Insignificant spectral information was excluded, which proved advantageous for classification of clay minerals in mixtures. The sensitivity of coherence to differences in curvature direction and choice of a small scale allowed to successfully retrieve clay occurrences also from subtle spectral features. The spectral shape driven SAM was not able to do so. Similar results as were retrieved with Wavelet Coherence Mapper could be generated by applying SAM to wavelet transformed spectra. The difference was SAM's focus on strong features, omitting the weaker curvature characteristics. Coherence instead weights features equally over the wavelength range irrespective of their amplitude. Coherence in the smallest scale component is very localised and susceptible to noise. Erroneously high coherence values were found in water and shadow spectra where random noise mimicked the targeted curvature characteristics.

Concluding, the Wavelet Coherence Mapper proved a useful method for deriving confident classifications of goethite, siderite and clay minerals, despite of the dominance of spectral mixing that modified the signatures. Coherence is a promising complement to shape matching methods like SAM. The suppression of border effects needs further investigation.

WAVELET TRANSFORM FEATURE ANALYSIS was used to produce quantitative mineral maps, combining spectroscopic laboratory measurements and normative modelled mineral fractions of 50 collected samples. Despite the suboptimal training data not optimised for the development of quantitative models and issues arising from adapting the models to be used on image spectra, the developed method proved successful.

The feature based approach of the Wavelet Transform Feature Analysis allowed a straightforward interpretation of the derived candidate predictor variables linking them to distinct spectral properties of the input spectra. Selective omission of inadequate features not relating to the targeted mineral or obscured by water absorptions was easily facilitated. This is an advantage over statistical methods like Partial Least Squares, Support Vector Machines or Linear Discriminant Analysis that optimise the model to the full spectral range or selected subsets.

Compared to the Correlation Scaleogram Analysis relying on transform values at fixed scale-wavelength points, it was shown that the wavelength occurrence of spectral features can play an important role in describing the abundance variability (e.g. goethite). Wavelet Transform Feature Analysis allowing the retrieval of such additional absorption feature describing parameters therefore is an important extension. Generally, the correlating spectral features identified by the Correlation Scaleogram Analysis were represented by the findings of the Wavelet Transform Feature Analysis.

Efforts to apply the Wavelet Transform Feature Analysis may appear high compared to the classical Spectral Feature Analysis. The semi-automatic clustering of the feature vectors into common vectors requires several parameters to be set. This clustering replaces the manual definition of wavelength ranges for feature searching and possible bias introduced by such an approach. If operating on few, pre-known and well separated spectral features, as often the case for classical Spectral Feature Analysis, this clustering step may be re-

placed by a simpler wavelength range based method. Here, by treating the four feature vector types separately, the proposed GDBSCAN clustering algorithm was successful.

The handling of nonexistent transform features in the training set also requires several thresholds to be set. This is necessary to automate feature search and correlation test performed by the Wavelet Transform Feature Analysis, contrary to the more rigid classical Spectral Feature Analysis that operates on predefined spectral features.

Automation comes with an increased complexity of the method. Values for most parameters and thresholds were determined from the data. Nevertheless, the clustering and thresholding are an issue and should be points for future improvements.

In summary, the Wavelet Transform Feature Analysis proved an applicable method despite limitations imposed by the relatively sparse training data and transition from laboratory to imagery. The processing of a real-world hyperspectral image mosaic was possible. As a semi-automatic feature search and correlation testing method that offers a higher flexibility in spectral feature describing parameters, it complements Correlation Scaleogram Analysis, Spectral Derivative Analysis and Spectral Feature Analysis.

THERMAL IN-FLIGHT CALIBRATION of the AHS data made use of the sensor description and sensor resampled atmospheric look-up-table, both created with ATCOR. Therefore the in-flight calibration module was easily adapted to the sensor and its output could be directly used for the further pre-processing. Plenty water and dense vegetation calibration targets could be identified close to nadir in the individual AHS flight lines. Their image-retrieved temperatures matched the climate data and temperature logs. A per-band regression of high linear correlation and stable over the individual flight lines could be established. As expected showed AHS bands #71, #79 and #80 a degraded performance due to strong atmospheric influences. The linear regression with free gain and offset however led to mathematically optimised yet physically impossible values (negative offsets) of opposing trend (negative offset & positive gain of proportional amplitude). This indicates remaining nonlinearities not explained by sensor laboratory calibration, atmospheric correction and temperature retrieval. Offset was forced to zero to counter this shortcoming of the linear fit. Therefore a correct laboratory calibration of the sensor gain is assumed. Stable per-band gain correction coefficients resulted that were physically sound.

In summary, the thermal in-flight calibration module worked as planned. It contributed to a significant improvement of the AHS derived emissivity spectra which were of appropriate quality for the successful quantification of kaolinite and quartz.

EMPIRICAL WORKFLOW VALIDATION worked well where there was independent reference data in adequate numbers available, such as for the spectroscopic data pre-processing. The approach naturally faces limitations when the number of independent reference data is small, such as for the MLR model transferability from laboratory to imagery. The value of the empirical validation was lowest, where no reference data was available for direct comparison. In case of the normative modelling and subsequent map products the validation was hence limited to intercomparisons and plausibility checks which can not reveal systematic bias.

Concluding, a proof-of-concept of the empirical workflow validation was possible. As expected, the validation was dependent on the available reference data. In this study, lack of such data at a central point - the normative modelling - led to a significant weakening of the expressiveness of the validation chain towards the higher mapping products. In future, this gap in the reference data should be avoided.

This can be achieved by avoiding the normative modelling step when sufficient quantitative mineralogical information is available or by including a representative set of reference analytics, e.g. quantitative XRD. The normative model can then be explicitly trained to the site's mineralogy by adapting the normative sequence and mineral formulas.

### 7.3. Recommended Application Scenario for Developed Concepts

From previous discussion, clear recommendations for using the main developments can be drawn.

#### **Wavelet Coherence Mapper**

The use of coherence is proposed when the evolution of the similarity measure over wavelength is of benefit, when spectral curvature changes are indicative for the presence of a material (e.g. in mixtures), or when features of different amplitude in a spectral signature are equally characteristic (e.g. a strong deep feature paired with a shallow one at same transform scale). The equal weighting of differently pronounced features and high sensitivity of the curvature comparison necessitates to limit the similarity investigations to characteristic wavelength regions. No benefit is provided in applications where exact shape matches are required. Normalisation of additive offsets, multiplicative factors and linear bias prohibits the application of wavelet coherence in studies where this information is of interest.

#### **Wavelet Transform Feature Analysis**

The semi-automated approach needs no knowledge of existence, shape or position of possible absorption features thus leading to a spectral feature detection and feature correlation testing methodology. Use of the second derivative ( $\text{DOG}_{m=2}$ ) allows for a direct translation of the found features back into reflectance space for interpretation. The method unifies the traditional Spectral Derivative and Spectral Feature Analysis plus parts of the Correlation Scaleogram Analysis into one method, while extending their flexibility to a greater number of feature describing parameters. The method is generic and applicable to reflectance, emissivity or absorbance spectra without modification. This benefits come at the price of a higher complexity of the method, increased computational expense and data volume. Use of the Wavelet Transform Feature Analysis therefore is recommended for the feature based investigative analysis of complex, mixed or subtle spectral features in hyperspectral signatures. Normalisation of local additive offsets and linear bias by the CWT prohibits its application where this information is of interest. Studies relying on pre-known distinct and isolated spectral features may consider using simpler quantification methods.

### **Thermal In-Flight Calibration**

An additional in-flight calibration is proposed for the pre-processing of all thermal imagery that makes use of a physical radiative transfer model. The developed approach estimates the calibration target's kinetic temperature from the imagery by a Planck curve fitting. Hence, it requires multi- or hyperspectral sensors that can adequately represent the shape of the radiance curve by suitably placed bands. The in-scene data estimated temperatures are of benefit when only a limited number of ground surface temperature measurements are available. The approach lessens the need for representative surface skin temperature measurements during sensor overflight and availability of respective emissivity spectra. Studies that have the previously mentioned reference data at hand are recommended to perform a direct calibration instead.

### **Empirical Workflow Validation**

A validation of the whole processing workflow is generally recommended to identify major uncertainty sources to judge the accuracy and stability of the chosen methods and models. The empirical product based validation is a practicable alternative to theoretical/statistical approaches when the workflow is complex, individual contributions to the aggregated error budget are of little interest and an adequate amount of independent reference data is available.

### **Applicability of Imaging Spectroscopy to Mineral Mapping**

Spatial and spectral resolution of hyperspectral imaging sensors are sufficient to identify and quantify the mineral composition and selected geochemical properties of soil, as demonstrated by numerous studies. Strictly, it is undue practise to deduce properties of deeper soil layers from the sensed surface (signal penetration some micrometres). Crusts, agglomerates, weathering, segregation & vegetation/biofilms alter the recorded signal. [Ben-Dor \*et al.\* \(1999\)](#) give a review.

In-situ reference measurements are collected to characterise this natural variability and to “train” the analysis models to it. The resulting quantification therefore is rather specialised and must be carefully re-validated when applied to a different site. In Sokolov, analysis was limited to barren areas. Where succession takes place, sub-pixel analysis techniques like unmixing can be applied to compensate for the influence of vegetation on the pixel's spectral signature or spectral bands/combinations little affected by vegetation can be used ([Rencz, 1999](#)).

There exist studies that simulate mineral assemblages by laboratory mixing experiments (e.g. [Krüger \*et al.\* \(1998\)](#); [Reinhäckel \(1998\)](#); [Richter \(2010a\)](#)). This allows to study influences of concentrations, moisture or grain size under controlled environment. Compared to natural mixtures, the number of components however is limited due to the exponentially growing effort. Also differences in packing, orientation, etc. exist that lead to inaccuracies when transiting from laboratory to the field. Laboratory mixing experiments with pure minerals hence are only an addition but can not replace in-situ spectroscopic measurements.

In summary, this work contributes to mineral classification and quantification by providing a practical concept to join reflective and thermal information. It becomes clear that the understanding of the spectral characteristics of the various influences are crucial and a field of ongoing research. Furthermore, a quantitative study necessitates a comprehensive set of reference information which should integrate geochemical, geophysical, spectroscopic and atmospheric data.

## 8. Outlook

The discussion in Chapter 7 highlighted several aspects that are avenues for future research.

The empirical workflow validation is a small step towards the envisioned ambition to establish a standardised quality and accuracy assessment of the whole product generating workflow, such that quantitative uncertainty measures can be given and the influences of individual errors on the result can be described (error budgeting).

The thermal in-flight calibration module did not take view angle dependencies of the atmospheric parameters into account. In a future extension it would be recommended to add view angle and ground topography effects. Both could be determined from the localisation of the calibration targets in the imagery and from a provided digital elevation model. This would drop the requirement of near-nadir calibration targets and would avoid inaccuracies in mountainous terrain. The ATCOR look-up-table could accommodate both degrees of freedom without a modification. A future technical aspect would be the direct integration of the developed thermal in-flight calibration module into ATCOR. Thermal and reflective sensors would then be equally supported by the software and standardised pre-processing would be promoted.

The employment of the developed wavelet based methods was technically limited by processing time and border effects. Use of fast CWT algorithms can greatly reduce computational expense from  $O(n \log_2 n)$  to  $O(n)$  per scale (Ho & Chan, 1999). They require orthogonal or biorthogonal wavelets. Mallat (2008) points out the possibility to construct smooth biorthogonal symmetric wavelets of compact support. Compact support means a narrow and well defined (finite) COI, which confines possible border effects. For the future, it is recommended to implement the CWT computation as fast filter bank operation (e.g. *algorithme à trous*) using spline biorthogonal wavelets. This would mitigate both limitations, while preserving most of the beneficial properties that led to use of the  $\text{DOG}_{m=2}$  wavelet in this research. Ho & Chan (1999) provide optimisations and comparisons of fast CWT algorithms. Signal padding should be a second focus of future research. An optimised smooth signal extension (e.g. splines, adaptation to reflectance or emissivity) would counter border effects and would give higher confidence in margin-near transform values. Setting coherence to zero in the COI, which especially limited Wavelet Coherence Mapper classifications at higher scales, could be possibly dropped. The difficulties in setting appropriate clustering parameters and thresholds could possibly be circumvented by a research on determining these from the data sets directly.

In retrospect has the potential of continuous wavelet transform for the analysis of hyperspectral imagery not been fully exploited yet. Availability of high-performance computation hardware may revive the interest of the community. Future research fields are the transform patterns in scale-wavelength space which can be used for a multi-scale hierarchical classification. Linear decomposition of the spectral signatures opens door for multi-scale unmixing techniques. Hadoux *et al.* (2013) conducted a spatial spectral multi-scale analysis (3d wavelets) on imagery recently.

Future availability of hyperspectral airborne sensors (HySpex, TASI, SEBASS) and satellite missions (EnMAP, HypSIPI) with largely varying spatial and spectral resolution will rise issues of spectral mixing and will be a promising testbed for the developed methods.



# Bibliography

- ADAMS, J. B., & GOULLAUD, L. H. 1978. Plagioclase feldspars-Visible and near infrared diffuse reflectance spectra as applied to remote sensing. *Lunar and Planetary Science Conference Proceedings*, **9**, 2901–2909.
- ADDISON, P. S. 2002. *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. Institute of Physics Publishing.
- ALFARO, L. P. 2009. *Monitoring minining rehabilitation according to methods derived from imaging spectroscopy: Case study of Sotiel-Migollas, Southern Spain*. Ph.D. thesis, Wageningen University and Research Centre, Netherlands.
- BACHMANN, M. 2007. *Automatisierte Ableitung von Bodenbedeckungsgraden durch MESMA-Entmischung*. Dissertation, Julius-Maximilians-Universität Würzburg.
- BACHMANN, M., HABERMEYER, M., HOLZWARTH, S., RICHTER, R., & MÜLLER, A. 2007. Including Quality Measures in an Automated Processing Chain for Airborne Hyperspectral Data. *In: Proceedings of the 5th EARSeL Workshop on Imaging Spectroscopy Workshop on Imaging Spectroscopy*.
- BÄNI, W. 2002. *Wavelets. Eine Einführung für Ingenieure*. 1st edn. R. Oldenbourg Verlag.
- BEEKHUIZEN, J., BACHMANN, M., BEN-DOR, E., BIESEMANS, J., GRANT, M., HEUVELINK, G., HUENI, A., KNEUBUEHLER, M., MALTHUS, T., DE MIGUEL, E., PIMSTEIN, A., PRADO, E., REUSEN, I., RUHTZ, T., & SCHAALÉ, M. 2009a. *DJ2.1.1 - Report on calibration, geometric and atmospheric error model/concept*. Tech. rept. EUFAR FP7 JRA2 - HYQUAPRO.
- BEEKHUIZEN, J., BACHMANN, M., BEN-DOR, E., BIESEMANS, J., GRANT, M., HEUVELINK, G., HUENI, A., KNEUBUEHLER, M., MALTHUS, T., DE MIGUEL, E., PIMSTEIN, A., PRADO, E., REUSEN, I., RUHTZ, T., & SCHAALÉ, M. 2009b. *DJ2.1.2 - Report on Full error propagation concept*. Tech. rept. EUFAR FP7 JRA2 - HYQUAPRO.
- BEN-DOR, E., IRONS, J. R., & EPEMA, G. F. 1999. Soil Reflectance. *Chap. 3, pages 111–188 of: RENCZ, A. N. (ed), Manual of Remote Sensing, Vol. 3: Remote Sensing for the Earth Sciences*, 3rd edn. Wiley & Sons.
- BERK, A., ANDERSON, G. P., ACHARYA, P. K., BERNSTEIN, L. S., MURATOV, L., LEE, J., FOX, M., ADLER-GOLDEN, S. M., CHETWYND, J. H., HOKE, M. L., LOCKWOOD, R. B., GARDNER, J. A., COOLEY, T. W., BOREL, C. C., LEWIS, P. E., & SHETTLE, E. P. 2006. MODTRAN5: 2006 Update. *Pages 508–515 of: SPIE 6233*, vol. 6233.
- BEST, M. G. 2002. Calculation of the CIPW norm. *Chap. Appendix B, pages 661–665 of: Igneous and Metamorphic Petrology*, 2 edn. Wiley-Blackwell.
- BLACKBURN, G., & FERWERDA, J. 2008. Retrieval of chlorophyll concentration from leaf reflectance spectra using wavelet analysis. *Remote Sensing of Environment*, **112**(4), 1614–1632.
- BLOWES, D. W., & PTACEK, C. J. 1994. Acid-neutralization mechanisms in inactive mine tailings. *Pages 271–292 of: JAMBOR, L. J., & BLOWES, D. W. (eds), Short Course Handbook on Environmental Geochemistry of Sulfide Mine Waste*. Mineralogical Association of Canada.
- BLOWES, D. W., PTACEK, C. J., JAMBOR, J. L., & WEISNER, C. G. 2003. The Geochemistry of Acid Mine Drainage. *Chap. 9.05, pages 149–204 of: HOLLAND, H. D., & TUREKIAN, K. K. (eds), Treatise on Geochemistry, Volume 9: Environmental Geochemistry*. Oxford: Elsevier.

- BOUŠKA, V., & PEŠEK, J. 1999. Quality parameters of lignite of the North Bohemian Basin in the Czech Republic in comparison with the world average lignite. *International Journal of Coal Geology*, **40**, 211–235.
- BROMILEY, P. A. 2003. *Products and Convolutions of Gaussian Distributions*. Tech. rept. Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester, Manchester.
- BROPHY, G. P., & SHERIDAN, M. F. 1965. Sulfate studies IV: The jarosite-natrojarosite-hydronium jarosite solid solution series. *The American Mineralogist*, **50**, 1595–1607.
- BRUCE, L. M., & LI, J. 2001. Wavelets for computationally efficient hyperspectral derivative analysis. *IEEE Transactions on Geoscience and Remote Sensing*, **39**(7), 1540–1546.
- BRUCE, L. M., MORGAN, C., & LARSEN, S. 2001. Automated detection of subpixel hyperspectral targets with continuous and discrete wavelet transforms. *IEEE Transactions on Geoscience and Remote Sensing*, **39**(10), 2217–2226.
- CACECI, M. S., & CACHERIS, W. P. 1984. Fitting Curves to Data. *Byte Magazine*, **9**(5), 340–362.
- CHANDER, G., HEWISON, T. J., FOX, N., WU, X., XIONG, X., & BLACKWELL, W. J. 2013. Overview of Intercalibration of Satellite Instruments. *IEEE Transactions on Geoscience and Remote Sensing*, **51**(3), 1056–1080.
- CHANG, C.-W., LAIRD, D. A., MAUSBACH, M. J., & HURBURGH JR., C. R. 2001. Near-infrared reflectance spectroscopy - principal components regression analyses of soil properties. *Soil Science Society of America Journal*, **65**, 480–490.
- CHOE, E., VAN DER MEER, F., VAN RUITENBEEK, F., VAN DER WERFF, H., DE SMETH, B., & KIM, K. W. 2008. Mapping of heavy metal pollution in stream sediments using combined geochemistry, field spectroscopy, and hyperspectral remote sensing: A case study of the Rodalquilar mining area, SE Spain. *Remote Sensing of Environment*, **112**(7), 3222–3233.
- CHP-HOMEPAGE. 2013. *The Chemistry Hypermedia Project [ONLINE]*. <http://www.files.chem.vt.edu/chem-ed/spec/beerslaw.html> [Accessed 22 August 2013].
- CHRISTENSEN, P. R., BANDFIELD, J. L., HAMILTON, V. E., HOWARD, D. A., LANE, M. D., PIATEK, J. L., RUFF, S. W., & STEFANOV, W. L. 2000. A thermal emission spectral library of rock-forming minerals. *Journal of Geophysical Research*, **105**(E4), 9735.
- CLARK, R. N. 1983. Spectral properties of mixtures of montmorillonite and dark carbon grains: Implications for remote sensing minerals containing chemically and physically adsorbed water. *Journal of Geophysical Research*, **88**(B12), 10635–10644.
- CLARK, R. N. 1999. Spectroscopy of Rocks and Minerals and Principles of Spectroscopy. *Chap. 1, pages 3–58 of: RENCZ, A. N. (ed), Manual of Remote Sensing, Vol. 3: Remote Sensing for the Earth Sciences*, 3rd edn. Wiley & Sons.
- CLARK, R. N., & ROUSH, T. L. 1984. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research*, **89**(B7), 6329–6340.
- CLARK, R. N., SWAYZE, G. A., GALLAGHER, A. J., KING, T. V. V., & CALVIN, W. M. 1993. *Digital Spectral Library: Version 1: 0.2 to 3.0 microns, U.S. Geological Survey Open File Report 93-592*. Tech. rept. U. S. Geological Survey.

- CLARK, R. N., SWAYZE, G. A., LIVO, K. E., RAYMOND, F. K., SUTLEY, S. J., DALTON, J. B., MCDUGAL, R. R., & GENT, C. A. 2003. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *Journal of Geophysical Research*, **108**(E12), 5-1 to 5-44.
- COHEN, D., & WARD, C. R. 1991. SEDNORM - a program to calculate a normative mineralogy for sedimentary rocks based on chemical analyses. *Computers & Geosciences*, **17**(9), 1235-1253.
- CONGALTON, R. G. 2001. Accuracy assessment and validation of remotely sensed and other spatial information. *International Journal of Wildland Fire*, **10**(4), 321-328.
- CUDAHY, T. J., OKADA, K., YAMATO, Y., MAEKAWA, M., HACKWELL, J. A., & HUNTINGTON, J. F. 2000. *Mapping skarn and porphyry alteration mineralogy at Yerington, Nevada, using airborne hyperspectral TIR SEBASS data. CSIRO Exploration and Mining Report 734R*. Tech. rept. CSIRO Exploration and Mining, Floreat Park, WA, Australia.
- DASGUPTA, A. 2008. Saddlepoint Approximations. *Chap. 14, pages 203-224 of: Asymptotic Theory of Statistics and Probability*. New York, USA: Springer.
- DASH, P., GÖTTSCHE, F. M., OLESEN, F. S., FISCHER, H., & GÖTTSCHE, F. M. 2002. Land surface temperature and emissivity estimation from passive sensor data: Theory and practice-current trends. *International Journal of Remote Sensing*, **23**(13), 2563-2594.
- DE CARVALHO, O. A., & MENESES, P. R. 2000. Spectral correlation mapper (SCM): An improvement on the spectral angle mapper (SAM). *Page 9 of: Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) 2000 Workshop Proceedings*.
- DEBBA, P. 2006. *Sampling scheme optimization from hyperspectral data*. Dissertation, Wageningen University.
- DI STASIO, R. J., & RESMINI, R. G. 2010 (Apr.). Atmospheric Compensation of Thermal Infrared Hyperspectral Imagery with the Emissive Empirical Line Method and the In-Scene Atmospheric Compensation Algorithms: A Comparison. *Page 76952B of: SHEN, SYLVIA S., & LEWIS, PAUL E. (eds), SPIE 7695, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVI*.
- DOLD, B. 2005. Basic concepts of environmental geochemistry of sulfide mine-waste. *Page 36 of: UNESCO-SEG-SGA Latin American Metallogeny Course*. Lima, Perú: Centre d'Analyse Minérale, Université de Lausanne, Switzerland.
- DOLD, B. 2010. Basic concepts in environmental geochemistry of sulfidic mine-waste management. *Chap. 10, pages 173-198 of: KUMAR, E. S. (ed), Waste management*. Rijeka, Croatia: InTech.
- DONOHO, D. L. 1995. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, **41**(3), 613-627.
- ELVIDGE, C. D. 1990. Visible and near infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing*, **11**(10), 1775-1795.
- EXELISVIS-HOMEPAGE. 2013. *Exelis VIS [ONLINE]*. <http://www.exelisvis.com/>.
- FENG, J., RIVARD, B., GALLIE, A., & SANCHEZ-AZOFEIFA, A. 2011. Rock type classification of drill core using continuous wavelet analysis applied to thermal infrared reflectance spectra. *International Journal of Remote Sensing*, **32**(16), 4489-4510.

- FERNÁNDEZ-RENAU, A., GÓMEZ, J. A., & DE MIGUEL, E. 2005. The INTA AHS system. *Pages 471–478 of: MEYNART, R., NEECK, S. P., & SHIMODA, H. (eds), Proceedings of the SPIE Sensors, Systems, and Next-Generation Satellites: IX.*
- GAFFEY, SUSAN J. 1987. Spectral reflectance of carbonate minerals in the visible and near infrared (0.35–2.55  $\mu\text{m}$ ): Anhydrous carbonate minerals. *Journal of Geophysical Research*, **92**(B2), 1429.
- GERIGHAUSEN, H., BORG, E., FICHELTMANN, B., VAJEN, H.-H., WLOCZYK, C., ZABEL, E., & MAASS, H. 2008. Validation and calibration of remote sensing data products on test site DEMMIN. *In: Proceedings on 43. Ziolkowski Conference.*
- GOUTIS, CONSTANTINO, & CASELLA, GEORGE. 1999. Explaining the Saddlepoint Approximation. *The American Statistician*, **53**(3), 216.
- GRINSTED, A., MOORE, J. C., & JEVREJEVA, S. 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, **11**(5/6), 561–566.
- GU, D., GILLESPIE, A. R., KAHLE, A. B., & PALLUCONI, F. D. 2000. Autonomous atmospheric compensation (AAC) of high resolution hyperspectral thermal infrared remote-sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **38**(6), 2557–2570.
- GUPTA, R. P. 1991. *Remote Sensing Geology*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- HAAR, R., & BART, M. 2003. Gaussian derivatives. *Chap. 4, pages 53–69 of: Front-End Vision & Multi-Scale Image Analysis*, vol. 7. Berlin: Springer.
- HABERMEYER, M., HOLZWARTH, S., MÜLLER, A., MÜLLER, R., RICHTER, R., SEITZ, K.-H., SEIFERT, P., STROBL, P., & BACHMANN, M. 2003. Developing a Fully Automatic Processing Chain for the Upcoming Hyperspectral Scanner ARES. *In: Proceedings of the 3rd EARSeL Workshop on Imaging Spectroscopy.*
- HABERMEYER, M., BACHMANN, M., HOLZWARTH, S., MÜLLER, R., & RICHTER, R. 2012. Incorporating a push-broom scanner into a generic processing chain. *In: IEEE Geoscience and Remote Sensing Symposium 2012.*
- HADOUX, X., GORRETTA, N., RABATEL, G., ROGER, J. M., & STRAUSS, O. 2013. Spectral spatial pre-processing using multi-resolution 3D wavelets for hyperspectral image classification. *Page 4 of: IEEE 5th Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing*, vol. 2.
- HECKER, C. 2013. *Mapping Feldspars from Above: a thermal infrared and partial least squares-based approach*. Ph.D. thesis, University of Twente, ITC.
- HECKER, C., HOOK, S. J., VAN DER MEIJDE, M., BAKKER, W., VAN DER WERFF, H., WILBRINK, H., VAN RUITENBEEK, F., DE SMETH, B., & VAN DER MEER, F. 2011. Thermal Infrared Spectrometer for Earth Science Remote Sensing Applications - Instrument Modifications and Measurement Procedures. *Sensors*, **11**(11), 10981–10999.
- HO, K. C., & CHAN, Y. T. 1999. Filter Design and Comparison for Two Fast CWT Algorithms. *IEEE Transactions on Signal Processing*, **47**(11), 3013–3026.
- HOLMQVIST, J., THELIN, G., ROSENGREN, U., STJERNQUIST, I., WALLMAN, P., & SVERDRUP, H. 2002. Biogeochemical Processes and Mechanisms. *Chap. 7, pages 91–196 of: SVERDRUP, H., & STJERNQUIST, I. (eds), Developing Principles and Models for Sustainable Forestry in Sweden*, 5 edn. Dordrecht: Kluwer Academic Publishers.

- HOLSCHNEIDER, M., KRONLAND-MARTINET, R., MORLET, J., & TCHAMITCHIAN, P. 1989. A real-time algorithm for signal analysis with the help of the wavelet transform. *Pages 289–297 of: COMBES, JEAN-MICHEL, GROSSMANN, ALEXANDER, & TCHAMITCHIAN, PHILIPPE (eds), Wavelets, Time-Frequency Methods and Phase Space*. Berlin: Springer.
- HOLZWARTH, S., WANG, X., RICHTER, R., BACHMANN, M., & FREER, M. 2010. *DN6.2.2 - Software test and comparison report. Documentation on best practices*. Tech. rept. 6. EUFAR FP7 N6SP - Standards and Protocols.
- HOLZWARTH, S., BACHMANN, M., & FREER, M. 2011. Standards for airborne hyperspectral image data. *Pages 1–7 of: Proceedings of the 7th EARSeL-SIG-IS*.
- HOOK, S. J., & KAHLE, A. B. 1996. The micro fourier transform interferometer ( $\mu$ FTIR) - A new field spectrometer for acquisition of infrared data of natural surfaces. *Remote Sensing of Environment*, **56**(3), 172–181.
- HOOK, S. J., GABELL, A. R., GREEN, A. A., & KEALY, P. S. 1992. A comparison of techniques for extracting emissivity information from thermal infrared data for geologic studies. *Remote Sensing of Environment*, **42**, 123–135.
- HSU, P. H., & TSENG, Y. H. 2000. Wavelet Based Analysis of Hyperspectral Data for Detecting Spectral Features. *Pages 61–68 of: International Archives of Photogrammetry and Remote Sensing*, vol. XXXIII.
- HUNT, G. R. 1977. Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics*, **42**(3), 501–513.
- HUNT, G. R., & SALISBURY, J. W. 1971. Visible and near infrared spectra of minerals and rocks: II. Carbonates. *Modern Geology*, **2**, 23–30.
- HUNT, G. R., SALISBURY, J. W., & LENHOFF, C. J. 1971a. Visible and near infrared spectra of minerals and rocks: III. Oxides and hydroxides. *Modern Geology*, **2**, 195–205.
- HUNT, G. R., SALISBURY, J. W., & LENHOFF, C. J. 1971b. Visible and near infrared spectra of minerals and rocks: IV. Sulphides and sulphates. *Modern Geology*, **3**, 1–14.
- IDL-COYOTE-HOMEPAGE. 2013. *IDL programs from Coyote's Guide to IDL Programming [ONLINE]*. <https://code.google.com/p/idl-coyote/>.
- JAMBOR, J. L., & BLOWES, D. W. 1998. Theory and applications of mineralogy in environmental studies of sulfide-bearing mine waste. *Pages 367–401 of: CABRI, L. J., & VAUGHAN, D. J. (eds), Short Course Handbook on Ore and Environmental Mineralogy*. Mineralogical Association of Canada.
- JCGM100:2008. 2008. *Evaluation of measurement data - Guide to the expression of uncertainty in measurement*. Tech. rept. Bureau International des Poids et Mesures.
- JIMÉNEZ-MUNOZ, J. C., SOBRINO, J. A., & GILLESPIE, A. R. 2011. Surface Emissivity Retrieval From Airborne Hyperspectral Scanner Data: Insights on Atmospheric Correction and Noise Removal. *IEEE Geoscience and Remote Sensing Letters*, 1–5.
- JOHNSON, P. E., SMITH, M. O., & ADAMS, J. B. 1992. Simple algorithms for remote determination of mineral abundances and particle sizes from reflectance spectra. *Journal of Geophysical Research*, **97**(E2), 2649–2657.

- KEMPENEERS, P., DE BACKER, S., DEBRUYN, W., COPPIN, P., & SCHEUNDERS, P. 2005. Generic wavelet-based hyperspectral classification applied to vegetation stress detection. *IEEE Transactions on Geoscience and Remote Sensing*, **43**(3), 610–614.
- KLEIN, A., SAUER, T., JEDYNAK, A., & SKRANDIES, W. 2006. Conventional and wavelet coherence applied to sensory-evoked electrical brain activity. *IEEE Transactions on Biomedical Engineering*, **53**(2), 266–272.
- KOGER, C. H., BRUCE, L. M., SHAW, D. R., & REDDY, K. N. 2003. Wavelet analysis of hyperspectral reflectance data for detecting pitted morningglory (*Ipomoea lacunosa*) in soybean (*Glycine max*). *Remote Sensing of Environment*, **86**(1), 108–119.
- KOPAČKOVÁ, V., BOURGUIGNON, A., CHEVREL, S., KOUBOVÁ, M., & ROJÍK, P. 2009. Effect of mineralogical and geochemical properties on reflectance properties of waste from Sokolov open pit lignite mine, Czech Republic. *Pages 569–580 of: FOURIE, A.B., & TIBBETT, M. (eds), Mine Closure 2009*. Perth: Australian Centre for Geomechanics.
- KOPAČKOVÁ, V., CHEVREL, S., & BOURGUIGNON, A. 2011 (Oct.). Spectroscopy as a tool for geochemical modeling. *Pages 818106–818106–13 of: MICHEL, U., & CIVCO, D. L. (eds), SPIE 8181, Earth Resources and Environmental Remote Sensing/GIS Applications II*, vol. 8181.
- KRAMER, H. J. (ed). 2002. *Observation of the Earth and Its Environment: Survey of Missions and Sensors*. 4th edn. Springer.
- KRÍBEK, B., STRNAD, M., BOHÁČEK, Z., SÝKOROVÁ, I., ČEJKA, J., & SOBALÍK, Z. 1998. Geochemistry of Miocene lacustrine sediments from the Sokolov Coal Basin (Czech Republic). *International Journal of Coal Geology*, **37**, 207–233.
- KRÜGER, G., ERZINGER, J., & KAUFMANN, H. 1998. Laboratory and airborne reflectance spectroscopic analyses of lignite overburden dumps. *Journal of Geochemical Exploration*, **64**(1-3), 47–65.
- KRUSE, F. A., & LEFKOFF, A. B. 1994. Knowledge-based geologic mapping with imaging spectrometers. *Remote Sensing Reviews*, **8**(1-3), 3–28.
- KRUSE, F. A., LEFKOFF, A. B., BOARDMAN, J. B., HEIDEBRECHT, K. B., SHAPIRO, A. T., BARLOON, P. J., & GOETZ, A. F. H. 1993. The spectral image processing system (SIPS) - Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of the Environment*, **44**, 145–163.
- KUNTZE, H., ROESCHMANN, G., & SCHWERDTFEGER, G. 1988. *Bodenkunde*. 4th edn. Stuttgart, Germany: UTB / Ulmer.
- LIANG, S. 2004. *Quantitative remote sensing of land surfaces*. Hoboken, NJ, USA: Wiley & Sons.
- LIANG, SHUNLIN, FANG, HONGLIANG, CHEN, MINGZHEN, SHUEY, CHAD J, WALTHALL, CHARLIE, DAUGHTRY, CRAIG, MORISSETTE, JEFF, SCHAAF, CRYSTAL, & STRAHLER, ALAN. 2002. Validating MODIS land surface reflectance and albedo products: methods and preliminary results. *Remote Sensing of Environment*, **83**(1-2), 149–162.
- LILLESAND, T. M., & KIEFER, R. W. 1994. *Remote sensing and image interpretation*. 3rd edn. Chichester: Wiley & Sons.
- LIU, M., LIU, X., WU, L., DUAN, L., & ZHONG, B. 2011. Wavelet-based detection of crop zinc stress assessment using hyperspectral reflectance. *Computers & Geosciences*, **37**(9), 1254–1263.

- LUNETTA, R. S., & LYON, J. G. (eds). 2004. *Remote Sensing and GIS Accuracy Assessment*. 1st edn. CRC Press.
- LYDER, D., FENG, J., RIVARD, B., GALLIE, A., & CLOUTIS, E. 2010. Remote bitumen content estimation of Athabasca oil sand from hyperspectral infrared reflectance spectra using Gaussian singlets and derivative of Gaussian wavelets. *Fuel*, **89**(3), 760–767.
- MALLAT, S. 2008. *A Wavelet Tour of Signal Processing: The Sparse Way*. 3rd edn. Academic Press.
- MALLAT, S., & HWANG, W. L. 1992. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, **38**(2), 617–643.
- MARS, J. C., & CROWLEY, J. K. 2003. Mapping mine wastes and analyzing areas affected by selenium-rich water runoff in southeast Idaho using AVIRIS imagery and digital elevation data. *Remote Sensing of Environment*, **84**(3), 422–436.
- MELKERUD, P. A., BAIN, D. C., JONGMANS, A. G., & TARVAINEN, T. 2000. Chemical, mineralogical and morphological characterization of three podzols developed on glacial deposits in Northern Europe. *Geoderma*, **94**(2-4), 125–148.
- MIDL-HOMEPAGE. 2013. *Meron Library, (c) 2002 of Mati Meron [ONLINE]*. <http://www.astro.washington.edu/docs/idl/htmlhelp/slibrary23.html>.
- MOROZ, L. V., & SCHADE, U. 2011. Effects of viewing geometry on thermal infrared spectra of planetary surfaces : The case of enstatite. *Pages 5–6 of: EPSC-DPS Joint Meeting 2011*, vol. 6.
- MORRIS, R. V., BLAKE, D., BISH, F. D., MING, D. W., AGRESTI, D. G., TREIMAN, A. H., STEELE, A., & AMUNDSEN, H. E. F. 2011. A terrestrial analogue from Spitsbergen (Svalbard, Norway) for the Comanche Carbonate at Gusev Crater, Mars. *Pages 7–8 of: 42nd Lunar and Planetary Science Conference*.
- MÜLLER, R., LEHNER, M., REINARTZ, P., & SCHROEDER, M. 2005. Evaluation of Spaceborne and Airborne Line Scanner Images Using a Generic Ortho Image Processor. *In: HEIPKE, C., JACOBSEN, K., & GERKE, M. (eds), ISPRS Workshop - High Resolution Earth Imaging for Geospatial Information*.
- MURAD, E., & ROJÍK, P. 2005. Iron mineralogy of mine-drainage precipitates as environmental indicators: review of current concepts and a case study from the Sokolov Basin, Czech Republic. *Clay Minerals*, **40**, 427–440.
- NERC-FSF-HOMEPAGE. 2013. *NERC Field Spectroscopy Facility (FSF) [ONLINE]*. <http://fsf.nerc.ac.uk/> [Accessed 31 May 2013].
- PERONA, P., & MALIK, J. 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**(7), 629–639.
- PIECH, M. A., & PIECH, K. R. 1987. Symbolic representation of hyperspectral data. *Applied Optics*, **26**(18), 4018–4026.
- PIECH, M. A., & PIECH, K. R. 1989. Hyperspectral interactions: invariance and scaling. *Applied Optics*, **28**(3), 481–489.
- PIECH, M. A., & PIECH, K. R. 1990. Fingerprints and fractal terrain. *Mathematical Geology*, **22**(4), 457–485.

- QUATTROCHI, D. A., & LUVALL, J. C. 2004. *Thermal remote sensing in land surface processes*. Boca Raton, FL, USA: CRC Press.
- RÄISÄNEN, M. L., TARVAINEN, T., & AATOS, S. 1995. NORMA - a program to calculate a normative mineralogy for glacial tills and rocks from chemical analysis. *GFF*, **117**(4), 215–224.
- RAMSEY, M. S., & CHRISTENSEN, P. R. 1998. Mineral abundance determination: Quantitative deconvolution of thermal emission spectra. *Journal of Geophysical Research*, **103**, 577–596.
- REINHÄCKEL, G. 1998. *Quantitative Analyse von Braunkohlenabraumkippen mit laborspektroskopischen Infrarot-Messungen (2.5-14  $\mu\text{m}$ ) und Fernerkundungsdaten (DAIS 7915/ASTER)*. Ph.D. thesis, Ludwig-Maximilians-Universität München.
- RENCZ, A. N. (ed). 1999. *Manual of Remote Sensing, Vol. 3: Remote Sensing for the Earth Sciences*. 3rd edn. Wiley & Sons.
- RIAZA, A., BUZZI, J., GARCÍA-MELÉNDEZ, E., CARRÈRE, V., & MÜLLER, A. 2011. Monitoring the Extent of Contamination from Acid Mine Drainage in the Iberian Pyrite Belt (SW Spain) Using Hyperspectral Imagery. *Remote Sensing*, **3**(10), 2166–2186.
- RICHARDS, J. A., & JIA, X. 2006. *Remote sensing digital image analysis. An introduction*. 4th edn. Vol. 46. Berlin: Springer.
- RICHTER, N. 2010a. *Pedogenic iron oxide determination of soil surfaces from laboratory spectroscopy and HyMap image data*. Dissertation, Humboldt-Universität zu Berlin.
- RICHTER, R. 1998. Correction of satellite imagery over mountainous terrain. *Applied Optics*, **37**(18), 4004–4015.
- RICHTER, R. 2010b. *Atmospheric/Topographic Correction for Airborne Imagery (ATCOR-4 User Guide)*. Tech. rept. German Aerospace Center.
- RICHTER, R., & COLL, C. 2002. Bandpass-resampling effects for the retrieval of surface emissivity. *Applied Optics*, **41**(18), 3523–3529.
- RIVARD, B., FENG, J., GALLIE, A., & SANCHEZ-AZOFEIFA, A. 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. *Remote Sensing of Environment*, **112**, 2850 – 2862.
- ROBERTS, D. A., YAMAGUCHI, Y., & LYON, R. J. P. 1985. Calibration of airborne imaging spectrometer data to percent reflectance using field spectral measurements. *Pages 679–688 of: Proceedings of the 19th International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan (ERIM)*.
- ROBERTS, D. A., QUATTROCHI, D. A., HULLEY, G. C., HOOK, S. J., & GREEN, R. O. 2011. Synergies between VSWIR and TIR data for the urban environment: An evaluation of the potential for the Hyperspectral Infrared Imager (HyspIRI) Decadal Survey mission. *Remote Sensing of Environment*, **117**(Sept.), 83–101.
- ROGGE, D., BACHMANN, M., RIVARD, B., & FENG, J. 2012. Hyperspectral flight-line leveling and scattering correction for image mosaics. *Pages 4094–4097 of: Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, vol. 1. IEEE.



- ROGGE, D. M., & RIVARD, B. 2010. Iterative spatial filtering for reducing intra-class spectral variability and noise. *Pages 1–4 of: 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*. Reykjavik, Iceland: IEEE.
- ROJÍK, P. 2004. New stratigraphic subdivision of the Tertiary in the Sokolov Basin in Northwestern Bohemia. *Journal of Czech Geological Society*, **49**(3-4), 173–185.
- SALISBURY, J. W. 1998. *Spectral measurements field guide*. Tech. rept. Earth Satellite Corporation; Published by the Defense Technology Information Center as Report No. ADA362372.
- SALISBURY, J. W., & DARIA, D. 1992. Emissivity of terrestrial materials in the 8-14  $\mu\text{m}$  atmospheric window. *Remote Sensing of Environment*, **42**(2), 83–106.
- SALISBURY, J. W., WALTER, L. S., VERGO, N., & D'ARIA, D. M. 1991. Infrared (2.1- 25 micrometers) Spectra of Minerals. *Johns Hopkins University Press*, 294 pp.
- SALMINEN, R., GREGORAUSKIENE, V., & TARVAINEN, T. 2008. The normative mineralogy of 10 soil profiles in Fennoscandia and north-western Russia. *Applied Geochemistry*, **23**(12), 3651–3665.
- SANDER, J., ESTER, M., KRIEGEL, H. P., & XU, X. 1998. Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and its Applications. *Data Mining and Knowledge Discovery*, **2**(2), 169–194.
- SAVITZKY, A., & GOLAY, M. J. E. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, **36**(8), 1627–1639.
- SCHEFFER, F., SCHACHTSCHABEL, P., BLUME, H.-P., HARTGE, K.-H., & SCHWERTMANN, U. 1984. *Lehrbuch der Bodenkunde*. 11th edn. Stuttgart, Germany: Enke.
- SCHLÄPFER, D., & RICHTER, R. 2002. Geo-atmospheric processing of airborne imaging spectrometry data - Part 1: parametric orthorectification. *International Journal of Remote Sensing*, **23**, 2631–2649.
- SCHODLOK, M. C. 2004. *Quantitative Analysen magmatischer Gesteine mittels reflexionsspektroskopischer Infrarot-Messungen*. Ph.D. thesis, Universität Potsdam.
- SCHOWENGERDT, R. A. 1997. *Remote Sensing, Models and Methods for Image Processing*. 2nd edn. Academic Press.
- SMITH, K. L., STEVEN, M. D., & COLLS, J. J. 2004. Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks. *Remote Sensing of Environment*, **92**(2), 207–217.
- STORCH, T., DE MIGUEL, A., MÜLLER, R., MÜLLER, A., NEUMANN, A., WALZEL, T., BACHMANN, M., PALUBINSKAS, G., LEHNER, M., RICHTER, R., BORG, E., FICHTELMANN, B., HEEGE, T., SCHROEDER, M., & REINARTZ, P. 2008. The future spaceborne hyperspectral imager EnMAP: its calibration, validation, and processing chain. *Pages 1265–1270 of: Proceedings of the 21th Congress of the International Society for Photogrammetry and Remote Sensing - ISPRS 2008*.
- SU, H., LIU, Q., & LI, J. 2011. Alleviating Border Effects in Wavelet Transforms for Nonlinear Time-varying Signal Analysis. *Advances in Electrical and Computer Engineering*, **11**(3), 55–60.
- SUNSHINE, J. M., PIETERS, C. M., & PRATT, S. F. 1990. Deconvolution of mineral absorption bands: An improved approach. *Journal of Geophysical Research*, **95**(B5), 6955.

- SWAYZE, G. A., SMITH, K. S., CLARK, R. N., SUTLEY, S. J., PEARSON, R. M., VANCE, J. S., HAGEMAN, P. L., BRIGGS, P. H., MEIER, A. L., SINGLETON, M. J., & ROTH, S. 2000. Using imaging spectroscopy to map acidic mine waste. *Environmental Science & Technology*, **34**(1), 47–54.
- TARVAINEN, T., AATOS, S., & RÄISÄNEN, M. L. 1996. A method for determining the normative mineralogy of tills. *Applied Geochemistry*, **11**, 117–120.
- TORRENCE, C., & COMPO, G. P. 1998. A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, **79**(1), 61–78.
- TORRENCE, C., & WEBSTER, P. J. 1999. Interdecadal changes in the ENSO-monsoon system. *Journal of Climate*, **12**(8), 2679–2690.
- TSAI, F., & PHILPOT, W. 1998. Derivative Analysis of Hyperspectral Data. *Remote Sensing of Environment*, **66**(1), 41–51.
- VAN DER MEER, F. 2004. Analysis of spectral absorption features in hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, **5**(1), 55–68.
- VAN DER MEER, F., VAN DER WERFF, H., VAN DER MEIJDE, M., VAN RUITENBEEK, F., HECKER, C., & DE JONG, S. 2008. Object extraction and attribution from hyperspectral images. *Chap. 15, pages 205–207 of: LI, Z., CHEN, J., & BALTSAVIAS, E. (eds), Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS Congress Book*. London: Taylor & Francis.
- VAUGHAN, R. G., CALVIN, W. M., & TARANIK, J. V. 2003. SEBASS hyperspectral thermal infrared data: surface emissivity measurement and mineral mapping. *Remote Sensing of Environment*, **85**(1), 48–63.
- VAUGHAN, R. G., HOOK, S. J., CALVIN, W. M., & TARANIK, J. V. 2005. Surface mineral mapping at Steamboat Springs, Nevada, USA, with multi-wavelength thermal infrared images. *Remote Sensing of Environment*, **99**(1-2), 140–158.
- VERPLANCK, P. L. (ed). 2008. *Understanding contaminants associated with mineral deposits: U.S. Geological Survey Circular 1328*. Reston, Virginia: U.S. Geological Survey.
- ŠEBESTOVÁ, E., MACHOVIČ, V., PAVLÍKOVÁ, H., LELÁK, J., & MINARÍK, L. 1996. Environmental impact of brown coal mining in Sokolovo basin with especially trace metal mobility. *Journal of Environmental Science and Health . Part A: Environmental Science and Engineering and Toxicology*, **31**(10), 2453–2463.
- WANG, X., & ANDREWS, L. 2006. Infrared spectra of M(OH)(1,2,3) (M = Mn, Fe, Co, Ni) molecules in solid argon and the character of first row transition metal hydroxide bonding. *The Journal of Physical Chemistry*, **110**(33), 10035–10045.
- WHITE, L. B., & BOASHASH, B. 1990. Cross spectral analysis of nonstationary processes. *IEEE Transactions on Information Theory*, **36**(4), 830–835.
- YITAGESU, F. A., VAN DER MEER, F., VAN DER WERFF, H., & HECKER, C. 2011. Spectral characteristics of clay minerals in the 2.5-14 $\mu$ m wavelength region. *Applied Clay Science*, **53**(4), 581–591.
- YOUNG, S. J., JOHNSON, B. R., & HACKWELL, J. A. 2002. An in-scene method for atmospheric compensation of thermal hyperspectral data. *Journal of Geophysical Research*, **107**(D24), 4774–4793.

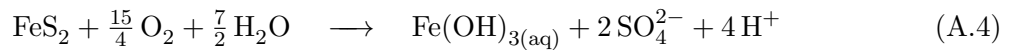
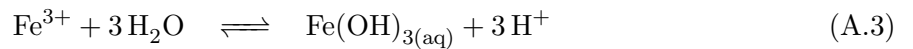
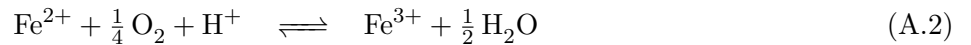
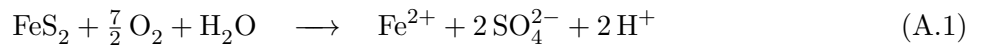
# A. Acid Mine Drainage Supplements

The effect of acid rock drainage (ARD) describes processes that lead to generation and seepage of acidic water enriched with metals due to the exposure of sulphur bearing minerals to oxygen and water. The most abundant sulphides pyrite and marcasite (both  $\text{FeS}_2$ ) commonly bear metals such as Mn, Co, Ni, Cu and Zn which replace the Fe to a certain fraction while As and Se can replace the S. Further common constituents are mixed forms like chalcopyrite ( $\text{CuFeS}_2$ ) and arsenopyrite ( $\text{FeAsS}$ ). On weathering these elements will be released. The effluent poses an environmental risk due to its load with heavy metals and sulphate, both having a deteriorating impact on aquatic life and vegetation (Murad & Rojik, 2005; Dold, 2010). The more specific term acid mine drainage (AMD) relates to similar effects occurring due to artificial exposure of sulphide rich materials in the course of mining operations.

This appendix presents the major abiotic processes that govern AMD in coal deposits and mining. A model scheme for precipitation-controlled climates is presented in Figure A.2. It visualises the minerals and reactions characteristic for the three general AMD processes, i) acid generation by sulfide oxidation, ii) neutralisation of the acidity, and iii) precipitation of secondary minerals. The given statements and formulae follow the reviews of Dold (2010) and Blowes *et al.* (2003).

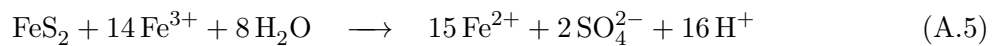
## A.1. Acid Production by Sulphide Oxidation

The predominant sulphide mineral is pyrite ( $\text{FeS}_2$ ). It is commonly associated with coal deposits. Its oxidation can be split into three steps. Initially, the pyritic sulphur is oxidised due to the presence of atmospheric oxygen forming ferrous ion (Equation A.1). The ferrous ion is oxidised forming ferric ion (Equation A.2). Finally, the ferric ion undergoes hydrolysis and various Fe(III) species may form, e.g. iron hydroxide (Equation A.3).



The net reaction of all three pyrite oxidation steps produces four moles of  $\text{H}^+$  per mole pyrite (Equation A.4). Hence, the final hydrolysis step is the main acidity producer.

Ferric ion, once produced, will become the primary oxidant of pyrite with 10 to 100 times faster reaction kinetics than oxidation by oxygen for  $\text{pH} < 3$  regimes (Equation A.5). Under abiotic conditions the pyrite oxidation rate is hence controlled by the rate of ferrous ion oxidation, latter reaction kinetic dwindling rapidly with decreasing pH.



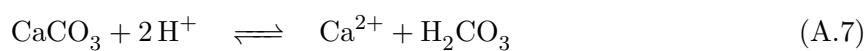
## A.2. Acid Neutralisation Processes

The neutralising processes vary strongly in reaction speed depending on the involved minerals. Table A.1 attempts a classification. Three important groups are detailed below.

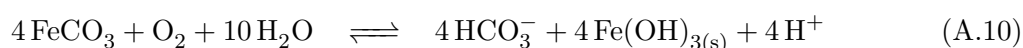
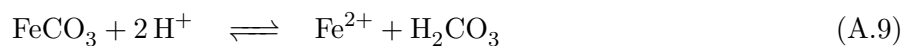
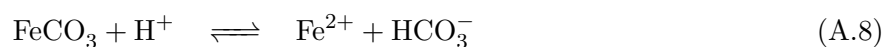
Table A.1.: Relative reactivity of acidity neutralising minerals, shortened after Jambor & Blowes (1998).

speed	typical minerals	rel. reactivity
dissolving	calcite, dolomite, anorthite	1.0
fast	nepheline, leucite	0.6
intermediate	biotite, chlorite	0.4
slow	plagioclase, kaolinite, montmorillonite, gibbsite	0.02
very slow	k-feldspar, muscovite	0.01
inert	quartz, rutile	0.004

**Carbonate Mineral Dissolution** Depending on the mineral Ca, Mg, Fe and Mn cations are released in fast neutralising reactions. The most common and fastest reacting carbonate is calcite. Its dissolution has differing neutralisation potential. At near neutral pH one mole calcite can neutralise 1 mole  $H^+$  (Equation A.6) while at lower pH this increases to two moles (Equation A.7). Calcite may precipitate as secondary mineral when pH increases.



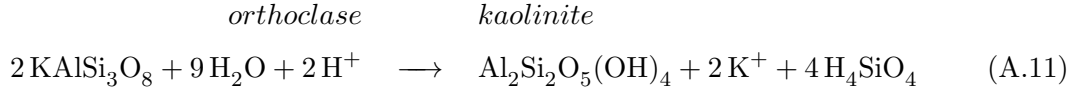
Siderite may act as acid neutraliser or producer. At pH around 5.0 - 5.5 the buffering reaction of Equation A.8 produces ferrous ion. Similarly, ferrous ion production occurs for Equation A.9 which buffers pH around 5. Together with oxidation of ferrous ion and precipitation of iron hydroxide an overall siderite oxidation reaction (Equation A.10) may take place producing a net acidity of one mole  $H^+$  per mole siderite when elevated pH conditions favour stability of  $HCO_3^-$ . If ferrous ion and carbonate are present in solution the precipitation of siderite may occur.



**Hydroxide Mineral Dissolution** Carbonate mineral dissolution maintains a near neutral pH which favours the precipitation of amorphous or crystalline metal hydroxides. In case of additional acidity being produced this secondary hydroxides themselves may subsequently dissolve contributing to neutralise the surplus acidity. Primarily, aluminium hydroxide (gibbsite) takes over the buffering followed by iron-oxyhydroxides, e.g. ferrihydrite or goethite at even lower pH.

**Silicate Mineral Dissolution** Dissolution of silicates liberates Ca, Mg, Fe, Na, K, Si and Al cations and consumes acidity. The reaction is much slower than that of carbonates and hydroxides therefore not able to buffer pH at a specific range. The release of Al may lead to the formation of secondary buffering minerals such as gibbsite.

Feldspar dissolution is dependent on pH value and the concentrations of K, Na, Ca and Si in the solution. K-feldspar may react forming kaolinite, at higher pH muscovite is preferred. Latter both may react to gibbsite. Plagioclase will similarly alter to kaolinite under low pH conditions while at higher values smectite is formed. Felsic minerals such as muscovite, pyroxene and amphibole weather to chlorite which itself may react forming sericite, kaolinite or montmorillonite.

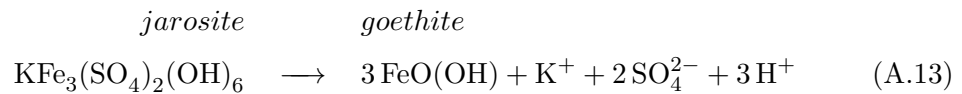
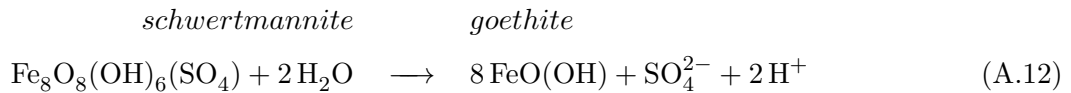


These strongly simplified reaction paths can not adequately describe the complex interactions. Details are given in [Dold \(2010\)](#), [Blowes \*et al.\* \(2003\)](#) and references.

### A.3. Formation of Secondary Precipitates

The hydrolysis of  $\text{Fe}^{3+}$  generates the bulk acidity during the sulphide oxidation process. The formation of stable Fe(III) species is dependent on the pH regime. At  $\text{pH} < 2$  no hydrolysis product is stable and thus the  $\text{Fe}^{3+}$  remains in solution. After hydrolysis has established an equilibrium in relatively short time, the ageing process takes considerably longer. More and more complex species form which may finally precipitate as secondary minerals. The precipitate strongly depends on the redox (Eh) and pH conditions, temperature, concentration of key elements such as dissolved metal cations and water availability.

Iron sulphate minerals are commonly found in the oxidising environments with excess sulphur. They are easily water soluble, with the exception of jarosite, and hence are meta-stable in respect to the more dissolution resistant iron hydroxides and oxides. This transformation from meta-stable to stable precipitates generates additional acidity.



### A.4. Buffering Sequence

Acidity and aqueous ferric ion of the solution are reacting with acid-neutralising minerals in the solution's pathway by dissolving minerals such as carbonates or silicates thus mobilising alkaline earth and metal cations. The neutralising reactions increase the pH value of the solution and metal hydroxide, oxyhydroxide and oxyhydroxide sulphate minerals are precipitated removing dissolved metals from the solution. Both neutralising minerals and secondary minerals buffer pH in specific ranges. A buffering sequence forms (Table [A.2](#)).

In tailings impoundments such ideal buffering sequences have been reported. Regions of relatively constant pH dominated by a single buffering reaction are separated by fronts of sharp pH value changes (Figure [A.1](#)). Along the solution's flow path near neutral pH may be reached with

Table A.2.: AMD relevant mineral pH buffering reactions, from Dold (2010) and Blowes et al. (2003).

mineral	reaction	pH buffer
jarosite	$3\text{Fe}^{3+} + \text{K}^+ + 2\text{SO}_4^{2-} + 6\text{H}_2\text{O} \rightleftharpoons \text{KFe}_3(\text{SO}_4)_2(\text{OH})_6 + 6\text{H}^+$	< 3.0
schwertmannite	$8\text{Fe}^{3+} + \text{SO}_4^{2-} + 14\text{H}_2\text{O} \rightleftharpoons \text{Fe}_8\text{O}_8(\text{OH})_6\text{SO}_4 + 22\text{H}^+$	2.8 - 3.5
ferrihydrite	$\text{Fe}(\text{OH})_{3(s)} + 3\text{H}^+ \rightleftharpoons \text{Fe}^{3+} + 3\text{H}_2\text{O}$	2.5 - 3.5
goethite	$\text{FeO}(\text{OH})_{(s)} + 3\text{H}^+ \rightleftharpoons \text{Fe}^{3+} + 2\text{H}_2\text{O}$	2.5 - 3.5
gibbsite	$\text{Al}(\text{OH})_3 + 3\text{H}^+ \rightleftharpoons \text{Al}^{3+} + 3\text{H}_2\text{O}$	4.0 - 4.5
siderite	$\text{FeCO}_3 + 2\text{H}^+ \rightleftharpoons \text{Fe}^{2+} + \text{H}_2\text{CO}_3$ $\text{FeCO}_3 + \text{H}^+ \rightleftharpoons \text{Fe}^{2+} + \text{HCO}_3^-$	5.0 - 5.5
calcite	$\text{CaCO}_3 + 2\text{H}^+ \rightleftharpoons \text{Ca}^{2+} + \text{H}_2\text{CO}_3$ $\text{CaCO}_3 + \text{H}^+ \rightleftharpoons \text{Ca}^{2+} + \text{HCO}_3^-$	6.5 - 7.5

pore water in equilibrium with respect to calcite. In direction to the oxidation zone, separated by a sharp drop in pH from the calcite-buffered zone, follows the siderite and dolomite buffering zone. When all carbonates are consumed another sharp drop in pH marks the transition to the gibbsite-buffered zone with Al enriched in the pore water. Nearest to the oxidising zone the pH value again drops sharply followed by ferrihydrite or goethite as buffer.

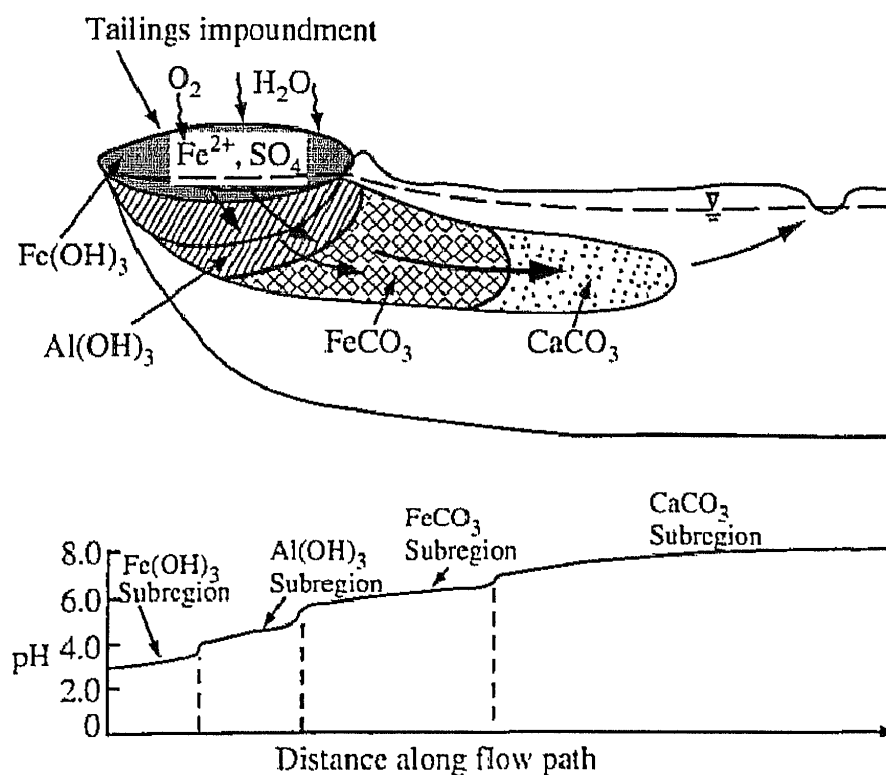


Figure A.1.: Idealised scheme of fully developed pH buffering sequence in a tailings impoundment, from Blowes & Ptacek (1994).

From above statements a schematic model describing the fundamental AMD reaction paths and associated element circulations for a precipitation-controlled climate can be developed (Figure A.2). The initial oxidation of sulphides releases ferrous ion ( $\text{Fe}^{2+}$ ), sulphate ions ( $\text{SO}_4^{2+}$ ) and protons ( $\text{H}^+$ ). Latter causes low pH conditions in the oxidation zone leveraging the alteration of carbonate, iron hydroxide and silicate minerals liberating metal cations. The oxidation of ferrous ion forms ferric ion ( $\text{Fe}^{3+}$ ) which subsequently may become the primary sulphide

oxidant or hydrolyse. Governed by the redox and pH conditions and the availability of free ions such as  $\text{SO}_4^{2+}$ ,  $\text{Fe}^{3+}$ ,  $\text{K}^+$  and  $\text{Al}^{3+}$  in solution, several secondary phases may form. Jarosite is favoured under acidic conditions close to the source of  $\text{K}^+$  hence its a dominant precipitate of the oxidation zone. The remaining ferric ion is mobile under acidic conditions and may precipitate as schwertmannite at higher pH environments encountered along the water flow path and when sulphate is available. Surplus ferric ion may reach the neutralisation zone and precipitate as ferrihydrite or goethite. Precipitation as well as transformation of the meta-stable jarosite and schwertmannite into goethite all produce additional acidity.

Many, but not all, of the minerals relevant for AMD are in principle accessible to remote sensing. The following section presents measuring concepts and characteristic spectral features for above minerals that allow their identification and mapping using passive optical sensors in the reflective and thermal spectral domains. Depending on the geochemical setting, climatic conditions and weathering rates only a subset of these key minerals are exposed at the surface. The spatial and spectral resolution of the sensor is another constraint which governs their detectability.

For the Sokolov demonstration site this limited the investigations to goethite, siderite, clays and quartz as reasoned in Section 2.2 - *Application Objective & Mineral Mapping Candidates*.

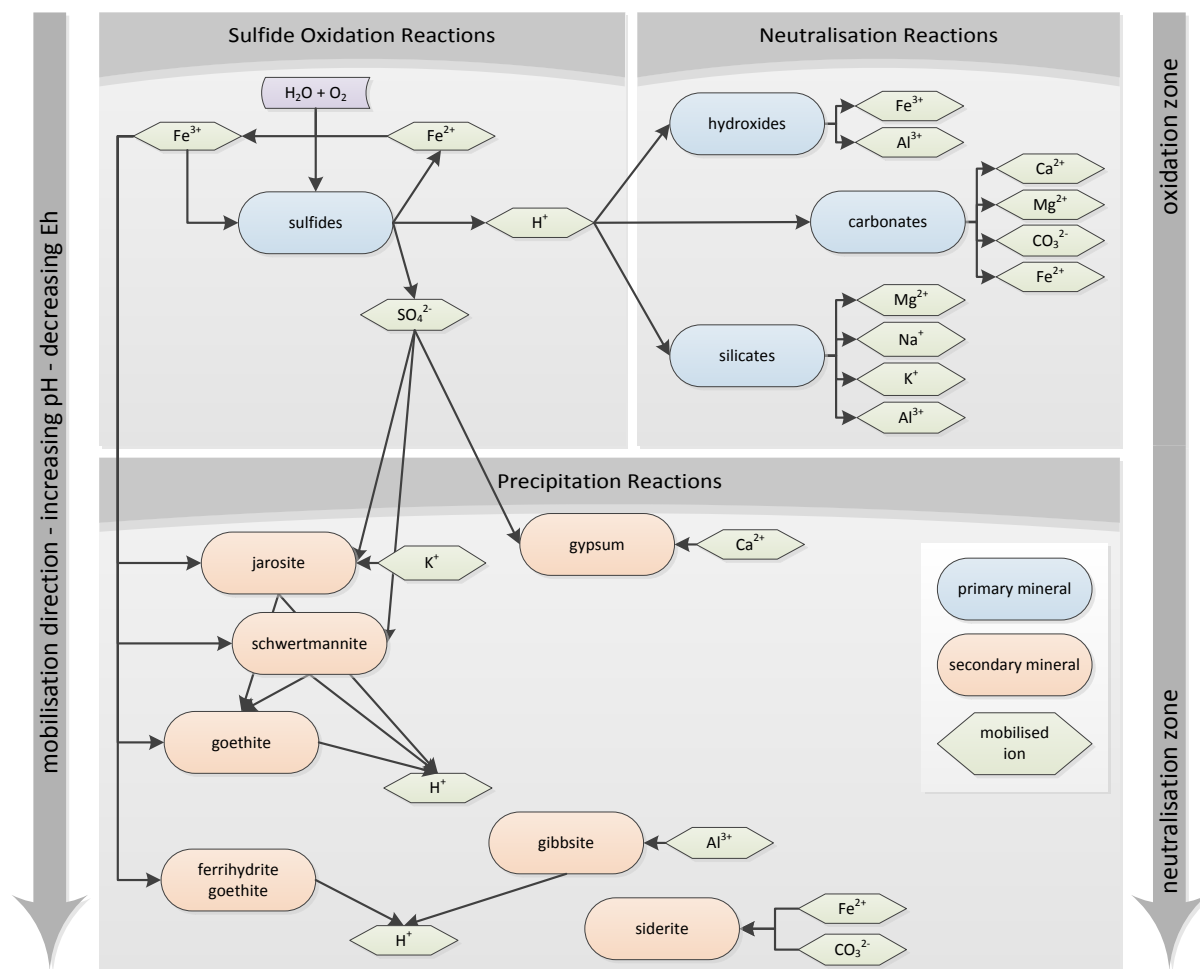


Figure A.2.: AMD reaction paths and element circulations model for precipitation-controlled climates, modified after Dold (2005).

## B. Normative Modelling Supplements

This appendix gives information about the stratigraphic setting of the Sokolov demonstration site, the adaptation of the normative models and the sample's geochemical laboratory analysis together with their normative modelling results.

### B.1. Sokolov Stratigraphy

Chrono-stratigraphy	Formation	Member	Characteristic rocks	Typical environment
Holocene			antropogenic deposits, fluvial loams, peats, porcelanites	antropogenic, fluvial
Pleistocene		nine local terrace benches	gravels, loams, loess loams, block fields, peats, porcelanites, mineralized faults	fluvial, colian, solifluction
unconformity				
Burdigalian <i>Ottangian-Karpatian</i>	Cypris Formation	Čankov Sand	sands, silty claystones	delta
			laminated claystones illite-montmorillonite-kaolinitic, admixtures of Ca-Mg-Fe-carbonates, Fe-sulfides, analcite and bitumen; diastems	lake-playas complex
			laminated kaolinitic clays, admixtures of siderite, sulfides or sulfates	permanent lake
Burdigalian <i>Eggenburgian</i>	Sokolov Formation	Antonín Member	humic coal; local diastems	mires
		Těšovice Member	basaltic rocks, tuffs, tuffites, sediments of gravity flows (altered)	volcanic, gravity flows
		Anežka Member	sapro- and liptodetritic coal	mires
		Habartov Member	sands, sandy and silty clays (bioturbated)	fluvial, swamps
disconformity				
Chattian – Aquitanian <i>Egerian</i>	Nové Sedlo Formation	Chodov Member	basaltic rocks, tuffs, tuffites (altered)	volcanic, swamps, lakes
Rupelian		Josef Member	sapropelitic and humic coal; local diastems	mires
Oligocene		Davidov Member	clayey sands / sandy clays with gravel admixture – unsorted, unbedded; local diastems	creep, fluvial
disconformity (tectonic movements and hydrothermal activity)				
Upper Eocene – Lower Oligocene	Staré Sedlo Formation		sands and sandstones, gravels and conglomerates, locally clays - sorted, bedded; diastems	fluvial
local disconformity				
Paleogene			kaolins and silica rezidues of metamorphites and granites	weathering
unconformity (tectonic fracturing and hydrothermal activity, Triassic - Cretaceous)				
Upper Proterozoic - Upper Carboniferous	Saxothuringian Unit: polymetamorphosed crystalline complexes – the Ohře, Krušné hory, Slavkovský les and Thüringen-Vogtland units and the Karlovy Vary pluton			

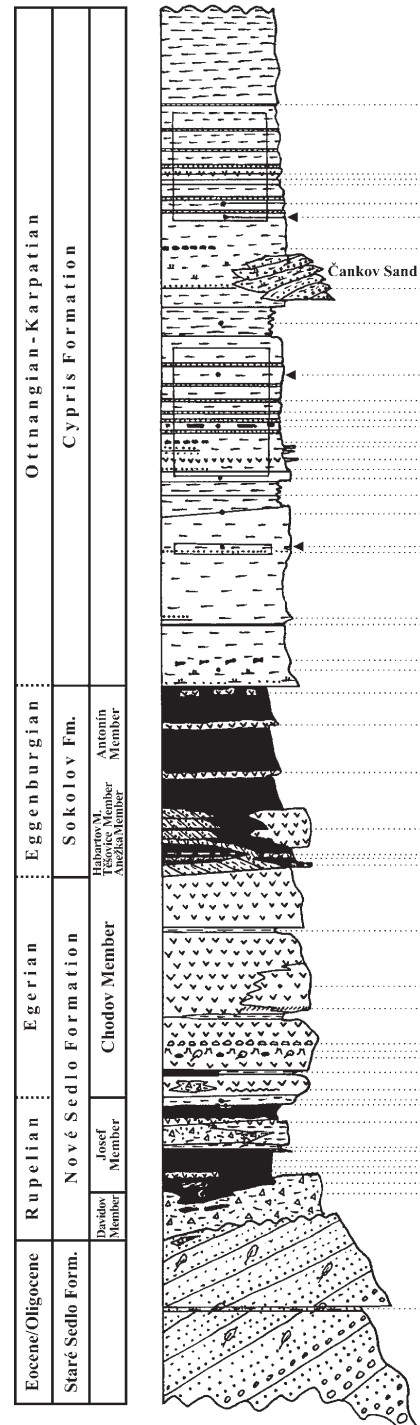


Figure B.1.: Stratigraphic scheme (left) and Tertiary column (right) of the Sokolov Basin (Roják, 2004).



## B.2. Adaptation of Normative Models

### B.2.1. SEDNORM

SEDNORM considers 18 minerals, that generally form the bulk of the most typical sediments (see Figure B.2). The considered minerals are not suitable for normative calculation of chemical sediments, magniferous deposits or iron ores. Cohen & Ward (1991) references a study of numerous sedimentary rock samples from eastern Australian coalfields which found good correlation between normative quartz and point count and XRD methods.

SEDNORM requires total elemental concentrations in form of the oxides  $\text{SiO}_2$ ,  $\text{TiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{Fe}_2\text{O}_3$ ,  $\text{MnO}$ ,  $\text{MgO}$ ,  $\text{CaO}$ ,  $\text{Na}_2\text{O}$ ,  $\text{K}_2\text{O}$ ,  $\text{P}_2\text{O}_5$ ,  $\text{H}_2\text{O}$ ,  $\text{CO}_2$ ,  $\text{SO}_3$  and  $\text{Cl}$ . No differentiation between ferric and ferrous ion can be made. Sulphur similarly is not differentiated between its variants S,  $\text{SO}_3$  or  $\text{SO}_4$ . Figure B.2 outlines the oxide allocation sequence. Various user options govern the normative calculations - such as the assignment of S to pyrite or gypsum, the incorporation of feldspars into the calculations, the distribution of K into muscovite, illite and feldspars as well as the inclusion of smectite with distribution ratios of Ca:Na and Mg:Fe into smectite and the distribution of Mg and Fe between dolomite, siderite and smectite.

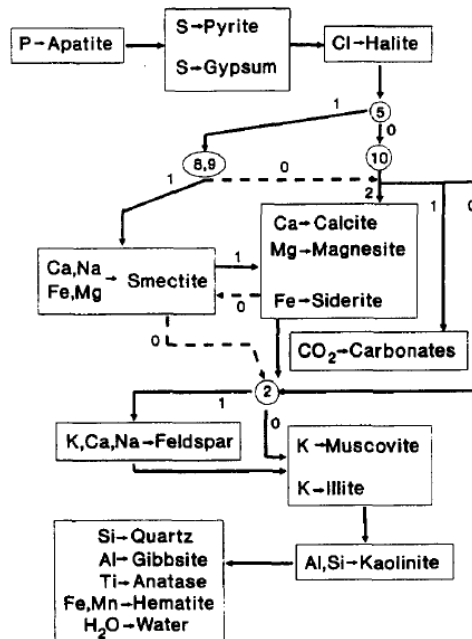


Figure B.2.: Normative mineral calculation sequence of SEDNORM taken from Cohen & Ward (1991).

For Sokolov, SEDNORM was set-up in the following way: All phosphorous is assumed to form apatite. Sulphur is nominated to be assigned to pyrite as sulphide. In case of a lack of iron the remaining sulphur is assigned to gypsum as sulphate. The available lab analysis gives no chlorine concentrations. Therefore it was fixed at zero for all samples effectively excluding halite. With the given  $\text{CO}_2$  concentrations, SEDNORM calculates carbonate minerals by first assigning  $\text{CO}_2$  to calcite, then magnesite and finally siderite. Any excess iron is assigned to smectite or finally haematite. SEDNORM allows potassium to be distributed into illite, muscovite and K-feldspar. In accordance with the XRD analysis was muscovite nominated the potassium bearing layer silicate resulting in no illite being formed. The selected inclusion of feldspars into the

normative calculations allows for the formation of K-feldspar and plagioclase. Also an initial distribution of potassium between muscovite and K-feldspar must be chosen and was selected 1:1 (out of 1:0, 1:1 and 0:1). Any excess potassium is used to form additional K-feldspar. Other clay minerals than illite and kaolinite are pooled by SEDNORM into a smectite material representing “an iron-magnesium-bearing three-layer clay mineral without interlayer water” (Cohen & Ward, 1991). Its formula  $(\text{Na,Ca})([\text{Mg,Fe}]\text{Al}_3)[\text{Si}_8\text{O}_{20}](\text{OH})_4$  constitutes “a hybrid of dehydrated montmorillonite, nontronite, vermiculite, and chlorite” (Cohen & Ward, 1991). Smectite was excluded from the norm given its low occurrence in the XRD analysis and highly variable nature interfering with other normative minerals that would need definition through element distribution ratios. All remaining calcium and sodium thus are assigned plagioclase in form of anorthite or albite respectively. Kaolinite is formed from leftover  $\text{Al}_2\text{O}_3$  and  $\text{SiO}_2$ . Any additional remaining  $\text{SiO}_2$  is assigned quartz, in case of remaining  $\text{Al}_2\text{O}_3$  gibbsite is formed.  $\text{MnO}$  is assumed to form haematite and all  $\text{TiO}_2$  is reported anatase.  $\text{H}_2\text{O}$  concentration was opted to not limit the normative mineral formation.

### B.2.2. NORMA

In its original version NORMA considers 22 minerals “predominantly found in acid and intermediate plutonic rocks and their weathering products” (Räisänen *et al.*, 1995). In order to differentiate secondary and easily weathered minerals from resistant components, NORMA handles two suites of element concentrations - minerals soluble in hot aqua regia by a leaching procedure and total element concentrations given by whole rock analysis. The hot aqua regia leaching is assumed to dissolve trioctahedral micas, clay minerals, salts and hydroxy-oxides while the main silicates quartz, feldspars, hornblende and resistant oxides are not. This allows for a more precise normative calculation of chemically similar primary and secondary minerals compared to traditional normative methods. Räisänen *et al.* (1995), Tarvainen *et al.* (1996) and Salminen *et al.* (2008) demonstrate the applicability of NORMA for a large number of soils developed on various parent materials and bedrocks spanning arctic to boreal climate zones in Finland and Russia.

NORMA requires 13 total and 10 leachable element concentrations encompassing  $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{Fe}_2\text{O}_3$  or  $\text{FeO}$ ,  $\text{MgO}$ ,  $\text{CaO}$ ,  $\text{Na}_2\text{O}$ ,  $\text{K}_2\text{O}$ ,  $\text{P}_2\text{O}_5$ ,  $\text{BaO}$ ,  $\text{TiO}_2$  and  $\text{ZrO}$  as well as  $\text{H}_2\text{O}$ , C and S. The normative calculation follows a strict sequence split into three stages: i) soluble pyrite, ii) other soluble minerals and iii) calculation of insoluble minerals.

First molar proportions (MP) are calculated by dividing the element oxide concentrations by their respective molar mass. In step i) normative pyrite is formed by assigning all sulphur to iron and reducing the MP of aqua regia soluble  $\text{FeO}$  by 0.5 times the MP of sulphur. For step ii) the normative minerals are calculated sequentially. The amount of mineral formed is limited by the availability of needed oxides as given by the oxide formula. The respective MP are subtracted and the next mineral is calculated. Water does not limit mineral formation as it is not given for aqua regia analysis. Instead, the required MP of  $\text{H}_2\text{O}$  are summed up for later deduction from the total MP of water. For step iii) the insoluble element MP are calculated by subtracting the initial soluble MP from the total element MP. Otherwise the same procedure as in step ii) is followed, however water concentrations now limiting the mineral formation.

Table B.1.: Sequence of normative mineral calculations of NORMA according to [Salminen et al. \(2008\)](#). Adaptation for Sokolov by adding bold typed minerals and omitting struck out ones.

step	mineral	oxide formula
1	Pyrite	FeS <sub>2</sub>
1b	<b>Gypsum</b>	<b>CaO · SO<sub>3</sub> · 2 H<sub>2</sub>O</b>
1c	<b>Jarosite</b>	<b>K<sub>2</sub>O · 3 Fe<sub>2</sub>O<sub>3</sub> · 4 SO<sub>3</sub> · 6 H<sub>2</sub>O</b>
2	Apatite	10 CaO · 3 P <sub>2</sub> O <sub>5</sub> · H <sub>2</sub> O
3	Titanite	CaO · TiO <sub>2</sub> · SiO <sub>2</sub>
4	Calcite	CaO · CO <sub>2</sub>
4b	<b>Siderite</b>	<b>FeO · CO<sub>2</sub></b>
5	Biotite	K <sub>2</sub> O · Al <sub>2</sub> O <sub>3</sub> · MgO · 5 FeO · 6 SiO <sub>2</sub> · 2 H <sub>2</sub> O
5b	<b>Muscovite</b>	<b>K<sub>2</sub>O · 3 Al<sub>2</sub>O<sub>3</sub> · 6 SiO<sub>2</sub> · 2 H<sub>2</sub>O</b>
6	Chlorite	5 MgO · Al <sub>2</sub> O <sub>3</sub> · 3 SiO <sub>2</sub> · 4 H <sub>2</sub> O
7	Weathered albite	Na <sub>2</sub> O · Al <sub>2</sub> O <sub>3</sub> · SiO <sub>2</sub>
8	Hydrous aluminosilicates	SiO <sub>2</sub> · Al <sub>2</sub> O <sub>3</sub> · 2 H <sub>2</sub> O
9	Goethite	2 Fe <sub>2</sub> O <sub>3</sub> · H <sub>2</sub> O
10	soluble residues	
11	Anatase	TiO <sub>2</sub>
12	Hornblende	Na <sub>2</sub> O · 2 CaO · 8 MgO · 4 FeO · Al <sub>2</sub> O <sub>3</sub> · 24 SiO <sub>2</sub> · 6 H <sub>2</sub> O
13	Potassium feldspar	K <sub>2</sub> O · Al <sub>2</sub> O <sub>3</sub> · 6 SiO <sub>2</sub>
14	Albite	Na <sub>2</sub> O · Al <sub>2</sub> O <sub>3</sub> · 6 SiO <sub>2</sub>
15	Anorthite	CaO · Al <sub>2</sub> O <sub>3</sub> · 2 SiO <sub>2</sub>
16	Tremolite	2 CaO · 5 MgO · 8 SiO <sub>2</sub> · H <sub>2</sub> O
17	Wollastonite	CaO · SiO <sub>2</sub>
18	Kaolinite	4 SiO <sub>2</sub> · 2 Al <sub>2</sub> O <sub>3</sub> · 2 H <sub>2</sub> O
19	Magnetite	FeO · Fe <sub>2</sub> O <sub>3</sub>
20	Zircon	ZrO <sub>2</sub> · SiO <sub>2</sub>
21	Quartz	SiO <sub>2</sub>
22	Carbon (graphite)	€
23	Non-soluble residuals	

For Sokolov the NORMA program as given in [Salminen et al. \(2008\)](#) was modified based on the available laboratory analyses and mineralogic setting (see Table B.1). A major deviation from the NORMA concept results from the lack of hot aqua regia leaching analysis, hence no differentiation between soluble and insoluble minerals could be made. As a result the NORMA principle of depleting element MP was generalized for any sequence of minerals given as oxide formulae. Due to the lack of the leaching data, weathered albite and hydrous aluminosilicates had to be skipped, only total residuals can be reported and mineral ambiguities arise, e.g. for magnetite and titanite. The soluble goethite now depletes all Fe<sub>2</sub>O<sub>3</sub> leaving none for the formation of the insoluble magnetite. The soluble titanite competes with the insoluble anatase for TiO<sub>2</sub> and with calcium-bearing minerals for CaO. Abundance of TiO<sub>2</sub> and SiO<sub>2</sub> would lead to the complete depletion of CaO. Based on the XRD analysis titanite was dropped in favour of anatase and calcite.

For Sokolov, the formation of normative calcite typically consumes all CaO. Together with the depletion of MgO by chlorite no hornblende, tremolite or wollastonite can form which is in accordance with the XRD analysis. Zircon was removed from the norm as no lab analysis for ZrO concentration was available. Lab analysis differentiates carbon into C<sub>total</sub> and CO<sub>2</sub>. The latter is directly used by the norm cancelling the necessity of collecting remaining carbon into graphite. The soluble normative mica biotite was replaced in accordance with XRD by muscovite. In doing so a conflict for K<sub>2</sub>O would result in no potassium feldspar to be formed. Following the SEDNORM approach a 50:50 distribution of K<sub>2</sub>O between muscovite and feldspar was chosen.

Incorporation of jarosite, siderite and gypsum is justified by their frequent XRD identification in the samples.  $\text{SO}_3$  is assumed to occur as jarosite and gypsum. Gypsum takes the precedence as CaO is assumed to first neutralise  $\text{SO}_3$  before normative calcite and jarosite can be formed. Brophy & Sheridan (1965) and Verplanck (2008) stress a natural solid solution series of three jarosite endmembers, with substitution of  $\text{K}^+$  in the regular jarosite by  $\text{Na}^+$  (natrojarosite) and  $\text{H}_3\text{O}^+$  (hydronium jarosite). The jarosite endmember interactions on XRD intensity peaks and reflectance spectra are complex requiring a dedicated laboratory analysis. Generally it can be concluded that a  $\text{K}^+$  over  $\text{Na}^+$  preference exists for low-temperature, low-pressure supergene jarosite with a considerable fraction of  $\text{H}_3\text{O}^+$  substituting. Formation of hydronium jarosite is further promoted by highly acid conditions such as mine waters because of low alkali concentrations. Concluding, the oxide formula of regular jarosite is chosen for NORMA which depletes the  $\text{SO}_3$  left after gypsum formation provided the availability of  $\text{K}_2\text{O}$ . In case  $\text{SO}_3$  remains after consumption of  $\text{K}_2\text{O}$  it is assigned to normative hydronium jarosite  $3\text{Fe}_2\text{O}_3 \cdot 4\text{SO}_3 \cdot 9\text{H}_2\text{O}$ . In the Sokolov adjusted implementation of NORMA water does not limit the formation of normative minerals by assuring an excess concentration of  $\text{H}_2\text{O}$ .

### B.3. Sample Mineralogy

The laboratory analysis of the mineral soil samples was conducted by the ČSN ISO/IEC 17025 accredited Prague-Barrandov laboratory of the Czech Geological Survey. It comprises of whole rock elemental analysis, qualitative X-ray diffraction (XRD) mineralogical identification and complementary analyses for total sulphur, total carbon, pH and total exchangeable acidity.

Table B.2.: Laboratory analysis methods of mineral soil samples as reported by Czech Geological Survey.

analysis target	detection limit	maximum error	method description
CO <sub>2</sub> [wt-%]	0.01%	sum to 97.5% - 101.3%	carbon dioxide released from reaction with concentrated H <sub>3</sub> PO <sub>4</sub> determined by infrared absorption using ELTRA CS-500 analyser; (ČSN EN 13137)
S <sub>pyrite</sub> [wt-%]	-		pyritic sulphur
SO <sub>3</sub> [wt-%]	0.05%		gravimetric determination of sulphate precipitated with barium (Ba <sup>2+</sup> ) and weighed as BaSO <sub>4</sub> ; (ČSN 72 0117)
Fe <sub>2</sub> O <sub>3</sub> [wt-%]	0.01%		ferric ion; (ČSN 72 0110-1)
FeO [wt-%]	0.03%		ferrous ion; (ČSN 72 0111)
SiO <sub>2</sub> [wt-%]	0.1%		total silica; (ČSN 72 0100)
Al <sub>2</sub> O <sub>3</sub> [wt-%]	0.01%		total aluminium; (ČSN 72 0109)
TiO <sub>2</sub> [wt-%]	0.01%		total titanium
MnO [wt-%]	0.001%		total manganese
MgO [wt-%]	0.01%		total magnesium; (ČSN 72 0114-1)
CaO [wt-%]	0.01%		total calcium; (ČSN 72 0113-1)
Li <sub>2</sub> O [wt-%]	0.001%		total lithium
Na <sub>2</sub> O [wt-%]	0.01%		total sodium; (ČSN 72 0119)
K <sub>2</sub> O [wt-%]	0.01%		total potassium; (ČSN 72 0120)
P <sub>2</sub> O <sub>5</sub> [wt-%]	0.05%		total phosphorus; (ČSN 72 0116-2)
H <sub>2</sub> O <sup>-</sup> [wt-%]	0.05%		loss on drying; (ČSN 72 0102)
H <sub>2</sub> O <sup>+</sup> [wt-%]	0.005%		loss on ignition; (ČSN 72 0103)
XRD	-	-	qualitative mineral identification by X-ray diffraction of whole-rock powder samples and oriented clay fraction (< 2μm) specimens; goethite and pyrite identified using voltammetry
S <sub>tot</sub> [wt-%]	0.01%	0.228%	total sulphur measured with ELTRA CS-500 analyser; (ČSN EN 13 137)
C <sub>tot</sub> [wt-%]	0.05%	0.0642%	total carbon measured with ELTRA CS-500 analyser; (ČSN EN 13 137)
pH [-]	-	0.15	soil pH measured using ion-selective electrode in 1M KCl solution
TEA [mmol/kg]	-	5%	total exchangeable acidity measured by titration of soil extracts (0.1M BaCl <sub>2</sub> ) to a pH of 8.2

The following figures show the elemental composition for all 80 samples sent for whole rock elemental analysis and the normative modelling results for this samples based on the SEDNORM and NORMA approaches. The plots are simplified to show only the most abundant elements or minerals and are sorted by descending SiO<sub>2</sub> or quartz fraction respectively. Note that in case of SEDNORM sums far below 100 wt-% result from mineral fractions not accounted for by this model, e.g. jarosite.

Table B.3.: Qualitative XRD analysis results of 50 mineral soil samples from 2010 as reported by Czech Geological Survey.

name	XRD main	XRD secondary	XRD accessory
DR10_02	Quartz, Kaolinite	Muscovite, Anatase, Goethite	Edisonite, Jarosite
DR10_03	Siderite, K-feldspar, Smectite	Kaolinite, probably Apatite(?)	Quartz
DR10_04A	Kaolinite, Quartz, Anatase	-	Jarosite, Siderite, probably Ilmenite and Edisonite
DR10_05	Kaolinite, Siderite, Anatase	-	Goethite, Quartz, probably Apatite
DR10_06	Kaolinite, Siderite, Anatase	Quartz, Goethite, Haematite	Edisonite, Smectite, probably Pyrite, Crandallite group mineral
DR10_07B	Quartz, K-feldspar, Albite, Muscovite	-	Pyroxene, Chlorite, probably Kaolinite, Smectite
DR10_08	Quartz, K-feldspar	Kaolinite, Pyroxene	Muscovite, Jarosite, probably Ilmenite
DR10_09	Quartz, K-feldspar	Kaolinite, Pyrite, Pyroxene	Gypsum, Jarosite, probably Ilmenite, Muscovite
KA10_01	Quartz, Kaolinite	Muscovite	-
KA10_02	Quartz	Kaolinite, Siderite, Muscovite	-
KA10_03A	Quartz	Muscovite, Kaolinite	K-feldspar, Jarosite, probably Siderite
KA10_04	Quartz	-	Kaolinite, Siderite
KA10_05	Quartz, Kaolinite, Muscovite	-	K-feldspar, Siderite, Gypsum
LI10_01B	Kaolinite, Quartz, Muscovite	Anatase	Jarosite, K-feldspar, Haematite, probably Siderite, Smectite(?)
LI10_04	Kaolinite, Quartz, Muscovite	Anatase	Jarosite, K-feldspar, Haematite, probably Siderite, Smectite(?), Pyroxene(?)
LI10_06	Kaolinite, Quartz, Muscovite, K-feldspar	-	Pyroxene, Chlorite, Goethite, probably Plagioclase (Albite) and Dolomite
LI10_08	Kaolinite, Quartz, Muscovite	Anatase	K-feldspar(?), Plagioclase, Jarosite, Chlorite(?), probably Gypsum
LI10_09	Kaolinite, Quartz, Muscovite	Anatase	Jarosite, Haematite, Gypsum, probably Chlorite
LI10_12	Quartz, Kaolinite	Muscovite, Anatase	Gypsum, Jarosite, probably Edisonite
LO10_01	Quartz, Kaolinite	Anatase	Gypsum, Pyroxene, Edisonite, traces of Muscovite and Jarosite
LO10_02A	Kaolinite, Quartz	Gypsum, Anatase, Edisonite, Jarosite	Muscovite, Pyroxene, Smectite, K-feldspar
LO10_02B	Kaolinite, Quartz	Muscovite	Edisonite, Gypsum, K-feldspar, Crandallite group mineral(?)
LO10_02C	Calcite	-	Gypsum, Pyrite, Quartz
LO10_03	Goethite	Jarosite	Gypsum, Quartz, Kaolinite
LO10_04	Kaolinite, Quartz	Anatase	K-feldspar, Jarosite, Gypsum, probably Ilmenite
LO10_05A	Jarosite, Goethite	-	Quartz, K-feldspar, probably Kaolinite
LO10_05B	Jarosite, Goethite	-	K-feldspar
LO10_08	Quartz, Kaolinite	Anatase, K-feldspar, Muscovite	Probably Edisonite, Pyroxene, Smectite, probably Hematite
MIX_DR_01	Quartz, Kaolinite	Jarosite, Goethite, Muscovite, Anatase	-
MIX_DR_02	Siderite, Quartz, Kaolinite	Anatase, Muscovite, Pyrite, Gypsum	Jarosite, Smectite group mineral
MIX_DR_03	Quartz, Kaolinite	Siderite, Anatase	-
MIX_KA_01	Kaolinite, Quartz	Muscovite	Anatase, Siderite, Gypsum, K-feldspar
MIX_KA_04	Quartz	Kaolinite, K-feldspar, Muscovite	Siderite, Amphibole
MIX_LO_01	Goethite	Quartz, Jarosite	Muscovite, probably Kaolinite
PVS10_01	Jarosite, Quartz, Kaolinite	Gypsum, Muscovite	K-feldspar
PVS10_02	Siderite, Kaolinite, Quartz	Anatase, Goethite	Haematite, Muscovite traces, Gypsum, Crandallite group mineral
PVS10_03	Anatase, Kaolinite, Quartz	-	Edisonite, Crandallite group mineral
PVS10_04	Kaolinite, Anatase, less Quartz	Goethite	Jarosite, Crandallite group mineral, probably Ilmenite
PVS10_05	Kaolinite, Quartz, Anatase, Goethite	-	Jarosite, Edisonite, Siderite, Gypsum, Crandallite group mineral, K-feldspar
PVS10_06	Quartz, Kaolinite	Anatase, Siderite, Muscovite, Goethite, probably Haematite	Crandallite group mineral, Edisonite, probably Pyrite
PVS10_07	Kaolinite, Quartz, Siderite, Anatase, Muscovite	Goethite	Gypsum, Pyrite, Edisonite, probably Hematite, Crandallite group mineral
PVS10_08	Kaolinite, Quartz, Anatase	-	Jarosite, Crandallite group mineral, Edisonite, Goethite, Gypsum(?)
PVS10_09	Kaolinite, Quartz, Anatase	Goethite	Jarosite, Crandallite group mineral, Edisonite, Goethite, Gypsum
PVS10_11	Kaolinite, Quartz, Anatase, Siderite	K-feldspar, Calcite	Crandallite group mineral, probably Analcime, Haematite
PVS10_13	Kaolinite, Quartz	Anatase, Muscovite, Goethite	Jarosite, Gypsum, Edisonite
PVS10_14	Kaolinite, Quartz, Muscovite	-	Anatase, very small amount of Jarosite, very small amount of Gypsum, Siderite
PVS10_15	Kaolinite, Quartz, Muscovite	-	Pyroxene, probably Siderite, Pyrite
PVS10_18	Kaolinite, Quartz, Muscovite	Anatase, Analcime, Smectite, Siderite, K-feldspar	Goethite, Gypsum
PVS10_20	Quartz, Plagioclase, K-feldspar, Muscovite	Kaolinite, Chlorite	Smectite, Anatase, probably Pyroxene, Dolomite, probably Siderite
VR10_01	Quartz, Anatase, Gypsum, Kaolinite	K-feldspar, Plagioclase, Mullite, probably Tridymite	Haematite, probably also Goethite

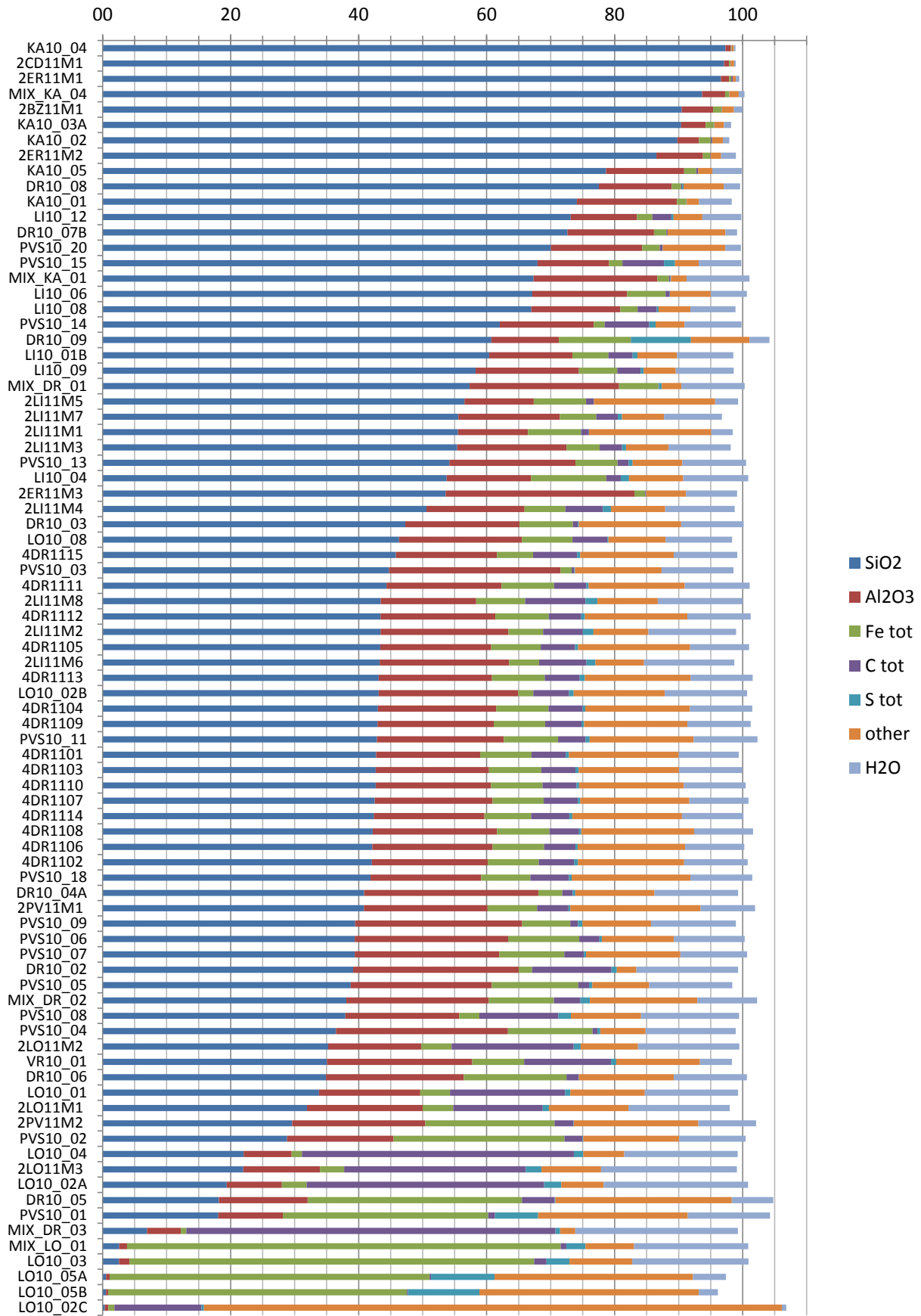


Figure B.3.: Simplified wet chemistry elemental concentrations for the 80 Sokolov samples given in wt-%. List ordered by SiO<sub>2</sub> fraction, all values given in weight percent.

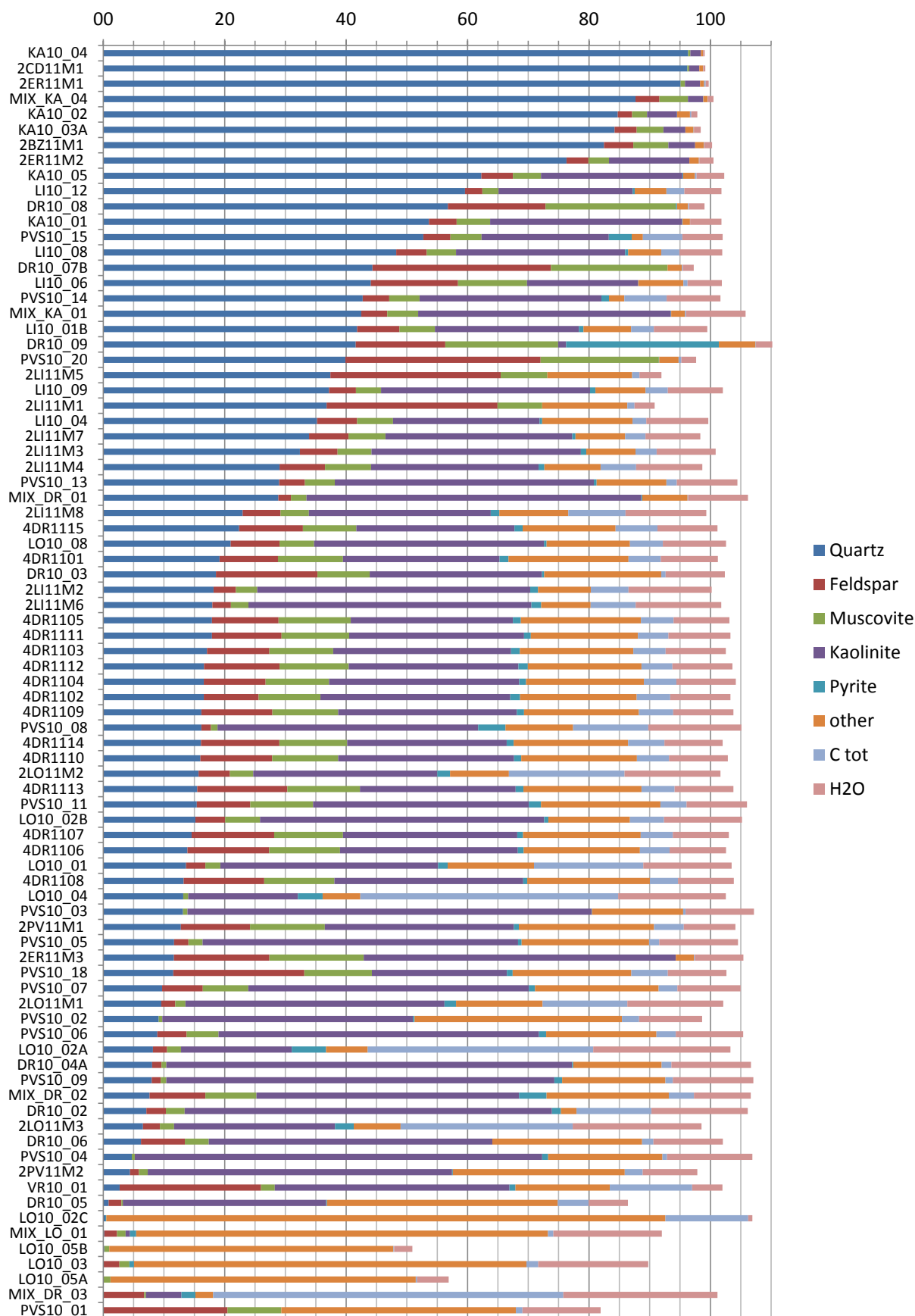


Figure B.4.: Simplified SEDNORM normative calculation result for the 80 Sokolov samples given in wt-%. List ordered by quartz fraction, all values given in weight percent.



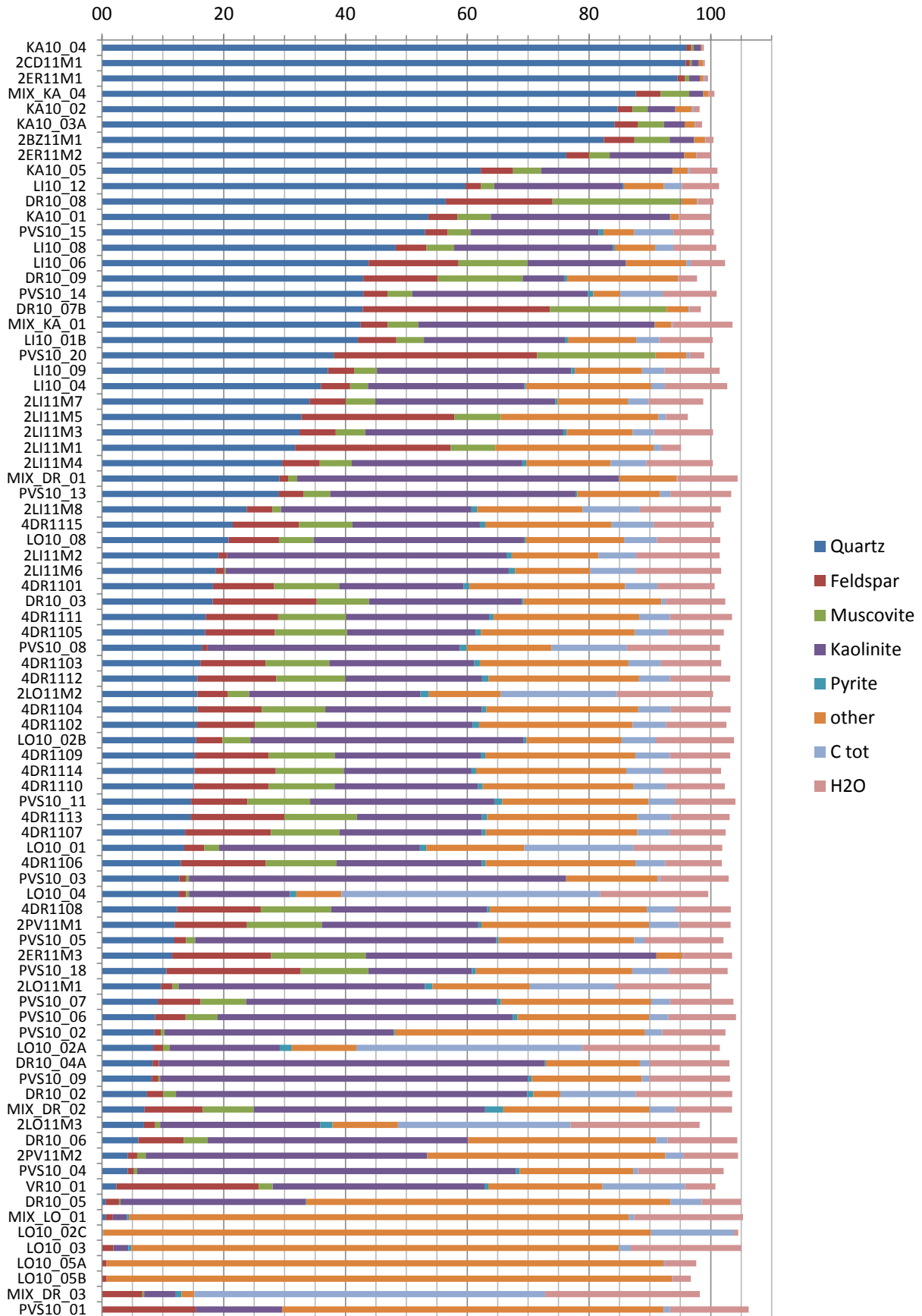


Figure B.5.: Simplified NORMA normative calculation result for the 80 Sokolov samples given in wt-%. List ordered by quartz fraction, all values given in weight percent.

## C. Sensor Descriptions

The appendix lists technical specifications of spectral sensors which data is used in the course of this study. Variable parameters specific to the prevailing setup/flight are marked with asterisk.

### C.1. Airborne Spectrometers

Given are summarised specifications for the HyMap sensor's VIS-NIR-SWIR1-SWIR2 ports and the AHS sensor's TIR port together with flight and acquisition parameters as used for this study.

property	HyMap	AHS (port 4)
spectral range	454 - 2,485 nm	8,310 - 12,952 nm
spectral bands	128	10
spectral sampling interval	13 - 17 nm	443 - 606 nm
FWHM bandwidth	15 - 20 nm	412 - 566 nm
FOV	61.3 °	90 °
IIFOV	2.0 & 2.5 mrad	2.1 & 2.5 mrad
GIFOV*	4 m	5 m
flight altitude*	2,480 m	2,438 m
date of acquisition*	21.08.2010	22.07.2011
time of acquisition*	10:12 - 11:10 UTC	21:45 - 23:30 UTC

### C.2. Field/Laboratory Spectrometers

Given portable field spectrometers have been used for reference measurements during the airborne sensor flight campaigns and in laboratory setup for the collected mineral soil samples.

property	ASD FieldSpec Pro 2	D&P FT-IR Model 102	CI Systems SR-5000
spectral range	400 - 2,500 nm	1,900 - 16,000 nm	7,533 - 12,992 nm
spectral bands	2,150	6,431	96
sampling interval	1.4 - 2.0 nm	5 - 18 nm	57 nm
FWHM bandwidth	3 - 12 nm	4 cm <sup>-1</sup>	<1.8% of wavelength
FOV*	8 ° (field) 25 ° (lab)	4.8 ° (field)	0.344 ° (lab)
GIFOV*	20 cm (field) 10 cm (lab)	10 cm (field)	3 cm (lab)
measure distance*	nadir view 1.5 m (field) 0.2 m (lab)	nadir view 1.1 m (field)	45 ° inclined view 5 m (lab)

## D. Mathematical Supplements

This appendix gives a brief summary on the Gaussian function, its derivatives and the convolution of two Gaussians. Notations of the Derivative of Gaussian (DOG) mother wavelet found in the literature are listed and the notation used in the thesis at hand is derived.

### D.1. Gaussians

The general Gaussian function is given by

$$f(t) = a \cdot e^{-\frac{(t-b)^2}{2c^2}} \quad (\text{D.1})$$

which parameters  $a$ ,  $b$  and  $c$  describe the amplitude, the peak location and the width of the bell shaped curve. The Gaussian function is related to the normal distribution

$$g(t) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad (\text{D.2})$$

with  $b$  the expectancy  $\mu$  and  $c$  the standard deviation  $\sigma$ . A scaling factor  $a$  of  $1/(\sqrt{2\pi}\sigma)$  performs a normalisation such that an integral of one is maintained.

### D.2. Derivatives of Gaussians

Suppose the general Gaussian function  $f(t)$  of amplitude  $A$ , mean  $\mu$  and standard deviation  $\sigma$ .

$$f(t) = A \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad (\text{D.3})$$

Since the exponential function  $e^t$  is identical to its first derivative applying the chain rule for differentiation results in

$$\frac{d}{dt} e^{f(t)} = f^{(1)}(t) \cdot e^{f(t)} \quad (\text{D.4})$$

which yields for first and second derivative of the Gaussian

$$\frac{d}{dt} f(t) = \frac{-(t-\mu)}{\sigma^2} \cdot f(t) \quad (\text{D.5})$$

$$\frac{d^2}{dt^2} f(t) = \frac{-\sigma^2 + (t-\mu)^2}{\sigma^4} \cdot f(t) \quad (\text{D.6})$$

[Haar & Bart \(2003\)](#) notes that the derivatives have the Gaussian function as common element. Dividing the Gaussian function into its derivative a polynomial is extracted which order is identical to the order of the derivative. This are the Hermite polynomials governing the properties of the respective derivative.

### D.3. Convolution of Gaussians

Bromiley (2003) repeats the convolution of two normalised Gaussian distributions of non-identical means and standard deviations. Here it is expanded to the case of two general Gaussian functions of arbitrary amplitude. Suppose the two Gaussian functions

$$f(t) = A_f \cdot e^{-\frac{(t-\mu_f)^2}{2\sigma_f^2}} \quad \text{and} \quad g(t) = A_g \cdot e^{-\frac{(t-\mu_g)^2}{2\sigma_g^2}} \quad (\text{D.7})$$

Using the convolution theorem above convolution problem can be conveyed into a multiplication operation in the Fourier space

$$\hat{h}^{-1}[\hat{f}(\omega) \cdot \hat{g}(\omega)] = (f * g)(t) \quad (\text{D.8})$$

with  $\hat{h}$  the Fourier transform

$$\hat{h}(\omega) = \int_{-\infty}^{\infty} h(t) \cdot e^{-2\pi i t \omega} dt \quad (\text{D.9})$$

and  $\hat{h}^{-1}$  the inverse Fourier transform. Applying Equation D.9 to  $f(t)$  and employing the substitution  $t' = t - \mu_f$  yields

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} A_f \cdot e^{-\frac{t'^2}{2\sigma_f^2}} \cdot e^{-2\pi i (t'+\mu_f)\omega} dt' \quad (\text{D.10})$$

which can be simplified according to Bromiley (2003), using Euler's formula and the fact that  $\sin(2\pi t' \omega)$  is odd and thus its integral over infinite range zeroes, to

$$\hat{f}(\omega) = A_f \cdot e^{-2\pi i \mu_f \omega} \int_{-\infty}^{\infty} e^{-\frac{t'^2}{2\sigma_f^2}} \cdot \cos(2\pi t' \omega) dt' \quad (\text{D.11})$$

Given the integral

$$\int_0^{\infty} e^{-at^2} \cdot \cos(2xt) dt = \frac{1}{2} \sqrt{\frac{\pi}{a}} \cdot e^{-\frac{x^2}{a}} \quad (\text{D.12})$$

with  $a = 1/(2\sigma_f^2)$  and  $x = \pi\omega$  the Fourier transform of  $f(t)$  can now be expressed as

$$\hat{f}(\omega) = A_f \sqrt{2\pi} \sigma_f \cdot e^{-2\pi i \mu_f \omega} \cdot e^{-2\pi^2 \sigma_f^2 \omega^2} \quad (\text{D.13})$$

The Fourier transform of  $g(t)$  is similar, hence the product of both Fourier transformed Gaussians gives

$$\hat{f}(\omega) \cdot \hat{g}(\omega) = 2\pi A_f A_g \sigma_f \sigma_g \cdot e^{-2\pi i (\mu_f + \mu_g) \omega} \cdot e^{-2\pi^2 (\sigma_f^2 + \sigma_g^2) \omega^2} \quad (\text{D.14})$$

Comparing Equation D.14 to D.13 it is obvious that the product of two Gaussians in Fourier space is another Gaussian. Its amplitude, mean and standard deviation are

$$A_{f*g} = \sqrt{2\pi} \cdot \frac{A_f A_g \sigma_f \sigma_g}{\sigma_{f*g}} \quad \text{and} \quad \mu_{f*g} = \mu_f + \mu_g \quad \text{and} \quad \sigma_{f*g} = \sqrt{\sigma_f^2 + \sigma_g^2} \quad (\text{D.15})$$

In the case of two normalised Gaussian functions with  $A = 1/(\sqrt{2\pi}\sigma)$  the above amplitude expression simplifies to  $A_{f*g} = 1/(\sqrt{2\pi}\sigma_{f*g})$ .

As the Fourier transform is an invertible operation one can deduct that the convolution of two Gaussian functions in real space is another Gaussian function with parameters given in Equation D.15, hence

$$(f * g)(t) = A_{f*g} \cdot e^{-\frac{(t-\mu_{f*g})^2}{2\sigma_{f*g}^2}} \quad (\text{D.16})$$

which expands for the generic case to

$$(f * g)(t) = \sqrt{2\pi} \cdot \frac{A_f A_g \sigma_f \sigma_g}{\sqrt{\sigma_f^2 + \sigma_g^2}} \cdot e^{-\frac{(t-(\mu_f+\mu_g))^2}{2(\sigma_f^2+\sigma_g^2)}} \quad (\text{D.17})$$

## D.4. DOG Wavelet Notations

In the literature various notations for the Derivative of Gaussian (DOG) mother wavelets exist. This appendix shows the equivalence of the notations to a general deduction. It focuses on the reference notation of the IDL wavelet toolbox found in [Torrence & Compo \(1998\)](#) and the mathematical reference found in [Mallat \(2008\)](#).

[Torrence & Compo \(1998\)](#) list several wavelet mother functions and their properties. Amongst is a generic formula for the m-th order derivative of Gaussian

$$\frac{(-1)^{m+1}}{\sqrt{\Gamma(m + \frac{1}{2})}} \cdot \frac{d^m}{dt^m} e^{-\frac{t^2}{2}} \quad (\text{D.18})$$

with  $\Gamma$  denoting the gamma function

$$\Gamma(m + \frac{1}{2}) = \frac{(2m)!}{4^m m!} \sqrt{\pi} \quad (\text{D.19})$$

[Mallat \(2008\)](#) gives a formula of the normalised Second Order Derivative of Gaussian ( $\text{DOG}_{m=2}$ ), also known as Mexican hat, Sombrero, Ricker or Marr wavelet depending on the field of application.

$$\frac{2}{\sqrt[4]{\pi} \sqrt{3}\sigma} \cdot \left( \frac{t^2}{\sigma^2} - 1 \right) \cdot e^{-\frac{t^2}{2\sigma^2}} \quad (\text{D.20})$$

The thesis at hand uses the second order derivative of Gaussian ( $\text{DOG}_{m=2}$ ) as wavelet mother function. An improved understanding of the literature sources can be achieved by comparing the wavelet definitions used within. To do so the definitions given by [Mallat \(2008\)](#) in Equation D.20 and by [Torrence & Compo \(1998\)](#) in Equation D.18 are compared to a generic second derivative of a Gaussian using the formulae of Appendix D.2 - *Derivatives of Gaussians*.

Revisiting the second derivative of a general Gaussian function given in Equation D.6

$$f^{(2)}(t) = A \cdot \frac{-\sigma^2 + (t - \mu)^2}{\sigma^4} \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad (\text{D.21})$$

a normalisation is necessary for use as mother wavelet. It ensures that the wavelet transforms are comparable between scales and between differing signals. Normalisation is achieved by choosing the amplitude  $A$  such that the wavelet maintains unit energy, i.e. its  $L^2$  norm  $\|f^{(2)}\|$  equals 1

$$\sqrt{\int_{-\infty}^{\infty} |f^{(2)}(t)|^2 dt} = 1 \quad (\text{D.22})$$

$$\int_{-\infty}^{\infty} \left| A \cdot \frac{-\sigma^2 + (t-\mu)^2}{\sigma^4} \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}} \right|^2 dt = 1 \quad (\text{D.23})$$

$$\frac{A^2}{\sigma^8} \cdot \int_{-\infty}^{\infty} \left( (t-\mu)^2 - \sigma^2 \right)^2 \cdot e^{-\frac{2(t-\mu)^2}{2\sigma^2}} dt = 1 \quad (\text{D.24})$$

$$\frac{3A^2\sqrt{\pi}}{4\sigma^3} = 1 \quad (\text{D.25})$$

The modulus operation in above deduction cancels for real-valued wavelets, e.g. the DOGs. Assuming  $\sigma > 0$  and  $A > 0$  then

$$A = \frac{2\sigma^2}{\sqrt[4]{\pi}\sqrt{3}\sigma} \quad (\text{D.26})$$

which yields a formula for the normalised general second derivative of Gaussian by inserting into Equation D.21

$$\frac{2}{\sqrt[4]{\pi}\sqrt{3}\sigma} \cdot \left( \frac{(t-\mu)^2}{\sigma^2} - 1 \right) \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad (\text{D.27})$$

Obviously Equation D.27 can be directly transformed into the definition D.20 given by Mallat (2008) when setting  $\mu = 0$ . For the definition D.18 given in Torrence & Compo (1998) the transformation is likewise straightforward. Solving D.18 for the case  $m = 2$  and inserting the relation D.19 for the gamma function yields

$$- \frac{2}{\sqrt[4]{\pi}\sqrt{3}} \cdot \frac{d^2}{dt^2} e^{-\frac{t^2}{2}} \quad (\text{D.28})$$

which can be expanded using the second derivative of the exponential function (cf. Equation D.6)

$$- \frac{2}{\sqrt[4]{\pi}\sqrt{3}} \cdot (t^2 - 1) \cdot e^{-\frac{t^2}{2}} \quad (\text{D.29})$$

This is equivalent to Equation D.27 when  $\sigma = 1$ ,  $\mu = 0$  and the negative sign is disregarded. Latter is skipped to yield an ‘‘upright’’ wavelet. This has the implication of the wavelet transform being swapped in sign compared to the mathematical sense when using the IDL wavelet toolbox.

## E. Validation Supplements

This appendix compiles the complete per-sample / per-target spectral comparisons of the pre-processed reference measurements and imagery that were previewed in the validation chapter. In addition validation related plots of the normative modelling are placed here.

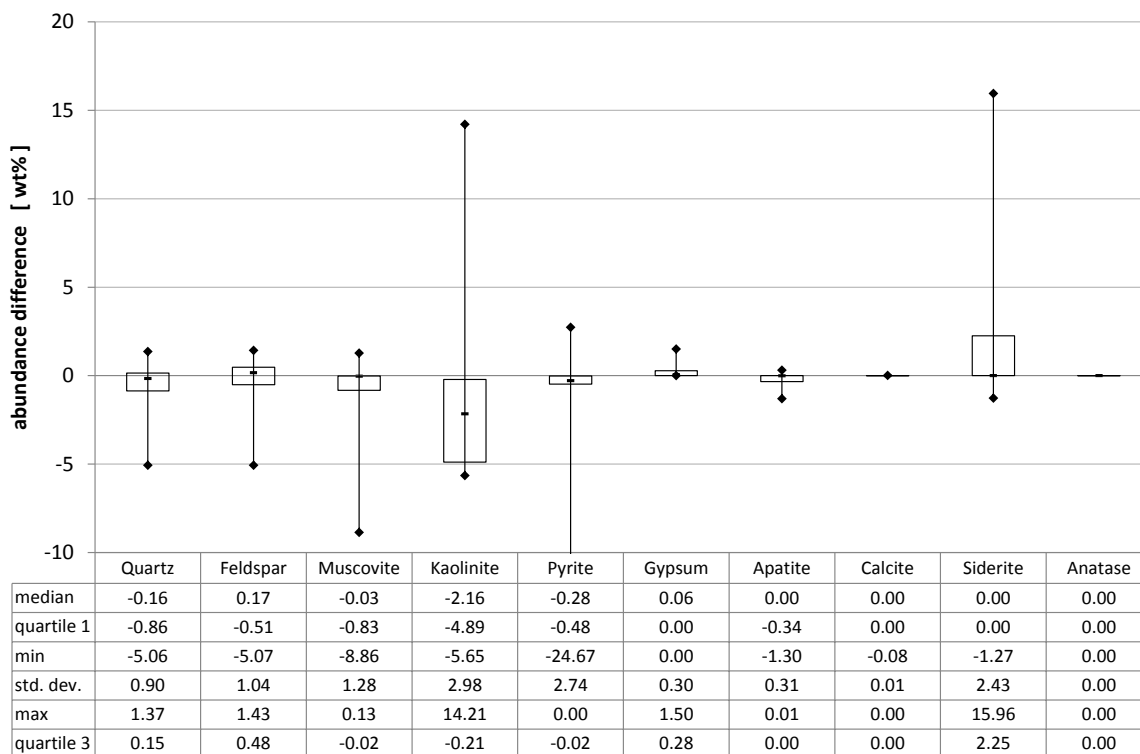


Figure E.1.: Descriptive statistics of the differences in abundance [wt-%] for minerals common to NORMA and SEDNORM.

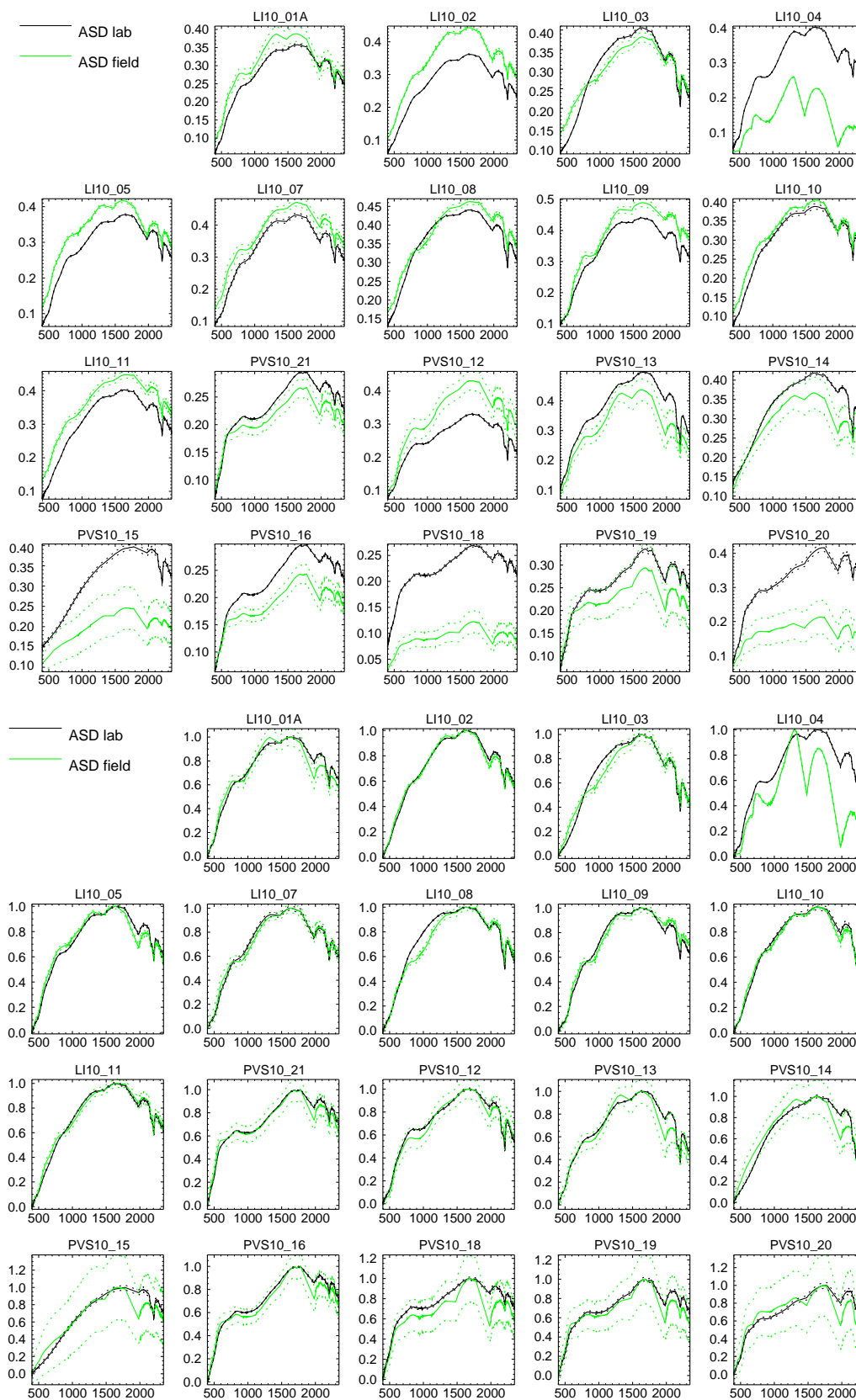


Figure E.2.: Validation of 19 corresponding field (green) and laboratory (black) reflectance spectra measured with ASD Field Spec Pro devices. Reflectance [-] (top) and normalised spectra [0...1] (bottom) over wavelength [nm] together with the standard deviations (dotted).



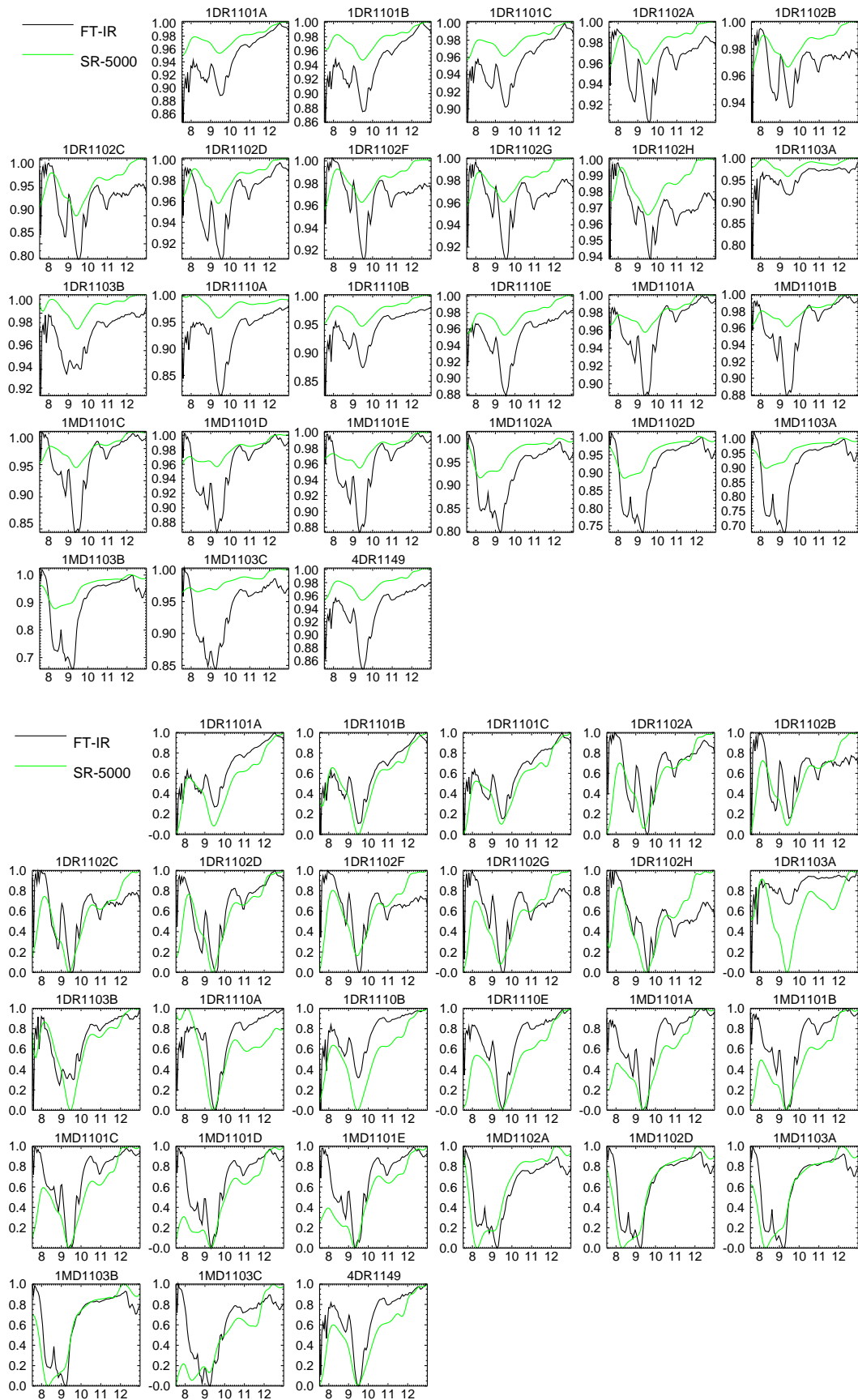


Figure E.3.: Validation of 26 corresponding field (black) and laboratory (green) emissivity spectra derived from FT-IR and SR-5000 measurements respectively. Emissivity [-] (top) and normalised spectra [0...1] (bottom) over wavelength [ $\mu\text{m}$ ]. No standard deviations shown as they were not provided for SR-5000 measurements.

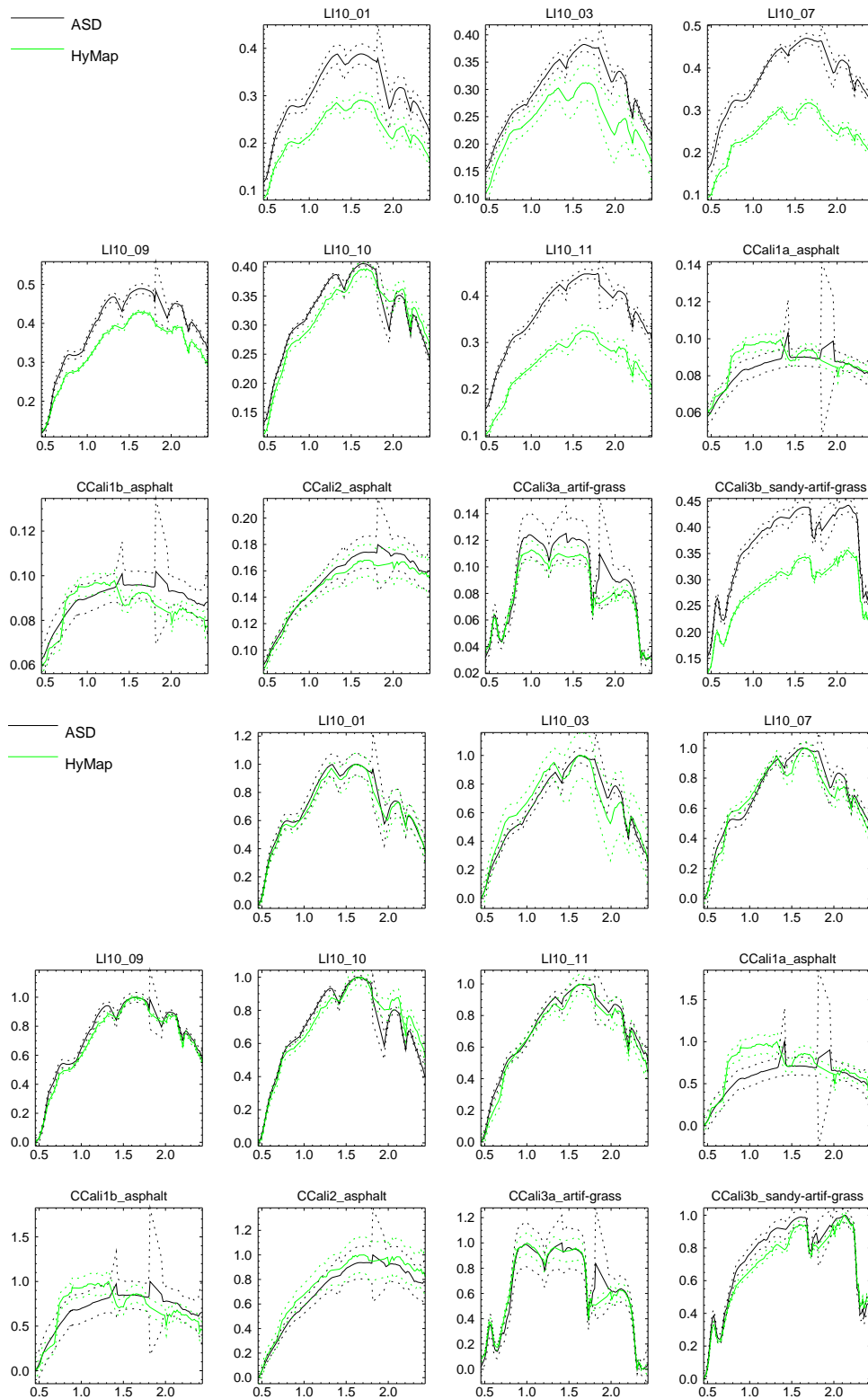


Figure E.4.: Validation of 11 corresponding field (black) and image (green) reflectance spectra derived from ASD and HyMap measurements respectively. Reflectance [-] (top) and normalised spectra [0...1] (bottom) over wavelength [ $\mu\text{m}$ ] together with the standard deviations (dotted).

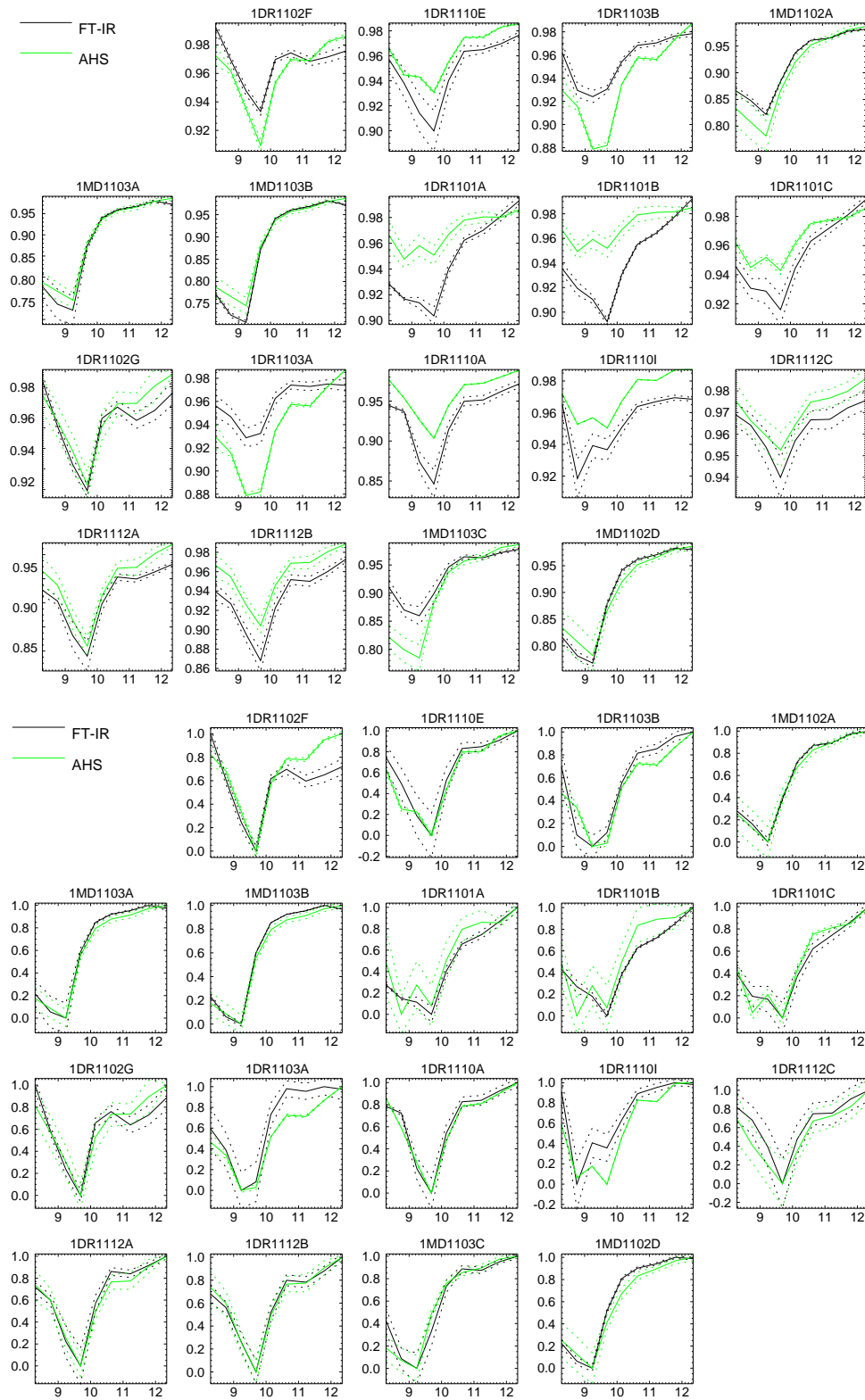


Figure E.5.: Validation of 18 corresponding field (black) and image (green) emissivity spectra derived from FT-IR and AHS measurements respectively. Emissivity [-] (top) and normalised spectra [0...1] (bottom) over wavelength [ $\mu\text{m}$ ] together with the standard deviations (dotted).

# Erklärung

Eidesstattliche Versicherung gemäß §6 Abs. 1 Ziff. 4 der Promotionsordnung des Karlsruher Instituts für Technologie für die Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften.

1. Bei der eingereichten Dissertation zu dem Thema „Scale-Wavelength Decomposition of Hyperspectral Signals: use for mineral classification & quantification in a reflective-thermal study of a lignite mining site“ handelt es sich um meine eigenständig erbrachte Leistung.
2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.
3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.
4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.
5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Ferner sind mir die Grundsätze des Karlsruher Instituts für Technologie (KIT) zur Sicherung guter wissenschaftlicher Praxis in der gültigen Fassung vom 05.05.2010 bekannt und wurden von mir beachtet.

Reichenbach, den 04.03.2014

Christoph Ehrler

# Acknowledgements

This work was conducted in the team Applied Spectroscopy of the Department of Land Surface, German Remote Sensing Data Center (DFD-LAX), German Aerospace Center (DLR). It would not have been completed without the help and support of many colleagues and friends. I would like to thank everyone who helped me along the way, in particular:

- Prof. Stefan Hinz from Karlsruhe Institute of Technology (KIT) for providing me the opportunity to conduct my dissertation under his patronage. Especially his efforts in guidance and for his encouragement whenever a problem appeared. Prof. Thomas Neumann from KIT for his absolute support in taking the task of second reviewer. Prof. Peter Reinartz and the members of the thesis committee for their interest in supporting this thesis and the scientific discussions that improved the manuscript.
- Dr. Christian Fischer for always keeping the big picture in focus, his willingness to establish first contact to his incredible network of experts and sharing his secret shelf library providing so much helpful literature. All excelled by his seemingly infinite patience in discussing scientific matters and letting me partake in his expertise. Also for finding time to always give encouraging advice and to critically review the thesis.
- Dr. Derek Rogge for his invaluable advice when it came to scientific details, especially in fields of geology and petrology. Not to forget his collection of data analysis tools saving so many days, and his patience in critically reviewing the thesis in several iterations that considerably improved the manuscript but also ate up many days.
- Dr. Eduardo de Miguel Llanes & Dr. José Antonio Gómez Sánchez and their teams at INTA for the joint work on the AHS data and the inspiring scientific exchange in Madrid. Dr. Rudolf Richter for the lessons in radiative transfer and atmospheric correction as well as granting me a peek “under the hood” of ATCOR.
- Dr. Martin Schodlok for sharing items of his spectral archive and, together with Maarit Middleton and Dr. Timo Tarvainen, for the valuable exchange on normative modelling.
- Dr. Anita Bayer who walked the “PhD path” about one year ahead for showing the light at the end of the tunnel, for the practical spectroscopy lessons and for sharing her extensive database for methodology testing.
- Grégoire Kerr for sharing his author signed wavelet book and all the math lessons learnt from him. Maybe even more important the regular supply with his French specialities and chocolate keeping my blood sugar level high and brain working.
- Also I would like to thank the whole Applied Spectroscopy team at DLR for their willingness to pass on their knowledge and to help with their expertise. Special thanks go to Dr. Martin Bachmann and Dr. Uta Heiden. In addition the EO-MINERS team for sharing data and providing helping hands during the field campaigns.
- Not to forget Tania and my family for unlimited support and constant encouragement.

# Curriculum Vitae

**Name** Dipl.-Geoinf. Christoph Ehrler  
**E-Mail** Christoph.Ehrler@outlook.com  
**Nationality** German  
**Year of Birth** 1984  
**Place of Birth** Reichenbach, Germany

## Education

---

**2010 - 2014** **German Aerospace Center in Oberpfaffenhofen, Germany**  
Ph.D. student in the “EO-MINERS” project,  
Team Applied Spectroscopy, Department of Land Surface

---

**2003 - 2009** **Technical University Bergakademie Freiberg, Germany**  
M.Sc. in geoinformatics with specialisation geomonitoring,  
Thesis “Development of JAVA Tools for RockFlow Mesh and Parameter  
File Generation tailored to Coal Seam Fire Models” at the  
German Aerospace Center, Department of Environment and Security

---

**1995 - 2003** **Secondary School Goethe Gymnasium in Reichenbach, Germany**  
Leaving certificate qualifying for university admission,  
Specialisation mathematics and physics

## Employment

---

**2009 - 2010** **German Aerospace Center in Oberpfaffenhofen, Germany**  
Research associate in the “Sino-German Coal Fire Research” project,  
Department of Environment and Security

---

**Jul. - Dec. 2007** **Joint Research Centre of the European Commission in Ispra, Italy**  
Student research project in the Institute for the Protection and Security  
of the Citizen, Nuclear Safeguards Unit

---

**Aug. - Oct. 2006** **Beak Consultants in Freiberg, Germany**  
Undergraduate assistant

## Scientific Publications

- EHRLER, C. 2009. *Development of JAVA Tools for RockFlow Mesh and Parameter File Generation tailored to Coal Seam Fire Models*. Diploma Thesis, Technische Universität Bergakademie Freiberg.
- EHRLER, C., FISCHER, C., & BACHMANN, M. 2011. Hyperspectral remote sensing applications in mining impact analysis. *In: Proceedings of 34th International Symposium on Remote Sensing of Environment*.
- FISCHER, C., EHRLER, C., & BAYER, A. 2010. Modelling Surface Energy Release to Validate ASTER Night Time Imagery. *Pages 205–213 of: Proceedings of Second International Conference on Coal Fire Research*.
- FISCHER, C., LI, J., EHRLER, C., & WU, J. 2011. Radiative energy release quantification of subsurface coal fires. *In: Proceedings of 34th International Symposium on Remote Sensing of Environment*.

## Related Unpublished Reports

- BEN-DOR, E., LIVNE, I., NOTESCO, G., ADAR, S., KOPACKOVA, V., ZIBERT, G., EHRLER, C., FISCHER, C., ROGGE, D., & KERR, G. 2013. *Thematic accuracy and validation*. Tech. rept. FP7 Project EO-MINERS, Deliverable D2.4.
- EHRLER, C. 2013. *Summary of AMD Processes Reviews*. Tech. rept. DFD-LAX, German Aerospace Center, Oberpfaffenhofen.
- EHRLER, C., & BAYER, A. 2012. *Report on Continuous Wavelet Transform Feature Analysis for Quantification of Clay Content and Iron Oxides in soil samples of the Thicket Biome, South Africa*. Tech. rept. DFD-LAX, German Aerospace Center, Oberpfaffenhofen.
- EHRLER, C., RICHTER, R., & DE MIGUEL, E. 2011. *Report on Atmospheric Modelling for DeMinTIR Imagery*. Tech. rept. DFD-LAX, German Aerospace Center, Oberpfaffenhofen.
- EHRLER, C., FISCHER, C., DE MIGUEL, E., & RICHTER, R. 2012a. *Report on In-flight Calibration and Atmospheric Correction of DeMinTIR Imagery*. Tech. rept. DFD-LAX, German Aerospace Center, Oberpfaffenhofen.
- EHRLER, C., DE MIGUEL, E., & RICHTER, R. 2012b. *Report on Inflight Calibration of MIR & TIR channels of DeMinTIR Imagery*. Tech. rept. DFD-LAX, German Aerospace Center, Oberpfaffenhofen.
- FALCK, W. E., LIVNE, I., ADAR, S., BEN-DOR, E., COETZEE, H., EHRLER, C., FISCHER, C., KERR, G., FLEMING, C., JORDAN, C. J., KOPAČKOVÁ, V., ZELENKOVÁ, K., OGILVY, R., & CHEVREL, S. 2012. *Status quo of monitoring environmental candidate indicators by EO service. Earth Observation for Monitoring and Observing Environmental and Societal Impacts of Mineral Resources Exploration and Exploitation*. Tech. rept. FP7 Project EO-MINERS, Deliverable D1.8.
- FISCHER, C., KERR, G., EHRLER, C., ROGGE, D., BEN-DOR, E., ADAR, S., LIVNE, I., NOTESCO, G., & CHEVREL, S. 2013a. *Atmospheric Correction Protocol and Validation Results*. Tech. rept. FP7 Project EO-MINERS, Deliverable D2.3.
- FISCHER, C., KERR, G., EHRLER, C., ROGGE, D., BEN-DOR, E., ADAR, S., LIVNE, I., NOTESCO, G., CHEVREL, S., BOURGUIGNON, A., & KOPACKOVA, V. 2013b. *Presentation of Algorithms and Novel Tools for the Characterisation, Mapping and Monitoring of Active Mines using combination of EO data*. Tech. rept. FP7 Project EO-MINERS, Deliverable D3.3.
- FISCHER, C., KERR, G., EHRLER, C., ROGGE, D., BEN-DOR, E., ADAR, S., LIVNE, I., NOTESCO, G., & CHEVREL, S. 2013c. *Raw and radiometric data quality indices pre and post correction*. Tech. rept. FP7 Project EO-MINERS, Deliverable D2.2.