Uncertainty Propagation of Atmospheric Correction Parameters

Nitin Bhatia

Graduation committee

Chair and Secretary
      Prof.dr.ir. A. Veldkamp   University of Twente
Promoter
      Prof.dr.ir. A. Stein   University of Twente
Co-promoters
      Dr. V.A. Tolpekin   University of Twente
      Dr. I. Reusen   VITO, Belgium
Members
      Prof.dr.ing. W. Verhoef   University of Twente
      Prof. dr. R. Zurita Milla   University of Twente
      Prof.dr. B. Somers   Catholic University of Leuven, Belgium
      Prof.dr. G.B.M. Heuvelink   Wageningen University

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

UNCERTAINTY PROPAGATION OF ATMOSPHERIC CORRECTION PARAMETERS

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
Prof.dr. T.T.M. Palstra,
on account of the decision of the graduation committee,
to be publicly defended
on Thursday 17 May, 2018 at 14.45 hrs

by

Nitin Bhatia
born on October 10, 1982
in New Delhi, India

This dissertation has been approved by:

Prof.dr.ir. A. Stein (promoter)
Dr. V.A. Tolpekin (co-promoter)
Dr. I. Reusen (co-promoter)

"Tasmai Shri Gurave Namah"

*To all teachers–especially to those
who treat the research with
particular respect, realizing that research is
both an art and science.*

*To my mother and Sivee–
I could not have done it without their
support.*

*To Shivniti, my delightful inspiration–
May you find the joy of exploring yourself as
I have in these years!*

Contents

Acknowledgements

A Ph.D to my mind is not just another degree that we earn to pursue a career but is an amalgamation of what we have learnt until now and how we are going to manifest that learning as a researcher. This reminds me of the kick-off meeting that we had for this Ph.D in Enschede, where prof.dr.ir. Alfred Stein asked me, "I had my story, Ils had her story, what is your story?" With this phrase ringing in my mind I began to write my story which was this Ph.D. Writing my 'story'was not easy. I had to spend considerable amount of time exploring myself, my strengths, my weaknesses that manifested my personality of a cheerful/sunny/nervous/confused/ disposition. And convert that into a meaningful research. Fortunately, along the way, I found great support from teachers, promoters, advisers, friends, and family. Their support did not just helped me survived but actually, enriched the journey. I remain deeply grateful for their contribution.

First and foremost, I thank my promoter prof. Stein for supporting me till the finish line of this Ph.D. One of my greatest strengths is that I have good analytical skills and can easily find out solutions to a scientific problem. At the same time, the biggest challenges was to express these ideas and solutions as a publishable and applicable output. I often encountered this weakness during my Ph.D. Yet, prof. Stein tirelessly worked with me on this weakness and helped me to become a better scientist. Often it was a bit embarrassing to get same editorial comments time and again from him but his trust and confidence allowed me to persist. The other challenge that I faced during this Ph.D was the location of my office which was at Flemish Institute for Technological Research (VITO), Belgium, far away from the university and prof. Stein's office. But the distance never deterred prof. Stein from constantly guiding me, he never hesitated to visit VITO just for me. Later, I had to move to far-far away land i.e. Australia. Yet, he vigorously replied to my emails, despite the distance and difference in time zones. This helped me to continue my research. I must admit that there were times, when I decided to quit the Ph.D, especially when the gaps between the work culture of VITO and academic expectations were increased. During those tough times, prof. Stein stood by my ideas and methods and gave me ample space to express myself to bridge those gaps. He provided me with the freedom to think and argue an idea even when we had a difference of opinion on a specific issue. I appreciate and acknowledge all his contributions of time, ideas, and funding to make my Ph.D experience productive and stimulating.

List of Figures

List of Tables

Summary

The main objective of this thesis was to quantify the propagation of uncertainty in a layered processing system as applied in Processing and Archiving Facilities (PAF). A framework is presented that is suited for the purpose. Practical uncertainty estimates are then obtained in estimating a reflectance product using three parameters: Column Water Vapour (CWV), Aerosol Optical Depth (AOD), and the adjacency range. We demonstrate the propagation of uncertainty from the reflectance product to application products, by focusing on unmixing i.e. retrieving materials and their proportional abundances present in each pixel. The thesis is divided into six chapters.

After the first introductory chapter, the second chapter introduces uncertainty and the framework.

The third chapter presents a generic method to quantify the sensitivity of reflectance to CWV concentration, AOD, and adjacency range parameters via the atmospheric correction modelling (AC). The approximate dispersion in reflectance estimates was related to the contribution of each parameter by performing a Sensitivity Analysis (SA) using a Fourier Amplitude Sensitivity Test (FAST). We studied the effects of surface albedo on Sensitivity Indices (SI) for three target surfaces in the spectral range 0.42–0.96 µm: a dark target (water), a bright target (bare soil), and a target with a low albedo in the visible and a high albedo in the near infrared range (vegetation). For AOD, high (≈ 0.9) SI values were observed at the non-water absorption wavelengths. For CWV concentration, high SI values were observed at wavelengths with strong absorption features and if the surface albedo was high. For the dark target, the effect of AOD was prominent throughout the spectral range. We found that the sensitivity of reflectance to CWV concentration and AOD is a function of the wavelength, strength of the absorption features, and surface albedo. Such information provided a greater insight into how to deal with absorption, scattering, and adjacency range type parameters.

The fourth chapter presents a generic method for a qualitative and quantitative analysis of uncertainty propagation from values of the CWV concentration and AOD to the fractional abundances derived from unmixing. Both Fully Constrained Least Squares (FCLS) and FCLS with Total Variation (FCLS-TV) were applied to estimate abundance maps. We used five simulated datasets contaminated by various noise levels. Three datasets cover two spectral scenarios with different endmembers. On those a

univariate and a bivariate analysis were carried out on CWV concentration and AOD. The other two datasets were used to analyse the effect of surface albedo. The analysis identified trends in performance degradation caused by the gradual shift in parameter values from their true value. The maximum achievable performance depends upon spectral characteristics of the datasets, noise level, and surface albedo. As expected, under noisy conditions FCLS-TV performs better than FCLS. This experiment helped in addressing various concerns pertaining to quantifying the propagation of uncertainty like identifying the best ways to report the propagation of uncertainty. Propagation of uncertainty was expressed both by measuring various quantities at pixel level and at scene level. In addition, we addressed the question on how to incorporate the effect of noise and surface albedo. We found that unmixing provided a greater insight into how to incorporate a wider range of applications to the propagation framework.

The fifth chapter presents a generic method to estimate and calibrate concentration of CWV under uncertainty. The method iteratively estimates the concentration of CWV from the pre-estimates of target surface reflectances. The method was free from assumptions, in contrast to at-sensor radiance based CWV concentration estimation methods. We considered two cases: (a) CWV concentration was incorrectly estimated in a processing chain; (b) CWV concentration was not estimated in a processing chain. To solve (a) we used the incorrect estimations as initial values to the proposed method during calibration, whereas for (b), CWV concentration was estimated without initial information. Next, we combined the two scenarios, resulting in a generic method to calibrate and estimate CWV concentration. We utilised the Hyperspectral Mapper (HyMap) and Airborne Prism EXperiment (APEX) instruments for the synthetic and real data experiments, respectively. Noise levels were added to the synthetic data to simulate real imaging conditions. For performance assessment, we compared the proposed method with two state-of-the-art methods. The developed method performed better than the two methods used for comparison. The developed method minimised the absolute error close to zero, within only 8–10 iterations. It thus suits existing PAFs where the number of iterations is an important consideration. Finally, the method is simple to implement and can be extended to address other atmospheric trace gases.

The sixth chapter presents a generic method to estimate AOD under uncertainty. AOD was estimated using the pre-estimates of surface reflectance. Assumptions concerning retrieval uncertainty and instrumental errors were less influential than for methods based upon the at-sensor radiance. Using simulated data from HyMap instruments and real data from Apex instruments, this resulted in an iterative pixelwise estimation of AOD from estimates of reflectance. Noise levels were added to the simulated data to simulate real imaging conditions. Results show that the proposed method requires 6–8 iterations. It thus suits existing PAFs where the number of iterations is an important consideration. Further, the method is free from assumptions for the at-sensor radiance based estimation methods. Finally, the method is simple to implement, it reduces the processing time in PAFs, and it can be extended to address other AC parameters.

Samenvatting

Het centrale doel van dit proefschrift was het ontwikkelen van een methode voor onzekerheidspropagatie in een gelaagde beeldverwerkingsketen zoals toegepast in Processing & Archiverings Faciliteiten (PAF). Een kader wordt gepresenteerd dat geschikt is voor dat doel. Onzekerheids schattingen worden verkregen als een reflectieproduct geschat wordt. Hierbij worden drie parameters gebruikt: de Kolom Water Damp (CWV), de Aerosol Optische Diepte (AOD) en de mate van aangrenzing (adjacency). Dit proefschift laat de voortplanting van onzekerheid zien waarbij de focus ligt op ontmrnging, d.w.z. het terugvinden van materialen en hun proportionele gehaltes die in ieder pixel voorkomen. Het proefschrift is verdeeld in zes hoofdstukken.

Na het eerste inleidende hoofdstuk beschrijft het tweede hoofdstuk het kader en de onzekerheden.

Het derde hoofdstuk presenteert een generieke methode om de gevoeligheid van reflectie voor CWV concentratie, AOT en de mate van aangrenzing te kwantficieren via atmosferische correctie modellering. De reflectie is in verschillende mate gevoelig voor de verschillende onzekerheden in AC parameters. Voor de bestudering van de impact van deze parameters op reflectanctie en applicatie producten, heb ik op basis van een gevoeligheidssanalyse prioriteiten toegekend aan de parameters. De Fourier Amplitude Sensitivity Test (FAST) is gebruikt voor de gevoeligheidsanalyse waarmee sensitiviteits indices (SI) bepaald kunnen worden. Hierbij is het effect van albedo op de parameters in het golflengtegebied 0.42 tot 0.96 µm bestudeerd voor drie oppervlakken: een donker oppervlak (water), een helder oppervlak (bodem) en een oppervlak met een laag albedo in het visuele en een hoog albedo in het nabij-infrarode golflengte gebied (vegetatie). Voor AOD zijn hoge SI waarden (≈ 0.9) geobserveerd bij de niet-water absorptie banden. Voor CWV concentratie zijn hoge SI waarden waargenimen met sterke absorptie verschijnselen en bij een hoog oppervlakte albedo. Voor het donkere oppervlak is een prominent AOD effect waargenoimen over het hele spectrale interval. De gevoeligheid van de reflectie voor CWV concentratie en AOT is daarmee een functie van de golflengte, de sterkte van de absorptie eigenschappen en het oppervlakte albedo. Zulke informatie geeft meer inzicht in hoe we om kunnen gaan met absorbtie, verstrooiing en parameters die de mate van aangrenzing beschrijven.

Het vierde hoofdstuk beschrijft een generieke methode voor een kwalitatieve en kwantitatieve analyse voor de voortplanting van onzekerejheid van de

waarden van CWV concentratie en AOD naar de fractionele gehaltes die zijn bepaald via ontmenging. Zowel kleinste kwadraten met een volledige bepering (FCLS) als FCLS met totale variatie zijn gebruikt om kaarten van de gehaltes te verkrijgen. We hebben vijf gesimuleerde gegevensbestanden gebruikt die voorzien zijn van verschillende mates van ruis. Drie bestanden hebben betrekking op twee spectrale scenarios met verschillende eindelementen. Hierop zijn een univariate en een bivariate analyse uitgevoerd op CWV concentratie en AOD. De twee andere bestanden zijn gebruikt zijn gebruikt om de effecten van het oppervlakte albedo te analyseren. De analyse toonde verschillende trends aan in de afname van het succes, veroorzaakt door de graduele verandering in de waarden van de parameters t.o.v. de ware waarden. Het hoogst haalbare prestaties zijn afhankelijk van de spectrale kenmerken van de bestanden, het ruis niveau en oppervlakte-albedo. Zoals verwacht presteert FCLS-TV presteert beter dan FCLS in situaties met veel ruis. Dit experiment heeft geholpen bij het adresseren van verschillende aandachtspunten bij het kwantificeren van de verspreiding van onzekerheid, zoals het identificeren van de beste manieren om de voortplanting van onzekerheid te rapporteren. Voortplanting van onzekerheid is enerzijds uitgedrukt door verschillende hoeveelheden op pixelniveau te meten en op sceneniveau. Anderzijds hebben we gekeken hoe het effect van ruis en oppervlakte albedo kan worden gentegreerd. We ontdekten dat onmenging een groter inzicht bood in hoe een breder scala aan toepassingen kan worden geingtegreerd binnen een voortplantings kader. Het vijfde hoofdstuk presenteert een generieke methode om de CWV concentratie te schatten en te valideren onder onzekerheid. De methode voert een iteratieve schatting uit van de CWV concentratie uit de pre-schattingen van het doeloppervlak reflecties. De methode is vrij van aannames, in tegenstelling tot de op-sensor-straling gebaseerde CWV-schattingsmethoden voor concentraties. We onderscheiden twee situaties: (a) CWV-concentratie is onjuist geschat in een verwerkingsketen; (b) CWV-concentratie is niet geschat in een verwerkingsketen. Voor het oplossen van (a) hebben we de onjuiste schattingen als initile waarden gebruikt bij de voorgestelde methode tijdens kalibratie, terwijl voor (b) de CWV-concentratie is geschat zonder initile informatie. Vervolgens hebben we de twee scenario's gecombineerd, resulterend in een generieke methode om de CWV-concentratie te kalibreren en te schatten. Voor de experimenten met de synthetische en de echte gegevens hebben we respectievelijk de Hyperspectral Mapper (HyMap) en Airborne Prism EXperiment (APEX) instrumenten gebruikt. Ruis van verschillend niveau is toegevoegd aan de synthetische gegevens om echte beeldvormingsomstandigheden te simuleren. Voor het beoordelen van de prestaties hebben we de voorgestelde methode vergeleken met twee standaard methoden. De ontwikkelde methode presteerde beter dan deze twee: de absolute fout is geminimaliseerd tot bijna nul, binnen 8 tot 10 iteraties. Het is geschikt voor de bestaande PAFs waar het aantal iteraties is een belangrijke overweging. Ten slotte is de methode eenvoudig te implementeren en kan deze worden uitgebreid in de richting van andere atmosferische sporengassen.

Het zesde hoofdstuk presenteert een generieke methode om AOD te

schattenin de aanwezigheid van onzekerheid. AOD is geschat met behulp van de pre-schattingen van het oppervlak reflectie. Aannames met betrekking tot de onzekerheid in het terugvinden en het instrument fouten zijn minder van invloed dan die op basis van de op-sensor-straling. Gebruik van gesimuleerde gegevens van HyMap-instrumenten en echte gegevens van Apex-instrumenten resulteerde in een iteratieve pixelwise schatting van AOD bij reflectie schattingen. Ruis van verschillend niveau is toegevoegd aan de synthetische gegevens om echte beeldvormingsomstandigheden te simuleren. Resultaten tonen aan dat de voorgestelde methode 6 tot 8 iteraties vereist. Het past dus binnen bestaande PAFs waar het aantal iteraties een belangrijke aandachtspunt vormt. Verder is de methode vrij van aannames voor de schattingsmethoden die gebaseerd zijn op-sensor-straling. Tenslotte is de methode eenvoudig te implementeren, het vermindert de verwerkingstijd in PAFs en kan worden uitgebreid in de richting van AC-parameters.

Introduction

1

1.1 General Introduction

Advanced hyperspectral instruments such as the Airborne Prism EXperiment (APEX) [1] are able to cover the wavelength region from 0.4 to 2.5 μm using more than 300 spectral channels at a spectral resolution of 4–10 nm (depending upon the spectral region). The wealth of spectral information available from such advanced hyperspectral imaging instruments has opened new perspectives in many applications and research domains, resulting in the development of several analytical tools that aid decision making. These apply to developing policies related to various applications such as social and economic, agriculture, and ecological activities. For instance, various methods and techniques have been developed for urban mapping and vegetation stress monitoring, envisaged under the large umbrella of on going Environmental Mapping and Analysis Program (EnMAP), a German hyperspectral satellite mission that aims at monitoring and characterising the Earth's environment on a global scale [2].

At the same time, a dedicated processing facility or operational chain is required to transform the wealth of spectral information to some meaningful products such as reflectance products, NDVI maps, and maps depicting an atmospheric condition parameter, such as aerosol optical depth.

An operational processing chain consists of a multitude of methods and models. An underlying assumption governing the operational chain is often that processing the raw data, i.e. sensor calibration to geometric and atmospheric correction, has successfully been corrected for all the anomalies present in the raw datacube. Each pre-processed pixel vector of a reflectance datacube thus comprises only the response of an underlying surface. From the lens of uncertainty the pixelwise true values of an estimate are unavailable because of limited cognition and the limited granularity of the surface reflectance. Uncertainty as defined by ISO guides to the expression of uncertainty [3,4] doubts the validity of measurements. It is an inherent property of any estimation and cannot be eliminated completely. The ISO definition is abstract and it has been elaborated in Chapter 2, to be used in this research. Further, a key principle of uncertainty is that it propagates. For instance, during atmospheric correction (AC), uncertainty from the AC parameters propagates to a reflectance product and from there to an application using reflectance as input. Quantifying the uncertainty offers confidence bounds for data analysis

and decision making, information that a user for example would rely on to analyse the data.

The missing uncertainty information in hyperspectral remote sensing products is a major scientific gap and well recognised by many international initiatives. In Europe, as identified during the Seventh Framework Program (FP7), Hyperspectral Remote Sensing in Europe [5], the lack of quality information associated with hyperspectral imagery hampers the use of hyperspectral data in Europe. The Committee on Earth Observation Satellites (CEOS) and the Quality Assessment Framework for Earth Observation (QA4EO) state that "*Quality Indicators (QIs) should be attributed to data and, in particular, to delivered information products, at each stage of the data processing chain from collection and processing to delivery. A QI should provide sufficient information to allow all users to readily evaluate a product's suitability for their particular application, i.e. its fitness for purpose*" [6].

Further, acknowledging and addressing the uncertainty is both useful to users and to the operational processing facilities. Image processing modules can be improved using knowledge about the sources of uncertainty. For example, in atmospheric correction, knowing the sources of uncertainty and their contribution, methods can be improved to obtain products of a higher quality.

1.2 Problem Statement

In the presence of the Earth's atmosphere the apparent reflectance differs from the target reflectance. This is primarily caused by the complex interaction of the surface reflected radiation with the atmospheric constituents while propagating along the path from the target surface to the sensor and by atmospheric scattered radiance path radiance, entering the Instantaneous Field Of View (IFOV) of the sensor. The interaction generates two main atmospheric effects: absorption by atmospheric gases, in particular water vapour and ozone, aerosols in the visible and near infrared spectral range and scattering by aerosols and molecules.

In radiative transfer based AC, used in this research, target surface reflectance is estimated assuming a plane parallel geometry of the atmosphere, whereas the viewing and illumination geometry and total optical depth of the atmosphere are known. For a reliable estimate of reflectance, the concentration of the atmospheric condition parameters, scatterers and absorbers, should be available at the time of imaging. Both absorbers, such as water vapour and scatterer i.e. aerosols and molecules, are highly varying in space and time. Thus, they are often estimated directly from satellite or airborne (remote) observations. For an area observed in several flight lines, multiple datacubes from each flight line are available because each flight observes the area with different illumination and viewing geometry. Applying an image based method on the observed datacubes results into multiple realisations of a parameter. Further, additional processing to mitigate say noise effect by averaging the data, say over 10×10 pixels, requires a significantly larger data volume for analysis. In addition,

the uncertainty induced by such processing is often ignored or impossible to model. Further, most estimation methods rely upon surface characteristics. For instance, the MODIS science team [7] has developed the dense dark vegetation (DDV) method to estimate AOD that is further developed in [8]. The suitability of such methods is limited to pixels with dense vegetation. For scenes with DDV pixels that are clustered at a few locations, pixelwise estimation of parameter is challenging. These limitations, reasonable as they are, cause uncertainty in the estimation of the parameters using image based methods, which likely propagates to reflectance estimates.

Acknowledging the uncertainty is a first step. The next basic question is how to quantify the propagation of uncertainty in an operational chain. A common practise is to quantify the propagation of uncertainty with a set of relevant parameters, as can be seen in the works of [9–18] that successfully used an uncertainty analysis. [11] showed the effect of the azimuthal look up table assumptions on the pre-processing of images, [12] showed the effect of sensor noise, aerosol and other pre-processing steps contributing to the uncertainty. [19] investigated how DEM uncertainty propagates to the geometrically corrected product and [16] focused on incorporating uncertainty in a decision tree. Finally, [13] focused on handling uncertainties in image mining for remote sensing studies.

An operational chain which deals with multitude of applications, requires a framework as a support structure providing a basis for implementation of uncertainty quantification methods, allowing interactions of the involved parameters, and help in analysing and presenting the propagation of uncertainty. In addition, an operational chain can not readily adapt to the pace with which new methods or models emerge. It is because the dependency structure of various applications on existing methods or models may be demolished as new methods or models do not necessarily developed as a replacement of the existing methods or models.

The Guide to the expression of Uncertainty in Measurement (GUM) [4], is a key document describing a set of frameworks to express and quantify the propagation of uncertainty. The frameworks described there, however, are abstract and provide a broad view on acknowledging and quantifying the uncertainty. For an operational chain, a detailed framework concerned with remote sensing and its applications are required. For instance, it should express how different surface types, categorised as dark and bright surfaces, affect the propagation of uncertainty. [18] provides a Data Uncertainty Engine (DUE), which gives insight into a requirement of having a framework for uncertainty assessment. Currently spatial and temporal patterns of uncertainty (autocorrelation), as well as cross correlations between related inputs, can be incorporated into an uncertainty analysis. Alongside expert judgement, sample data may be used to help estimate uncertainties, and to reduce the uncertainty of the simulated output by ensuring that each realisation reproduces the sample data. The DUE serves as an external support where a limited uncertainty assessment can be performed within a Processing and Archiving Facilities (PAF). For an operational chain, however, a framework integrated with an operational chain is required. The integration ensures that the framework is scalable to incorporate various

scene scenarios, noise effects, and sources of uncertainty, whereas it allows optimisation of parameters under uncertainty.

The focus of this thesis was to develop a framework facilitating the quantification of uncertainty as an integral part of a PAF. The main objective was to develop a methodology for uncertainty propagation in a layered processing system. This should enable the assessment of the uncertainty from an application point of view.

1.3 Scope of this work: the opted operational chain and the parameters

This research is performed using the processing and archiving facility of the Flemish Institute for the Technological Research (VITO) i.e. Central Data Processing Center (CDPC) [20]. The airborne data processing facility of the CDPC estimates hemispherical directional reflectance factor (HDRF) product. The HDRF product is a suitable option as the IFOV of airborne sensors is usually very small, for instance, 0.129° for the Hyperspectral Mapper (HyMap) airborne hyperspectral scanner [21].

Within the CDPC, the MODerate resolution atmospheric TRANsmission version 4.1 (MODTRAN 4) [22] radiative transfer code is used to determine the sources of radiations and the MODTRAN 4 interrogation technique is used for applying the atmospheric correction. In radiative transfer modelling, the surface reflectance can be estimated assuming a plane parallel geometry of the atmosphere, whereas the viewing and illumination geometry and the total optical depth of the atmosphere are known. For a reliable estimate of reflectance, the concentration of the atmospheric scatterers and absorbers should be available at the time of imaging.

The viewing and illumination geometry is an outcome of the orthorectification process. In the CDPC, for the estimation of the atmospheric scatterers and absorbers, key parameters are the visibility and amount of column water vapour. To estimate these parameters, estimation of visibility and column water vapour were integrated [8, 23].

Hyperspectral datacubes are geometrically corrected using accurate and precise Differential Global Positioning System (DGPS)/Inertial Measurement Unit (IMU) measurements. With the current camera systems, and the current quality of Digital Elevation Model (DEM) products, the uncertainty introduced by the geometric correction is limited and has no severe impact on further processing for flat terrain [19].

The impact of the uncertainty of atmospheric condition parameters on the reflectance product is not well known. This concerns especially uncertainty related to the adjacency effect, amount of water vapour, and aerosol optical depth. We quantify the propagation of uncertainty that originates from these three atmospheric condition parameters to the reflectance product via atmospheric correction modelling. Further, to demonstrate the propagation of uncertainty from the reflectance product to application products, we used unmixing as an application where the focus is on retrieving fractional

4

abundances of materials present in each pixel. As unmixing requires a high quality of spectral information, this application is well suited to study the impact of uncertainty in the three atmospheric condition parameters.

When using atmospheric correction modelling, we generally have hundreds of atmospheric condition parameters representing the imaging conditions. We have selected three parameters: column water vapour, aerosol optical depth, and adjacency range. The goal was to develop a generic framework that quantifies the propagation of uncertainty in an operational processing chain. To develop such a framework, this thesis focused on different types of parameters instead of having multiple parameters of the same type.

A basic hypothesis is that an analysis with, say, column water vapour renders an insight to deal other atmospheric absorption parameters. For instance, an absorption parameter distorts a spectrum at absorption features where strength of the absorption feature, a function of wavelength, is important. To quantify the propagation of uncertainty requires a probability distribution for the parameter. Because of wavelength wise variation in the strength of the absorption feature some strong absorption features are used to obtain realistic probabilistic information. For simplicity, an absorption feature with a low number of channels is better than a channel with a large number of channels. Such information, which can be extracted by analysing one absorption type parameter, is useful to build a framework that quantifies the propagation of uncertainty from an absorption parameter to reflectance. For instance, the method which can quantify uncertainty due to column water vapour is also applicable for ozone. For an airborne sensor, however, not all trace gases are important.

Uncertainty analysis informs users about the quality of the products. Two ways of uncertainty propagation are distinguished: Taylor Series (TS) and Monte Carlo simulations. Radiative transfer based AC, includes factors like the internal sources of variation and relations between parameters that are too complex to fully explore. Thus, TS simulations are unsuitable. Monte Carlo simulations treat a model as black box, ans is thus better suited to quantify uncertainty propagation for radiative transfer based AC. Further, Monte Carlo simulations allow an explicit probability distribution to all input quantities based upon information concerning these quantities. In this way the impact of larger variations of AC parameters can be studied.

1.4 Research Objectives

The main objective of the thesis was to develop a methodology for uncertainty propagation originating from atmospheric condition parameters to reflectance and unmixing products in a layered processing system.

To achieve the main objective the thesis is divided into four sub-objectives as below:

1. To quantify the importance of the absorption effect due to water vapour concentration and the importance of the scattering effect due to aerosols in terms of aerosol optical thickness in estimating reflectance via the atmospheric correction modelling (AC).

2. To provide a qualitative and quantitative analysis of uncertainty propagation from values of the AC parameters to the fractional abundances derived from unmixing.

3. To estimate and calibrate column water vapour under uncertainty.

4. To estimate aerosol optical depth under uncertainty.

This study was performed using the Central Data Processing Center (CDPC) at VITO and thus to assess the propagation of uncertainty under real operational conditions.

1.5 Thesis Structure

The dissertation comprises seven chapters including four technical chapters, which have been published or have been submitted to peer reviewed ISI journals. Chapter 1 gives an introduction and the scope of the dissertation. It provides the relevance of the research, the problem statement, research objectives, and thesis structure.

Chapter 2 presents a theoretical background of this research. Several overlapping components are included in this chapter. In addition, datasets used in this research are presented in this research.

Chapter 3 presents a sensitivity analysis to quantify the importance of the absorption effect due to water vapour concentration and the importance of the scattering effect due to aerosols in terms of aerosol optical thickness in estimating reflectance via Atmospheric Correction (AC).

Chapter 4 illustrates the propagation of uncertainty from column water vapour (CWV) and aerosol optical depth (AOD) towards abundance maps obtained by means of spectral unmixing. Both Fully Constrained Least Squares (FCLS) and FCLS with Total Variation (FCLS-TV) are applied. We use five simulated datasets contaminated by various noise levels.

Chapter 5 presents a method to estimate CWV iteratively from target surface reflectances. The method is free from assumptions for at-sensor radiance based CWV estimation methods. We consider two cases: (a) CWV is incorrectly estimated in a processing chain; (b) CWV is not estimated in a processing chain. To solve (a) we use the incorrect estimations as initial values to the proposed method during calibration. To solve (b), CWV is estimated without initial information. Next, we combined the two scenarios, resulting in a generic method to calibrate and estimate CWV.

Chapter 6 explores the use of pre-estimates of surface reflectance where assumptions pertaining to surface, sensor noise, and instrument characterisation are less influential than for methods based upon at-sensor radiance. Using simulated and real data, we present an iterative pixelwise estimation of AOD from estimates of reflectance.

Chapter 7 synthesises the results produced during this research, a discussion on their specific relevance to the research problem. It also links the main conclusion of the research and its relevance to scientific knowledge. Finally, an outlook to new research topics is provided.

Theoretical background and datasets

2

This chapter introduces a number of key terms, their definitions, and their mathematical relations, which will be referred to in most of the subsequent chapters. Section 2.1 defines and describes the term uncertainty for the three atmospheric condition parameters. It is followed by a section on the framework i.e. Section 2.2 as the backbone of many methodological choices in this thesis. Section 2.3 describes basic atmospheric effect modelling and presents some equations for the transfer of the radiation through the Earth's atmosphere. A relationship between aerosol optical depth and visibility is explained in Section 2.4. This relationship is essential when setting the atmospheric scattering conditions due to aerosols in the radiative transfer code as used in this thesis. In Section 2.5 various types of datasets are described that have different spatial configurations and spectral characteristics of the underlying surface.

2.1 Uncertainty

The aim of this section is to present the definition of uncertainty for the three atmospheric condition parameters. In the literature, the term uncertainty is defined in a number of ways for various applications [3, 6, 9, 14–19, 24–28]. The definition of uncertainty adopted in this research is from the guide to the expression of uncertainty [3]. According to this guide, uncertainty means doubt, and thus in its broadest sense uncertainty of measurement means doubt about the validity of measurement.

2.1.1 Random error, bias, and precision

All measurements have measurement errors. An error is defined as the difference between the true value of a physical quantity and its measurement. Thus, if one has the true value, an error can be measured using individual measurement without requiring a statistical or probability model. An error is a quantity that has a particular sign and magnitude and can be removed by correction. If the true value is not known, then an error has an unknown sign and magnitude. For the unknown error, a data analyst may wish to determine the range defining the confidence bounds within which the true

value of a quantity falls. An analysis of unknown errors is broadly characterised as uncertainty analysis. It defines a range within which the actual but unknown value of the quantity lies [26].

A measurement error has two components: a systematic error and a random error. The systematic error or bias is the constant error i.e. for repeated measurement, each measurement has the same bias. Unlike error, bias is a statistical expectation of over- or under-estimation of the true value based upon a statistical model and is often estimated by the mean error. Random error or noise, is seen in repeated measurements that have variation due to various sources [27].

A key element in the quality of the data is precision, which is related to the random error and is determined using a measure of the spread of errors around the mean error, for example, the standard deviation of the error [29]. A measurement is considered as accurate when it is precise and there is no bias.

The uncertainty of a physical quantity implies that the true value of the quantity is not known and instead it lies within a confidence bounds. This definition comprises the uncertainty due to the random error. It may be the case that the confidence interval does not contain the true value. This occurs if the bias is larger than the width of the confidence interval. The bias cannot be determined unless the measurements are compared with the true value of a quantity. For atmospheric correction, the true values of the involved physical quantities are often unknown for individual pixels, even if a highly accurate device is used to measure a quantity on ground. It is because ground measurements are limited to a few pixels of an image whereas for pixelwise measurement of the physical quantity, numerous ground measurements are required. Alternatively, image based methods can be used. For image based methods also the granularity of the quantity causes uncertainty. This applies for instance, to the amount of water vapour measured over a scene which is observed by two different satellites. The sensors altitude, the spatial and spectral resolution of the sensors, and the viewing and illumination geometry can cause variation in the measurement of amount of water vapour over the area. It is impossible to determine which satellite measurement of water vapour concentration should be considered as true measure and alternative ways have to be chosen as is done in this thesis.

2.1.2 Sources of uncertainty

The source of uncertainty relevant in this thesis originates mainly from column water vapour, aerosol optical depth, and adjacency range parameters. Those are among hundreds of parameters used to define the state of atmosphere at the time of imaging. It is not feasible to measure column water vapour concentration and aerosol optical depth values for each pixel coinciding with image acquisition. As an alternative, image based methods are used. The adjacency range parameter determining the background contribution will be based upon the atmospheric scattering and on the semi variogram. An estimation of each parameter is based upon either the at-sensor radiance or the estimates of reflectance. As a source of uncertainty, processing noise,

random noise, and illumination and viewing geometry are identified.

For an area observed from several flight lines, multiple datacubes from each flight line are obtained because each flight observes the area with different illumination and viewing geometry. Applying an image based method on the observed datacubes results in multiple realizations of an atmospheric condition parameter, say aerosol optical depth. Multiple realisations of AOD result in multiple estimates of reflectance via atmospheric correction. The question is which reflectance estimate is closest to the true value and should be used as input to an application.

This example can be extended to reflectance estimates where other atmospheric condition parameters, such as column water vapour, are estimated using an image based method resulting in multiple realisations. In this context, the uncertainty of a parameter at each pixel is caused by a dispersion in the parameter's value due to different viewing and illumination geometry. This measurement is further influenced by sensor noise and processing noise.

Structure uncertainty considers the question, what atmospheric correction structure should be used. This is another important source of uncertainty. The parameter uncertainty and structure uncertainty can be grouped under model uncertainty. In this thesis, we use Monte Carlo Simulations to quantify the propagation of uncertainty, where uncertainty arising from model structure is not specifically analysed. Parameter uncertainty and uncertainty due to noise will be used to estimate uncertainty in reflectance.

Further, the definition of uncertainty is elaborated to incorporate two scenarios: scene based parametrisation and pixelwise parametrisation. Scene based parametrisation is useful to the methods that depend upon surface characteristics to estimate a parameter value. For instance, the dense dark vegetation (DDV) method to estimate visibility relies upon the presence of DDV pixels in a scene. For non-DDV pixels, visibility cannot be estimated. For a scene with a low number of DDV pixels clustered at a few locations, pixelwise estimation of visibility is determined through an interpolation method. Interpolation methods, however, can induce additional uncertainty, requiring a separate analysis. In this research, for such methods, uncertainty in a parameter is defined at the scene level. The challenge to handle scene wise uncertainty is in dealing with the spatial variation of a parameter. The spatial variation in a parameter is attributed to two components: (a) the natural spatial variation due to the physical processes and (b) the inherent noise. To handle the natural spatial variability, a sub scene covering a small area of the full scene is considered. From our experience working with the real images in the CDPC, we found that the spatial variability of a parameter within a sub scene of size 75×75 pixels or 100×100 pixels with spatial resolution of 4 m is lower than the uncertainty caused by the noise and viewing and illumination geometry. Thus, the images considered in this research are of those sizes.

For pixelwise parametrisation, the spatial variability of a parameter is not considered as relevant. The spatial variability becomes relevant for pixels when interpolation is required to estimate a parameter's value at other pixels. To handle the spatial variability for such scenarios, like scene based

parametrisation, a small sub scene is used. In addition, an interpolation method is developed that is based upon the adjacency range parameter. Because uncertainty for the adjacency range parameter is analysed here, uncertainty due to the developed interpolation method is tractable.

Scenario based approaches to define uncertainty ensure that the methodologies to quantify the propagation of uncertainty are scalable and inclusive. Under such scenarios a large number of methods employed in parameters estimation can be included.

2.1.3 The importance of uncertainty

In uncertainty propagation, within a larger framework of PAFs, we work with models and methods used for correction and calibration, parameter(s) estimation, and for various applications.

By their intrinsic nature, the outcomes of these models and methods are uncertain. An uncertainty analysis of outcomes offers confidence bounds for data analysis and decision making, information that a researcher for example would rely on to use the data. It is then often necessary to know if a model or a method is uncertain as concerns its output. Such questions are of fundamental concern in this research and have been addressed in this thesis. Further, the structure of models and methods is often considered as certain i.e. we generally have a specific estimate of parameters resulting in a specific outcome. The net implication is that the models and methods are operating in different settings and often do not acknowledge uncertainty. For this reason it is interesting to see how close models and methods are to acknowledge uncertainty. In view of this, the scope of the research is to maintain the existing models and methods. It is more important to develop a practical methodology to reason about uncertainty in existing models and methods within PAFs. In addition, the thesis develops a methodology to transform the existing methods to acknowledge uncertainty.

2.2 The importance of framework

Remote sensing products generated using methods and models within a PAF are rarely certain. This causes propagation of uncertainty within a layered processing chain where an outcome of methods and models from one level are used as input to methods and models at another level. For instance, the outcome of atmospheric correction, e.g. a reflectance product is used in unmixing algorithms for extracting materials present in a scene. An acknowledgement and quantification of uncertainty is important in such layered processing of data for measuring the quality of decisions that are informed by data. It is also important for establishing the causes of uncertainty in PAF and for directing resources towards improving data quality. In recent years, various approaches have emerged to quantify the propagation of uncertainty. From the lens of PAF, applying these approaches implies quantifying the propagation of uncertainty within a specific level of a PAF. For a PAF, it is important to quantify the propagation of uncertainty
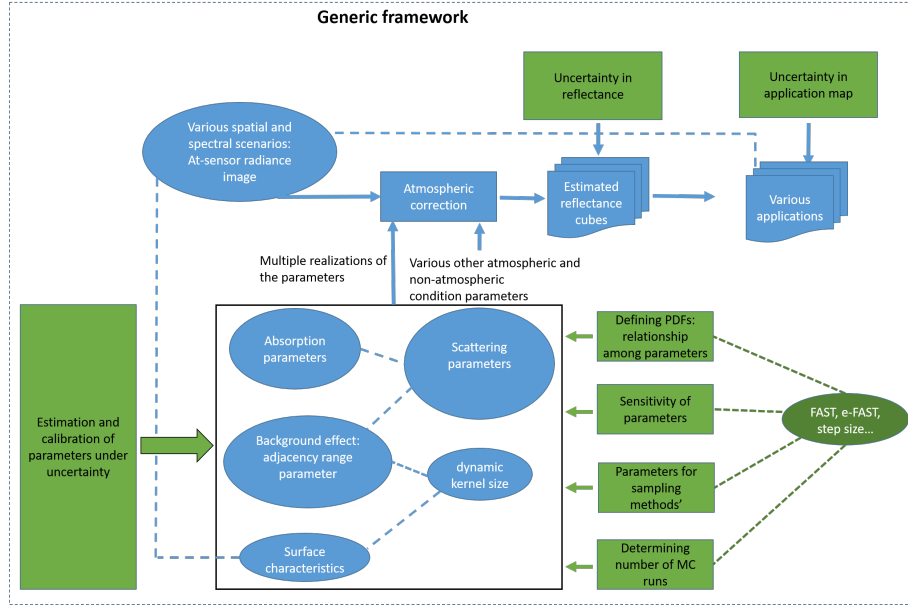
Figure 2.1: A generic form of the developed framework.

originating at a physical quantity to methods and models at various levels and to the final product. For instance, quantifying the propagation of uncertainty from the amount of column water vapour to reflectance product to abundance map of materials. This involves methods, models, and physical quantities of different types. Quantifying the propagation of uncertainty for PAF is, therefore, challenging.

These scientific gaps has motivated us to develop a framework that works as a support structure providing a basis for implementation of uncertainty quantification methods, allowing interactions of the involved parameters, and analysing and presenting the propagation of uncertainty in a PAF. The framework can be applied at a range of different levels of processing from correction and calibration of images to application levels.

Figure 2.1 illustrates a generic framework for estimating the propagation of uncertainty from atmospheric condition parameters to reflectance products and further, to application maps. The boxes in green indicate the various modules developed in the thesis. For a Monte Carlo based assessment of uncertainty, the number of runs, the probability distribution of the parameters, and sampling methods are crucial. Further, the relationships between the three atmospheric condition parameters are shown as dotted lines. The modules related to these techniques are discussed in Chapter 3.

There are various methods to represent the propagation of uncertainty at reflectance level and at abundance maps level. These methods are discussed in Chapter 4. The framework has a module to quantify the sensitivity of each parameter to the reflectance product. This module aids in prioritising the relative contribution of the parameters to the reflectance estimates. Further,

the module describing estimation and calibration of the parameters values under uncertainty is discussed in the Chapters 5 and 6.

Varying spectral characteristics and spatial configuration of datasets have been used to perform different types of experiments in the thesis. The links between the characteristic of data with the parameters and the application are shown in Figure 2.1. Details about these datasets are given in Section 2.5. Further, Figure 2.1 shows that of particular interest in this framework are the three atmospheric condition parameters, the reflectance product, and the abundance maps. This specific framework is a generic framework that provides a general and integrated view to quantify the propagation of uncertainty involving similar parameters and applications. For instance, the method developed to quantify the propagation of uncertainty from column water vapour to abundance map is also applicable for other absorption gases such as ozone gas.

2.3 Basic atmospheric effect modelling

The surface reflects a fraction $\rho_{\rm t}$ of the total irradiance at the surface $E_{\rm g}$. This fraction depends upon the type of surface, the sun zenith and azimuth angle $\theta_{\rm s}$ and $\psi_{\rm s}$ respectively, the viewing zenith and azimuth angle $\theta_{\rm v}$ and $\psi_{\rm v}$ respectively and wavelength λ. On the path of the beam to the sensor other radiation components are added to the radiance reflected by the surface ($L_{\rm t}(\lambda)$) due to atmospheric scattering. We distinguish four contributions to the at-sensor radiance ($L_{\rm rs,t}(\lambda)$):

$$L_{\rm rs,t}(\lambda) = L_{\rm t}(\lambda) + L_{\rm pa}(\lambda) + L_{\rm pb}(\lambda) + L_{\rm b}(\lambda). \tag{2.1}$$

$L_{\rm t}(\lambda)$ contains the target surface information, $L_{\rm pa}(\lambda)$ and $L_{\rm pb}(\lambda)$ are path radiance and background path radiance, respectively, that enter the IFOV of the sensor due to scattering and $L_{\rm b}(\lambda)$ is the background radiance, or adjacency effect, being the average radiance of the surrounding surface.

For a target surface with reflectance $\rho_{\rm t}(\lambda)$ and background reflectance $\rho_{\rm bck}(\lambda)$, the path radiance, the background path radiance, background radiance, and target radiance are:

$$L_{\rm pa}(\lambda) = \frac{1}{\pi} R(\theta_{\rm v}, \theta_{\rm s}, \psi_{\rm v} - \psi_{\rm s}) \cos(\theta_{\rm s}) F \tag{2.2}$$

$$L_{\rm pb}(\lambda) = \frac{1}{\pi} \rho_{\rm bck}(\lambda) T_{\rm diff}^{+}(\tau, \theta_{\rm v}, \lambda) E_{\rm g}(\lambda), \tag{2.3}$$

$$L_{\rm b}(\lambda) = \frac{1}{\pi} \rho_{\rm bck}(\lambda) T_{\rm dir}^{+}(\tau, \theta_{\rm v}, \lambda) E_{\rm g}(\lambda), \tag{2.4}$$

$$L_{\rm t}(\lambda) = \frac{1}{\pi} \rho_{\rm t}(\lambda) T_{\rm dir}^{+}(\tau, \theta_{\rm v}, \lambda) E_{\rm g}(\lambda), \tag{2.5}$$

where function $R(\theta_{\mathrm{v}}, \theta_{\mathrm{s}}, \psi_{\mathrm{v}} - \psi_{\mathrm{s}})$ describing the reflection of light by the atmosphere, T_{diff}^{+} and T_{dir}^{+} expresses upward diffuse and direct transmittances respectively, and F is the extraterrestrial solar irradiance [30]. Let the residual terms in (2.1) be denoted as:

$$L_{\mathrm{rs,b}}(\lambda) = L_{\mathrm{pa}}(\lambda) + L_{\mathrm{pb}}(\lambda) + L_{\mathrm{b}}(\lambda). \tag{2.6}$$

Then the background reflectance can be retrieved as

$$\rho_{\mathrm{bck}}(\lambda) = \frac{L_{\mathrm{rs,b}}(\lambda) - L_{\mathrm{pa}}(\lambda)}{\frac{\cos(\theta_{\mathrm{s}}) \; T_{\mathrm{tot}}^{+}(\tau, \; \theta_{\mathrm{v}}, \; \lambda) \; T_{\mathrm{tot}}^{-}(\theta_{\mathrm{s}}, \; \lambda) \; F}{\pi} + S\left[L_{\mathrm{rs,b}}(\lambda) - L_{\mathrm{pa}}(\lambda)\right]} \tag{2.7}$$

Here S is the spherical albedo for illumination from below of the atmosphere, $T_{\mathrm{tot}}^{+}(\tau, \; \theta_{\mathrm{v}}, \; \lambda)$ is the sum of $T_{\mathrm{diff}}^{+}(\tau, \theta_{\mathrm{v}}, \; \lambda)$ and $T_{\mathrm{dir}}^{+}(\tau, \theta_{\mathrm{v}}, \; \lambda)$, likewise T_{tot}^{-} expresses total downward transmittance from the sun to surface. Substituting the expression for $\rho_{\mathrm{bck}}(\lambda)$, the target reflectance equals

$$\rho_{\mathrm{t}}(\lambda) = \frac{L_{\mathrm{rs,t}}(\lambda) - L_{\mathrm{pa}}(\lambda) + \left[L_{\mathrm{rs,t}}(\lambda) - L_{\mathrm{rs,b}}(\lambda)\right] \frac{T_{\mathrm{dir}}^{+}(\tau, \; \theta_{v}, \; \lambda)}{T_{\mathrm{diff}}^{+}(\tau, \; \theta_{\mathrm{v}}, \; \lambda)}}{\frac{\cos(\theta_{\mathrm{s}}) \; T_{\mathrm{tot}}^{+}(\tau, \; \theta_{\mathrm{v}}, \; \lambda) \; T_{\mathrm{tot}}^{-}(\theta_{\mathrm{s}}, \; \lambda) \; F}{\pi} + S\left[L_{\mathrm{rs,b}}(\lambda) - L_{\mathrm{pa}}(\lambda)\right]}. \tag{2.8}$$

The basic atmospheric effect model is well described in [30–34]. We use MODerate resolution atmospheric TRANsmission version 4.1 (MODTRAN 4) [22] to estimate the radiance components in (2.8). It represents the computing of absorption and scattering in the terrestrial atmosphere at high spectral resolution [35] and is treated in this thesis as a black box. It allows one to pixelwise solve the DIScrete Ordinate Radiative Transfer (DISORT) [36] for accurate computations of atmospheric multiple scattering. In an operational processing chain, however, the considerable execution time to do so is a problem. Thus, MODTRAN 4 is executed for a uniform Lambertain surface reflectance with a spectrally flat surface albedo of $A_{\mathrm{pp}} = 0$, $A_{\mathrm{pp}} = 0.5$, and $A_{\mathrm{pp}} = 1.0$, to determine the various radiance components for a given atmospheric state and angular geometry. This is called the MODTRAN interrogation technique that has been used in operational processing chains to derive the same radiance component as in (2.8) [37]. MODTRAN 4 provides four radiance components

1. the total radiance as measured by the sensor, $L_{\mathrm{rs,t}}(\lambda)$,

2. the total path radiance $L_{\mathrm{path}}(\lambda)$ that consists of the light scattered in the path,

3. the total ground radiance that consists of all the light reflected by the surface and travelling directly towards the sensor, $L_{\mathrm{gnd}}(\lambda)$,

4. the direct ground reflectance, $L_{\mathrm{dir}}(\lambda)$ as a fraction of $L_{\mathrm{gnd}}(\lambda)$ resulting from direct illumination of the ground surface.

The four components are then linked to various radiance components in (2.8).

2.4 The relation between AOD and visibility

In an operational processing chain, true aerosol optical depth values that coincide with image acquisition are unavailable [38]. As an alternative, image
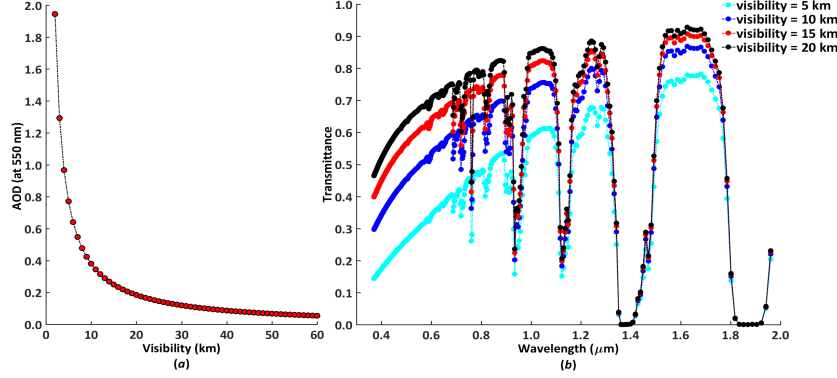
Figure 2.2: Relation between visibility and aerosol optical depth in (a) and effect of visibility on atmospheric transmittance in (b). Standard MODTRAN 4 mid latitude summer model with the rural aerosol model is used while sensor altitude is approximately 3 km and column water vapour equals 2.0 g cm^{-2} and other MODTRAN 4 parameters are at their default values.

based methods measuring visibility are often used to set aerosol optical profiles [32,39] where visibility measure in km, is the horizontal distance over which contrast transmittance at 550 nm reduces to 2%. We specify aerosol optical depth values and their corresponding visibility values so that it is useful to the remote sensing communities dealing with both quantities.

In MODTRAN 4, visibility scales the aerosol content in the atmosphere, whereas AOD specifies extinction due to aerosols at wavelength λ and is the product of the extinction coefficient $\text{Ext}(\lambda)$ and the path length. Aerosol optical depth decreases with increasing visibility. At 550 nm the contributions of molecular depth, ozone depth, and trace gases usually are small and aerosol optical depth is the main contributor to the total optical depth of the atmosphere i.e. $\text{Ext}(550)$ is directly related to aerosol optical depth. A high precision in aerosol optical depth can be achieved by subtracting the Rayleigh scattering coefficient and a very small trace gas depth from a known total optical depth. In MODTRAN 4, for $\lambda = 550$ nm and for a geographic region at mid latitude during the summer season, visibility is related to aerosol optical depth (Figure 2.2) as,

$$\text{visibility} = \frac{\ln(50)}{\text{Ext}(550) + 0.01159}, \tag{2.9}$$

where, 0.01159 km^{-1} is the surface Rayleigh scattering coefficient for $\lambda = 550$ nm. Therefore, visibility is given as a parameter for transmittance simulations for a given illumination and viewing geometry. This leads to a look up table of visibility against transmittance. The relation between visibility and aerosol optical depth, jointly with the effect of visibility on atmospheric transmittance, are illustrated in Figure 2.2.

2.5 Datasets used

Hyperspectral remote sensing and its applications, deals with multitude of scenes depicting enormous spatial configuration of materials with their distinct and similar spectral characteristics. The spatial configurations and the spectral characteristics of the underlying surface are very crucial for the analysis of the propagation of radiation through the atmosphere, which is the central idea of this Ph.D. One of the prominent characteristics of datasets required for such analysis are datasets comprising scenes that are juxtaposing spatially and spectrally to cover relevant cases. For instance, to analyse the absorption and scattering of the radiation reflected from dark and bright materials, respectively, a set of scenes comprising bright and dark materials are required. Such scenes may exist in nature but remote observation of only bright or dark targets in a scene through an airborne campaign is quite rare. Thus, we generated a few datasets with physically plausible materials to cover such important cases.

In addition, we selected some real scenes that are relevant to understand the impact of atmospheric parameters on estimates of reflectance and unmixing of materials with those estimates under real imaging conditions.

To generate and select a set of suitable datasets, we characterised scenes comprising: 1) spectrally similar and distinct targets, 2) dark and bright surfaces, 3) homogeneous and heterogeneous background, 4) structured versus more natural (continuous) class configurations, 5) inland surfaces for the image based estimations of column water vapour and visibility.

2.5.1 Simulated dataset: the geological scenario datasets

We generated two datasets for a geological scenario using two Spectral Scenarios (SS): SSa with spectrally distinct endmembers in a scene; SSb with spectrally similar endmembers in a scene. The two geological datasets are generated using the MATLAB Hyperspectral Imagery Synthesis tools, available online [40]. They contain two distinct sets of five endmembers, collected from the USGS spectral library [41]. We use the following spectrally distinct endmembers to generate the first geological dataset: Buddingtonite, Alunite, Montmorillonite, Jarosite, and Rivadavite (Figure 2.3). The spectrally similar endmembers, used to generate the second geological dataset, are signatures of a single mineral, Nontronite, to address endmember variability (Figure 2.4). All endmember spectra were resampled to the central wavelengths of the HyMap airborne hyperspectral sensor [42]. Each dataset contains 128×128 pixels on 126 spectral bands with the spatial resolution of 2.5 m along track and 2.0 m across track.

2.5.2 Simulated dataset: the vegetation scenario datasets

We generated two datasets for a scenario containing vegetation. The first dataset is a hyperspectral image cube of 75×75 with 126 spectral bands. There are five endmembers present in the scene: leather oak, dry grass, sandy loam, construction concrete, and asphalt. These endmembers were obtained
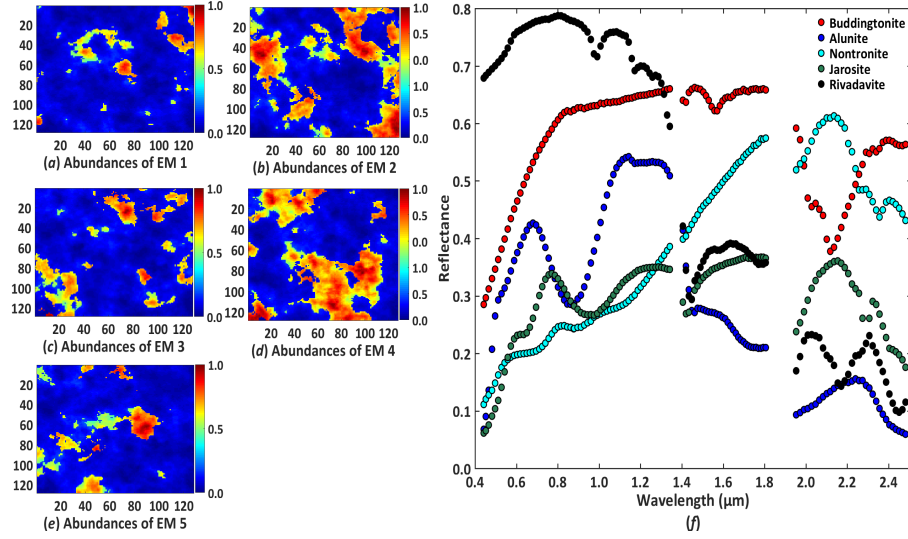
Figure 2.3: Pixelwise fractional abundances of the five (spectrally distinct) endmembers (a)–(e) and their corresponding spectral signatures (f) generating the first geological dataset.
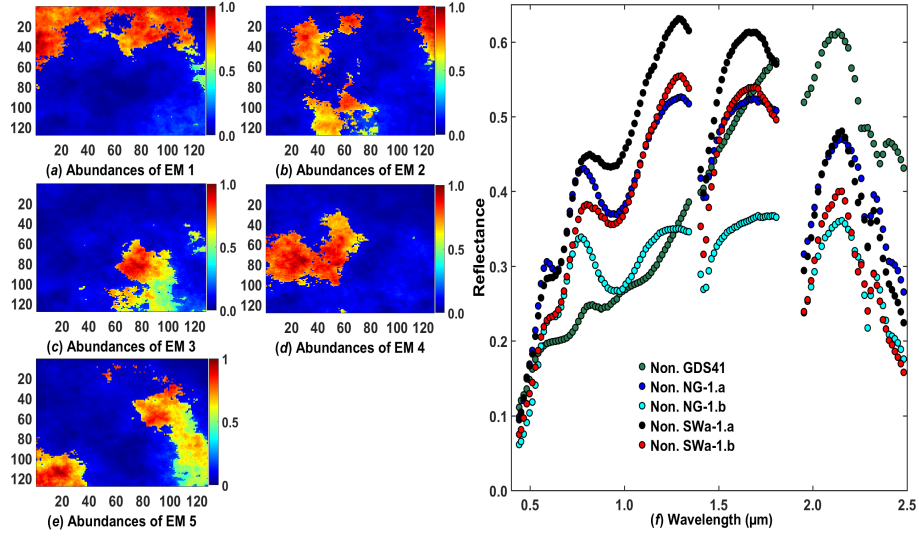
Figure 2.4: Pixelwise fractional abundances of the five (spectrally similar) Nontronite (Non.) types (a)–(e) and their corresponding spectral signatures (f) generating the second geological dataset.

as follows: dry grass and leather oak spectra from the database of Jasper Ridge, spectral library [43] and construction, sandy loam, and asphalt from the database of the Johns Hopkins University Spectral Library [44]. All endmember spectra were resampled to the central wavelengths of the HyMap
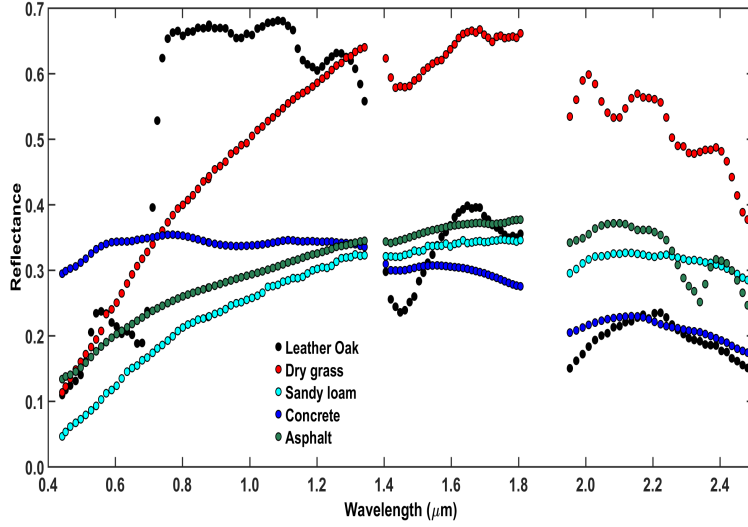
Figure 2.5: Spectral signatures of the five materials used to generate the first vegetation dataset.

airborne hyperspectral sensor (Figure 2.5). This vegetation dataset contains square regions of 5×5 pixels, which are spectrally homogeneous, and all pixels within a square region have the same reflectance values. Figure 2.6 depicts spatial arrangement of the homogeneous regions and ground truth fractional abundance maps.

The second vegetation dataset is a hyperspectral image that contains 100×100 pixels generated using nine endmembers, namely: leather oak, sandy loam, concrete, dry grass, lime stone, pine wood, red brick, terracotta tiles, and tumble weeds. These endmembers were selected as follows: dry grass and leather oak spectra were obtained from the database of Jasper Ridge, spectral library [43] available in ENVI software [45] whereas construction, sandy loam, and asphalt spectra were obtained from the database of the Johns Hopkins University Spectral Library [44]. All endmember spectra were resampled to the central wavelengths of the HyMap airborne hyperspectral sensor [42]. Hence, the data contain 126 spectral bands (Figure 2.7a). The two vegetation datasets obey the linear mixture model, which satisfies both the sum to one and the non-negativity constrains and are piecewise smooth, i.e., they are smooth with sharp transitions, as shown in Figure 2.7b. The resulting observations exhibit spatial homogeneity as described in Figure 2.8, which shows the true abundances of the endmembers.

2.5.3 Variability of surface albedo (dark and bright surfaces)

The surface albedo influences the quality of the spectra retrieved via atmospheric correction (AC), independently from the effects of the uncertainty in the AC parameters. Aerosol optical depth is particularly important in this respect, as it influences the amplitude of the spectra in a
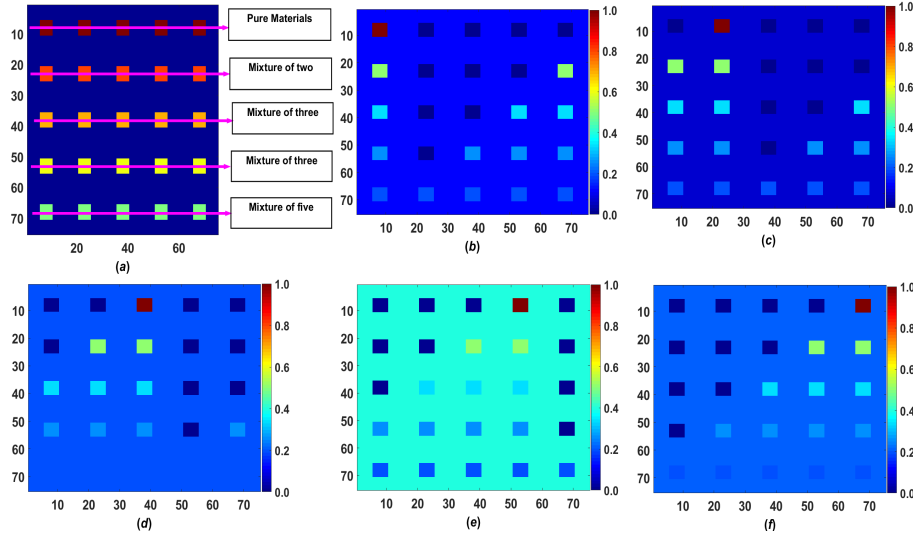
Figure 2.6: Spatial arrangement of the homogenous regions (a) and ground truth fractional abundance maps (b)–(f). The square regions at the top are pure and correspond to the five endmembers. The other square regions are mixed with the number of endmembers ranging between two and five. The remaining pixels (image background) contain a mixture of the five endmembers in which the abundance fractions were randomly fixed to 0.5130, 0.1476, 0.1158, 0.1242 and 0.0994, respectively.

large range of wavelengths. This case study analyses the effect of surface albedo on reflectance and abundance estimates with two datacubes comprising dark and bright targets separately. The bright target datacube is generated using spectra of five minerals: Albite, Ammonio Illite, Ammonio Alunite, Muscovite, and Topaz. The dark surface is generated mainly using man made materials: Black tar paper, Cinders, Construction asphalt, Reddish asphalt, and Bornoite (mineral). Each dataset contains 128 × 128 pixels with 126 spectral bands and the spatial resolution of 2.5 m along track and 2.0 m across track. The spectra of these materials are presented in Figure 2.9.

2.5.4 Addition of noise to the simulated datasets

Sensor noise and processing noise are major sources of distortion. Sensor noise refers to the random electronic noise like dark current, processing noise occurs due to the final pre-processing steps on the reflectance datacube, e.g. spectral smoothing. We added the two types of noise to the datasets at two stages: sensor noise to the radiance cube and processing noise to the estimated reflectance cube. In order to observe the effect of the different noise levels we considered three levels of the correlated processing noise, with signal to noise ratios (SNR): 30, 40, and 50 dB, respectively. All correlated noise levels were generated from independent, normally distributed noise by
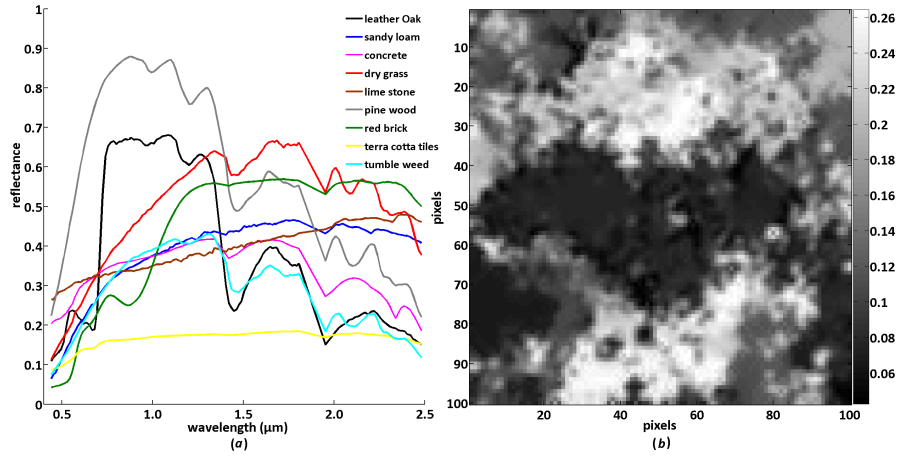
Figure 2.7: Spectral profiles of the endmembers (a) used to generate the second vegetation dataset (b).
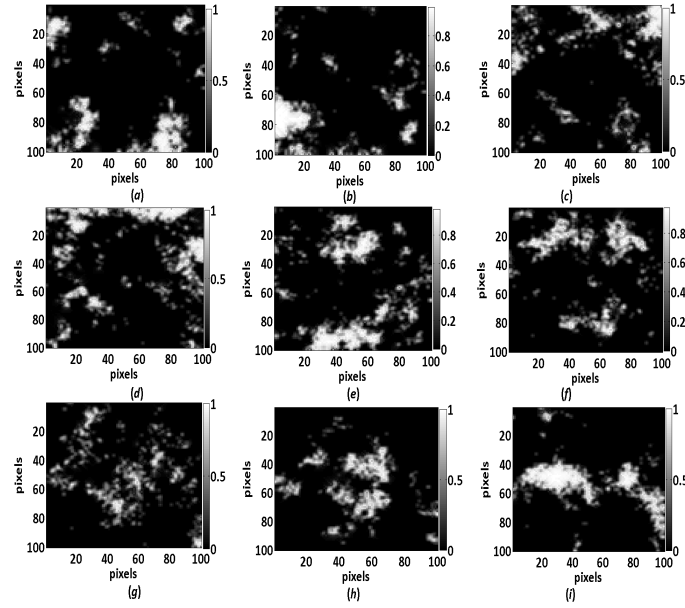
Figure 2.8: True fractional abundances of the endmembers in the second vegetation dataset.
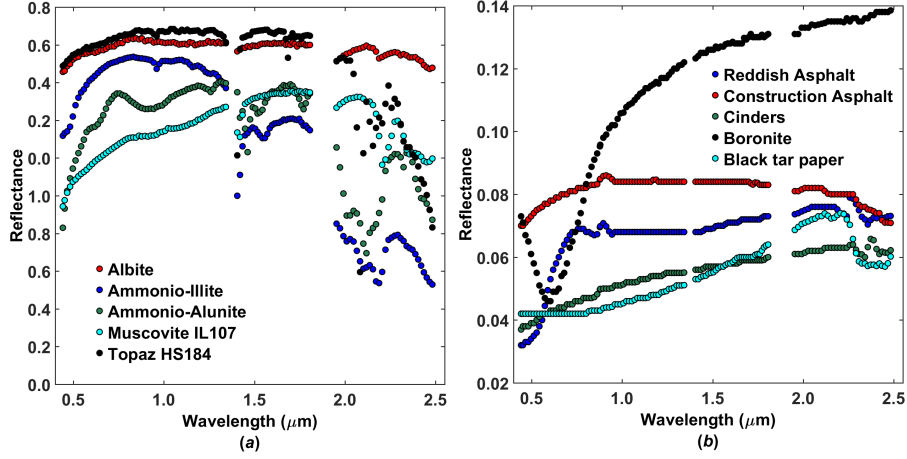
Figure 2.9: Spectra of the materials used to generate the bright surface in (a) and spectra of the materials used to generate the dark target surface in (b).
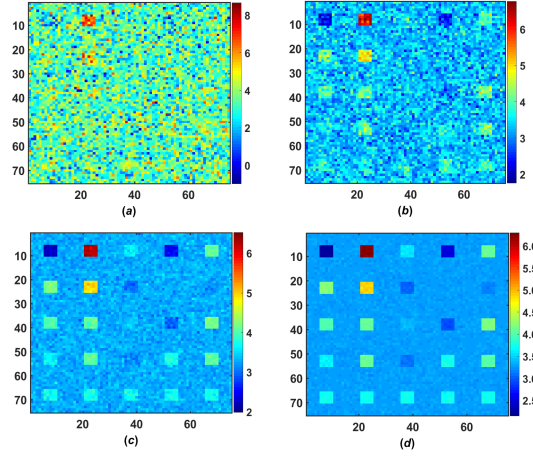
Figure 2.10: Band at (2.2 µm) of the pixelwise reflectance estimate with 30 dB (a), 40 dB (b), 50 dB (c), and 60 dB (d) noise levels. The noise levels strongly distort the reflectance image.

low pass filtering with a normalised cut off frequency of $\frac{8 \cdot \pi}{B}$ for each SNR, where B is the number of channels.

Figure 2.10 shows four estimates of band 100 (2.2 µm) of the reflectance cube when random (white) sensor noise with SNR: 30, 40, 50, and 60 dB is added to the at-sensor radiance. True AC parameters were employed and no correlated noise was added to the reflectance cube. From visual interpretation we conclude that sensor noise with SNR = 6 dB is a realistic choice as the other noise levels severely distort the images when other errors are missing.
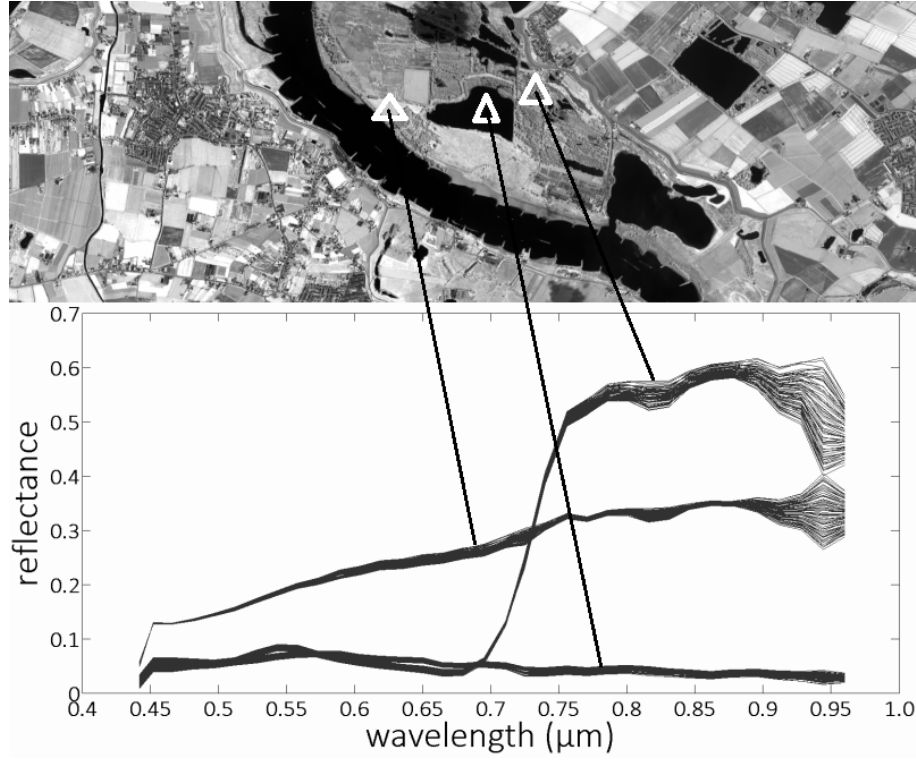
Figure 2.11: HyMap sensor image of the Millingerwaard area in the Netherlands acquired in 2004. Three target surfaces: water (dark target), bare soil (bright target), and forest (dark in visible and bring in the NIR spectral region). The mean value, from the 16 pixels (4 × 4 window), was used for each target surface.

2.5.5 Real Data

We used three real datasets in our experiments. Two datasets are used as input to AC to illustrate propagation of uncertainty. As first dataset, we used at-sensor radiance obtained from the airborne hyperspectral HyMap sensor data of the Millingerwaard area in The Netherlands acquired in 2004, shown in Figure 2.11. The spatial resolution of the image is 2.5 m along track and 2.0 m across track. The 128 spectral bands are divided into four wavelength ranges with each 32 bands (0.45–0.89 µm, 0.89–1.35 µm, 1.4–1.8 µm and 1.95–2.48 µm). The second dataset is a scene of the Airborne Prism EXperiment (APEX) sensor [1] datacube (2014) over the Liereman area (within 51.33816° N, 4.984975° E and 51.30653° N, 5.013669° E) in Belgium, shown in Figure 2.12(a). To approximate the size of the simulated image the real scene is cropped into a sub scene shown in Figure 2.12(b), which is exactly the size of the simulated image (100 × 100).

The image of the Liereman area comprises 302 spectral bands between 0.4 and 2.5 µm with a spectral resolution between 4–10 nm depending upon
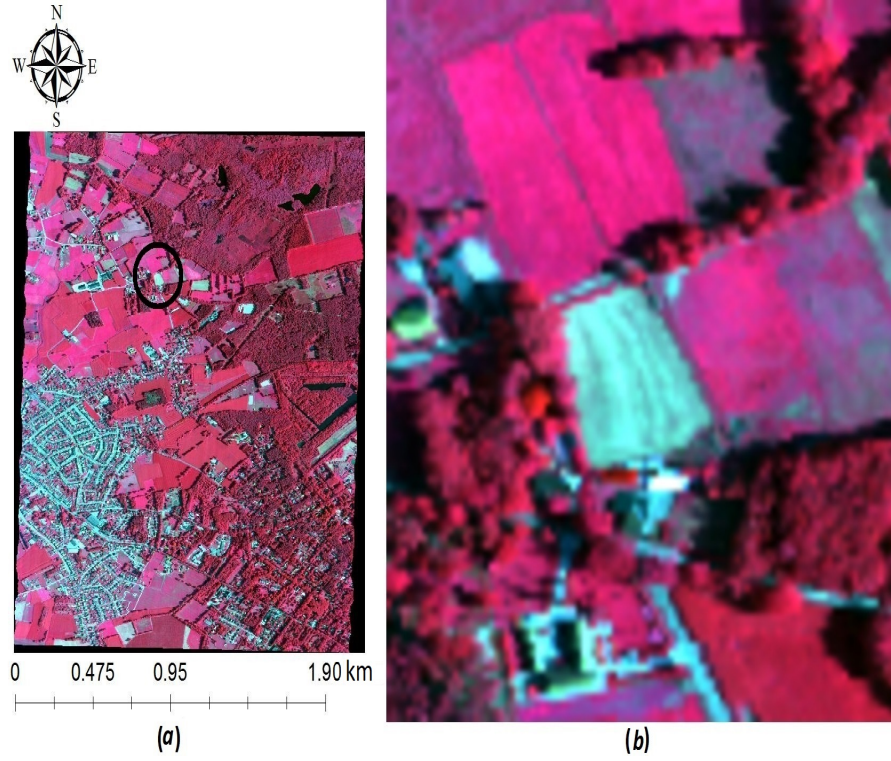
Figure 2.12: True color representation of the real APEX RGB image (a) and the sub scene used in the experiment (b) for the Liereman area, Belgium. The full scene covers 7.5 km^2 of area.

the spectral region. Prior to the analysis, bands contaminated by water absorption and bands with low SNR were removed, resulting in 227 spectral bands.

Sensitivity of reflectance to water vapour and aerosol optical thickness

3

Abstract

The atmospheric condition parameters used in the radiative transfer based atmospheric correction (AC) are often uncertain. This uncertainty propagates to the estimated reflectance. The reflectance, is, however, not equally sensitive to all the parameters. A Sensitivity Analysis (SA) helps in prioritising the parameters. The objective of this study was to perform a SA of reflectance to water vapour concentration (CWV) and Aerosol Optical Thickness (aerosol optical depth) (set by visibility). SA was performed using the Fourier Amplitude Sensitivity Test (FAST) method, which computes sensitivity indices (SI) of these parameters. Besides variation in the two parameters, we also studied the effect of surface albedo on the SI by quantifying SI for three target surfaces (in the spectral range 0.44–0.96 µm: a dark target (water), a bright target (bare soil) and a target having low albedo in the visible and high albedo in near infrared range (forest).

For aerosol optical depth, high (≈ 0.9) SI values were observed at the non-water absorption wavelengths. For CWV, high SI values were observed at wavelengths where strong absorption features are located and when the surface albedo was high. For the dark target, the effect of aerosol optical depth was prominent throughout the spectral range. We found that the sensitivity of reflectance to CWV and aerosol optical depth is a function of wavelength, strength of the absorption features, and surface albedo. We conclude that aerosol optical depth is a more important parameter for dark targets than CWV even at the principal absorption feature. For bright targets the importance of CWV and aerosol optical depth depends upon the strength of the absorption feature.

23

3. Sensitivity of reflectance to water vapour and aerosol optical thickness

3.1 Introduction

Airborne hyperspectral imaging sensors record radiance, called at-sensor radiance, in narrow contiguous spectral bands as a hypercube. In the absence of the Earth's atmosphere, the at-sensor radiance allows to perform spectral analysis of a target surface in the Instantaneous Field Of View (IFOV). In the presence of the Earth's atmosphere, however, the spectral analysis is challenging because of scattering and absorption of radiation caused by the atmospheric gases and aerosols. When radiation interacts with an atmospheric particle, electric charges in the particle start oscillating, which causes the electric charges to radiate in all directions [46]. It is this radiation which is called scattered radiation. The scattering reduces the radiant energy of radiation passing through the atmosphere and thus distorted surface reflected radiation reaches to the sensor. Also, due to scattering, the diffuse solar radiation enters into the IFOV of a sensor which is called path radiance. Furthermore, the scattering causes radiation reflected from the background of the target surface to enter into the IFOV of a sensor [47]. In the presence of the atmosphere, at-sensor radiance is composed of three radiance components: distorted surface reflected radiance, path radiance, and background radiance. Another process that distorts the at-sensor radiance is absorption. Like scattering, absorption is a function of wavelength. However, unlike scattering, absorption represents a transformation of the radiation into another form of energy [48]. The spectral ranges at which radiation is absorbed by atmospheric constituents are known as absorption features. Light in a strong absorption feature cannot penetrate the atmosphere, thus, can not reach to a sensor.

An atmospheric correction process is used to retrieve surface reflected radiation from at-sensor. Our focus is on a radiative transfer based atmospheric correction (AC). Radiative transfer based AC, first, simulates transmission of radiation through the atmosphere. These simulations are then used to estimate atmospheric correction parameters such as path and background radiances. The estimated path and background radiance are used to estimate surface reflectance from the at-sensor radiance [30]. In AC, the transmission simulations are based on a state of the atmosphere represented by the atmospheric condition parameters which define scattering and absorption. Here the challenge is to know the amount and type of the condition parameters present in the atmosphere at the time of imaging. As these parameters are variable in space and time they can not be measured locally. Here, uncertainty refers to dispersion in the values of the condition parameters. Uncertainty in the condition parameters propagates to the estimated target reflectance via the AC of a processing chain (i.e. a set of atmospheric correction processes that produce reflectance product from at-sensor radiance).

This study focuses on two atmospheric condition parameters and measures their importance in the atmospheric correction process. The two condition parameters are: scattering due to aerosols in terms of aerosol optical thickness (aerosol optical depth) that measures the degree to which aerosols impede the transmission of radiation, and absorption due to water vapour. Water vapour is an absorbing gas that effects the transmission of the radiance by

absorbing the radiation at water absorption features. Alternatively, values of these condition parameters are estimated by image based methods [8, 23]. These methods often use interpolation and mathematical assumptions to infer pixel-wise values. Therefore, the two condition parameters are uncertain. The uncertainty in aerosol optical depth and water vapour refers to the degree to which their true value and concentration are known at the time of imaging, respectively.

The objective of this study is to quantify the importance of the absorption effect due to water vapour concentration (CWV) and the importance of the scattering effect due to aerosols in terms of aerosol optical thickness (aerosol optical depth) in estimating reflectance via the AC.

In the literature, [16] focused on incorporating uncertainty at various levels of the decision tree for image classification. In [49], the estimation of uncertainties in satellite derived inherent optical properties was investigated, whereas [19] investigated the propagation of Digital Elevation Model uncertainty to the geometrically corrected product. In [11] the effect of uncertainty propagation from aerosol optical thickness, due to a simplification of azimuth angle, to the surface reflectance was studied. In [50] the dependence of aerosol optical thickness on wavelength was studied. In [51], the retrieval of aerosol optical thickness and its sensitivity to surface albedo were studied. In [52], a sensitivity analysis was performed using synthetic MERIS data to investigate the impact of atmospheric state, the target elevation or the surface roughness on aerosol optical depth, and column water vapour. To our knowledge, previous studies did not provide quantitative information on the importance of the condition parameters in the AC.

3.2 Methods

The importance of the two condition parameters was quantified using a sensitivity analysis (SA). A SA provides information on the contribution of sources of variation to dispersion in the output. A SA is used as well to investigate the importance of parameters to a model [53]. Two types of SA are commonly distinguished: a local and a global SA. A local SA perturbs each parameter with respect to its baseline value. Usually, a local SA is carried out by determining partial derivatives and evaluating them within an interval. The interval is usually small which does not cover a range of uncertainty in a parameter. Thus, this method partially covers parameter uncertainty. Further, a local SA is unfeasible to implement for a processing chain because of complexity of the processing chain, it is unfeasible to compute its partial derivatives. On the contrary to a local SA, a global SA perturbs parameters by exploring their probability distributions that cover a whole range of parameter uncertainty. A global SA allows a simultaneous perturbation of the parameters. Besides, a global SA facilitates a vast range of implementation methods and sampling strategies, like Monte Carlo Simulation (MCS) and latin hypercube simulation. Further, a global SA can be implemented without knowing the structure of a model. This makes a global SA more robust and useful for complex modelling.

For MODTRAN 4, factors, such as the internal sources of variation, relations between parameters are too complicated to fully explore. Therefore, we consider MODTRAN 4 as a black box model. For such a model, a SA methods such as a local SA or a global SA such as [54] and [55] can not be used because these methods assume that a model structure is known. In [56], ANOVA based SA techniques were discussed which assume no model structure, thus, we found them the most suitable for MODTRAN 4. Two classes of ANOVA based SA are found in the literature [53]: the Sobol method [57] and the Fourier Amplitude Sensitivity Test (FAST) method [58]. The disadvantage of the Sobol method is its computational inefficiency. It requires a large number of model evaluations that is a serious concern for the complex MODTRAN 4 model. In contrast, the FAST method is computationally more efficient, it computes the sensitivity index (SI) of parameters to a model as indicators of the parameters importance. Thus, in this research we used the FAST method to measure the importance of the two condition parameters.

The two condition parameters, column water vapour and aerosol optical depth (set by visibility (in km), refer to Section 2.4), are derived in the CDPC from image based methods [8, 23].

3.2.1 Dataset used

As an input to the atmospheric correction process we used at-sensor radiance obtained from the airborne hyperspectral HyMap sensor data of the Millingerwaard area in The Netherlands acquired in 2004, shown in Figure 2.11.

3.2.2 SA Implementation

The FAST method was originally developed within the Monte Carlo Simulation (MCS) framework. Thus, an initial step in the FAST method is to develop a joint distribution of parameters and then sample multiple (m) parameters values from the joint probability distribution. For the m parameters values m outputs are generated. A Fourier analysis is applied to a set of model output to compute the proportion of the output uncertainty contributed by each parameter. This proportion of uncertainty is defined as the sensitivity index (SI). In this study, the simulation of multiple atmospheric transmittances from m samples of CWV and visibility were used to estimate m realisations of the target reflectance $(\rho_{t,1}(\lambda), \ldots, \rho_{t,m}(\lambda))$. We have summarised the above process in Figure 3.1. We are now going to explain the methodology in detail.

Step 1: **Develop a joint distribution P(CWV, visibility)**: Here, we studied two test cases that were based on the possibility of two types of joint distribution for the condition parameters. The goal was to study the effect of the joint distribution on the FAST sensitivity index. The two cases are as follows:

- Case 1: Uniform joint distribution. We obtained ranges of CWV and visibility derived from the image based methods. Using the
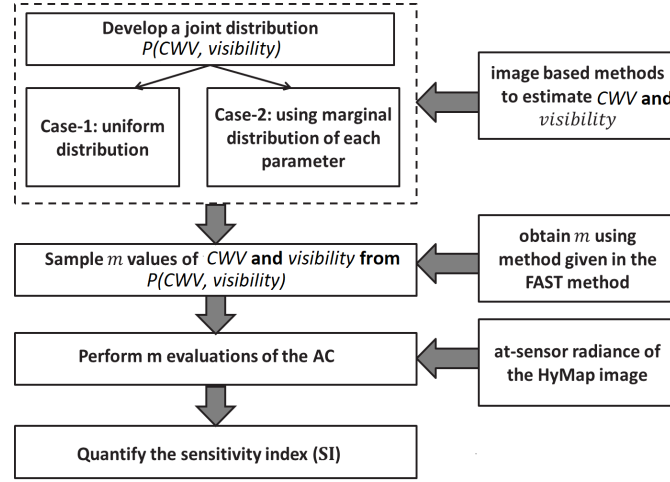
Figure 3.1: The process of quantifying the sensitivity indices is depicted in the block diagram.

ranges of the condition parameters we developed a joint uniform distribution.

- Case 2: The probability densities of CWV and visibility, and their statistical relation were used to derive their joint distribution. We developed marginal distribution of each parameter. Besides the marginal distributions, we also studied the statistical relation between the condition parameters. Aerosols and water vapour interact in the atmosphere. Many studies have explained the effect of water vapour on the size distribution and optical characteristics of the atmospheric aerosols [59, 60]. The variation in CWV, thus, affects aerosol optical depth. Therefore, we considered a statistical relation between these parameters while developing their joint distribution. For this purpose, we analysed their scatter plot, measured the correlation and the p values.

Step 2: **Sample m values of CWV and visibility from P(CWV, visbility)**: For sampling, simultaneous perturbation to the condition parameters is used. This was achieved by using the FAST search curve. The basic idea of the FAST search curve is to transform a n-dimensional parameters space into a 1-dimensional s space, such that variation in s generates a search curve that traverses P(CWV, visibility) and generates multiple CWV and visibility samples.

Here, we represent the two condition parameters as (k_l) (with $l = 1$ for CWV and $l = 2$ for visibility). For the two parameters, the equation of the search curve can be written as

$$k_l = \frac{1}{2} + \frac{1}{\pi} \arcsin(\sin \omega_l s), \tag{3.1}$$

Table 3.1: Values of Ω_N and d_h used to calculate frequencies in the FAST method are shown for $N = 1, 2, 3, 4$. The N refers to an interference factor that determine the number of Fourier coefficients retained in calculating partial variance (SI) due to k_1 and k_2

N	Ω_N		h	d_h
1	**0**		**1**	**4**
2	**3**		**2**	**8**
3	**1**		**3**	**6**
4	**5**		**4**	**10**

where ω_l is a frequency assigned to each parameter. The frequencies used in the FAST search curve were obtained following [61],

$$\begin{cases} \omega_l = \Omega_N & \text{with } N = 4 \text{ for } l = 1, \\ \omega_l = \omega_{l-1} + d_h & \text{with } h = (n+1) - l \text{ for } l = 2. \end{cases} \tag{3.2}$$

We used $N = 4$, as recommended by [61] and [62], which is an interference factor that determines the number of Fourier coefficients retained in calculating the partial variance (SI) due to k_1 and k_2. Thus, according to Table 3.1, $\omega_l = 5$ for $l = 1$ and $\omega_l = 9$ for $l = 2$.

The number of samples is a critical parameter in MCS based SA as it determines the number (m) of times a model is evaluated. The calculation of m in the FAST method is based on the Nyquest criterion and gives the relation between the maximum frequency assigned to parameters and m. Here, for reference, we have included the formula to calculate m [58],

$$m = (2 \cdot N \cdot \omega_{l_{max}}) + 1, \tag{3.3}$$

with $\omega_{l_{max}}$ the maximum of ω_l. For the FAST method,

$$\omega_{l_{max}} = 9, \text{ thus, } m = 73.$$

Thus, using the FAST search curve we obtained 73 samples of the two condition parameters, which we used to propagate the two condition parameters to the reflectance via AC.

Step 3: **Perform m evaluations of an AC for m values of CWV and visibility to obtain $\rho_{t,1}(\lambda), \ldots, \rho_{t,m}(\lambda)$.** This resulted in 73 realizations of the estimated reflectance.

To study the effect of the surface albedo and wavelength on the SI, three target surfaces were utilised. The three targets were: 1) a low reflectance, dark target (water), 2) a high reflectance, bright target (bare soil), and 3) a surface with low reflectance in the visible and high reflectance in the near infrared range (forest). The location of the pixels

used to represent the three surfaces is depicted in the Figure 2.11. We used the mean value of pixels in a 4×4 window for each of the target surfaces.

Step 4: **Quantify the sensitivity indices of $\rho_{t,1}(\lambda), \ldots, \rho_{t,m}(\lambda)$ to variation in CWV and visibility**: The multiple target reflectance were Fourier analysed using the frequencies (ω_l) assigned to k_l to calculate the Fourier coefficients. From the Fourier coefficients sensitivity indices of k_l were calculated using the method [58],

$$SI_{\omega_l} = \frac{\sigma_{\omega_l}}{\sigma}, \tag{3.4}$$

where σ is the total variance of the output indicating the dispersion in $\rho_{t,1}(\lambda), \ldots, \rho_{t,m}(\lambda)$ due to uncertainty in both the parameters k_1 and k_2, σ_{ω_l} indicates the partial variance due to parameter k_l. The indices SI_{ω_l} are the first order sensitivity indices that correspond to the main effect. They represent the contribution of CWV and visibility to $\rho_{t,1}(\lambda), \ldots, \rho_{t,m}(\lambda)$.

In this section we have presented the SA methodology that is based on the FAST method. The FAST method can compute the first order indices (SI) that corresponds to the main effect, however, it can not compute the total sensitivity index (TSI). The TSI is considered as a better measure of sensitivity of a parameter than SI as it also measures the effect of parameters interaction [53]. In [62], the extended FAST (e-FAST) technique is proposed that preserves the computational efficiency of the FAST and can quantify the TSI. Therefore, we also performed the SA using the e-FAST method and quantified the TSI to compare the FAST and e-FAST methods.

3.3 Results And Discussion

In this section, we present and discuss the results achieved for the two cases.

Step 1: **Develop the joint distribution P(CWV,visibility)**: For Case 1, we obtained the ranges of the condition parameters from the image based methods. For the HyMap image of the study area, in total, 1420 observations of visibility were obtained, however, CWV was observed at each pixel (7.87×10^5). The ranges of the two parameters observed were 1.6–2.9 g cm^{-2} for CWV and 20–120 km for visibility. The sampling region for Case 1 is the entire range of these parameters.

For Case 2, we first analysed the scatter plot and measured the correlation coefficient and p value to investigate a relation between the two parameters. Here we only used 1420 observations of CWV that corresponds to pixels for which visibility was measured. Therefore, the range of CWV shown in Figure 3.2 is different from the range used in Case 1. The scatter plot in Figure 3.2, suggested no direct evidence for any relations between the parameters. From the correlation test we found a low value (−0.0184) of the correlation between the two

parameters. From the p value test we obtained $p = 0.5144$. From the scatter plot and the statistical tests we observed no relation between the parameters. Because no statistical relation between the condition
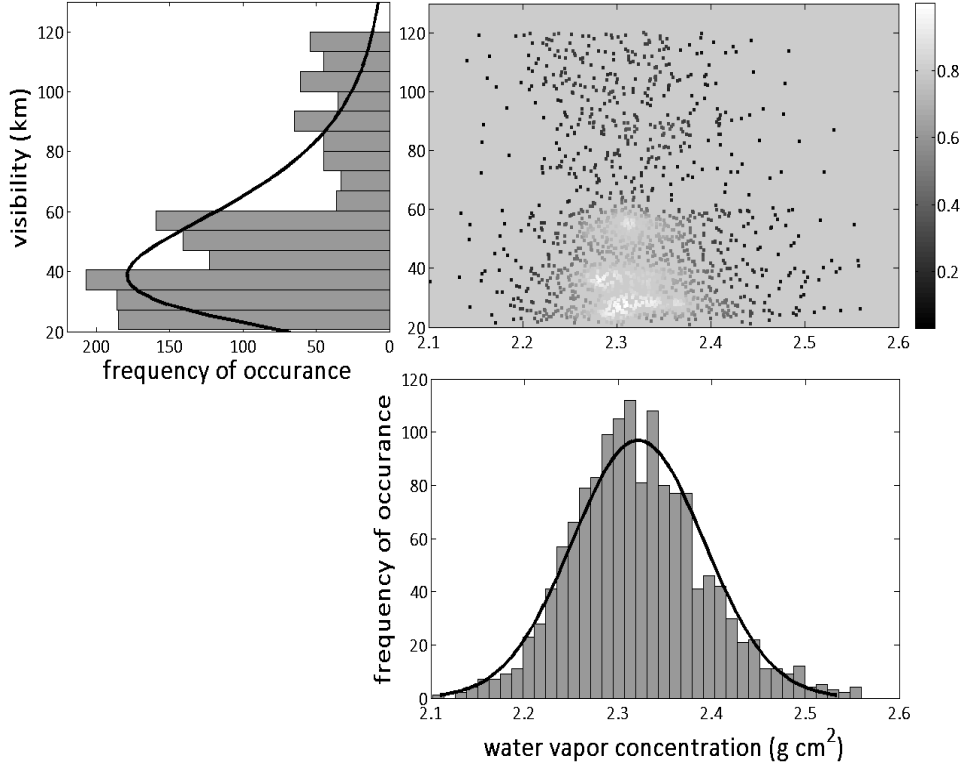
Figure 3.2: The plot depicts three informations: scatter plot between estimated visibility and the corresponding CWV values, histogram of CWV and visibility, and the sampling region as indicated by a high density region.

parameters was found, we focused on marginal distributions of the condition parameters to develop the joint distribution. For this purpose, we first examine the histogram of the two parameters. In Figure 3.2, the histograms of the two parameters are shown. The histogram of CWV (Figure 3.2 (x-axis)) indicates that a normal Gaussian distribution closely represents the marginal distribution of CWV. The histogram of visibility (Figure 3.2 (y-axis)) suggests that visibility is having a large number of occurrences far from the mean of the distribution. Statistically, this indicates a long tail distribution. Thus, we used a log-normal distribution to represent the marginal distribution of the visibility. In Figure 3.2, the sampling region for Case 2 is indicated by the high density region.

We now discuss the relevance of considering the two cases for this study. Case 1 represents a situation when one does not have information about the probability density or the statistical relation of CWV and

visibility. In this situation, one can assume a uniform distribution in a reasonable range. This range can be presumed on the basis of historic data, expert's opinion on similar areas, weather conditions etc. Here, we obtained the range of CWV and visibility from the image based methods. Whereas, Case 2 considers a situation in which one has a detailed knowledge about the probability densities and statistical properties of CWV and visibility. The knowledge about CWV and visibility is used to develop P(CWV, visibility). The two cases provide ample opportunity to investigate the significance of probability densities of CWV and visibility in quantifying their importance in the AC.

Step 2: **Sample m values of CWV and visibility from P(CWV, visibility)**: In Figure 3.2, the sampling region for Case 2 is indicated by the high density region. In this step, we discuss the performance of the search curve in terms of how effectively it can sample the sampling region of the joint distribution.

An important characteristic of the search curve is that it perturbs the parametric space simultaneously and passes through each point in the parametric space such that the path of the search curve corresponds to a joint probability density of parameters [58]. This can be achieved if the frequencies chosen are incommensurate. For implementation point of view commensurate frequencies (integer frequencies) are used instead of incommensurate frequencies. Due to the use of integer frequencies two type of errors arise: 1) the search curve is not space filling, which means it does not pass through each point in P(CWV,visibility), 2) the interference effect—the Fourier coefficients. We, however, observed that the reflectance was not sensitive to small variation in CWV and visibility. The fact that the search curve is not space filling, therefore, does not effect the SA and the density of points sampled by the search curve is appropriate for the SA. This implies that the use of integer frequencies does not affect the joint probability density of the parameters. The parameter N in (3.3) is used to avoid the interference of the Fourier coefficients.

In Figure 3.3, the scatter plot and the histogram of the sampled values of CWV and visibility are for Case 1 and Case 2. Comparing the range of the sampled values with the range of the sampling region of P(CWV,visibility) indicated with high density region in Figure 3.2 we observed that the sampling is effectively performed from the required region of the joint probability of the two parameters. Also, comparing the histogram of the measured CWV and visibility with the histogram of the sampled values of these parameters, we observed that they are similar. This shows that the FAST samples match to the two parameters probability distribution and that the FAST search curve is an effective sampler.

Step 3: **Perform 73 evaluations of an AC for CWV and visibility to obtain** $\rho_{t,1}(\lambda), \ldots, \rho_{t,m}(\lambda)$ and

Step 4: **Quantify the sensitivity indices of** $\rho_t(\lambda)$ **to variation in CWV and visibility**:
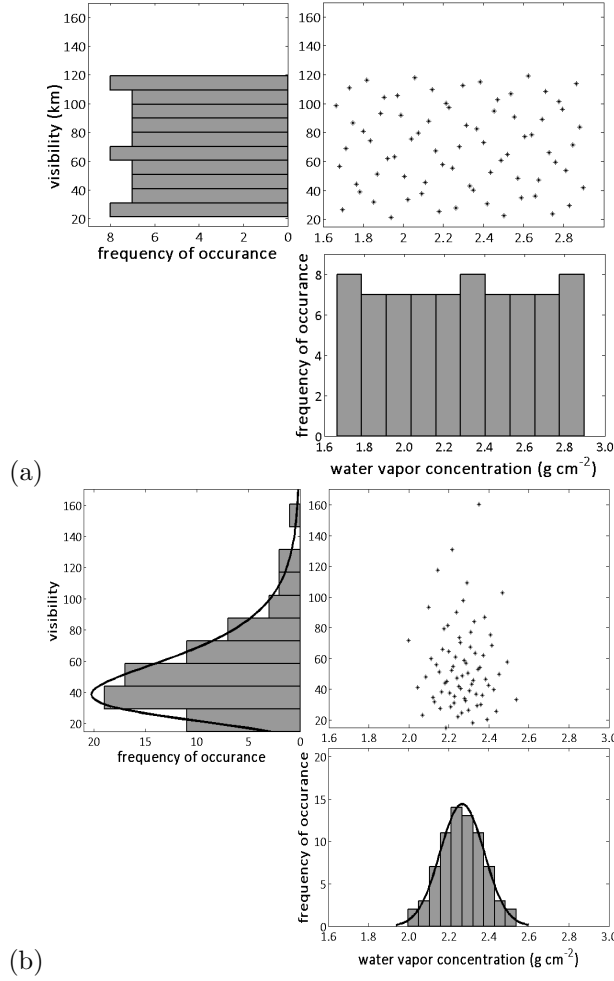
Figure 3.3: The samples obtained from the FAST search curve for the two cases: scatter and histogram plots of FAST samples for Case 1 (a) and scatter and histogram plots of FAST samples for Case 2 (b). From these samples, we observed that the sampling is effectively performed from the required region as per P(CWV, visibility). Also, comparing the histogram of the measured CWV and visibility with the histogram of the sampled values, we observed that they are similar.

The atmospheric correction is performed on the HyMap images for the spectral range 0.44–0.96 μm. The HyMap bands close to the water vapour absorption features are located at 0.58, 0.65, 0.72, 0.82, 0.94 μm. The equations given in Section 2.3 express the relation between the atmospheric state, the correction parameters, and the estimated reflectance. The outcome of 73 atmospheric correction simulations are depicted in Figure 2.11 for the three surface albedo. From these results,

we observed that the estimated reflectance at 0.94 µm showed largest variations. In the design of our experiments sources of variation other than CWV and aerosol optical depth were considered as constant, thus, the variation in the reflectance at the principal absorption feature can be attributed to column water vapour and aerosol optical depth. This variation near the principal water absorption feature (0.94 µm) is often termed as over and under estimation of reflectance. The over and under estimation occur because of uncertainty in column water vapour and aerosol optical depth. This indicates that the reflectance is sensitive to variations in column water vapour and aerosol optical depth. These results, however, are only a qualitative indicator of the sensitivity.
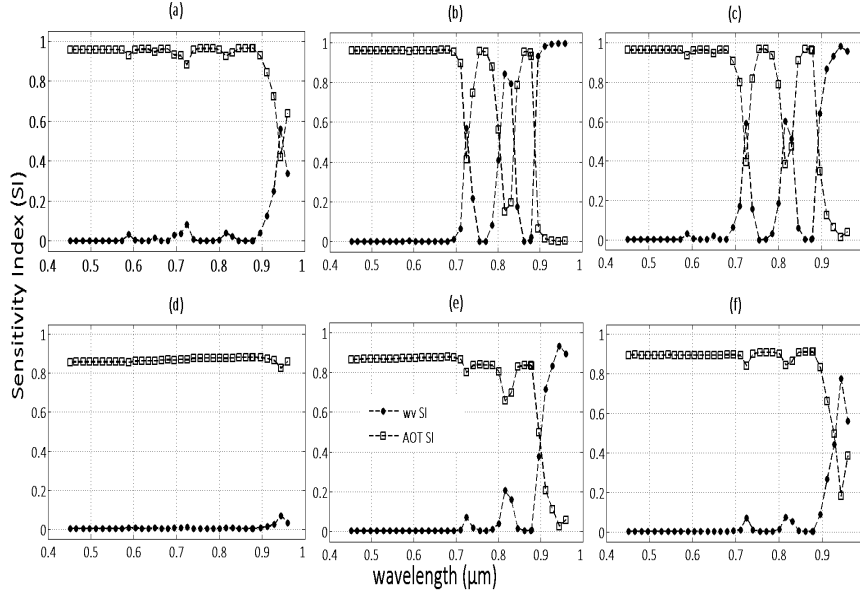
Figure 3.4: The sensitivity indices for the three target surface for the two cases: a)–c) depicts results from Case 1 for water, forest, and bare soil targets, respectively; d)–f) illustrate results from Case 2 for water, forest, and bare soil targets, respectively.

We now discuss the results of the importance of the two parameters quantified by their sensitivity index (SI) presented in Figure 3.4. The results in Figure 3.4 (a–c) are results from Case 1. The results in Figure 3.4 (d–f) are results from Case 2. Here, we are discussing the two cases separately.

- Case 1: From the indices shown in Figure 3.4 (a–c) we observe that in the spectral range 0.44–0.96 µm, aerosol optical depth remains an important parameter, as its sensitivity indices are high (>0.9). The aerosol optical depth indices, however, vary at the water absorption features (0.58, 0.65, 0.72, 0.82, 0.94 µm).

Likewise, indices of CWV were high at the absorption features. At the weak absorption features such as at 0.58 and 0.65 µm the aerosol optical depth indices only slightly decreased whereas the CWV indices showed very small increase. From these results we concluded that the influence of aerosol optical depth in the spectral range 0.44–0.96 µm, also depends on the strength of the absorption feature. This implies that in the non-absorbing wavelengths and at the weak absorption features, the aerosol optical depth plays an important role, which can be observed from high sensitivity index values for aerosol optical depth in these wavelengths. Likewise, for strong absorption features the importance of CWV was larger than the scattering effect, resulting in low indices for the aerosol optical depth.

Besides the strength of the absorption features, another important factor that affects the sensitivity of the two parameters is the surface albedo. From Figure 3.4(a–c), we observed that for the low albedo target the effect of scattering was more prominent in the entire spectral range. Even at the principal absorption feature (0.94 µm), the sensitivity indices for aerosol optical depth and CWV are almost equal. The difference in sensitivity indices over bright and dark targets is primarily due to the difference in radiation energy reflected by the bright and the dark targets. The bright target reflects more radiation energy than the dark target. Thus, for the bright target the majority of the at-sensor radiance consists of photons that are not scattered. In contrast, above the dark target most of the photons are scattered, which results in strong sensitivity to aerosol optical depth [63]. For the bare soil and the forest pixels which are bright (reflectance in the range of 0.4–0.7) at 0.94 µm, the CWV sensitivity indices were high and the SI of aerosol optical depth was low. The forest pixel, has low reflectance in the visible region (0.45–0.69), therefore, the effect of CWV at the absorption features 0.58 and, 0.65 µm was weak compared to the bare soil CWV sensitivity indices. The forest and water pixels sensitivity indices are, therefore, similar in the visible region. At bands with moderate absorption features, 0.72 and 0.82 µm, the dark target radiance is dominated by the scattering effect and the effect of CWV is low. For the bright target, however, at these wavelengths the scattering and absorption effects are nearly equal. This implies that at these wavelengths for bright target the scattering and absorption effect are both important. From these results we conclude that the influence of aerosol optical depth on the reflectance in the spectral range (0.44–0.96) µm, also depends on the surface albedo.

- Case 2: Here we only discuss those results which are relevant for the comparison between Case 1 and Case 2. From the indices shown in Figure 3.4 (d–f) we observed that for Case 1, the sensitivity indices of aerosol optical depth in the spectral range 0.44–0.96 µm are high (>0.9). The aerosol optical depth indices, however, vary

at the water absorption features (0.58, 0.65, 0.72, 0.82, 0.94 µm). This variation, however, is lower than in Case 1. Even at the principal absorption feature, SI indices for CWV remain lower than aerosol optical depth SI for all the surface albedo as in Case 1. Like, Case 1, the effect of surface types is noticeable. The reason for the difference in SI for the surface types has already been discussed for Case 1.

In both cases, the objective was to quantify the importance of the CWV and aerosol optical depth. We, however, observed that when the joint distribution was built using the probability densities and the relation between CWV and visibility (Case 2), the importance of the parameters, measured in terms of SI, is influenced by aerosol optical depth. The primary cause of this influence is high variability in visibility values. To measure the variability of CWV and visibility we used the Coefficient of Variation (CV). The CV represents the ratio of the standard deviation to the mean. The CV is a useful statistic to measure the relative variability for comparing different data series. From the image based methods, in total, 1420 observations of visibility were obtained, however, CWV was observed at each pixel (7.87×10^5). The CV of CWV and visibility was 4.83% and 50.07%, respectively. The CV of the two parameters indicates that variability in visibility values is larger than CWV. There can be two reasons for the large variability in visibility measurements. First, compared to CWV, visibility was observed at lower number of pixels. Second, the performance of the visibility estimation method is influenced by the viewing and illumination geometry, and spatial heterogeneity. This happens because the image based method is based on dense dark vegetation technique (DDV) and uses radiance values of pixels in red and NIR bands to search for the dense dark vegetation pixels.

To summarise, in this section we have presented and discussed the sensitivity of the two atmospheric condition parameters as a function of wavelength, strength of the absorption feature, parameters uncertainty, and surface albedo. In the Method section, we indicated that for a comparison purpose we also quantified total sensitivity indices (TSI) using the e-FAST method. The TSI is considered as a better estimation of sensitivity of parameters. However, on the basis of the results obtained from the e-FAST method, we observed that the difference between the two methods, as measured from their respective indices, were minor (<0.05). A possible explanation is the sensitivity of the condition parameters is captured by the first order Fourier coefficients, thus, the higher order terms that are calculated to quantify TSI does not carry much information about sensitivity to the parameters. This is because in this study only two parameters were used. For more parameters the effect of higher order indices might be more useful. In [62], describing the e-FAST method, the dependence of the interaction terms (higher order) to the number of factors is mentioned. The larger the number of

factors, the higher the likelihood of substantial higher order terms. We have, therefore, not reported the results from the e-FAST method in this chapter.

3.4 Conclusions

In this chapter, we have presented a methodology to quantify sensitivity of the estimated surface reflectance to water vapour concentration (CWV) and aerosol optical thickness (aerosol optical depth) in the atmospheric correction process. Besides quantifying the sensitivity indices, we also focused to analyse the effect of parameter's uncertainty, wavelength and surface albedo on these indices. Our approach was based on the Fourier Amplitude Sensitivity Test (FAST) method that was implemented in a Monte Carlo Simulation (MCS) framework. The methodology was applied to a hyperspectral image (HyMap) of the Millingerwaard area, The Netherlands. The atmospheric correction was performed with the Central Data Processing Center (CDPC) which is an airborne processing chain at the Flemish Institute for the Technological Research (VITO).

The results of the SA, expressed as sensitivity indices, were presented for three surface albedo in the spectral range 0.44–0.96 µm. The sensitivity of the estimate reflectance to CWV and aerosol optical depth depends on the wavelength, strength of the water absorption feature, parameters uncertainty, and on surface albedo.

1. Surface albedo: for the dark target, the estimated reflectance was more sensitive to aerosol optical depth compared to CWV over the 0.44–0.96 µm spectral range with exception of the bands near the 0.94 µm water absorption feature, where the effect of CWV and aerosol optical depth is almost similar.

2. Strength of the absorption feature: for the bright targets the sensitivity of the estimated reflectance to CWV increases with the strength of the water absorption feature. For a dark target, however, except for the principal absorption feature near band 0.94 µm CWV indices did not increase much compared to indices for bright targets.

3. Parameter uncertainty: if one parameter is more uncertain than the other, the SI is influenced by the more uncertain parameter. For instance aerosol optical depth's uncertainty was higher than CWV's.

Therefore, aerosol optical depth is an important parameter for dark targets for atmospheric correction, more important than CWV even at the principal absorption feature. For bright targets the importance of CWV and aerosol optical depth depends upon the strength of the absorption feature. aerosol optical depth sensitivity indices were high for the non-absorption water bands. We conclude that at the non-absorption water bands aerosol optical depth is a more important parameter compared to CWV.

As an outlook from the study, we recommend to include the impact of uncertainty in the aerosol type in the sensitivity analysis. Like CWV and aerosol optical depth, knowledge about the type of atmospheric aerosol at

the time of imaging is challenging to obtain. Often, expert judgement is used to select the appropriate aerosol type for the atmospheric correction process. Therefore, it would be interesting to quantify the effect of the choice about the aerosol type on the estimated reflectance. Also, it might be more interesting to collectively study the impact of uncertainty in aerosol type, aerosol optical depth, and CWV by simultaneously varying them.

Propagation of uncertainty in atmospheric parameters to hyperspectral unmixing

4

Abstract

Atmospheric correction (AC) is important in pre-processing of airborne hyperspectral imagery. AC requires knowledge on the atmospheric state expressed by atmospheric condition parameters. Their values are affected by uncertainties that propagate to the application level. This study investigates the propagation of uncertainty from column water vapour (CWV) and aerosol optical depth (AOD) towards abundance maps obtained by means of spectral unmixing. Both Fully Constrained Least Squares (FCLS) and FCLS with Total Variation (FCLS-TV) are applied. We use five simulated datasets contaminated by various noise levels. Three datasets cover two spectral scenarios with different endmembers. A univariate and a bivariate analysis are carried out on CWV and AOD. The other two datasets are used to analyse the effect of surface albedo. The analysis identifies trends in performance degradation caused by the gradual shift in parameter values from their true value. The maximum achievable performance depends upon spectral characteristics of the datasets, noise level, and surface albedo. As expected, under noisy conditions FCLS-TV performs better than FCLS. Our research opens new perspectives for applications where estimation of reflectance is so far considered to be deterministic.

4.1 Introduction

Hyperspectral imaging sensors record the at-sensor radiance reflected from a surface, for hundreds of narrow contiguous spectral bands. A recorded image can thus be seen as a three dimensional cube with two spatial dimensions

and one spectral dimension. A pixel in such a cube usually covers an area comprising several endmembers. These mixed pixels are in contrast with pure pixels that cover a single endmember. The occurrence of mixed pixels is due to two main reasons: i) the spatial resolution of a hyperspectral sensor is relatively low, thus, several endmembers share the spatial extent of a pixel, and ii) the underlying surface is a mixture of several materials.

As important information about the scene might reside in mixed pixels, extraction of quantities of interest at the subpixel level is needed. Spectral unmixing is a popular extraction method at the subpixel level. It exploits spectral information to derive the endmembers in the scene, their spectral signatures, and their fractional abundances, i.e. areas occupied by each endmember in each pixel. For a comprehensive review of unmixing techniques, see [64] and the references therein. In this study, we rely on the Linear Mixture Model (LMM) [65]. It expresses the observed spectrum of a pixel as a linear combination of the spectra of the endmembers weighted by their fractional abundances.

Spectral unmixing using the recorded radiance is challenging in the presence of the Earth's atmosphere. This is primarily because of the interaction of the surface reflected radiation with the atmospheric constituents while propagating along the path from the target surface to the sensor [33]. The interaction generates two main atmospheric effects: absorption by atmospheric gases in particular water vapour and ozone and aerosols in the visible and near infrared spectral range and scattering by aerosols and molecules [66]. In addition, on the path of the beam to the sensor, reflection by the surrounding area of the target pixel and radiance backscattered by the atmosphere that did not interact with the surface distorts the at-sensor radiance.

An Atmospheric Correction (AC) algorithm retrieves the surface reflectance from the at-sensor radiance. AC algorithms can be divided into scene based empirical algorithms and algorithms based on radiative transfer modelling. We use the latter, as it is a mature approach for routine processing of hyperspectral image data [67].

In radiative transfer modelling, the target radiance can be derived assuming a plane parallel geometry of the atmosphere, whereas the viewing and illumination geometry and total optical depth of the atmosphere are known. For a reliable estimate of reflectance, the concentration of the atmospheric scatterers and absorbers, i.e. the optical parameters, should be available at the time of imaging. In this chapter, we analyse the effect of uncertainty in estimations of atmospheric aerosol optical depth (AOD) and column water vapour (CWV). Both CWV and AOD are highly varying in space and time. Thus, they are estimated directly from satellite or airborne (remote) observations. With knowledge of CWV and AOD, transmission of radiation through the atmosphere can be simulated.

Estimation of CWV from at-sensor radiance consists of identification of the measurement channels, identification of reference channels, and using a relation between reference and measurement channels [68]. These methods are limited with respect to several assumptions. First, surface reflectance is assumed to vary with wavelength in a linear way; second, the effect of

sensor noise is often not considered, and third, uncertainty emerging from instrument characterisation is ignored [69, 70].

Estimation of AOD consists of determining aerosol radiative properties characterised by their shape, their size, their chemical composition, and total amount [71]. The MODIS science team [7] has developed the dense dark object method to estimate AOD that is further developed in [8]. The limitation of such methods is their suitability for pixels with dense vegetation. For scenes with dark pixels that are clustered at a few locations, pixelwise estimation of AOD is challenging. Besides, at-sensor based inference of AOD is adversely affected by noise of at-sensor radiance.

These assumptions and limitations, reasonable as they are, cause uncertainty in the estimation of CWV and AOD which likely propagates to reflectance estimates.

The objective of this chapter is to analyse the impact of uncertainty in unmixing caused by CWV and AOD, given their specific influence on the estimated reflectance spectra. A basic hypothesis of unmixing is that the estimated reflectance spectra are free from atmospheric artefacts. By ignoring uncertainty in the AC parameters, however, it is likely that this hypothesis is violated. The chapter specifically focuses on an operational processing chain. The operational processing chain is implemented in the multi mission Processing, Archiving, and distribution Facility (PAF) for earth observation products [39]. Experiments in this chapter are performed using the PAF incorporated in the Central Data Processing Center (CDPC) [20] at the Flemish Institute for Technological Research.

4.2 Theoretical Background: the linear mixture model (LMM) and unmixing methods

Let $y \in \mathbb{R}^B$ be the reflectance spectrum of one pixel, where B is the number of spectral bands. According to the LMM, it can be expressed as a linear combination of the spectra of the endmembers, weighted by their fractional abundances:

$$y = A \cdot x + n. \tag{4.1}$$

Here, $A \in \mathbb{R}^{B \times m}$ is the set of endmembers in the scene serving as a spectral library containing m pure spectra, $x \in \mathbb{R}^m$ is the vector of corresponding fractional abundances compatible with A, and $n \in \mathbb{R}^B$ is a noise vector. In this paper, we assume that A is available a priori. Unmixing thus aims at identifying the atoms of A which are active in each pixel and their respective abundances. To solve (8), we consider the classical least-squares solution obtained by solving:

$$\min_x \frac{1}{2} \left\| A \cdot x - y \right\|_2^2. \tag{4.2}$$

Two constraints, arising from the physical meaning of the fractional abundances, can be imposed on (9): 1) the non-negativity constraint (ANC), and 2) the sum-to-one constraint (ASC) [72].

In addition, we take intrinsic spatial smoothness into account as an important characteristic of natural scenes. We use the Sparse Unmixing via Variable Splitting, Augmented Lagrangian and Total Variation (SUnSAL-TV) [73] to obtain a solution with piece-wise smooth transitions of the abundance fractions in neighbouring pixels. SUnSAL-TV solves the optimisation problem:

$$\min_X \frac{1}{2}\|AX - Y\|_F^2 + \lambda_1\|X\|_{1,1} + \lambda_{TV}\,\mathrm{TV}(X), \tag{4.3}$$

where

$$\mathrm{TV}(X) \equiv \sum_{\{i,j\}\in\,\mathcal{E}} \|x_i - x_j\|_1 \tag{4.4}$$

is a vector extension of the non-isotropic TV [74] and ε denotes the set of neighbours in the image. In (4.3), $Y \in \mathbb{R}^{L\times n}$ is the observed data matrix with each column containing the observed spectrum at a pixel, $X \in \mathbb{R}^{m\times n}$ is the matrix of fractional abundances, $\|X\|_F \equiv \sqrt{\mathrm{trace}\{XX^T\}}$ represents the Frobenius norm of X and $\|X\|_{1,1} \equiv \sum_{i=1}^{n}\|x_i\|_1$, with x_i denoting the ith column of X. The first term in (4.3) measures the data misfit, the second term forces the matrix of fractional abundances to be sparse, and the last term accounts for spatial homogeneity of the abundance maps. The parameters $\lambda_1 \geq 0$ and $\lambda_{TV} \geq 0$ are regularisation parameters. SUnSAL-TV introduces a set of new variables per regularizer and then uses the Alternating Direction Method of Multipliers (ADMM) [75] to solve the resulting constrained optimisation problem. In our experiments, we neglect sparsity as the spectral libraries employed will contain a small number of endmembers. We use SUnSAL-TV applying both ANC and ASC to solve both optimisation problems (4.2) and (4.3), as follows:

- Fully constrained least squares (FCLS): no sparsity is enforced and no spatial information is considered. This situation is encountered if the regularisation parameters $\lambda_1 = 0$ and $\lambda_{TV} = 0$;

- FCLS with Total Variation (FCLS-TV): Only spatial information is considered in SUnSAL-TV by setting $\lambda_1 = 0$ and $\lambda_{TV} > 0$.

The ASC, often ignored due to signature variability [76] is used in this work as we assumed the set of image endmembers to be known and hence no sparsity is enforced on the vectors of fractional abundances.

4.3 Datasets

Three synthetic datasets are used where two datasets represent a geological scenario and one dataset represents a scenario containing vegetation. Further, we use two Spectral Scenarios (SS): SSa with spectrally distinct endmembers in a scene; SSb with spectrally similar endmembers in a scene. The two spectral scenarios are linked to the three synthetic datasets presented in Figures(2.3–2.6), respectively.

4.4 Experiments

Atmospheric correction, as a mathematical model of an atmosphere, returns an estimate of reflectance for any physically realisable values of column water vapour, aerosol optical depth and other parameters. In our experiments, the other parameters are treated as constants. Column water vapour and aerosol optical depth, however, should also have unique values during atmospheric correction. In simulations, it may be necessary to vary them over a range of values, where the distribution of values reflects the uncertainties of the parameters. We use one at a time variation in aerosol optical depth and column water vapour, i.e. univariate analysis, and their joint variation in a bivariate analysis for the uncertainty analysis. A bivariate analysis permits large deviations from the nominal parameter values, thus allowing evaluation of atmospheric correction for various combinations of column water vapour and aerosol optical depth. We use uniform probability distributions for the two atmospheric parameters, resulting in equally probable samples.

4.4.1 The forward modelling

At-sensor radiance cubes are simulated using the forward radiative transfer modelling using MODTRAN 4. Standard MODTRAN 4 mid latitude summer model with the rural aerosol model is used and the sensor altitude is approximately 3 km above sea level. In univariate and bivariate analysis applied to the FCLS-TV algorithm, the regularisation parameter λ_{TV} is set to values between 0 and 0.06 with a step size of 0.01, and the best performance among all values is retained. Average performances over ten runs are taken.

4.4.2 Experimental setup for univariate analysis

The univariate analysis experiments for aerosol optical depth consist of three ranges: high to moderate (visibility: 4–26 km), moderate to low (visibility: 16–45 km), and low (visibility: 45–72 km). From our experience with real images processed in the CDPC, we observed that the reflectance is not sensitive to visibility values above 60 km. Thus, the maximum value of visibility is set to 72 km. Each visibility range is associated with an experiment given in Table 4.1, where column water vapour is kept fixed to the value used in the forward modelling. Likewise, the experiments for column water vapour consist of two column water vapour ranges: low to moderate (0.1–1.0 g cm^{-2}) and moderate to high (1.0–2.25 g cm^{-2}). Each column water vapour range is associated with an experiment given in Table 4.2. In these experiments, visibility is kept fixed to the value used in the forward modelling. We used the same atmospheric settings to analyse the propagation of uncertainty to the abundance maps. The unmixing performed using the FCLS and FCLS-TV is repeated for noiseless data and correlated noise with SNR of 50 dB, 40 dB and 30 dB.

Table 4.1: visibility (km) and corresponding aerosol optical depth ranges to generate various atmospheric combination to perform univariate analysis of aerosol optical depth for geological scenario datasets with distinct and similar endmembers, respectively. The sampling rate is 2 km and column water vapour is fixed at 1.5 g cm^{-2}.

scattering condition	aerosol optical depth range	true aerosol optical depth	visibility range	true visibility
high to moderate	1.42–0.28	0.47	4–26	15
moderate to low	0.45–0.18	0.28	16–45	28
low	0.18–0.11	0.14	45–72	60

Table 4.2: Column water vapour ranges (g cm^{-2}) to generate various atmospheric combination to perform univariate analysis of column water vapour for the geological datasets with distinct and similar endmembers. The sampling rate is 0.05 g cm^{-2} and the visibility is fixed at 15 km (i.e. aerosol optical depth equals 0.48).

absorption condition	column water vapour range	true column water vapour
low to moderate	0.1–1.0	0.5
moderate to high	1.0–2.25	1.5

4.4.3 A univariate analysis of dark and bright targets

This experiment analyses the joint influence of aerosol optical depth uncertainty and albedo amplitude on the spectral quality of the estimated reflectance cube and on the fractional abundances derived from it.
For the forward modelling of dark and bright target surfaces, low scattering (aerosol optical depth = 0.14) and high scattering (aerosol optical depth = 0.48) are considered, whereas column water vapour is fixed to 1.5 g cm^{-2}, resulting in four at-sensor radiance cubes. AC is evaluated for two discrete sets of samples of the aerosol optical depth in the ranges (1.42–0.28) and (0.16–0.13), corresponding to visibility ranges 6–27 km and 51–69 km, respectively.

4.4.4 Experimental setup for a bivariate analysis

The bivariate analysis applied to the vegetation dataset analyses the joint effect of column water vapour and aerosol optical depth on atmospheric correction. Column water vapour and visibility are varied within the ranges 1.35–1.65 g cm^{-2} and 15–45 km, corresponding to aerosol optical depth values between 0.48 and 0.18, with reference column water vapour and visibility

set to 1.5 g cm^{-2} and 30 km, respectively. These experiments are repeated for data without correlated noise and for data perturbed by correlated noise with SNR of 50 dB, 40 dB, and 30 dB. This atmospheric setting is also used to analyse the propagation of uncertainty to abundance mapping.

4.4.5 Uncertainty propagation

To quantitatively assess the propagation of uncertainty we use the signal to reconstruction error SRE$\equiv \frac{E\left[\|x\|_2^2\right]}{E\left[\|x-\hat{x}\|_2^2\right]}$ for both the estimated reflectance and fractional abundances. Here, x is the reference signal and $\hat{x}$ is an estimation of x. SRE provides more information on the power of the signal w.r.t. the power of the error than, e.g., the Root Mean Squared Error (RMSE) [77]. In all experiments, we report SRE measured in dB: SRE(dB)= $10\log_{10}$(SRE). In addition, to quantify the pixelwise uncertainty in the estimates of reflectance and abundance maps, we measure two other quantitative errors: Normalised Root Mean Square Error (NRMSE)= $\frac{\text{RMSE}}{\max(x)-\min(x)}$ and Mean Absolute Percentage Error (MAPE) $= \sum_{t=1}^{n}|\frac{x_t-\hat{x}_t}{x_t}| \times 100$.

4.5 Experimental Results

4.5.1 Univariate analysis of aerosol optical depth

Figure 4.1 shows results of univariate analysis of aerosol optical depth for the geological dataset with distinct endmembers with different degrees of correlated noise added to the dataset. The maximum performance is achieved if visibility is close to the true value. The rate of degradation in performance is higher, shown by the steeper slope, if visibility is underestimated and is smaller if overestimated. This performance trend is valid for all visibility scenarios. The maximum achievable performance varies among the datasets and depends upon the true value of visibility. For the true atmospheric conditions, i.e. the same atmospheric conditions for forward and inverse radiative modelling, the minimum performance is achieved if true visibility equals 15 km, whereas the maximum performance is achieved if true visibility equals 60 km. Ideally, the performances should not vary if atmospheric conditions are kept constant for the forward and inverse radiative modelling. This difference is attributed to surface albedo and is further investigated in Section 4.5.3.

From Figure 4.1, it can be seen that the performance degrades if the data are affected by noise. This observation is valid for all datasets, and the decrease in the performance is stronger if the noise level increases. To quantify this effect, we compute the ratio (R) between the maximum performance for SNR = 30 dB and the maximum performance without correlated noise (Table 4.3). A low ratio indicates a high degradation. Each noise level limits the maximum achievable performance even at the true value of visibility. This affects the slope between the performance achieved at the true value of
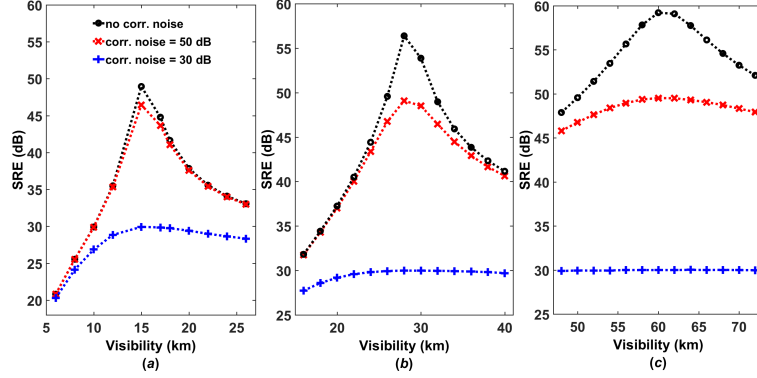
Figure 4.1: SRE (dB) values for univariate analysis of aerosol optical depth for the geological dataset with distinct endmembers with and without correlated (corr.) noise, when the true visibility is set to: a) 15 km; b) 28 km, and c) 60 km.

Table 4.3: Performance ratios for the geological dataset with distinct endmembers ($G_{dist.}$) and the geological dataset with similar endmembers ($G_{sim.}$). P_1 corresponds to noisy data with SNR = 30 dB, while P_2 corresponds to noiseless data. The lower the ratio values, higher the performance degradation during the univariate analysis of the reflectance product.

label	true visibility (km)	P_1	P_2	ratio
$G_{dist.}$	15	29.94	48.95	0.612
$G_{sim.}$	15	29.98	56.87	0.527
$G_{dist.}$	28	29.99	56.39	0.532
$G_{sim.}$	28	30	58.88	0.51
$G_{dist.}$	60	30	59.21	0.507
$G_{sim.}$	60	29.98	59.70	0.502

visibility and at all other points. For instance, if the true visibility is 15 km and without correlated noise the maximum performance is close to 49 dB and the minimum performance is close to 20 dB, thus the slope is approximately 3.2. With 30 dB correlated noise, the maximum achievable performance is limited to 30 dB and the slope is approximately 1.3. For the two geological scenario datasets, performance of FCLS and FCLS-TV is quantified using reference abundance cubes i.e. ground truth abundances used to generate those datasets.

Figure 4.2 shows the performance of FCLS for noiseless data. Results reveal that for both FCLS and FCLS-TV (not presented), the performance trend is similar to that of reflectance estimates, but its magnitude is lower than the reflectance estimates. The different spectral compositions of the scenes may cause differences in performance between the geological datasets with distinct
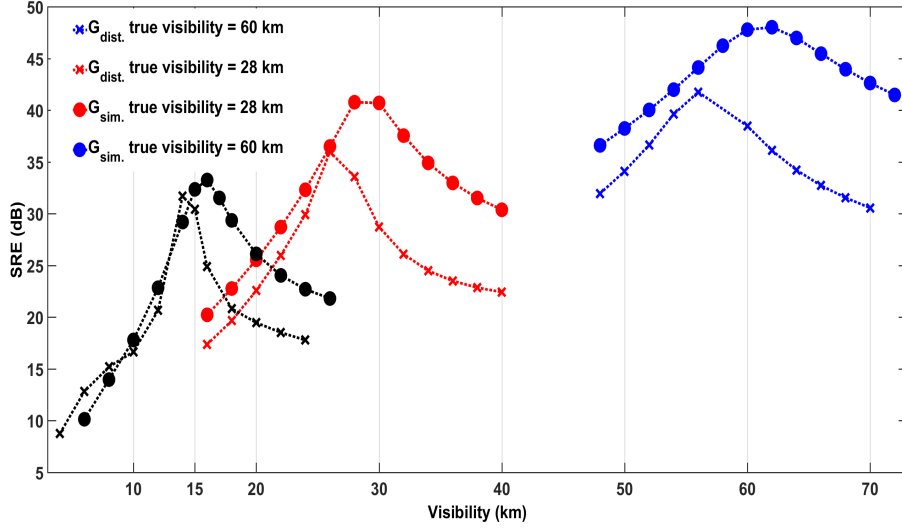
Figure 4.2: SRE(dB) values for FCLS with the geological dataset with distinct endmembers ($G_{dist.}$) and the geological dataset with similar endmembers ($G_{sim.}$) without correlated noise.

and similar endmembers. Another explanation may be the difference in surface albedo between the datasets. Further, peak performances are observed at the true value of visibility and at the next higher value of visibility i.e. the performance is insensitive to visibility values between these two points. For the geological dataset with similar endmembers, peak performances are not available at the true values and appear at the next lower value of visibility.

Figure 4.3 shows the variation of SRE (dB) with respect to λ_{TV} for the geological dataset with distinct endmembers for all scattering scenarios affected by various levels of noise. For no correlated noise, 50 dB correlated noise and 40 dB correlated noise, the performances of FCLS and FCLS-TV are comparable, whereas FCLS-TV performs better with increasing noise level.

To explore propagation of the uncertainty to the abundance maps, we quantify NRMSE to generate various other results. First, in Figure 4.4 a bar graph representing NRMSE of the geological dataset with distinct endmembers is shown. Note, the larger errors are occurring for low visibility values.

Figure 4.5 shows propagation of uncertainty in terms of mean NRMSE estimates of endmember-4 abundance maps obtained by varying visibility between 6–26 km and their corresponding standard deviation maps. The true visibility is set to 15 km. The mean NRMSE corresponds to a map with the mean values at every pixel computed by an average of NRMSE at each visibility value between 6 and 26 km, with a sampling rate of 1 km.
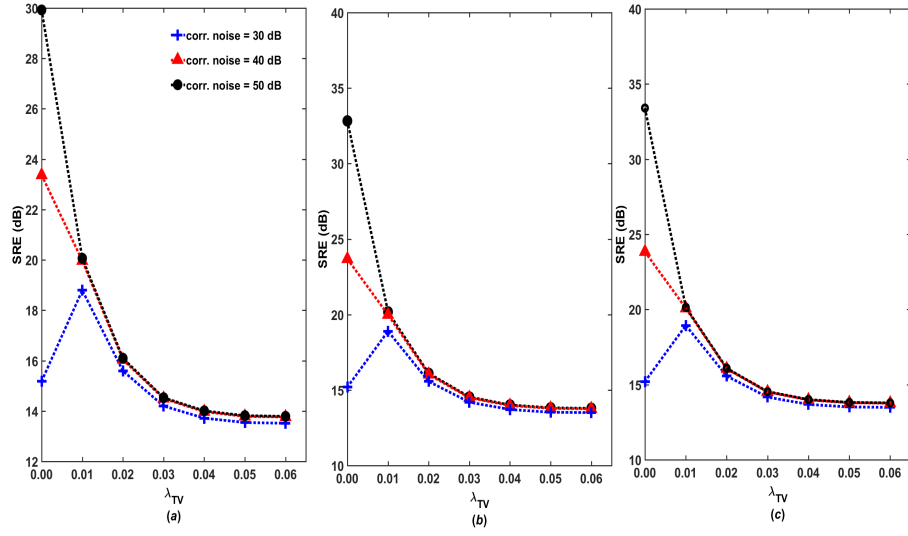
Figure 4.3: Variation of SRE (dB) with respect to λ_{TV} values for FCLS-TV in geological dataset with distinct endmembers for the true visibility equals 15 km (a), 28 km (b), and 60 km (c). The optimal value of λ_{TV} corresponds to the maximum value of SRE (dB).
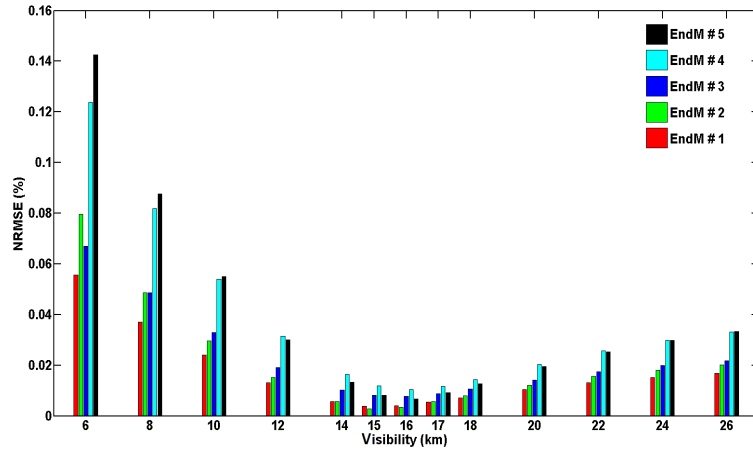
Figure 4.4: NRMSE for each endmember of the geological dataset with distinct endmembers for a visibility range 6–26 km when the true visibility is set to 15 km.
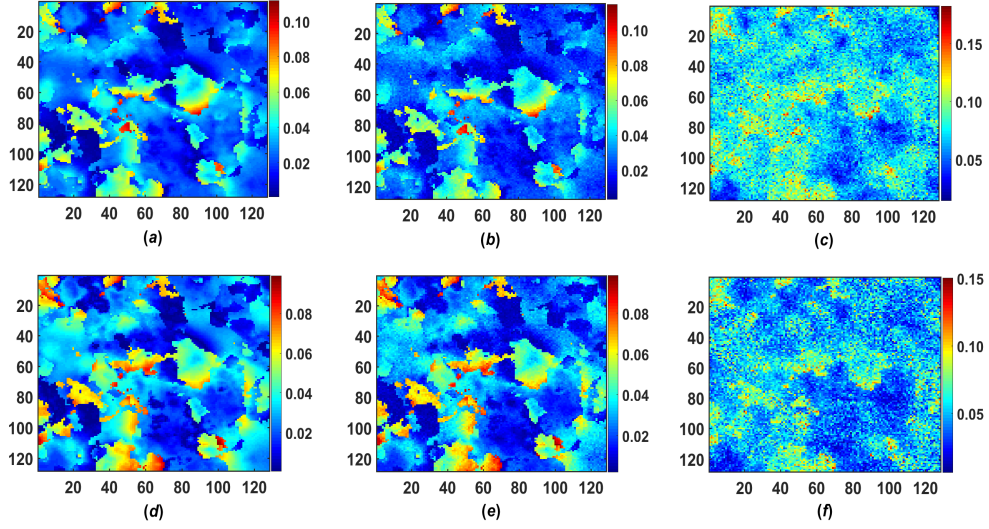
Figure 4.5: Pixelwise mean NRMSE estimates of endmember 4 for no noise (a), 50 dB noise in (b) and 30 dB noise in (c) obtained due to variations in visibility between 6–26 km. The corresponding standard deviation are shown in plots (d)–(f) for no noise, 50 dB, and 30 dB noise levels, respectively.

4.5.2 Univariate analysis for column water vapour

As for aerosol optical depth, the maximum performance is obtained at the true value of column water vapour. Unlike for aerosol optical depth, however, the performance degradation is sharp and less dependent upon column water vapour overestimation or underestimation. For higher column water vapour (> 2.0 g cm^{-2}) the performance tends to saturate. Further, the maximum attainable performance is independent of the true column water vapour values. This is as expected and different from aerosol optical depth where differences in performance are observed for visibility values of 15 km and 60 km. Performances are higher for geological scenario dataset with similar endmembers. With the increase in noise level, the performance degrades but the peak of the performance remains at the true values i.e. no shift is observed.

For the two geological scenario datasets with distinct and similar endmembers, and for both FCLS and FCLS-TV, the performance trend is similar to that of reflectance estimate, whereas the magnitude of the performances is lower. Unlike aerosol optical depth, no significant variation in performance is observed due to changes in the spectral property i.e. changes of the datasets.

Figure 4.6 shows a reference spectrum and various spectra of the same material, derived via AC that are distorted due to the uncertainty in visibility and column water vapour. These results illustrate the effect of a parameters's deviation on the spectral quality and depict how deviation in the parameter affect the shape of the spectra, which is an important input for unmixing.

49
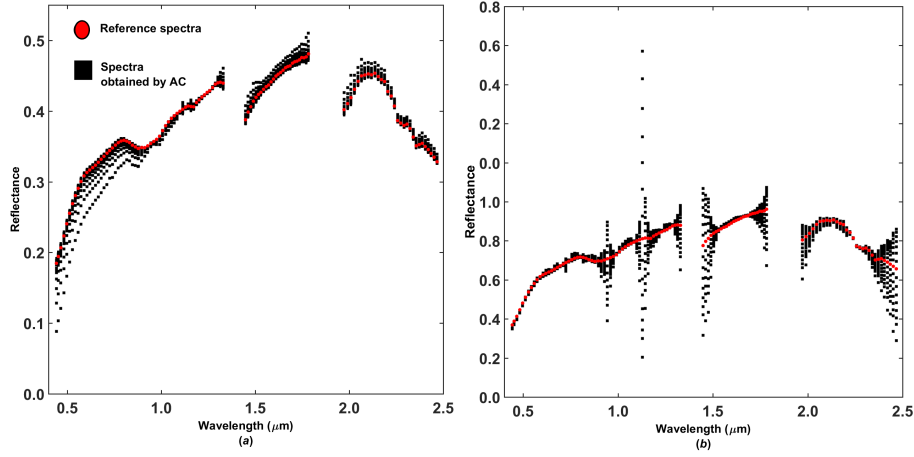
Figure 4.6: Deviation in the spectral quality due to the uncertainty in visibility in (a) and CWV in (b). The visibility is varying between 6–26 km with a true visibility at 15 km and CWV is varying between 1.0–2.25 g cm^{-2} with a true value at 1.5 g cm^{-2}.

Table 4.4: Effect of dark and bright targets on the performance with true visibility conditions for reflectance estimates.

Albedo	visibility (km)	aerosol optical depth	SRE (dB)	NRMSE	MAPE (%)
Dark	60	0.14	64	$4.2e^{-04}$	$3.8e^{-02}$
Bright	60	0.14	72	$3.0e^{-04}$	$9.4e^{-03}$
Dark	15	0.47	52	$1.7e^{-03}$	$1.6e^{-01}$
Bright	15	0.47	58	$1.5e^{-03}$	$5.7e^{-02}$

4.5.3 Univariate analysis of dark and bright targets

Table 4.4 shows results of the experiment related to the dark and bright targets if no noise is added to the reflectance cubes. For illustration, only performances for the true atmospheric conditions of 15 km and 60 km are shown. At reflectance level, we noted that the performance for the dark targets is lower than the one for the bright targets irrespective of the atmospheric conditions. This trend propagates to the abundance level estimation as well, as shown in Figure 4.7. As the unmixing performance follows the trend of the SRE (dB) computed for reflectance, it can be concluded that the abundance maps obtained via unmixing for bright scenes are less affected by uncertainty in the atmospheric correction parameters than those obtained for dark scenes.

4.5.4 The bivariate analysis for column water vapour and aerosol optical depth
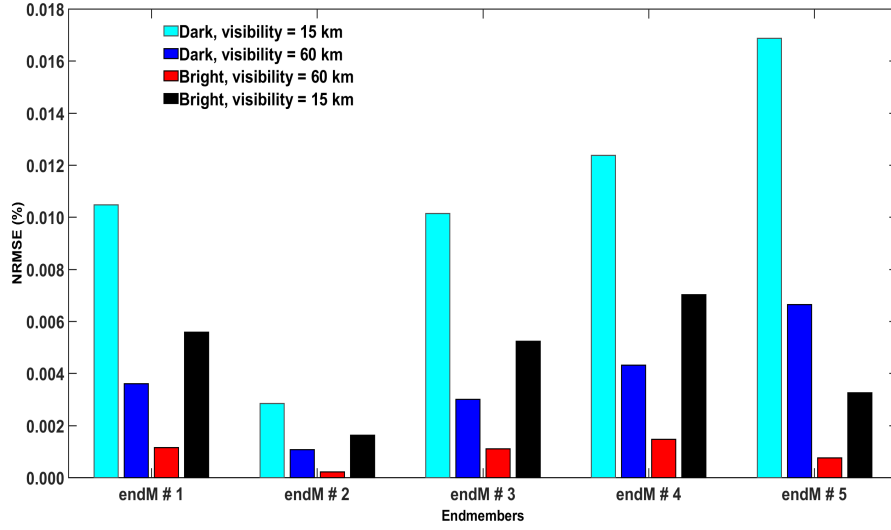
Figure 4.7: The NRMSEs (%) obtained for the five endmembers quantify the effect of dark and bright targets with true visibility and column water vapour conditions for abundance estimates.

The performance analysis for reflectance and abundance maps is given in Figure 4.8. As observed from the univariate analysis experiments, high performance around the true values of column water vapour and visibility is obtained. If the values of the two parameters deviate from the true values, the performance degrades. A general pattern is that the performance tends to saturate, after a relatively low decrease, if visibility is higher than 30 km (overestimation), whereas, if visibility is lower than 30 km (underestimation), the performance sharply declines. For column water vapour, however, performance declines sharply for both overestimation and underestimation.
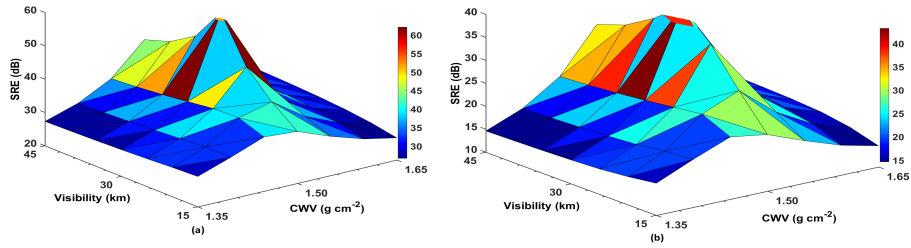
Figure 4.8: SRE (dB) values for bivariate analysis of column water vapour and visibility for reflectance estimates (a) and abundance estimates (b) without correlated noise.

Figure 4.9 shows SRE (dB) of reflectance estimates and abundance estimates under high noise conditions, i.e. 30 dB correlated noise. This noise level strongly influences the performance at both estimates in two ways: 1) the overall amplitude of the performance degrades as compared to the performances observed at 40 dB and 50 dB noise levels; 2) comparing the trend of the performances with the noiseless case, the distinct peak of the performances is not visible anymore i.e. the highest performance can be obtained around the true values of the parameters and not necessarily when the parameters are correctly set. These findings are more clearly visible for unmixing estimates.
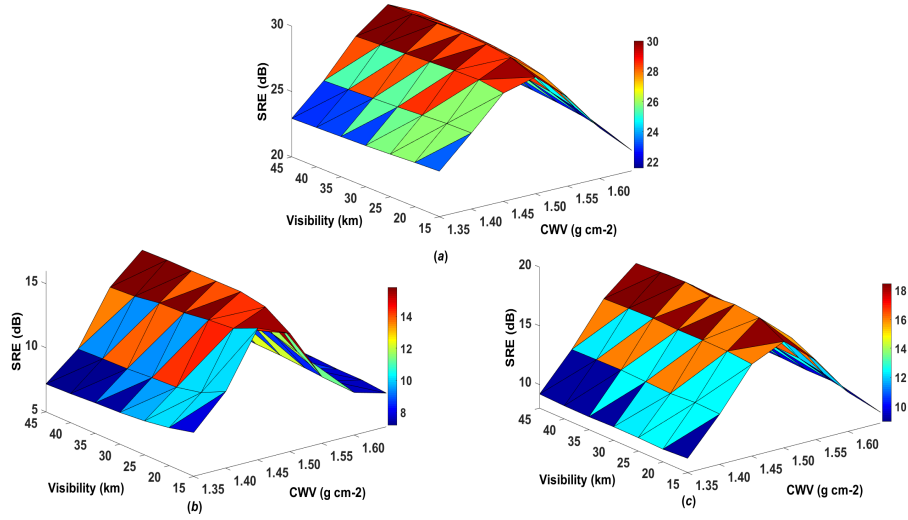
Figure 4.9: SRE (dB) for the vegetation dataset with correlated noise (SNR = 30 dB) for reflectance estimate (a), for the abundance estimate with FCLS solution (b), and with FCLS-TV solution (c) for variations in column water vapour and visibility.

For correlated noise with SNR of 40 and 50 dB, the decrease in the performances is smoother if column water vapour is overestimated as compared to if it is underestimated. Further, the distinct peak of the performances is also not achieved for 40 and 50 dB as compared to the noiseless case.

A comparison between the unmixing solutions reveals that FCLS-TV is more performant than FCLS in noisy scenes. The superior performance of FCLS-TV is less important if noise decreases, confirming the previous results. This is valid for all combinations of atmospheric parameters. The optimal λ_{TV} values under various noisy conditions are given in Table 4.5.

Table 4.5: Optimal λ_{TV} for different noise levels and for various combinations of column water vapour and visibility.

Noise level (dB)	column water vapour (g cm^{-2})	visibility (km)	Optimal λ_{TV}
No noise	1.35–1.65	15–45	0
30	1.35	15	0.03
30	1.40–1.65	20–45	0.01
40	1.35	15	0.03
40	1.40–1.65	20–45	0.01
50	1.35–1.45	15	0.01
50	1.50–1.65	20–45	0

4.6 Discussion and Conclusions

This chapter aimed to quantify the uncertainty propagation from aerosol optical depth (AOD) and column water vapour to the fractional abundance estimates via reflectance estimates. It was applied to the HyMap hyperspectral sensor and there is no reason why it should also not be effective on other hyperspectral sensors. The method is simple to implement and can be extended to encompass other atmospheric trace gases and other atmospheric condition parameters.

One of the important contributions of this study is that the effect of uncertainty was analysed and quantified. The results shown provide a useful insight into how a given uncertainty affects the performance at reflectance level and at subsequent applications (spectral unmixing). For instance, 30% degradation in performance (SRE) was observed under the high scattering conditions. This degradation came down to 6 % under the moderate scattering conditions. The steepness in performance degradation significantly changes when the noise corrupts the data, especially under the high noise conditions. This indicates that the atmospheric settings, noise, and the parameter's deviation have a combined effect on the propagation of the uncertainty. Thus, an in depth analysis of performance degradation due to the uncertainty in aerosol optical depth and column water vapour was performed for different atmospheric scenarios, covering low, moderate, and high scattering and absorption conditions for various noisy setups. The analysis identified performance degradation trends that were actually caused by the gradual shift in parameter values from their true value within the uncertainty limits (ranges) defined in the experiments.

We noted that the variability in aerosol optical depth influenced the amplitude of the retrieved spectra. Its influence was not linear, in the sense that deviations from the reference spectra depend upon the surface reflectance of the target, the wavelength, and the scattering conditions [51]. This observation is consistent with our previous work [78]. For high aerosol optical depth values and high surface reflectance (> 0.3), the estimated reflectance was higher than the reference reflectance, whereas for low surface reflectance (<0.3) estimated spectra were lower than reference spectra. This consistently

happened for spectra of all considered endmembers. An opposite pattern was observed for low scattering conditions. In [79], a threshold $P_c = \frac{\text{AOD}}{\rho_t(\lambda)}$ was introduced as a critical parameter to determine how the estimated reflectance changes due to aerosol scattering. Applying this threshold we found that if $P_c > 1$, then the estimated surface reflectance was underestimated, otherwise it was overestimated. The changes in the estimated reflectance were negligible when aerosol optical depth approaches the reflectance value.

Variation in column water vapour influenced the retrieved spectra mainly in the bands around the atmospheric water vapour absorption features. Uncertainty in column water vapour affected the performance more severely than uncertainty in aerosol optical depth, because variation in column water vapour significantly distorted the spectral shape. This was evident from the low performance obtained for maximum deviation in aerosol optical depth as compared to maximum deviation in column water vapour. The uncertainty in aerosol optical depth, however, strongly influenced the quality of the abundance maps if the endmembers were spectrally similar, especially if high atmospheric scattering occurred.

Further, the performance of reflectance estimates was always higher than that of abundance estimates. This might be due to the inherent limitations of the unmixing method and their numerical approximations. Without noise FCLS and FCLS-TV are comparable because spatial homogeneity in the datasets introduces only negligible improvements of the fractional abundance values in noiseless scenes.

A difference in maximum attainable performance observed for high scattering conditions could be attributed to the presence of bright and dark targets in a scene. Bright targets reflect more radiation energy than dark targets. Thus, the majority of the at-sensor radiance consists of photons that are not scattered. In contrast, most of the photons for the dark targets are scattered [63]. This contribution amplifies under high scattering conditions resulting in performance degradation. In this chapter, we do not explicitly explore the effect of albedo on the parameters estimation. From our experience working with the real images in the CDPC, we found that surface albedo affects the parameters estimation by widening the parameters uncertainty. The surface albedo becomes a source of the parameters uncertainty. As an outlook of this work, a sensitivity analysis can be performed to apportion the contribution of sources of uncertainty to reflectance estimates.

Effective ranges of column water vapour and aerosol optical depth were within the 90% confidence interval from their true values. Thus, for the bivariate analysis, we varied visibility and column water vapour values within relatively narrow ranges: 15–45 km and 1.35–1.65 g cm^{-2}, respectively. Outside this interval, the performance strongly degraded and saturated to a low value. Such effective ranges could be a useful measure for calibrating the value of the two parameters in optimizing the performance under uncertainty. Under high noise conditions (SNR = 30 dB), both spectral quality and unmixing performance are strongly degraded, whereas the effect of scattering becomes less important. A high noise level therefore has a weaker influence on the performance trend for column water vapour than for aerosol optical depth,

resulting in a distinct peak of the performance. Thus, in the presence of high noise in airborne data, uncertainty in atmospheric correction parameters and data noise affect the final products jointly.

Care should be taken when using visibility as a substitute to aerosol optical depth, especially under high scattering conditions where aerosol optical depth is highly sensitive to visibility values (see Figure 2.2). Ideally, estimation of aerosol optical depth can be obtained from the AErosol RObotic NETwork (AERONET) stations [80]. However, most of the airborne campaigns and many satellite images do not include an AERONET site location. Alternatively, aerosol optical depth measurements can be taken from handheld instruments. This requires ground measurements at the time of satellite or airborne image acquisition, which is often hard to collect and is impossible to obtain for archived imagery. Two measures that are often used for visibility are horizontal extinction coefficients derived from the horizontal visibility measured at airport stations and image based methods. Horizontal based visibility estimation leads to inaccurate values because of the dependency of estimates to the viewing direction. Lower values of visibility are reported when looking in the direction of the sun due to the strong forward scattering radiation [32]. It results in errors in horizontal visibility that are significantly higher than those derived from image based methods and standard satellite products. Airborne campaigns are also normally not acquired under high scattering conditions. Therefore, errors seen at 0–10 km visibility in [38] are unlikely to be experienced with operational remote sensing.

The present study can be further extended to take other atmospheric correction parameters into account. Effects of adjacency can be investigated in scenes where heterogeneity and/or topographic effects are important to include. It would be interesting to compare different unmixing methods while considering the reflectance uncertainty due to atmospheric condition parameters. Moreover, as uncertainties were also introduced by the applied unmixing method, a comparison between unmixing methods can be extended with inclusion of model uncertainty. Experiments with spectral libraries containing a large number of spectral signatures are recommended, as those were beyond the scope of the current study.

An Optimisation Approach to Estimate and Calibrate Column Water Vapour for hyperspectral airborne data

5

Abstract

The chapter describes a novel approach to estimate and calibrate Column Water Vapour (CWV), a key parameter for atmospheric correction of remote sensing data. CWV is spatially and temporally variable, and image based methods are used for its inference. This inference, however, is affected by methodological and numeric limitations, which likely propagate to reflectance estimates.

In this chapter, a method is proposed to estimate CWV iteratively from target surface reflectances. The method is free from assumptions for at-sensor radiance based CWV estimation methods. We consider two cases: (a) CWV is incorrectly estimated in a processing chain; (b) CWV is not estimated in a processing chain. To solve (a) we use the incorrect estimations as initial values to the proposed method during calibration. In (b), CWV is estimated without initial information. Next, we combined the two scenarios, resulting in a generic method to calibrate and estimate CWV.

We utilised the Hyperspectral Mapper (HyMap) and Airborne Prism EXperiment (APEX) instruments for the synthetic and real data experiments, respectively. Noise levels were added to the synthetic data to simulate real imaging conditions. The real data used in this research are cloud free scenes acquired from the airborne campaigns. For performance assessment, we compared the proposed method with two state of the art methods. Our method performed better as it minimises the absolute error close to zero, only within 8–10 iterations. It thus suits existing operational chains where the number of iterations is considerable. Finally, the method is simple to implement and can be extended to address other atmospheric trace gases.

5.1 Introduction

A pixel of a three dimensional datacube recorded by a hyperspectral sensor comprises radiation energy measured at-sensor level, at hundreds of wavelengths. In the absence of the atmosphere, a reflectance obtained from the recorded radiation would be the spectral signature that characterises the underlying surface within the Instantaneous Field of View (IFOV) of the sensor. In the presence of Earth's atmosphere, however, the apparent reflectance differs from the target reflectance. This is primarily because of the complex interaction of the surface reflected radiation with the atmospheric constituents while propagating along the path from the target surface to the sensor. The interaction generates two main atmospheric effects: absorption by atmospheric gases (in particular water vapour and ozone) and aerosols (in the visible and near infrared spectral range) and scattering by aerosols and molecules.

In addition, on the path of beam to the sensor two major scattering components distort the at-sensor radiance: reflection by the surrounding area of the target pixel and the radiance back scattered by the atmosphere that did not interact with the surface.

An Atmospheric Correction (AC) algorithm is commonly applied to retrieve the radiance reflected at the surface from the at-sensor radiance. AC algorithms can be classified into scene based empirical algorithms and algorithms based on radiative transfer modelling. A comprehensive review is given in [34]. As radiative transfer modelling is mature for routine processing of hyperspectral image data [67], we will use its algorithms in this chapter.

In radiative transfer modelling, the target reflected radiance can be derived assuming a plane parallel geometry of the atmosphere, whereas the viewing and illumination geometry and total optical depth of the atmosphere are known.

For a reliable estimate of reflectance, the concentration of the atmospheric scatterers and absorbers should be available at the time of imaging. We consider estimation of the atmospheric Column Water Vapour (CWV) as the principal absorbent in the atmosphere. As CWV is highly variable in space and time, it is estimated from at-sensor radiance using image based methods. Such inference is adversely affected by methodological or numeric limitations. For instance, noise of at-sensor radiance is often poorly modelled. Besides, it is often assumed that the spectral response of a target surface across water absorption features exhibits a linear relation. Due to such limitations errors in CWV estimates likely propagate to reflectance estimates. Even though components of error have been evaluated and the appropriate corrections have been applied, uncertainty remains about the correctness of CWV estimates i.e. how well a CWV estimate represents the value of the actual CWV at the time of imaging.

The aim of this chapter is to present a new method for obtaining pixelwise CWV estimates. It specifically focuses on estimating CWV for an operational processing chain. The operational processing chain is implemented in the multi mission Processing, Archiving, and distribution Facility (PAF) for earth observation products [20, 39, 47]

Two scenarios are addressed: (a) CWV is incorrectly estimated in the chain due to the methodological or numeric limitations of the existing methods; (b) CWV is not estimated in the chain. This chapter presents a generic method such that for (a) a CWV estimate is used as an initial estimate. For this scenario, the CWV has to be calibrated. For (b), CWV is estimated without initial estimate. Combining the two scenarios, this results into a generic method to calibrate and estimate CWV as an augmented approach to existing processing chains.

5.1.1 Water vapour absorption

The propagation of surface reflected radiation and its interaction with water vapour in the atmosphere results in both absorption and scattering. During absorption, a photon transfers its energy to an atom or a molecule and is eventually removed from the radiation field.

For atmospheric correction of the image recorded by the sensor, CWV has to be provided to the radiative transfer code to simulate transmittance due to water vapour. The simulated water vapour transmittance is then used to correct for the influence of water vapour gas absorption.

A popular technique to estimate CWV is to use the sensor recorded radiance on a pixel by pixel basis. The technique consists of three steps: 1) identification of the spectral location of water absorption features, the so called measurement channels; 2) identification of the fraction of sensor recorded radiance without absorption features from any trace gas, the so called reference channels; 3) determination of CWV using a relation between reference and measurement channels. See [23, 32, 68, 81–83].

These methods are limited with respect to several assumptions. First, surface reflectance is assumed to vary with wavelength in a linear way; second, the effect of sensor noise is often not considered, and third, uncertainty is ignored that emerges from instrument characterisation. Such characterisation is generally performed in a laboratory prior to flight and includes linearity of various detectors, gains and offset of the sensor, and spectral response of the sensor channels. If any of the above assumptions fails then the resulting estimation of CWV produces residual effects in the absorption features [70]. This in turn prevents correct estimation of $\rho_t(\lambda)$. In [69], some of these limitations are highlighted. These assumptions are, however, quite reasonable as including all the influencing factors to model at-sensor radiance is analytically too complex and virtually impossible.

In this chapter, we estimate CWV from iterative estimations of target surface reflectance instead of from the recorded sensor radiance. The primary hypothesis of our approach is that a relation between estimates of CWV and strength of an absorption feature is linear while using the reflectance spectra. This linear relation facilitates pixelwise estimation of CWV while solving the objective function.

5.2 Estimation and Calibration Methodology

The methodology to estimate and calibrate CWV is based upon the estimates of reflectance spectra. In a hyperspectral datacube each pixel is represented with a vector of length B denoting the number of channels. Further, the actual reflectance ρ_t of the underlying surface within the IFOV is affected by atmospheric absorption and scattering. Then its estimate $\hat{\rho}_t$ obtained via AC is expressed as a linear model:

$$\hat{\rho}_t = \mathbf{C} \cdot \rho_t + n, \tag{5.1}$$

where the matrix $\mathbf{C} \in \mathbb{R}^{B \times B}$, is assumed to be diagonal, stores coefficients that model the deviation of $\hat{\rho}_t$ from ρ_t and n is the noise effect.

If the coefficients of $\mathbf{C}$ approach 1 then the estimates of CWV approach the actual CWV at the time of imaging, indicating a perfect match. If, however, the coefficients of $\mathbf{C}$ deviate from 1, then the error in the estimates of CWV increases.

Fitting the linear model (5.1) by least squares is equivalent to an l_2-norm optimisation problem where the deviation in the coefficients of $\mathbf{C}$ is minimised. In fact, it minimises the sum of the squares of the differences between $\hat{\rho}_t$ and ρ_t:

$$\min_{\mathbf{C}} \frac{1}{2} \left\| \mathbf{C} \cdot \rho_t - \hat{\rho}_t \right\|_2^2, \tag{5.2}$$

where the $l_2$norm equals $\|x\|_2 = \sqrt{\Sigma_i |x_i|^2}$.

While minimising the coefficients of $\mathbf{C}$ iteratively, indexed with k, we derive an offset $D_{\mathrm{CWV}_{i,k,\lambda}}$ for a pixel i at an absorption feature λ such that

$$D_{\mathrm{CWV}_{i,k,\lambda}} = 1 - \mathbf{C}(i,k,\lambda). \tag{5.3}$$

Note that $D_{\mathrm{CWV}_{i,k,\lambda}} < 1$ indicates that the estimate of CWV is lower than the actual CWV value i.e. an underestimate of CWV, whereas $D_{\mathrm{CWV}_{i,k,\lambda}} > 1$ implies overestimation. We use this offset to derive a CWV estimate at the next iteration as

$$\mathrm{CWV}_{i,k+1,\lambda} = \mathrm{CWV}_{i,k,\lambda} + D_{\mathrm{CWV}_{i,k,\lambda}}. \tag{5.4}$$

The $\mathrm{CWV}_{i,k+1}$ is used in turn to provide a new estimate $\hat{\rho}_{t,k+1}$. A fitting coefficient of any feature is specific for each absorption feature. Therefore, the proposed method may estimate multiple or incorrect values of CWV following global least squares optimisation. To avoid this, we used local least squares minimisation, i.e. limited to one specific absorption feature, namely the principal water absorption feature located at a wavelength of 0.944 µm. This feature affects only one channel, which makes the minimisation problem relatively simple as compared to using absorption features which involve multiple channels. In particular, (5.2) is iteratively solved until the coefficients of $\mathbf{C}$ at water absorption wavelengths are close to one. We use the coefficient range between 0.99–1.01 as a convergence criterion. The assumption here

is that the sensor is sensitive to the CWV values within $\pm$ 0.01 g cm^{-2}. This is quite a strict assumption and will be invalid for most of the existing hyperspectral sensors. The narrow range, however, allows estimating the finest achievable CWV value of a sensor. Depending upon the end product and sensor configurations, wider convergence ranges can be explored.

The parameter ρ_t is not known beforehand. Therefore, an estimate of ρ_t that serves as a reference to solve (5.2) is required. The slope of the reflectance spectra at the absorption feature located at 0.944 µm can be reconstructed using spectra value at reference channels by means of interpolation. We used cubic spline interpolation with cross validation between each pair of adjacent points to set the degree of smoothing. For a simple dataset this is more accurate than other polynomial interpolation methods [84]. A channel is considered as the reference channel if its signal is not being influenced by any atmospheric species and the signal to noise ratio is large.

5.3 Datasets

5.3.1 Synthetic and real surface reflectance data

The synthetic reflectance datacube is a hyperspectral image that contains 100×100 pixels shown in Figures 2.7–2.8 are used. In addition to the synthetic dataset, real dataset used here is shown in Figure 2.12a. To approximate the size of the simulated image the real scene is cropped into a sub scene shown in Figure 2.12b, at the size of the simulated image (100×100).

5.4 Experimental Setup

5.4.1 Profile of the atmospheric condition parameters for the forward modelling

To transform the synthetic surface reflectance image to at-sensor radiance, the atmospheric scattering and absorption conditions must be specified to the MODTRAN 4 radiative code. For this purpose we obtain Aerosol Optical Depth (AOD) and CWV using image based methods [8, 23] applied on a real APEX sensor datacube [1] flight. As an input to these methods, we used three radiance cubes of the Coast of Belgium. The reason for using the coastal area is its diversity, as the scene is covered by both sea and land. These types of scenes are interesting in terms of determining the uncertainty bounds of CWV and AOD for uncertainty exploration as over land and sea these parameters show high variation. This gives us an opportunity to validate the robustness of the proposed methodology for such complexity in the scene.

Using the three radiance cubes, for various illumination and viewing geometry, realistic uncertainty estimates of CWV and AOD are obtained. For CWV we obtained a range of 1.35–2.25 g cm^{-2}.

5.4.2 Pixelwise parameter configuration for the forward modelling

From the ranges of visibility, CWV, and r, we now derive their spatial pattern or variability over each pixel as we may and in real scenes for the forward modelling. For the simulated dataset, we assumed a gradual increase in CWV as we move row wise from the top of the simulated dataset to the bottom. We assigned a value CWV = 1.35 g cm^{-2} to the central pixel of the first row. The other pixels of this row received randomly sampled values from the $\mathcal{N}(1.35, \ 0.02)$ distribution. Likewise, with a step of 0.00912 g cm^{-2} we assigned values to pixels of each row. The resultant spatial variability of CWV is shown in Figure 5.1a.
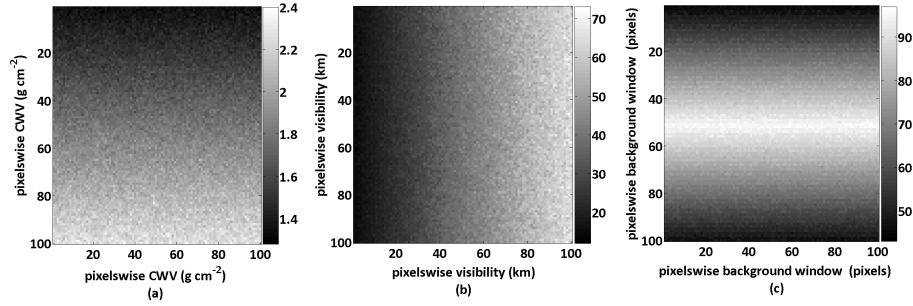
Figure 5.1: Spatial variability of CWV in (a), visibility in (b), and background window in (c) of the simulated image.

A similar procedure was applied to visibility which varies between 15 km and 65 km corresponding to an AOD range 0.47–0.13. The direction of variation is taken from the left of the simulated image to the right, i.e. column wise, with a step size of 2 km. To each pixel in a column, randomly sampled values from the $log\mathcal{N}(\mu, \ 0.02)$ with μ = 15 km for the first column are assigned. A log-normal distribution is used to approximate the actual distribution of visibility, being a skewed normal distribution (Figure 5.2).

For the adjacency effect, we considered the spatial variability of the classes present in the simulated image. From visual interpretation we found it spectrally more homogeneous at the center of the image than at the edges. Moving away from the center of the image, this homogeneity weakens (Figure 2.7). As spectrally homogeneous pixels cause a low adjacency effect, a significant influence on the center pixels can only come from pixels that are far away from the center. To address this spatial variability we set a large r = 91 pixels at the center of the image. The range gradually decreases as we move away from the center of the image and to a minimum of 41 pixels (Figure 5.1c). This value was determined from the variogram analysis.

5.4.3 Performance Discriminators

The quantitative assessment of the propagation of uncertainty originating from the AC parameters is measured by the signal to reconstruction error SRE
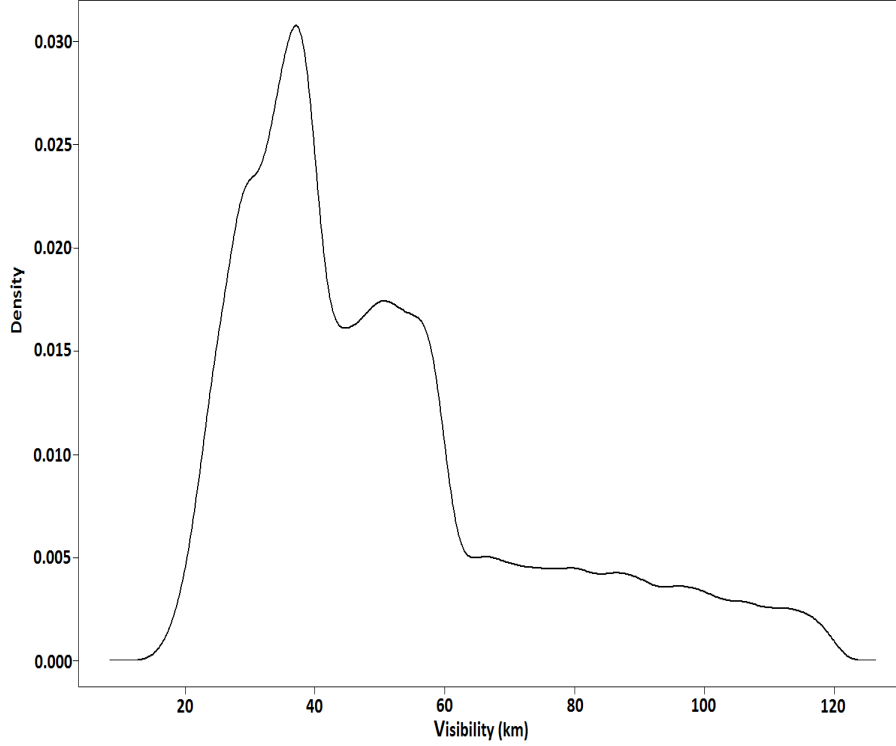
Figure 5.2: Marginal kernel density of visibility estimated using all (≈ 80 thousand) measurements from the image based method. A bandwidth of 1.56 is used as a parameter for kernel density estimation.

$\equiv \dfrac{E\left[\|x\|_2^2\right]}{E\left[\|x-\hat{x}\|_2^2\right]}$, where x is the reference signal and $\hat{x}$ represents its estimation. SRE provides information on the power of the signal with respect to the power of the error [77]. In all experiments, we report SRE measured in dB: SRE (dB)= $10\log_{10}$(SRE). Together with the SRE, we also measured the Root Mean Squared Error (RMSE) of the estimated and reference data.

5.5 Experimental Results

5.5.1 Experiments with simulated data

We first present the estimation results for CWV without considering noise, visibility, and adjacency. To achieve this, reference values of visibility and adjacency are used for each pixel (Figure 5.1b and Figure 5.1c).
First, the CWV is set equal to 1.6 g cm^{-2}. At those pixels where this value is used in forward modelling, reflectance estimation at absorption features leads to a correct estimate (Figure 5.1a). For other pixels, CWV = 1.6 g cm^{-2} will either underestimate or overestimate the reflectance at water

absorption features, depending upon whether it is smaller or larger than the value used in forward modelling. In Figure 5.3 a spectral plot of the class concrete from three different image regions is depicting correct, over-, and under-estimation scenarios. The pixelwise coefficient values indicate correct
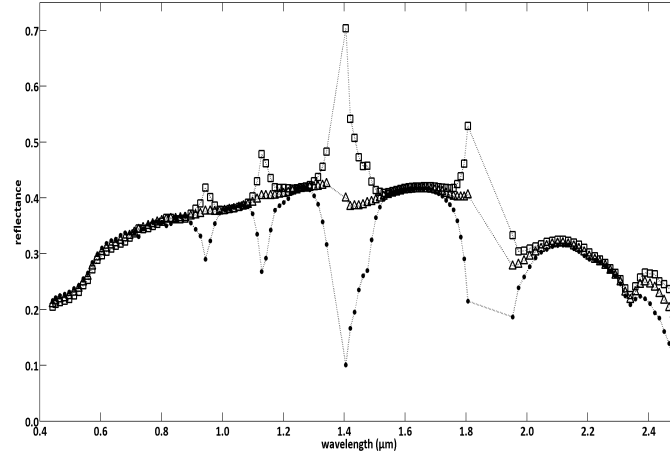
Figure 5.3: Overestimation (□), underestimation (●), and correction estimation (△) of reflectance for pixels where the class *concrete* dominates the initial water vapour value (1.6 g cm^{-2}).

estimation, overestimation, and underestimation of reflectance at absorption features as shown in Figure 5.4a. We observe coefficient values close to 1 at those pixels for which 1.6 g cm^{-2} is used in forward modelling i.e. the correct estimation of reflectance at the absorption feature. The robustness of our method can be observed in Figure 5.4b where small or no changes in CWV values are observed if their corresponding coefficients are close to one. For the other pixels, CWV values lower and higher than 1.6 g cm^{-2} are set for which the reflectance is over- or under-estimated, respectively.

From Figure 5.4c–d we observe small absolute errors between the previous and new CWV values, and between the new and reference CWV values if the reflectance is approximately correctly estimated. We observe large absolute errors if the reflectance is either over- or under-estimated. Iteratively estimating the CWV, we observe that the coefficients for each pixel are approaching one and that errors tend to be small. This is confirmed from Figure 5.5a–d, showing the CWV after 15 iterations.

Estimates converged after 20 iterations, as can be seen in Figure 5.6. As is evident from these, CWV estimation improved after each iteration. This implies that the solution is leading towards convergence. We already observed convergence after the 15$^{\text{th}}$ iteration, followed by small improvements for higher iterations.

Figure 5.7 presents the impact of CWV estimates on the reflectance estimation. We observe convergence of estimation of reflectance after 10 iterations. Continuing CWV estimation for 12 iterations did not result in any further improvement.

65

Figure 5.4: Pixelwise coefficient values at 0.944 µm (a), new estimates of CWV for each pixel (b), absolute difference between the new and previous estimations of CWV (c), and absolute difference between the new and reference CWV values (d) obtained after the first iteration.
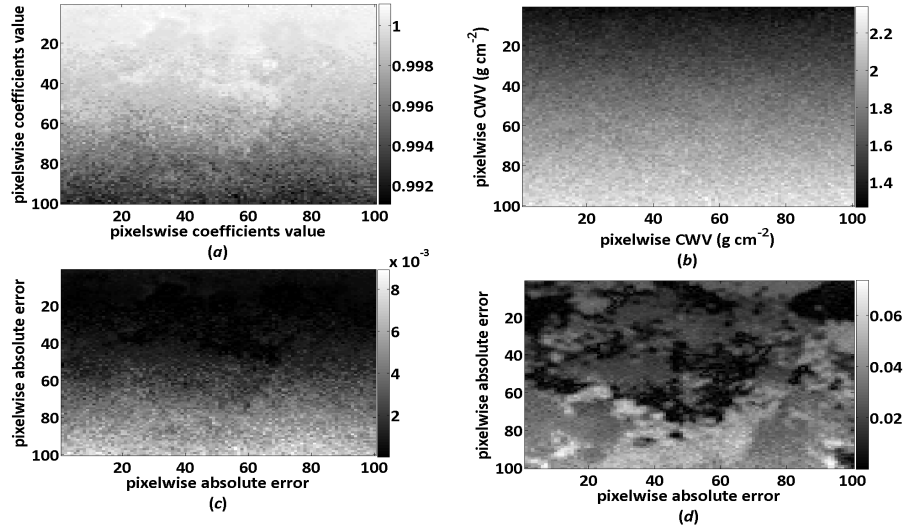
Figure 5.5: Pixelwise coefficient values at 0.944 µm (a), new estimates of CWV over each pixel (b), absolute difference between the new and previous estimations of CWV (c), and absolute difference between the new and reference CWV values (d) obtained after 15 iterations.

The objective of the second experiment was to include noise to the data and to perturb the spatial variability of visibility and adjacency while estimating
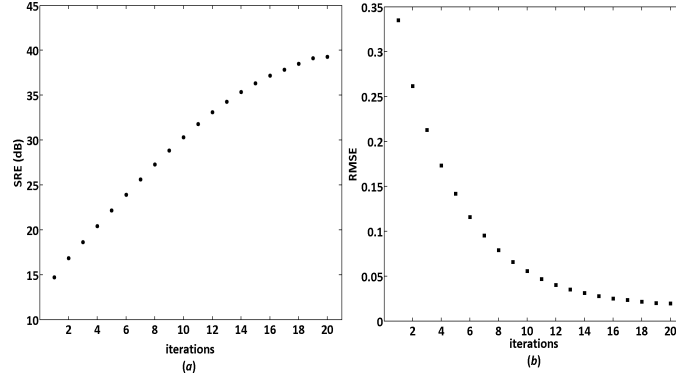
Figure 5.6: SRE (dB) and RMSE obtained during CWV estimation at successive iterations.
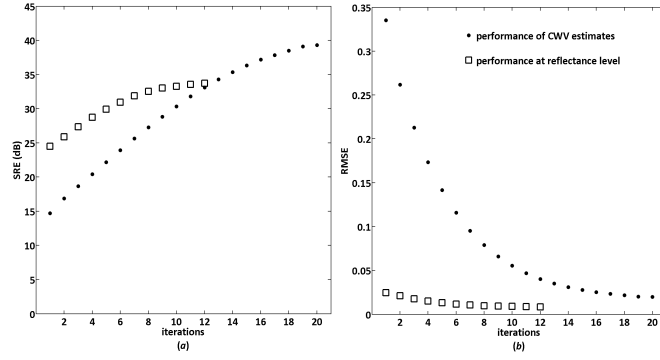
Figure 5.7: SRE (dB) and RMSE computed for CWV estimates at reflectance level for each iteration.

CWV. This scenario is similar to real imaging conditions. We applied rotation to the reference values of visibility and the background window size defined for each pixel (Figure 5.1b and Figure 5.1c). We also applied spatial smoothing to visibility values at each pixel after the rotation (Figure 5.9) and measured visibility and background window with diagonal spatial variability Figure 5.10–5.12. The reason to rotate and smooth the reference data and to realise diagonal variability, is to realise different ways to arrange visibility and adjacency in the spatial domain.

From Figure 5.13, we observe that noise and the arbitrary setting of visibility and the background window do not influence the estimation of CWV.

This experiment showed that the method presented in this chapter estimates CWV correctly if no prior information about the CWV is present, using an arbitrary value of 1.6 g cm^{-2} for CWV to start the estimation.
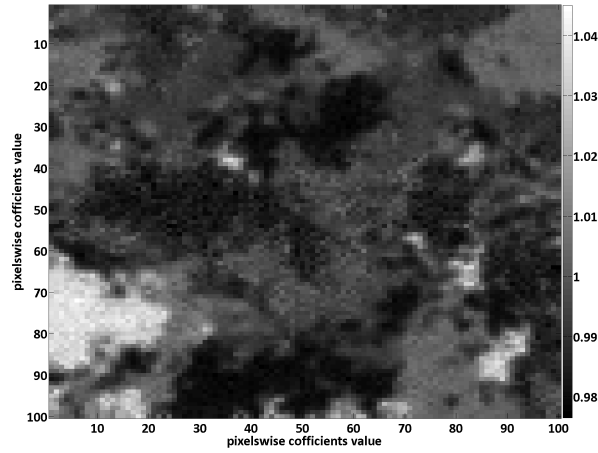
Figure 5.8: Coefficients values for CWV estimation after 10 iterations.
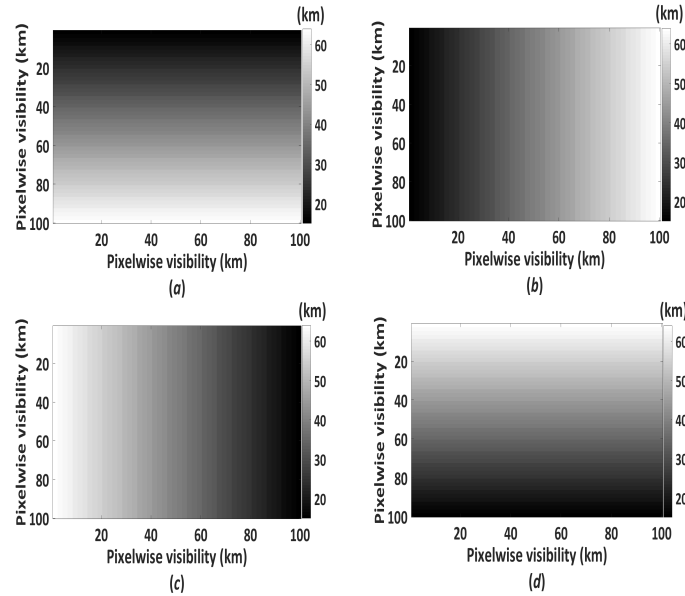
Figure 5.9: Pixelwise perturbation to the spatial variability of visibility with (a) 90 ° rotation, (b) no rotation, (c) 180 ° rotation, and (d) 270 ° rotation.

5.5.2 A comparison study

For performance assessment, we compared the new method with two state of the art methods. Method 1 (M_1) [23] is currently implemented in the CDPC and it estimates CWV using at-sensor radiance. Method 2 (M_2) [69] uses variations of a second order derivative algorithm (SODA) to assess the impact of atmospheric residual features on calculated surface reflectance spectra after
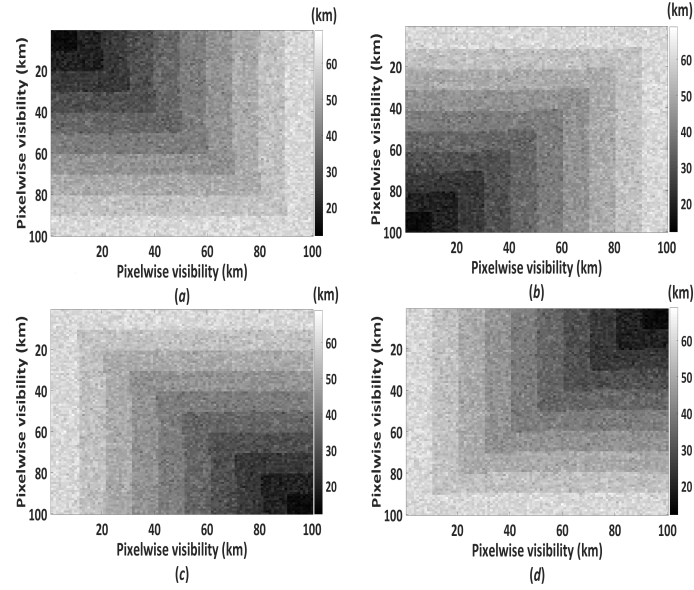
Figure 5.10: Pixelwise perturbation to the spatial variability of visibility with diagonal variation (a) top left to bottom right, (b) bottom left to top right, (c) bottom right to top left, (d) top right to bottom left.
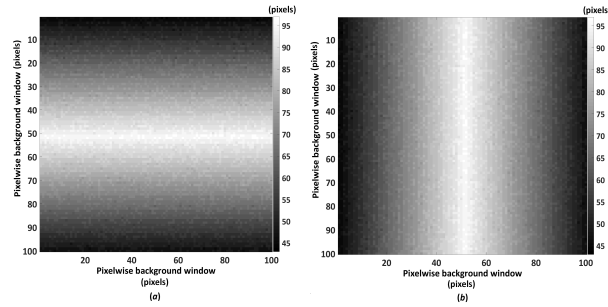
Figure 5.11: Pixelwise perturbation to spatial variability of the background window with no rotation (a) and 90 ° rotation (b).

AC. Performance is assessed in two ways. First, computational efficiency of the methods is compared with the proposed method, which is critical in operational processing chains. Second, the accuracy of each method is measured in terms of the absolute error.

M_1 estimates CWV in five steps. In each step it generates a number of synthetic at-sensor spectral signals each with a different amount of atmospheric CWV. After applying scaling to the signals and to the actual at-sensor radiance, it iteratively determines CWV values. These steps are repeated until convergence is reached. Applying M_1 on the synthetic image, results in convergence in 5–6 iterations. The accuracy in terms of absolute

69
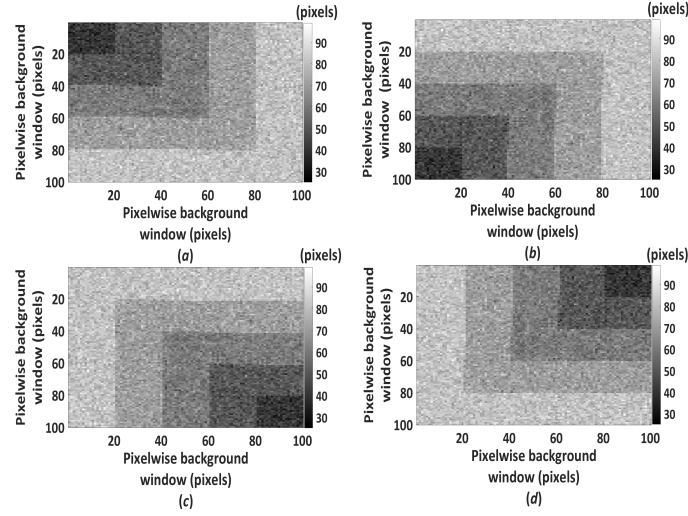
Figure 5.12: Pixelwise perturbation to the spatial variability of the background window with diagonal variation (a) top left to bottom right, (b) bottom left to top right, (c) bottom right to top left, (d) top right to bottom left.
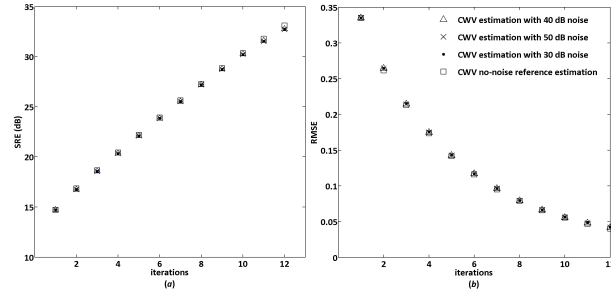
Figure 5.13: SRE (dB) and RMSE values obtained during CWV estimation for successive iterations with noise applied to the data and visibility and background window incorrectly set.

error is shown in Figure 5.14 where the maximum absolute error equal to 0.42. This is a large value as compared to 0.015 of the proposed model (Figure 5.15).

Method 2 calculates the surface reflectance multiple times using several CWV values to obtain a complete set of SODA values corresponding to the N surface reflectance values. This set is searched for the value that yields the lowest SODA, which is assumed to be the true value of the reflectance. Its corresponding CWV value is the estimated value. The minimum recommended value of N for M_2 is 50, e.g. M_2 estimates CWV with at least 50 iterations. The proposed method, however, requires 8–10
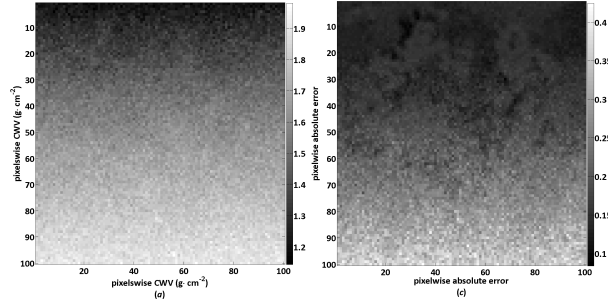
Figure 5.14: Pixelwise estimation of CWV using the method described in [23] (a) and in (b) the absolute error between the prior estimates of CWV in (a) and the reference CWV is shown.

iterations for pixelwise estimate of CWV with the absolute error minimised close to zero. Iteration times of M_2 and the proposed method are comparable because both use MODTRAN 4 to estimate various radiance components for AC (2.8). Also, M_2 first estimates CWV for a super pixel and in a later stage only it calculates pixelwise CWV. This leads to additional computational time that is critical within a processing chain.

5.5.3 Calibration of CWV

To analyze the performance of the proposed method while calibrating CWV, we first applied the method described in [23] to obtain initial estimate on CWV, shown in Figure 5.14. We then used these pre-estimates to our method. In Figure 5.15, CWV values are shown after 8 iterations. We observe that the proposed method successfully calibrates CWV.

An important observation during calibration and CWV estimation is that small improvements in CWV estimation do not further improve reflectance estimation after a relatively low number of iterations. The acceptable degree of change, i.e. the tolerance limit, is quantified by assessing the coefficients values at the 10th iteration (Figure 5.8). The CWV coefficients range between 0.9765 and 1.045. Using (5.3), we obtained a range of 0.0235–0.045 g cm^{-2}, specifying the tolerance limit in CWV estimation at the reflectance level.

5.5.4 Experiments for the real scene

For the real scene (Figure 2.12b), the range for CWV is between 1.5 and 2.4 g cm^{-2}. To check robustness of the proposed method, we selected at random CWV values within a narrow range of 1.5–1.6 g cm^{-2} for the first iteration. This pixelwise assignment avoids any prior knowledge about the scene. In Figure 5.16, the CWV values at each pixel are shown. For the real scene, reference data are not available. Thus, to validate the results we computed SRE (dB) and the RMSE at coefficient level for each pixel. The resultant performance plot is shown in Figure 5.17. From Figure 5.17, we observe that our method works well for the real dataset, as perceived from the coefficient's
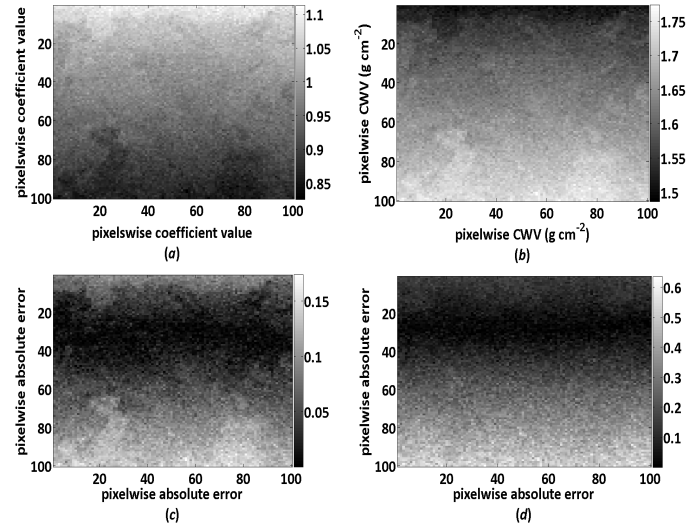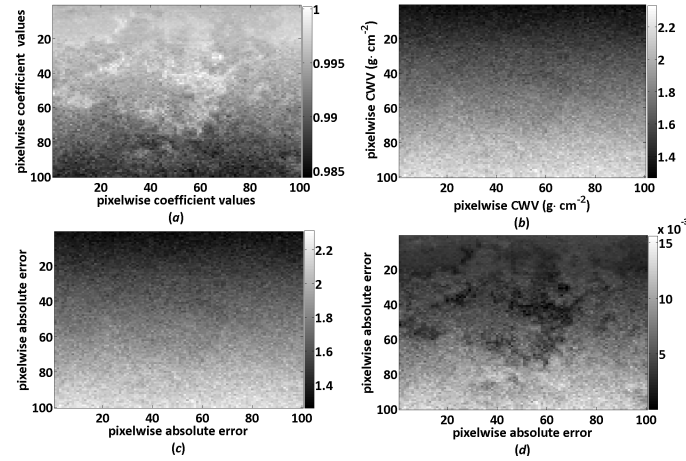
Figure 5.15: Pixelwise coefficient values at 0.944 µm (a), new estimates of CWV for each pixel (b), absolute difference between the new and previous estimations of CWV (c), and absolute difference between the new and reference CWV values (d) obtained after the 8th iteration.
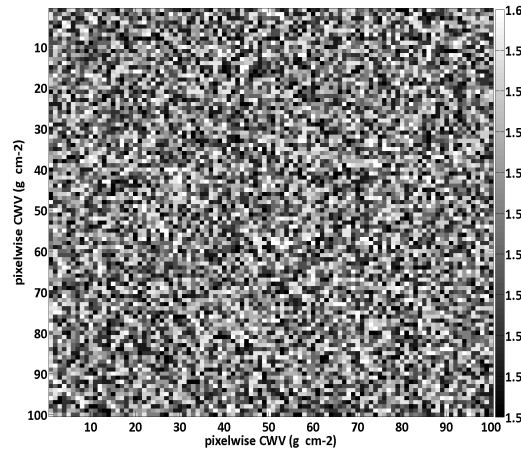

Figure 5.16: Absence of spatial variability of CWV for each pixel during the first iteration.

SRE and RMSE values. Also, we observe that the coefficients can be further improved i.e. estimation of CWV can be improved with more iterations. Performing seven iterations was sufficient, as applying more iterations show no substantial improvement in the coefficients values.
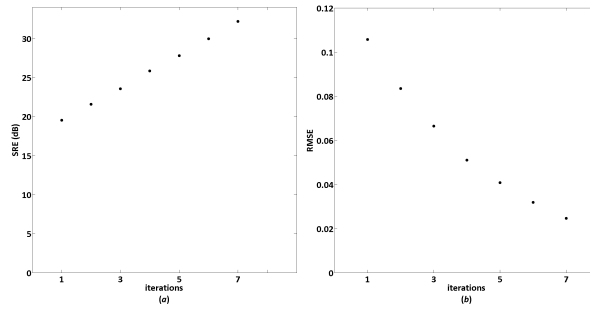
Figure 5.17: SRE (dB) and RMSE values using reference coefficients and estimated fitting coefficients obtained after seven iterations.

5.6 Discussion and Conclusion

This chapter presents a method that iteratively estimates CWV from reflectance spectra. It was applied to two hyperspectral sensors and there is no reason why it should also not be effective on other hyperspectral sensors. The method is simple to implement and can be extended to encompass other atmospheric trace gases. Its main benefit is that it is free from any assumptions that are usually made for at-sensor radiance based CWV estimation methods. For instance, it does not require the surface reflectance to be linear across spectral absorption features, nor does the spectral reflectance require pre-smoothing. A water profile obtained using the first assumption is only applicable to scenes that show a linear surface behaviour across different wavelengths. Such a linearity assumption is not valid for a reflectance spectrum of a complex and variable surface. The study further shows that the method works well under noisy conditions, i.e. if the second assumption does not hold.

With the proposed method, noise and incorrect estimation of AOD and the background window have little to no influence on CWV estimation. The reason is that these perturbations affect the amplitude of the spectra of a pixel and do not contribute to over- and underestimation of reflectance at the absorption feature. These findings are, however, only valid for high visibility values (<18 km). Under lower visibility conditions the effect of AOD on CWV retrieval is significant [85] and might have influenced the retrievals. Airborne campaigns, however, occur on clear days, that is when visibility is sufficiently high. This is for example evident from the visibility range observed for the real scene. These findings, well known from the literature, were helpful in validating our method. Thus, for strong water absorption features, the quality of CWV estimation is dominant over visibility, the adjacency effect and noise correlation. Also, for higher visibility values (>18 km), estimation of CWV can be performed without setting visibility and the background window to their actual values. This is consistent with our previous work [78] in which we found that CWV is the most important parameter at strong absorption features.

We further observed that our method works well when calibrating the prior

CWV estimates from the in built method to estimate CWV. From theoretical considerations this is what we might expect assuming that our method works well if prior estimates are available. We found that the number of iterations required for calibration is less as compared those needed for CWV estimation. For at-sensor based CWV estimation, influences of sensor noise and uncertainty arising from instrument characterisation are generally estimated in a laboratory prior to flight. Such estimation requires linearity of the detectors, gains and offset of the sensor, and spectral response of the sensor channels. If any of the above assumptions is not fulfilled, then the estimated CWV contains residual effects in the absorption features. Such effects cause the relation between CWV and the absorption feature to become non linear. Due to non linearity, an absorption feature cannot be related with CWV, as otherwise it will be prone to errors. In contrast to at-sensor based estimation of CWV, our method does not require a separate modelling to account for errors induced due to other influencing factors.

Estimating CWV based on reflectance spectra results into a linear relation between residuals at absorption features and CWV. This is because reflectance spectra are obtained after atmospheric corrections. It implies that all errors are already modelled and show their effects as over- or under-estimation at an absorption feature.

The pivot of the methodological choices in this chapter is its use in an operational chain. From this point of view, we note that the processing time to estimate CWV is low. In particular, the number of iterations to pixelwise estimate CWV is between 8 and 10. The reason why we could estimate CWV with less iterations is because for each iteration the new estimates are fed back into atmospheric correction procedures. This is different from estimation based upon reflectance spectra for a set of CWV values, followed by value optimisation. See e.g. the SODA method described in [69]. Also, solving CWV for each pixel for a set of CWV values in an operational chain environment requires a much higher computing time.

Estimation of AOD under uncertainty: an approach for hyperspectral airborne data

6

Abstract

A key parameter for atmospheric correction (AC) is Aerosol Optical Depth (AOD), which is often estimated from at-sensor radiance $(L_{\text{rs,t}}(\lambda))$. Noise, the dependency on surface type, viewing and illumination geometry cause uncertainty in AOD inference. We propose a method that determines pre-estimates of surface reflectance where effects associated with $L_{\text{rs,t}}(\lambda)$ are less influential. The method then identifies pixels comprising pure materials, the reference pixels, from the pre-estimates of reflectance. AOD values at the reference pixels are iteratively estimated by solving l_2-norm optimisation problem. Using the adjacency range function AOD is estimated at each pixel. We utilised the Hyperspectral Mapper (HyMap) and Airborne Prism EXperiment (APEX) instruments for the synthetic data and real data experiments. To simulate real imaging conditions, noise were added to the data. Our method minimises the estimation error in terms of AOD to 0.06–0.08 with signal to reconstruction error (SRE) equals 35 dB within 8–10 iterations. As AC is not sensitive for such small difference in AOD. It thus suits existing operational chains where the number of iterations are considerable. We compared the proposed method with dense dark vegetation based state-of-the-art method. With this reference method, large variability in AOD estimates were observed at DDV pixels, resulting in low SRE between 5–10 dB. For per pixel estimation of AOD, the performance further degrades as the reference method depends upon the number of DDV pixels in the scene. Finally, the proposed method is simple to implement and can be extended to address other AC parameters.

6.1 Introduction

A pixel of a three dimensional datacube recorded by a hyperspectral sensor comprises radiation measured by the sensor, at hundreds of wavelengths. In the absence of the Earth's atmosphere, a reflectance obtained from the recorded radiation is the spectral signature that characterises the underlying surface within the Instantaneous Field of View (IFOV) of the sensor. In the presence of Earth's atmosphere, however, the apparent reflectance differs from the target reflectance. This is primarily because of the complex interaction of the surface reflected radiation with the atmospheric constituents while propagating along the path from the target surface to the sensor. The interaction generates two main atmospheric effects: absorption by atmospheric gases (in particular water vapour and ozone) and aerosols (in the visible and near infrared spectral range) and scattering by aerosols and larger atmospheric gas molecules.

In addition, on the path of beam to the sensor two major scattering components distort the at-sensor radiance: reflection by the surrounding area of the target pixel and the radiance backscattered by the atmosphere.

An Atmospheric Correction (AC) algorithm is commonly applied to retrieve the radiance reflected at the surface from the at-sensor radiance. AC algorithms can be classified into scene based empirical algorithms and algorithms based on radiative transfer modelling. A comprehensive review is given in [34]. As radiative transfer modelling is mature for routine processing of hyperspectral image data [67], we will use its algorithms in this chapter.

In radiative transfer modelling, the target reflected radiance can be derived assuming a plane parallel geometry of the atmosphere, whereas the viewing and illumination geometry are known while some inference of the total optical depth of the atmosphere is required. For this inference, the concentration of the atmospheric scatterers and absorbers should be available at the time of imaging. Both absorbers, water vapour, and scatterer, aerosols and gas molecules, are highly varying in space and time. Thus, they are often estimated directly from satellite or airborne (remote) observations.

Despite substantial research efforts, pixelwise estimation of Aerosol Optical Depth (AOD) is still a challenge within Processing and Archiving Facilities (PAF). Most estimation methods use at-sensor radiance where inference of AOD is affected by retrieval uncertainties of the observed surface, atmospheric parameters, and instrumental errors. Additional processing to mitigate those errors are required to have a reliable reflectance product [86]. For instance, the use of reference pixels and averaging the data, say over 10×10 pixels, to reduce the influence of noise, requires a significantly larger data volume for analysis. In addition, the uncertainty induced by such processing is often ignored or impossible to model. Further, those estimation methods rely upon some surface characteristics. For instance, the dense dark vegetation (DDV) method to estimate AOD that is further developed in [8] is limited to pixels with dense vegetation. For scenes with DDV pixels that are clustered at a few locations, pixelwise estimation of AOD is challenging. These effects and limitations cause uncertainty in the estimation of AOD which likely propagates to reflectance estimates.

Uncertainty, acknowledging that the pixelwise true values of AOD are unavailable because of limited cognition and of limited granularity, is an inherent property of AOD estimation and cannot be eliminated completely. With more information about estimated spectra over a scene and its estimation conditions, however, we can minimise the uncertainty about AOD. This chapter focuses on such minimisation and is a continuation to our previous works [78], [87], [88]. We estimate AOD from estimates of surface reflectance at pure pixels, covering a single endmember and iteratively compare them with corresponding spectra stored in a library. Besides, the other novelty of this research is that the proposed method estimates AOD considering sensitivity of reflectance to AOD estimates resulting in less processing time. This is in contrast to those methods which focus on the estimation accuracy outside its sensitivity regime, such as the method of visibility estimation in [39] which is implemented in Atmospheric and Topographic Correction for Airborne Imagery (ATCOR 4) PAF of German Aerospace Center (DLR).

6.2 Datasets: synthetic and real datasets

The synthetic reflectance datacube is a hyperspectral image shown in Figure 2.7. The scene used in real data experiments is a scene of the APEX sensor [1] datacube (2014) over the Liereman area in Belgium, shown in Figure 2.12b. In order to observe the effect of the different noise levels we considered adding two types of noise to the datasets at two stages, refer to Section 2.5.4.

6.3 Experimental Setup

6.3.1 Profile of the atmospheric condition parameters for the forward modelling

To transform the synthetic surface reflectance image to at-sensor radiance, the atmospheric scattering and absorption conditions must be specified to the MODTRAN 4 radiative code. For this purpose, column water vapour value equals 2.0 g cm^{-2} are used and the sensor altitude is approximately 5 km above sea level resulting in pixel size of 4 m. We obtain Aerosol Optical Depth (AOD) using [8] applied on a real APEX sensor datacube [1] flight. As an input to these methods, we used three radiance cubes of the Coast of Belgium. The reason for using the coastal area is its diversity, as the scene is covered by both sea and land. These types of scenes are interesting in terms of determining the uncertainty bounds of AOD for uncertainty exploration as over land and sea AOD shows high variation. This gives us an opportunity to validate the robustness of the proposed methodology for such complexity in the scene.
Estimating the range of AOD is, however, not straightforward, because the method in [8] provides visibility (in km) as an output. In an operational

processing chain, true aerosol optical depth values that coincide with image acquisition are unavailable [38]. As an alternative, image based methods measuring visibility are often used to set aerosol optical profiles [8,20,32,37,39]. Here, we specify aerosol optical depth values and their corresponding visibility values so that it is useful to the remote sensing communities dealing with both quantities. Further, determining the range of visibility is challenging because the method in [8] is based on the dense dark vegetation (DDV) technique. One of its limitation is that visibility can only be estimated for those pixels. This implies that interpolation is required to estimate visibility for non-DDV pixels. Interpolation methods, however, can induce additional uncertainty in the parameter estimation, requiring a separate analysis. To avoid such uncertainty in further analysis, we first recorded all estimations of visibility for the DDV pixels for the entire scene. We then estimated a marginal density of those visibilities.

We used kernel methods to estimate the marginal density of visibility [89]. It is implemented using the kde function of the ks package [90] in R [91]. Parameters, e.g. bandwidth and number of modes, used to estimate the marginal densities, are obtained from the h.crit and nr.modes functions as illustrated in the silvermantest package [92,93]. In Figure 5.2 (Chapter 5) the estimated marginal density of visibility is shown.

From the marginal density, we observe a skewed distribution, which covers the range between 15 and 120 km. From our experience with the optical parameters in the CDPC, we realised that visibility >60 km has little impact on the estimate of reflectance. Thus, we limit the possible range of visibility to 15 to 65 km, which is approximately equivalent to the AOD range 0.47–0.13.

To simulate the average background radiance for the forward modelling i.e. for the adjacency effect, the spatial relationship is built upon the adjacency range (r) values corresponding to the background effect modelling. The average background radiance is as a function of: 1) the extent of the neighbourhood ($EN = (2r + 1, 2r + 1)$) of pixels that can spatially influence the target pixel and 2) a weight function within EN. Scattering conditions, spatial cross correlation, and uncertainty in estimating r have an influence on pixel wise estimation of EN. The optimal value of EN is determined in three steps. First, the minimum value of EN (pre-estimates) is determined by means of variogram fitting using the function vgram.matrix of the fields package in R. Second, a Monte Carlo simulation of photon scattering yielded the relative contribution of EN as a function of AOD, wavelength, aerosol Ångstrom coefficient, anisotropy factor of the Henyey-Greenstein aerosol phase function, the pre-estimated range, and sensor altitude. The photon simulation calculates the cumulative histogram of the adjacency contribution as a function of the distance. Third, using the cumulative histogram, the final EN is determined. To determine the Ångstrom coefficient average AERONET measurements of the Bruges area over the period 2014–2017 are used. Applying this method to the simulated and real scenes, the minimum value of EN was obtained as (41, 41) pixels. The maximum value was set to (91, 91) pixels. This limit was enforced by the spatial extent of the simulated and the real sub image (100 × 100 pixels).

6.3.2 Pixelwise parameter configuration for the forward modelling

From the ranges of visibility and EN, we now derive their spatial pattern or variability over each pixel as we may find in real scenes for the forward modelling. For the simulated dataset, a gradual increase in visibility which varies between 15 km and 65 km corresponding to an AOD range 0.47–0.13 was realised. The direction of variation is taken from the left of the simulated image to the right, i.e. column wise, with a step size of 2 km. To each pixel in a column, randomly sampled values from the log–normal(μ, 0.02) with μ = 15 km for the first column are assigned. A log-normal distribution is used to approximate the actual distribution of visibility, being a skewed normal distribution (Figure 5.2). The resultant spatial variability of visibility is shown in Figure 5.1b.

For the adjacency effect, we considered the spatial variability of the classes present in the simulated image. From visual interpretation we found it spectrally more homogeneous at the centre of the image than at the edges. Moving away from the centre of the image, this homogeneity weakens (Figure 2.7). As spectrally homogeneous pixels cause a low adjacency effect, a significant influence on the centre pixels can only be caused by pixels that are far away from the centre. To address this spatial variability we set a large EN = 91 pixels at the centre of the image. The range gradually decreases as we move away from the centre of the image and to a minimum of 41 pixels (Figure 5.1c).

6.3.3 Performance Discriminators

The quantitative assessment of the propagation of uncertainty originating from the AC parameters is measured by the signal to reconstruction error SRE $\equiv \frac{E\left[\|x\|_2^2\right]}{E\left[\|x-\hat{x}\|_2^2\right]}$, where x is the reference signal and $\hat{x}$ represents its estimation. SRE provides information on the power of the signal with respect to the power of the error [77]. In all experiments, we report SRE measured in dB: SRE (dB) = $10\log_{10}$(SRE).

6.4 Estimation Methodology

6.4.1 Linear model for calibration

In a hyperspectral datacube each pixel is represented with a vector of length B denoting the number of channels. Further, the true reflectance ρ_t of the underlying surface within the IFOV is affected by atmospheric absorption and scattering. Then its estimate $\hat{\rho}_t$ obtained via AC is expressed as a linear model:

$$\hat{\rho}_t = \mathbf{C} \cdot \rho_t + n, \tag{6.1}$$

where the matrix $\mathbf{C} \in \mathbb{R}^{B \times B}$, is assumed to be diagonal, stores coefficients that model the deviation of $\hat{\rho}_t$ from ρ_t and n is the noise effect.

79

If the coefficients of $\mathbf{C}$ approach 1 then the estimates of AOD approach the actual AOD at the time of imaging, indicating a perfect match. If, however, the coefficients of $\mathbf{C}$ deviate from 1, then the error in the estimates of AOD increases. Here, we assumed that the effect of other atmospheric parameters such as column water vapour is absent.

Fitting the linear model (6.1) by least squares is equivalent to an l_2–norm optimisation problem where the deviation in the coefficients of $\mathbf{C}$ is minimised. In fact, it minimises the sum of the squares of the differences between $\hat{\rho}_t$ and ρ_t:

$$\min_{\mathbf{C}} \frac{1}{2} \left\| \mathbf{C} \cdot \rho_t - \hat{\rho}_t \right\|_2^2, \tag{6.2}$$

6.4.2 Searching for reference library spectra

To generate pre-estimates of reflectance, we executed AC for k different visibility values that generated k reflectance cubes. As ρ_t is not known a priori, we need its estimation that serves as a reference to solve (6.2). Here, we assume that the ρ_t contains materials that are available in a spectral library, which is known a priori. Spectral unmixing of the k reflectance cubes provides k abundance maps indicating the fractional coverage of material present in each pixel. For unmixing realised through the linear mixture model, ρ_t can be expressed as a linear combination of the spectra of the endmembers, weighted by their abundances:

$$\rho_t = y = \mathbf{A} \cdot x + N. \tag{6.3}$$

Here, $\mathbf{A} \in \mathbb{R}^{B \times m}$ is the set of endmembers in the scene serving as a spectral library containing m pure spectra, $\boldsymbol{x} \in \mathbb{R}^m$ is the vector of corresponding fractional abundances compatible with $\mathbf{A}$, and $N \in \mathbb{R}^B$ is a noise vector. In this chapter, we assume that $\mathbf{A}$ is available a priori. Unmixing thus aims at identifying the atoms of $\mathbf{A}$ which are active in each pixel and their respective abundances. From the works of [73, 87], we found that Sparse Unmixing via Variable Splitting, Augmented Lagrangian and Total Variation (SUnSAL-TV) [73] a better option than SUnSAL without TV [77] for unmixing in the presence of noise. SUnSAL-TV takes intrinsic spatial smoothness into account as an important characteristic of natural scenes. We thus use SUnSAL-TV to obtain a solution with piece wise smooth transitions of the abundance fractions in neighbouring pixels. SUnSAL-TV solves the optimization problem:

$$\min_{X} \frac{1}{2} \|AX - Y\|_F^2 + \lambda_1 \|X\|_{1,1} + \lambda_{TV} \text{TV}(X), \tag{6.4}$$

where

$$\text{TV}(X) \equiv \sum_{\{i,j\} \in \varepsilon} \|x_i - x_j\|_1 \tag{6.5}$$

is a vector extension of the non-isotropic TV [74] and ε denotes the set of neighbours in the image. In (6.4), $\mathbf{Y} \in \mathbb{R}^{L \times n}$ is the observed data matrix with each column containing the observed spectrum at a pixel, $\mathbf{X} \in \mathbb{R}^{m \times n}$ is

the matrix of fractional abundances, $\|\mathbf{X}\|_F \equiv \sqrt{\mathrm{trace}\{XX^T\}}$ represents the Frobenius norm of $\mathbf{X}$ and $\|\mathbf{X}\|_{1,1} \equiv \sum_{i=1}^{n} \|x_i\|_1$, with x_i denoting the ith column of $\mathbf{X}$. The first term in (6.4) measures the data misfit, the second term forces the matrix of fractional abundances to be sparse, and the last term accounts for spatial homogeneity of the abundance maps. The parameters $\lambda_1 \geq 0$ and $\lambda_{TV} \geq 0$ are regularisation parameters. SUnSAL-TV introduces a set of new variables per regularizer and then uses the Alternating Direction Method of Multipliers (ADMM) [75] to solve the resulting constrained optimisation problem. In our experiments, we neglect sparsity as the spectral libraries employed will contain a small number of endmembers. We use SUnSAL-TV applying both the non-negativity constraint (ANC) and the sum to one constraint (ASC) [72] to solve the optimisation problem (6.4). Only spatial information is considered in SUnSAL-TV by setting $\lambda_1 = 0$ and $\lambda_{TV} > 0$. The ASC, often ignored due to signature variability [76] is used in this work as we assumed the set of image endmembers to be known and hence no sparsity is enforced on the vectors of abundances. The sparsity indicates that only a small number of endmembers are present inside a pixel and implies that only a few endmembers contribute to y, which is true for high resolution images.

6.4.3 Generating first visibility estimate

The methods then identifies the locations of pure pixels in the k reflectance cubes using the unmixing method discussed in Section 6.4.2. The retrieved abundance maps are, however, contaminated by retrieval uncertainties of the observed surface, atmospheric parameters, and instrumental errors. To accommodate for these uncertainties we select pure pixels in the k reflectance cubes with an abundance value between $1 - t_1$ and 1.0, where t_1 indicates a threshold value that can be set according to the target application and the sensor configurations. The spectra of each pure pixel are then evaluated for distortion in terms of shape as can be determined by the spectral angle mapper. The smaller the spectral angle the more similar a pixel is to a given library spectra. The corresponding visibility value providing the minimum spectral angle value and minimising (6.2) in least square terms is considered as estimates of visibility for the reference pixel.

6.4.4 Iterations

Using the pre-estimates of visibility, we re estimate $\hat{\rho}_t$ by minimising the $\mathbf{C}$ iteratively. We derive a new visibility value $D_{\mathrm{vis}_{i,e+1,\lambda}}$ for pixel i as

$$D_{\mathrm{vis}_{i,e+1,\lambda}} = \begin{cases} D_{\mathrm{vis}_{i,e,\lambda}} + q, & \text{if } C(\lambda, i) < 1 - t_2 \\ D_{\mathrm{vis}_{i,e,\lambda}} - q, & \text{if } C(\lambda, i) > 1 + t_2 \\ D_{\mathrm{vis}_{i,e,\lambda}}, & \text{otherwise} \end{cases} \tag{6.6}$$

where e is the iteration, $D_{\mathrm{vis}_{i,e,\lambda}}$ is incremented or decremented by q depending upon the conditions whether $\mathbf{C}$ is greater or less than the threshold value set by t_2. In this way, (6.2) is iteratively solved unless coefficient of $\mathbf{C}$ are close

to 1.0. We use the coefficient range between $(1-t_2)$–$(1+t_2)$ as a convergence criterion.

6.4.5 Interpolation of visibility

Following step 6.4.4, we obtained visibility values for the reference pixels, a spatial estimation method is applied to pixels that are spatially related to the reference pixels. A choice of spatial estimation method depends upon the type of surface, number of reference pixels, and size of the scene. The optimal number of neighbouring points to be used for spatial estimation depends upon the adjacency range (r) values corresponding to the background effect modelling. Three cases are distinguished to take account of the spatial locations of the reference pixels and to select spatial estimation method:

- Case 1: No overlapping of the neighbourhood. This case accounts for the scenario when within the neighbourhood of the target reference pixel EN_{target}, no other reference pixel fall.

- Case 2: Overlapping of the neighbourhood. This case accounts for scenarios when the neighbourhood of two or more reference pixels overlaps with EN_{target}. This occurs when two or more pixels falls within EN_{target}.

- Case 3: No reference pixel is available for spatial prediction. This case account for the scenario when to estimate visibility at the unsampled location, no reference pixel can be spatially linked using EN.

To solve Case 1, stationarity is assumed within EN_{target} implying the parameters, such as mean and variance, are same in all parts of EN_{target} and in all directions. Randomly sampled values from $P(\mu, \sigma)$ distribution are used to assigned values within EN_{target}. Here, μ is the value of visibility at the reference pixel and σ is used to accommodate for the insignificant spatial variability in visibility.

To solve Case 2, if w number of neighbourhoods $EN_1, \ldots, EN_w$ are overlapping then the pixels within $EN_1, \ldots, EN_w$ are combined to create a large field accommodating all the overlapping reference pixels. An interpolation is then applied on this large field using values of visibility at all the accommodated reference pixels as data points.

To solve Case 3, for a pixel that is spatially unrelated to any reference pixel, visibility value is estimation using a method, which is called binning. Binning comprises the following steps:

- step 1, find the spatial variation of visibility using values estimated for Case 1 and Case 2. We distinguished three types of spatial variation in visibility: a) vertical, b) horizontal, and c) diagonal. The specific variation is obtained by evaluating histograms of visibility values from line transects in vertical, horizontal, and diagonal directions. A bimodality or multimodality of the histogram from a specific line transect shows a strong variation of the visibility. For instance, if visibility is vertically varying (column wise) then the histogram obtained from the horizontal line transect shows bimodality or multimodality,

- step 2, the histogram that shows unimodality, that is direction of non-variation of visibility, is used to generate the distribution $P(\mu, \sigma)$. For instance, if visibility is vertically varying then histogram generated from the values from vertical line transect are used to generate the $P(\mu, \sigma)$,

- step 3, for vertical and horizontal variation, randomly sampled value from $P(\mu, \sigma)$ is assigned to spatially unrelated pixel. If visibility is diagonally varying then nearest neighbourhood method [94]S is used to interpolate the missing value.

6.5 Experimental Results

6.5.1 Specific choices of methods and values

The value of k is set to three and visibility values, 35, 45, and 55 km that generated three reflectance cubes are used. The values between 35 and 55 km (AOD values between 0.22 and 0.14) corresponds to moderate scattering conditions, which suits to the imaging conditions of airborne campaigns. The value of t_1 is set to 0.05 resulting in fraction between 0.95 and 1.0 as a criterion to select a pure pixel. From our experience with real images processed through CDPC we learnt that a reflectance estimate is not sensitive to visibility values that are less than 5 km. Therefore, to solve Case 1, we set $\sigma = 2.5$. To solve Case 2, we used a weighted average inverse distance to power method with power equal to 2 [94, 95]. To solve (6.6), we set $q = 5$ km as reflectance is sensitive to a difference of 5 km in visibility value. The convergence criterion is set between 0.99 and 1.01 by setting t_2 to 0.01. Such a narrow range allows estimating the finest achievable visibility value of a sensor. Depending upon the end product and sensor configurations, wider convergence ranges can be explored.

6.5.2 Visibility estimation for the simulated dataset

In Figure 6.1, visibility estimations after six iterations with missing visibility values and interpolated values using binning are shown. The spatial variability of estimated visibility values is comparable to the reference spatial variability as shown in Figure 5.1b.

We then estimated reflectance using visibility values for six iterations and observed a significant improvement in the estimations from iteration 1 to iteration 6. This is confirmed by the range of estimated visibility values in Figure 6.1. We observed a performance (SRE) of 35 dB after six iterations. After more iterations, however, further changes in visibility do not improve reflectance estimations. The reason is that a distinction between 5 and 7 km of visibility, which corresponds to AOD values between 0.06 and 0.08, is difficult to observe from the coefficient values because of the relatively low sensitivity of reflectance to variability in visibility. This is consistent with our previous work [87]. The optimal value of λ_{TV} corresponds to the maximum value of SRE (dB) is 0.01 for 30 dB noise case.
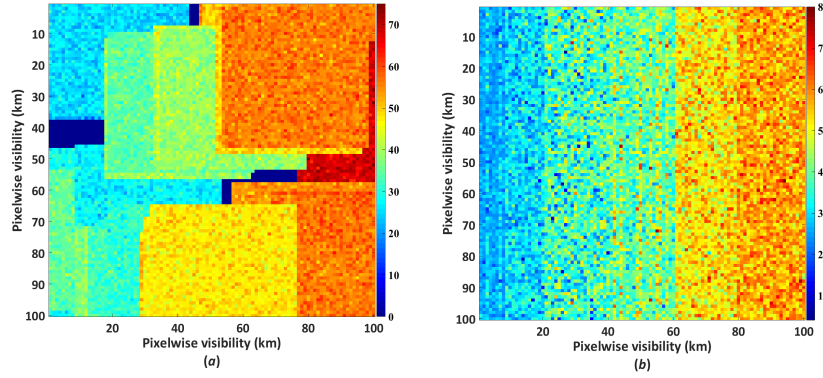
Figure 6.1: Interpolated values of visibility for 30 dB correlated noise case after the sixth iteration for the related pixels (a) and after binning for the related pixels.

6.5.3 Visibility estimation for the real dataset

For the real dataset, we cannot evaluate the performance of the method as there is no reference dataset available. Thus, we considered generating a pseudo reference dataset. To achieve this, we perform additional pre-processing on the available real reflectance cube (Figure 2.12b). The procedure to generate the reference reflectance cube comprises the following steps:

1. From the real at-sensor radiance cube, a set of endmembers is manually selected, whereas the corresponding pixels are called reference pixels.

2. Reflectance are estimated by performing AC using various pre-selected combinations of the atmospheric condition parameters.

3. From the estimated reflectance cubes spectra of the reference pixels are evaluated to check distortion in terms of shape (Section 6.4.2).

4. The reflectance cube that corresponds to the optimal combination serves as the reference cube.

5. The reference cube is transformed to obtain a pseudo at-sensor radiance cube using the MODTRAN 4 forward modelling simulations with the same atmospheric conditions as used in the forward modelling of the simulated dataset.

6. To the pseudo reference dataset we added 60 dB white Gaussian noise at-sensor level and 30 dB correlated noise to the estimated reflectance cube, as we did for the simulated dataset.

Figure 6.2 shows estimations of visibility after the second and fifth iterations. After five iterations we observed a saturation in performance and observed a performance (SRE) of 35 dB. We observed results of estimating visibility which are similar to findings with the simulated dataset. The optimal value of λ_{TV} corresponds to the maximum value of SRE (dB) is 0.01 for 30 dB noise case.

For performance assessment, we compared our method with the DDV pixels based visibility estimation method in [8]. Applying the DDV method on both
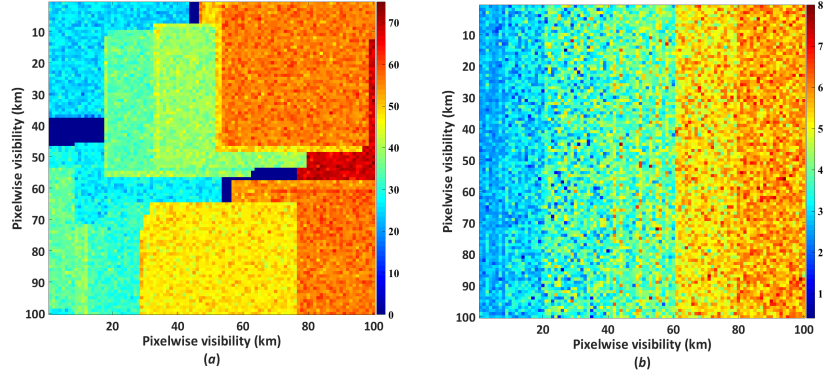
Figure 6.2: Interpolated values of visibility with a correlated noise 30 dB after the second iteration (a) and the fifth iteration (b) for the real dataset.

the datasets we obtained visibility at DDV pixels only. Further, the DDV based method is highly effected by illumination and viewing geometry even for the same flight line. We observed large variation in visibility estimation for adjacent pixels. For instance, at one pixel the estimated value was 32 km whereas at the next pixel it was 80 km. Because of large variation, low SRE (between 5 and 10 dB) was observed at the DDV pixels. From these results we conclude that the proposed method performs better than the reference method.

6.5.4 Experiment with spatial variability

This experiment aims to test the proposed method for different spatial variability of visibility and adjacency range. The spatial variations in the parameters are achieved by applying rotation to the reference values of visibility and the background window size defined for each pixel in Figure 5.1b and Figure 5.1c. We also applied spatial smoothing to visibility values at each pixel after the rotation. In addition, visibility and background window are varied diagonally, see Figures 5.9–5.12. The reason to rotate and smooth the reference data and to realise diagonal variability, is to generate different ways to arrange visibility and adjacency in the spatial domain. These variations of the parameters cover some relevant variation scenarios as can be found in real imaging conditions. These spatial variations are used in the forward modelling as explained in Section 6.3.2 to obtain various at-sensor radiance cubes. The proposed method is then applied on the at-sensor radiance cubes. The measured SRE values are close to 35 dB. In addition, the number of iterations to achieve convergence is between 5 and 7. This experiment showed that the method presented in this chapter estimates visibility correctly for various spatial variation of the parameters.

6.6 Discussion and Conclusion

This chapter presents a method that iteratively estimates aerosol optical depth (AOD) from the estimates of reflectance spectra. It was applied to a hyperspectral sensors and there is no reason why it should also not be effective on other hyperspectral sensors. The method is simple to implement and can be extended to encompass other atmospheric parameters. In particular, the number of iterations for a pixelwise estimation is between 6 and 8. The reason why we could use less iterations is that for each iteration the new estimates are fed back into atmospheric correction procedures where estimation converges when reflectance becomes insensitive to changes in visibility.

The method was applied on a small sub scene to differentiate between uncertainty and variability of the scene. This way larger real scenes divided into smaller sub scenes can be processed. The only challenge there would be to estimate background contribution for adjacency correction. It is because in larger scenes background radiance can originate from pixels that are outside the sub scene, especially for spaceborne data. This challenge can be resolved by first estimating the background contribution over the full scene.

To generate first estimates of reflectance, we considered 35, 45, and 55 km values of visibility for AC that generated three reflectance cubes. The values between 35 and 55 km (AOD values between 0.22 and 0.14) is a good preliminary range because airborne campaigns are often performed under low to moderate scattering conditions. This value can, however, be adjusted for space borne data or when the data is acquired under high scattering conditions. The proposed method is flexible for such changes.

The interpolation method used in this study is weighted average inverse distance to power method. This choice is based on the assumption that for a small scene pixels that are close to the target pixel are contributing more to the background radiation than those that are farther apart. Moreover, for small area the choice of the interpolation method is not critical. Further, any uncertainty arising from the interpolation method is tractable. It is because the spatial extent of neighbouring pixels is based upon an adjacency range, which is calculated using the three step process defined in Section 6.3.1. While performing experiments for different spatial variations of the two parameters, low performance was expected when visibility was diagonally varying. It is because to solve Case 3, when visibility is diagonally varying we used nearest neighbour method which could effect the performance of the method for various diagonal variations (as shown in Figures 5.9–5.12. No notable variation in the performance was observed. This is attributed to small area where choice of the interpolation method was not critical. More experiments can be performed to test the effect of spatial variation on the performance for larger scenes. As for small area the choice of interpolation method is not critical, the choice for the interpolation method made in this chapter is based upon the speed of the methods. For instance, the nearest neighbour interpolation method is faster than cubic or spline based interpolation methods.

The criterion to select the pure pixels is based upon the value of t_2 and the spectral angle values. To further automate this process factors such as scattering conditions, scene heterogeneity, path length from surface to the

sensor can be considered and more sophisticated methods that identifies reference pixels under uncertainty can be developed.

Further, the performance reported in this chapter are globally calculated. It is because the scattering conditions are set to moderate a suitable option for airborne data. For high scattering conditions, AOD between 1.42 and 0.47 which is equivalent to visibility between 4 and 15 km, it is important to report performance at each pixel. It is because for high scattering conditions, the presence of bright and dark targets in a scene becomes relevant [87]. Bright targets reflect more radiation energy than dark targets. Thus, the majority of at-sensor radiance consists of photons that are not scattered. In contrast, most photons for the dark targets are scattered. This contribution amplifies under high scattering conditions, resulting in performance degradation for dark targets.

Care should be taken when using visibility as a substitute to AOD, especially under high scattering conditions where AOD is highly sensitive to visibility. As long as airborne campaigns are not acquired under high scattering conditions, errors seen at 0–10 km visibility in [38] are unlikely to be experienced with operational remote sensing.

The present study can be further extended to take other atmospheric correction parameters into account. Deeper effects of the Ångstrom coefficient can be investigated. It would be interesting to use the proposed method in an application where various methodological choices can be fine tuned to suit the application. Experiments with spectral libraries containing a large number of spectral signatures are recommended but those are beyond the scope of the current study.

Synthesis

The goal of this chapter is to synthesise the key findings from this research and put them into a broader perspective. The chapter begins with positioning the developed framework. It is followed by answers on the aims and the objectives and the conclusions. Finally, suggestions and recommendations for future extension of this research are presented.

7.1 Positioning the research

The reflected radiance of a target surface measured by a sensor mounted on a distant platform comprises of surface reflected radiation and radiation components such as path radiance and background radiance. Measuring the desired surface reflected radiance using at-sensor radiance can then be formulated as an inverse problem. This consists of using the at-sensor radiance to infer the values of the atmospheric condition parameters. These characterise the interaction of surface reflected radiation with atmospheric constituents to estimate surface reflected radiance. The inverse problem is implemented by means of radiative transfer based atmospheric correction.

In an operational processing and archiving facilities (PAF) of satellite earth observation systems, atmospheric correction has been important since the products are typically used for the global and multi temporal estimation as a time series analysis of the radiation budget, environmental parameters or quantitative atmospheric remote sensing applications.

Airborne remote sensing has been used for recurrent conventional photogrammetric mapping tasks resulting in mixed sensor type time series of airborne imagery. Examples are agricultural acreage estimation, dry matter productivity estimation and mapping of vegetation stress. The analytical tools generating such products often need quantitative measurements. Also, in applications where data from different sensor systems are to be fused, surface reflected radiance is often used. As a consequence, atmospheric correction has been increasingly important in airborne remote sensing PAFs. It reduces the effects of scattering and absorption by gases and aerosols in the atmosphere between the Earth's surface and the sensor to estimate surface reflectance from the observed signal. Depending upon the sensor platform and product requirements, various types of reflectance products can be generated using atmospheric correction. The airborne data processing

facility of the CDPC estimates the hemispherical directional reflectance factor (HDRF) product. The HDRF product is suitable to airborne sensors as their IFOV is usually small, for instance, 0.129° for the HyMap airborne hyperspectral scanner.

When applying radiative transfer based atmospheric correction, reflectance is estimated assuming a plane parallel geometry of the atmosphere, whereas the viewing and illumination geometry and total optical depth of the atmosphere are assumed to be known. For a reliable estimate of any reflectance product, the concentration of the atmospheric scatterers and absorbers, i.e. the atmospheric condition parameters, should be available at the time of imaging.

The methodology implemented in the CDPC to quantify pixelwise water vapour concentration is based upon [23]. It estimates column water vapour concentration in five steps. Each step generates a number of synthetic at-sensor spectral signals each with a different amount of atmospheric column water vapour concentration. After applying scaling to the signals and to the actual at-sensor radiance, it iteratively determines column water vapour concentration values. These steps are repeated until convergence is reached. The results of this methodology, are then used in the subsequent atmospheric correction.

In the CDPC and other PAFs such as the Atmospheric and Topographic Correction for Airborne Scanner Data (ATCOR) of the German Aerospace Centre (DLR), the methodology implemented to quantify aerosol optical depth is based upon [8]. It measures a parameter called visibility for characterisation of aerosol optical depth. The method first retrieves at-sensor radiance of dense dark vegetation pixels. It then estimates visibility over those pixels using a relation between at-sensor radiance observed at the shortwave infrared (SWIR) band, the red band, and the blue band. Visibility is related with the aerosol extinction coefficient at 550 nm via (2.9). At 550 nm the contributions of molecular depth, ozone depth, and trace gases usually are small and aerosol optical depth is the main contributor to the total optical depth of the atmosphere, i.e. the extinction coefficient at 550 nm is directly related to aerosol optical depth. The sources of uncertainty in the atmospheric condition parameters are discussed in Section 2.1.

Another important parameter is the adjacency range. This parameter determines the contribution of the radiance originating from the background of a target pixel in the observed target spectrum by means of atmospheric scattering. The adjacency effect is prominent if the reflectance of neighbouring surface deviates strongly from the reflectance of the target surface. The additive effect of adjacency can be simulated by taking into account a band specific average radiation of some neighbouring pixels. Here, the average background radiance is a function of: 1) the extent of the neighbourhood of pixels that can spatially influence the target pixel and 2) a weight function within the neighbourhood. Scattering conditions, spatial cross correlation, and uncertainty in estimating the adjacency range all influence the pixelwise estimation of the extent of the neighbourhood. The parameter with the highest uncertainty is given by the target specific kernel size of the filter applied on the image. The adjacency kernel size can be

adjusted by the CDPC operators. As a rule of thumb, the adjacency effect can be simulated by a kernel covering a spatial extent of half the flight height. Further, this kernel size can be fine-tuned if in-situ spectroradiometer measurements are available. In Section 6.3.1, an approach to determine the extent of the neighbourhood is presented. Uncertainty in the atmospheric condition parameters, as formulated in Section 2.1, propagates to reflectance estimates. As the uncertainty in the reflectance estimates further propagates to application products, confidence bounds are required for a further data analysis. The MODTRAN interrogation technique is used in the CDPC to estimate various radiative components of (2.8) for atmospheric correction. To execute this technique, atmospheric condition parameters such as concentration of trace gases and amount of aerosol are required. In addition, various other types of parameters are used to specify the imaging conditions. For instance, 12 ambient band model species are included in MODTRAN 4, such as water vapour, carbon dioxide, ozone, and oxygen and several heavy molecular gases. Likewise, single scattering and multiple scattering options and aerosol types that are related to aerosol settings have multiple options to model radiative transfer. These parameters are linked to the background processes through physical radiative transfer models. In this research, an uncertainty free physical structure of the radiative transfer is assumed.

In MODTRAN 4, one way to model the aerosol loading is by specifying visibility, selecting a specific aerosol model, and specifying how water vapour influences aerosol. By selecting an aerosol type a pre-defined aerosol model is used in the background processes assuming it is present in the atmosphere at the time of imaging. For example, a rural aerosol model in MODTRAN 4 represents regions of the atmosphere not impacted by urban or industrial processes. It is assumed to be composed of 70% water soluble material and 30% aerosol for a specific water vapour concentration. One may include uncertainty about the ratio of water soluble material and dust by quantifying it with Monte Carlo methods and using a probability distribution of each aerosol component. To do so, either regional data measured in a study area, or global data are required. With such data the aerosol loading can be directly set in many radiative transfer models, e.g. using a multimodal log-normal distribution. Besides, a separate uncertainty analysis of the ground measured data would be required. A full uncertainty analysis might be interesting for the future but is at present out of the scope of this research.

The three atmospheric condition parameters studied in this research i.e. column water vapour, aerosol optical depth, and adjacency range cover the absorption, the scattering, and the adjacency range type parameters. For the developed framework, a large number of atmospheric condition parameters can be processed to quantify the propagation of uncertainty, allowing scalability. This implies that an analysis with, say column water vapour, provides insight into handling other trace gases such as ozone and carbon dioxide. In this way, the thesis allows to develop a framework that is scalable to incorporate additional parameters and applications.

A peculiar aspect of this thesis is that it is concentrated on processing of images within a PAF. A PAF cannot readily adapt to new methods or models at the pace with which these emerge. In particular, existing relations between

applications and methods or models may change or even disappear. The framework presented in this thesis provides a basis for existing methods to allow interactions with the involved parameters, and assists in analysing and representing uncertainty propagation. In this way, such a framework is useful for a PAF in the long run.

The developed framework comprises of methods to optimise uncertainty of the parameters as described in Chapter 5 and Chapter 6. For a PAF acknowledging and quantifying the uncertainty has been a challenge. One of the important reasons has been the large data volume processed through a PAF. Application of Monte Carlo based methods on a large data volume is computationally intensive and time consuming. The methods developed in this research to optimise the parameter's uncertainty are, therefore, important in a PAF.

7.2 General conclusions

The thesis has four components: atmospheric correction (AC), retrieval of atmospheric condition parameters, unmixing as an application, and propagation of uncertainty from the atmospheric condition parameters to abundance maps. AC requires knowledge on the atmospheric state expressed by parameters. Their values are affected by uncertainties that propagate to the abundance maps via reflectance estimates. The thesis identifies trends in performance degradation caused by the gradual shift in parameter values from their true value. This relationship leads to quantifying the propagation of uncertainty from the parameters to the abundance maps via reflectance estimates. The general conclusions are the following.

1. Sensitivity analysis (SA) gives information on the contribution of sources of variation to dispersion in the output. In Chapter 6 it was shown that a distinction between 5 and 7 km of visibility, which corresponds to AOD values between 0.06 and 0.08, is difficult to observe from the coefficient values because of the relatively low sensitivity of reflectance to variability in visibility. From these experiments, I conclude that sensitivity of an application product to parameter uncertainty provides a useful convergence criterion while estimating the involved parameters.

2. From the lens of uncertainty propagation, an estimate of an atmospheric condition parameter using reflectance is more efficient than using at-sensor radiance. I found that it is possible to establish a linear relationship between reflectance and an atmospheric condition parameter. This implies that the parameter's estimation problem can be solved with a linear equation as shown in Chapters 5 and 6.

3. Sources of uncertainty and propagation of uncertainty are useful measures to determine the performance of a method or a model. As shown in Chapter 4, deviation of a parameter value from its true value due to uncertainty can cause deflection in the estimates of reflectance. From this viewpoint, uncertainty analysis should be integrated to assess the performance of a method or a model that uses reflectance as

an input.

4. Surface albedo is an important source of uncertainty while quantifying the propagation of uncertainty. Chapter 4 concludes that the surface albedo influences the quality of the spectra retrieved via AC, independently from the effects of the uncertainty in the AC parameters and scattering conditions.

5. A SA of the sensor noise and processing noise is as important as a SA of model parameters. Such an SA provides useful information whether noise has any impact on the outcome or not, this further helps in building an appropriate noise model. As shown in Chapter 5, for a strong water absorption feature the effect of water vapour uncertainty is higher than the effect of the processing noise. For such scenarios, measuring the water vapour should be carried out in a more intensive way than building a noise model.

7.3 Detailed conclusions

The main objective of this research was to develop a methodology for uncertainty propagation originating from atmospheric condition parameters to reflectance and to unmixing products in a layered processing system. To achieve the main objective of this research, it is divided into four sub-objectives.

7.3.1 Sub-objective 1

The first sub-objective aimed to quantify the importance of the absorption effect due to water vapour concentration and the importance of the scattering effect due to aerosols in terms of aerosol optical thickness in estimating reflectance via the atmospheric correction modelling (AC).

The sensitivity of the estimated reflectance to water vapour and aerosol optical depth depends upon the wavelength, strength of the water absorption feature, parameter uncertainty, and surface albedo. Further, aerosol optical depth is an important parameter for atmospheric correction. It is more important than column water vapour even at the principal absorption feature. For bright targets, the importance of column water vapour and aerosol optical depth depends upon the strength of the absorption feature. Sensitivity indices for aerosol optical depth were high for the non-absorption water bands.

To quantify the uncertainty propagation, a probability distribution for each parameter was defined. To do so, I incorporated the Fourier Amplitude of Sensitivity Test (FAST) and the extended FAST (e-FAST) methods. These methods facilitate the use of a large range of probability distributions, a search curve based sampling method, and calculation of Monte Carlo simulations. This suits to scenarios that require expert knowledge on a parameter. They also facilitate a sensitivity analysis of the output to each parameter. For scenarios that require expert knowledge, the framework enables probability distributions and sampling schemes defined by experts.

These findings help in setting various components of the framework. For instance, if an effect of an absorption parameter is a function of wavelength, strong absorption features can be used to calculate probabilistic information. This information is useful to build a method for another absorption parameter. In addition, using the search curve, a method in the framework is devised that provides interaction among the atmospheric condition parameters. For instance, interaction between aerosols and water vapour in the atmosphere can be modelled using the probability distribution of the parameters.

7.3.2 Sub-objective 2

The second sub-objective aimed to provide a qualitative and quantitative analysis of uncertainty propagation from values of the AC parameters to the fractional abundances derived from unmixing. It considered three atmospheric condition parameters. The uncertainty propagation method was applied to the HyMap hyperspectral sensor and there is no reason why it should also not be effective on other hyperspectral sensors. The method is simple to implement and can be extended to encompass other atmospheric trace gases and other atmospheric condition parameters.

A useful insight was found into how a given uncertainty affects the performance at reflectance level and at subsequent applications (spectral unmixing). I found that the atmospheric settings, noise, and the parameter's deviation have a combined effect on the propagation of the uncertainty.

I further noted that the variability in aerosol optical depth influenced the amplitude of the retrieved spectra. Its influence was not linear, in the sense that deviations from the reference spectra depend upon the surface reflectance of the target, the wavelength, and the scattering conditions. This observation is consistent with the findings for sub-objective 1.

Uncertainty in column water vapour affected the performance more severely than uncertainty in aerosol optical depth, because variation in column water vapour significantly distorted the spectral shape. The uncertainty in aerosol optical depth, however, strongly influenced the quality of the abundance maps if the endmembers were spectrally similar, especially if high atmospheric scattering occurred.

Further, the propagation of uncertainty was always higher for abundance estimates than reflectance estimates. This might be due to the combined uncertainty i.e. propagated uncertainty from the reflectance estimates and the inherent limitations of the unmixing method and their numerical approximations.

I found a difference in uncertainty propagation for high scattering conditions due to the presence of bright and dark targets in a scene. Bright targets reflect more radiation energy than dark targets. Thus, the majority of the at-sensor radiance consists of photons that are not scattered. In contrast, most of the photons for the dark targets are scattered. This contribution amplifies under high scattering conditions resulting.

Effective ranges of column water vapour and aerosol optical depth were within the 90% confidence interval from their true values. Outside this interval, the performance strongly degraded and saturated to a low value. Such effective

ranges could be useful for calibrating the value of the two parameters in optimising the performance under uncertainty. Under high noise conditions (SNR = 30 dB), both spectral quality and unmixing performance are strongly degraded, whereas the effect of scattering becomes less important. A high noise level therefore has a weaker influence on the performance trend for column water vapour than for aerosol optical depth, resulting in a distinct peak of the performance. Thus, in the presence of high noise in airborne data, uncertainty in atmospheric correction parameters and data noise affect the final products jointly.

These finding helped in addressing various fundamental concerns pertaining to quantifying the propagation of uncertainty. For instance, finding the best ways to report the propagation of uncertainty. To benefit various types of applications and researchers, involved in a PAF, the propagation of uncertainty was reported by measuring various quantities at pixel level and at scene level. In addition, incorporating the effect of noise, surface albedo in the propagation analysis, and including the unmixing application provided a greater insight into how to incorporate an application to the propagation framework.

7.3.3 Sub-objective 3

The third sub-objective was to estimate and calibrate column water vapour under uncertainty. It aimed to develop a method to iteratively estimate column water vapour from pre-estimates of reflectance spectra under uncertainty. I found several benefits of estimating the column water vapour using pre-estimates of reflectance than estimations based upon at-sensor radiance. This is because a pre-estimated reflectance spectrum is obtained after atmospheric correction where all errors are already modelled and show their effects as over- or under-estimation at an absorption feature.

Further, I found that under the influence of noise and incorrect estimation of aerosol optical depth and the background window have little to no influence on column water vapour estimation. The reason is that these perturbations affect the amplitude of the spectra of a pixel and do not contribute to over- and under-estimation of reflectance at the strong absorption feature. These findings are, however, only valid for low to moderate scattering conditions (>15 km). Under high scattering conditions the effect of aerosol optical depth on column water vapour retrieval is significant and have influenced the retrievals. Airborne campaigns, however, occur on clear days, that is when scattering conditions are moderate to low. Thus, for strong water absorption features, the quality of column water vapour estimation is dominant over visibility, the adjacency effect and the correlated noise.

7.3.4 Sub-objective 4

The fourth sub-objective was to estimate aerosol optical depth under uncertainty. The objective aimed to iteratively estimate aerosol optical depth from the estimates of reflectance spectra. I noted that the processing time to estimate aerosol optical depth under uncertainty was low. It is

95

because for each iteration the new estimates are fed back into atmospheric correction procedures where estimation converges when reflectance becomes insensitive to changes in visibility.

The method was applied on a small sub-scene to differentiate between uncertainty and variability of the scene. This way larger real scenes divided into smaller sub-scenes can be processed. The only challenge there would be to estimate background contribution for adjacency correction. It is because in larger scenes background radiance can originate from pixels that are outside the sub scene, especially for spaceborne data. This challenge can be resolved by first estimating the background contribution over the full scene. To generate first estimates of reflectance, we considered 35, 45, and 55 km values of visibility for AC that generated three reflectance cubes. The values between 35 and 55 km (AOD values between 0.22 and 0.14) is a good preliminary range because airborne campaigns are often performed under low to moderate scattering conditions. This value can, however, be adjusted for space borne data or when the data is acquired under high scattering conditions. The proposed method is flexible for such changes.

The interpolation method used in this study is weighted average inverse distance to power method. This choice is based on the assumption that for a small scene pixels that are close to the target pixel are contributing more to the background radiation than those that are farther apart. Moreover, for small area the choice of the interpolation method is not critical. Further, any uncertainty arising from the interpolation method is tractable. It is because the spatial extent of neighbouring pixels is based upon an adjacency range, which is calculated using the three step process defined in Section 6.3.1. While performing experiments for different spatial variations of the two parameters, low performance was expected when visibility was diagonally varying. It is because to solve the case, when visibility is diagonally varying we used nearest neighbour method which could effect the performance of the method for various diagonal variations. No notable variation in the performance was observed. This is attributed to small area where choice of the interpolation method was not critical. More experiments can be performed to test the effect of spatial variation on the performance for larger scenes. As for small area the choice of interpolation method is not critical, the choice for the interpolation method made in this thesis is based upon the speed of the methods. For instance, the nearest neighbour interpolation method is faster than cubic or spline based interpolation methods.

Care should be taken when using visibility as a substitute to AOD, especially under high scattering conditions where AOD is highly sensitive to visibility. As long as airborne campaigns are not acquired under high scattering conditions, errors seen at 0–10 km visibility in [38] are unlikely to be experienced with operational remote sensing.

The findings from sub-objective 3 and sub-objective 4 answered the question whether uncertainty can be minimised. I found that with more information about estimated spectra over a scene and its estimation conditions uncertainty about the parameters can be minimised, called optimisation. In addition, for uncertainty propagation, parameter estimates from the pre-estimates of reflectance is more useful than using at-sensor radiance

based parameter estimates. At-sensor based inference of a condition parameter is affected by retrieval uncertainties and instrumental errors. Additional processing to mitigate those errors is required to have a reliable reflectance product. For instance, the use of reference pixels and averaging the data, say over 10×10 pixels, to reduce the influence of noise, requires a significantly larger data volume for analysis. In addition, uncertainty induced by such processing is often ignored or impossible to model. Further, most estimation methods rely upon some surface characteristics. Finally, I found that the uncertainty propagation with the categorical parameters such as aerosol type is challenging, because for a Monte Carlo analysis, sampling of a categorical parameter is difficult to implement. The analysis in this research is restricted to the absorption and scattering type parameters. Nevertheless, I tested three of the four aerosol type aerosol models available in MODTRAN 4: maritime, rural, and urban. An univariate analysis with the three aerosol models under various atmospheric conditions gives nearly identical uncertainty propagation results in the visible/NIR wavelength region. This implies that the propagation of uncertainty under various aerosol types and models does not influence the trend of the propagation of uncertainty.

7.4 Outlook

This research can be useful to PAFs for integrating uncertainty propagation into their processing chains. Within the scope of this research I found some limitations of this research, which are discussed below. Some limitations are due to limitations of the methods adopted while others are induced by the scope of this research work within the given time frame.

In our experiments, the adjacency range has a negligible influence on uncertainty propagation in the current setup. This might have been a consequence of the size of the real and simulated images used. From our experience, with real images non-negligible reflectance variations are mainly detected for large variations in adjacency range e.g. in the order of hundreds of pixels, for hyperspectral data with a spatial resolution similar to the HyMap and APEX sensors. These variations are mainly triggered by spectral heterogeneity or topographic characteristics of a scene. Integrating these factors with uncertainty propagation is out of the scope of our research. Thus, the adjacency effect can be further integrated with the uncertainty propagation for scenes where heterogeneity and/or topographic effects are important.

Ignorance of mutual interaction of atmospheric constituents i.e. how variation (due to uncertainty) in the concentration of CWV affects the concentration and behaviour of aerosol optical depth and adjacency range. I addressed such interactions in all experiments by deriving a joint probability distribution function (PDF) of column water vapour concentration and aerosol optical depth. The method is explained in detail in Chapter 3. In short, I used column water vapour and aerosol optical depth as obtained from the image based methods to derive the joint PDF. Also, in MODTRAN 4, setting the

parameter H2OAER (in Card 1A) to true, aerosol optical properties are modified to reflect the changes from the original relative humidity profile arising from the scaling of the water column. As the measurements of aerosol optical depth and column water vapour concentration obtained from the image based methods were not showing significant correlation, the actual physical interaction between the parameters could not be realised. Thus, more investigations to study the parameters interaction are required.

The presented study can be easily extended to take other atmospheric correction parameters into account, like oxygen and ozone absorption. Further, uncertainty arising from other pre-processing steps such as removing shadow can be included in assessing uncertainty in reflectance products.

Factors like the internal sources of variation and relations between parameters are too complex to fully explore. Thus, model uncertainty, i.e. uncertainty due to the internal structure of atmospheric correction, is not considered. Instead, I assumed that the physical structure of the background processes in a radiative transfer code and methods used to estimate the condition parameters are certain. For further studies it may be important to include the uncertainty associated with the physical structure of the background processes.

It would be interesting to compare different unmixing methods while considering the reflectance uncertainty. Moreover, as uncertainties are also introduced by the specific unmixing method used, a comparison between unmixing methods can be extended while considering the model uncertainty at the unmixing level. Experiments with spectral libraries containing a large number of spectral signatures are recommended to be carried out. Apart from unmixing, other applications, e.g. environmental modelling, precision agriculture and land cover monitoring, can benefit from the study presented in this work.

Remote sensing products are used in many applications. To proliferate the capabilities of the developed framework further, methods should be developed to incorporate uncertainty arising from a time series of data for various applications should be developed.

Bibliography

[1] M. E. Schaepman, M. Jehle, A. Hueni, P. D'Odorico, A. Damm, J. Weyermann, F. D. Schneider, and et. al., "Advanced radiometry measurements and Earth science applications with the Airborne Prism Experiment (APEX)," *Remote Sensing of Environment*, vol. 158, pp. 207–219, 2015. [Online]. Available: Link

[2] T. Stuffler, K. Forster, S. Hofer, M. Leipold, B. Sang, H. Kaufmann, B. Penne, and et. al., "Hyperspectral imaging-An advanced instrument concept for the EnMAP mission (Environmental Mapping and Analysis Programme)," *Acta Astronautica*, vol. 65, no. 7–8, pp. 1107–1112, 2009.

[3] JCGM, "JCGM 100: 2008 Evaluation of measurement data Guide to the expression of uncertainty in measurement," Joint Committee for Guides in Metrology, Standard Guidelines 100, 2008.

[4] JCGM-guidelines, "JCGM 101: 2008 evaluation of measurement data Supplement 1 to the Guide to the expression of uncertainty in measurement Propagation of distributions using a Monte Carlo method," Joint Committee for Guides in Meteorology, Standard Guidelines 101, 2008.

[5] I. Reusen, M. Bachman, J. Beekhuizen, E. Ben-Dor, J. Biesemans, J. L. Brenguier, P. Brown, and et. al., "Eufar goes hyperspectral in FP7," in *2009 First workshop on hyperspectral image and signal processing: evolution in remote sensing*, August 2009, pp. 1–4.

[6] N. Fox, "A guide to expression of uncertainty of measurements," Quality Assurance framework for Earth Observation, Technical Report QA4EO-QAEO-FEN-DQK-006, 2010, last Accessed on 23/01/2018. [Online]. Available: Link

[7] L. A. Remer, Y. J. Kaufman, D. Tanre, S. Mattoo, D. A. Chu, J. V. Martins, R. R. Li, and et. al., "The MODIS aerosol algorithm, products, and validation," *Journal of the Atmospheric Sciences*, vol. 62, no. 4, pp. 947–973, 2005.

[8] R. Richter, D. Schlapfer, and A. Muller, "An automatic atmospheric correction algorithm for visible/NIR imagery," *International Journal of Remote Sensing*, vol. 27, no. 10, pp. 2077–2085, 2006.

[9] G. Heuvelink, *Error propagation in Environmental Modelling with GIS*. London: Taylor & Francies Ltd., 1998.

[10] G. M. Foody and P. M. Atkinson, Eds., *Uncertainty in Remote Sensing and GIS*. Chichester: J. Wiley & Sons inc., 2002.

[11] C. H. Liu, "Error propagation in atmospheric correction due to azimuthal angle simplification of lookup tables," *International Journal of Remote Sensing*, vol. 30, no. 3, pp. 275–282, 2010.

[12] M. S. Salama and A. Stein, "Error decomposition and estimation of inherent optical properties," *Applied Optics*, vol. 48, no. 26, pp. 4947–4962, 2009.

[13] A. Stein, N. A. S. Hamm, and Q. Ye, "Handling uncertainties in image mining for remote sensing studies," *International Journal of Remote Sensing*, vol. 30, no. 20, pp. 5365–5382, 2009.

[14] A. Lucieer, A. Stein, and P. Fisher, "Multivariate texture based segmentation of remotely sensed imagery for extraction of objects and their uncertainty," *International Journal of Remote Sensing*, vol. 26, no. 14, pp. 2917–2936, 2005.

[15] N. Castoldi, L. Bechini, and A. Stein, "Evaluation of the spatial uncertainty of agro-ecological assessments at the regional scale: The phosphorus indicator in northern Italy," *Ecological Indicators*, vol. 9, pp. 902–912, 2009.

[16] D. E. van de Vlag and A. Stein, "Incorporating Uncertainty via Hierarchical Classification Using Fuzzy Decision Trees," *IEEE Transaction on Geoscience And Remote Sensing*, vol. 45, no. 1, pp. 237–245, 2007.

[17] J. Beekhuizen, G. B. M. Heuvelink, I. Reusen, and J. Biesemans, "Uncertainty Propagation Analysis of the airborne hyperspectral data processing chain," in *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, August 2009, pp. 1–4.

[18] J. D. Brown and G. B. M. Heuvelink, "The Data Uncertainty Engine (DUE): A software tool for assessing and simulating uncertain environmental variables," *Computers & Geosciences*, vol. 33, no. 2, pp. 172–190, 2007.

[19] J. Beekhuizen, G. B. M. Heuvelink, J. Biesemans, and I. Reusen, "Effect of DEM Uncertainty on the Positional Accuracy of Airborne Imagery," *IEEE Transaction on Geoscience And Remote Sensing*, vol. 49, no. 5, pp. 1567–1577, 2011.

[20] J. Biesemans, S. Sterckx, E. Knaeps, K. Vreys, S. Adriaensen, J. Hooyberghs, K. Meuleman, and et. al., "Image processing work flows for airborne remote sensing," in *Proceedings 5th EARSeL Workshop on Imaging Spectroscopy*. European Association of Remote Sensing Laboratories (EARSeL), April 2007.

[21] G. Schaepman-Strub, M. Schaepman, J. Martonchik, T. Painter, and S. Dangel, *The SAGE Handbook of Remote Sensing*. London: SAGE publication LTD., 2009, ch. 15. Radiometric and Reflectance: From terminology Concepts to Measured Quantities.

[22] A. Berk, G. P. Anderson, P. K. Acharya, J. H. Chetwynd, L. S. Bernstein, E. P. Shettle, M. W. Matthew, and et. al., "MODTRAN 4 user's Manual," Air Force Research Laboratory, Hanscom AFB, MA, USA and Naval Research Laboratory, Washington, DC, USA and Spectral Sciences, Burlington, MA, USA, Tech. Rep., 2000.

[23] A. Rodger and J. M. Lynch, "Determining atmospheric column water vapour in the 0.4–2.5 m spectral region," in *Proceedings of the AVIRIS Workshop Pasadena, California*. Jet Propulsion Laboratory (JPL) publication, USA, March 2001.

[24] R. J. Moffat, "Contributions to the Theory of Single-Sample Uncertainty Analysis," *Engineering-Transactions of the ASME*, vol. 104, pp. 250–258, 1982.

[25] R. Abernethy, "Handbook on Uncertainty in Gas-Turbine Measurement," USAF AEDC-TR-73-5, Technical Report, 1973.

[26] H. Coleman and W. Steele, *Experimentation, Validation, and Uncertainty Analysis for Engineers*. New Jersey: A John Willey and Sons, Inc., 2009.

[27] G. Foody and P. Atkinson, Eds., *Uncertainty in Remote Sensing and GIS*. West Sussex: John Wiley & Sons Ltd., 2002.

[28] G. B. M. Heuvelink, J. D. Brown, and E. E. van Loon, "A probabilistic framework for representing and simulating uncertain environmental variables," *International Journal of Geographical Information Science*, vol. 21(5), no. 5, pp. 497–513, 2007.

[29] J. Burt, G. Barber, and D. Rigby, *Elementry Statistics for Geographer*. London: The Guilford Press, 2009.

[30] J. F. Haan and J. M. M. Kokke, "Remote sensing algorithm development Toolkit I: Operationalization of atmospheric correction methods for tidal and inland waters," Netherlands Remote Sensing Board, Development Toolkit NRSP-2 96-16, 1996.

[31] W. Verhoef, "Theory of radiative transfer models applied in optical remote sensing of vegetation canopies," Ph.D. dissertation, University of Wageningen, Netherlands, 1998.

[32] D. Schlapfer, "Differential absorption methodology for imaging spectroscopy of atmospheric water vapor," Ph.D. dissertation, Remote Sensing Laboratories, Department of Geography, Universiy of Zurich, 1998.

[33] W. Verhoef and H. Bach, "Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative

transfer models," *Remote Sensing of Environment*, vol. 87, no. 1, pp. 23–41, 2003.

[34] B. C. Gao, M. J. Montes, C. O. Davis, and F. H. Goetz, A., "Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean," *Remote Sensing of Environment*, vol. 113, pp. S17 – S24, 2009, imaging Spectroscopy Special Issue.

[35] K. Staenz, J. Secker, C. Gao, B., C. Davis, and C. Nadeau, "Radiative transfer codes applied to hyperspectral data for the retrieval of surface reflectance," *Journal of Photogrammetry and Remote Sensing*, vol. 57, no. 3, pp. 194 – 203, 2002.

[36] K. Stamnes, S.-C. Tsay, W. Wiscombe, and K. Jayaweera, "Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media," *Applied Optics*, vol. 27, no. 12, pp. 2502–2509, June 1988.

[37] S. Sterckx, K. Vreys, J. Biesemans, M. D. Iordache, L. Bertels, and K. Meuleman, "Atmospheric correction of APEX hyperspectral data," *Miscellanea Geographica Regional Studies Development*, vol. 20, no. 1, 2016.

[38] R. T. Wilson, R. J. Milton, and J. M. Nield, "Are visibility-derived aot estimates suitable for parameterizing satellite data atmospheric correction algorithms?" *International Journal of Remote Sensing*, vol. 36, no. 6, pp. 1675–1688, 2015.

[39] R. Richter and D. Schlapfer, "Atmospheric/Topographic Correction for Airborne Imagery (ATCOR-4 User Guide)," DLR-German Aerospace Center, User Guide, 2007.

[40] Matlab-Toolbox, "Hyperspectral imagery synthesis (eias) toolbox," Grupo de Inteligencia Computacional, Universidad del Pas Vasco / Euskal Herriko Unibertsitatea (UPV/EHU), Spain, The MathWorks, Inc., Massachusetts, United States, Tech. Rep., 2012.

[41] R. F. Kokaly, R. N. Clark, G. A. Swayze, K. E. Livo, T. M. Hoefen, N. C. Pearson, , and et. al., "USGS spectral library version 7," U.S. Geological Survey Data Series 1035, Tech. Rep., 2017. [Online]. Available: Link

[42] T. Cocks, T. Jenssen, A. Stewart, I. Wilson, and T. Shields, "The hymaptm airborne hyperspectral sensor: The system, calibration and performance," in *In Proceedings of the 1st EARSeL Workshop on Imaging Spectroscopy, Zurich*. European Association of Remote Sensing Laboratories (EARSeL), October 1998, pp. 37–42.

[43] ENVI-Team, "ENVI Classic Tutorial: Vegetation Hyperspectral Analysis," Exelis Visual Information Solutions, Inc., Tech. Rep., 2014.

[44] A. M. Baldridge, S. J. Hook, C. I. Grove, and G. Rivera, "The ASTER spectral library version 2.0," *Remote Sensing of Environment*, vol. 113, no. 4, pp. 711 – 715, 2009.

[45] ENVI-Guide, "ENVI EX Users Guide," ITT Visual Information Solutions, Tech. Rep., 2009.

[46] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles.* Weinheim: Wiley & Sons, 2004.

[47] D. Burazerovic, R. Heylen, B. Geens, S. Sterckx, and P. Scheunders, "Detecting the Adjacency Effect in Hyperspectral Imagery With Spectral Unmixing Techniques," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 3, pp. 1070–1078, 2013.

[48] S. Chandrasekhar, *Radiative Transfer.* New York: Dover Publication Inc., 1960.

[49] M. S. Salama and A. Stein, "Error decomposition and estimation of inherent optical properties," *Applied Optics*, vol. 48, no. 26, pp. 4947–4962, 2009.

[50] D. G. Kaskaoutis, H. D. Kambezidis, A. D. Adamopoulos, and P. A. Kassomenos, "On the characterization of aerosols using the Angstrom, exponent in the Athens area," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 68, no. 18, pp. 2147–2163, 2006.

[51] F. C. Seidel, A. A. Kokhanovsky, and M. E. Schaepman, "Fast retrieval of aerosol optical depth and its sensitivity to surface albedo using remote sensing data," *Atmospheric Research*, vol. 116, pp. 22–32, 2012.

[52] L. Guanter, L. Gmez-Chova, and J. Moreno, "Coupled retrieval of aerosol optical thickness, columnar water vapor and surface reflectance maps from envisat/meris data over land," *Remote Sensing of Environment*, vol. 112, no. 6, pp. 2898 – 2913, 2008.

[53] A. Saltelli, K. Chan, and E. M. Scott, *Sensitivity Analysis.* Chichester: John Wiley & Sons Ltd., 2008.

[54] A. Saltelli, T. H. J. Andres, and T. Homma, "Sensitivity analysis of model output An investigation of new techniques," *Computational Statistics & Data Analysis*, vol. 15, no. 2, pp. 211–238, 1993.

[55] A. Saltelli and J. Marivoet, "Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques," *Reliability Engineering and System Safety*, vol. 28, no. 2, pp. 229–253, 1990.

[56] G. E. B. Archer, A. Saltelli, and I. M. Sobol, "Sensitivity measures,anova-like Techniques and the use of bootstrap," *Journal of Statistical Computation and Simulation*, vol. 58, no. 2, pp. 99–120, 1997.

[57] I. M. Sobol, "Sensitivity Estimates for Nonlinear Mathematical Models," *Mathematical Modelling & Computational Experiment*, vol. 1, no. 4, pp. 407–414, 1993.

[58] R. I. Cukier, H. B. Levine, and K. E. Shuler, "Nonlinear Sensitivity

Analysis of Multiparameter Model Systems," *Journal of Computational Physics*, vol. 26, no. 1, pp. 1–42, 1978.

[59] E. Shettle and R. Fenn, "Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties," Air Force Geophysics Laboratory, USA, Technical Report, 1979.

[60] G. A. d'Almeida, P. Koepke, and E. P. Shettle, *Atmospheric Aerosols Global Climatology and Radiative Characteristics*. Hampton: A. Deepak Publishing, 1991.

[61] G. J. McRay, W. J. Tilde, and J. H. Seinfeld, "Global sensitivity analysis-A computational implementation of the Fourier Amplitude Sensitivity Test (FAST)," *Computers & Chemical Engineering*, vol. 6, no. 1, pp. 15–25, 1982.

[62] A. Saltelli, S. Tarantola, and K. P. S. Chan, "A quantitative Model-independent method for global sensitivity analysis of model output," *Technometrics*, vol. 41, no. 1, pp. 39–56, 1999.

[63] R. Lindstrot, R. Preusker, H. Diedrich, L. Doppler, R. Bennartz, and J. Fischer, "1D-Var retrieval of daytime total columnar water vapour from MERIS measurements," *Atmospheric Measurement Techniques*, vol. 5, no. 3, pp. 631–646, 2012.

[64] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, 2012.

[65] N. Keshava, "A Survey of Spectral Unmixing Algorithms," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 55–78, 2003.

[66] J. Lenoble, *Atmopsheric Radiative Transfer*. Virginia: A. Deepak Publishing, 1998.

[67] B. C. Gao, C. Davis, and A. Goetz, "A review of atmospheric correction techniques for hyperspectral remote sensing of land surfaces and ocean color," in *2006 IEEE International Symposium on Geoscience and Remote Sensing*, July 2006, pp. 1979–1981.

[68] V. Carrere and J. Conel, "Recovery of atmospheric water vapor total column abundance from imaging spectrometer data around 940 nm – sensitivity analysis and application to airborne visible/infrared imaging spectrometer (AVIRIS) data," *Remote Sensing of Environment*, vol. 44, no. 2, pp. 179–204, 1993.

[69] A. Rodger, "SODA: A new method of in-scene atmospheric water vapor estimation and post-flight spectral recalibration for hyperspectral sensors: Application to the hymap sensor at two locations," *Remote Sensing of Environment*, vol. 115, no. 2, pp. 536–547, 2011.

[70] Z. Qu, B. C. Kindel, and A. F. H. Goetz, "The high accuracy atmospheric

correction for hyperspectral data (HATCH) model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1223–1231, June 2003.

[71] J. Diner, D., V. Martonchik, J., R. A. Kahn, B. Pinty, N. Gobron, D. L. Nelson, and B. N. Holben, "Using angular and spectral shape similarity constraints to improve MISR aerosol and surface retrievals over land," *Remote Sensing of Environment*, vol. 94, no. 2, pp. 155–171, 2005.

[72] D. C. Heinz and C. I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.

[73] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total Variation Spatial Regularization for Sparse Hyperspectral Unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 484–4502, 2012.

[74] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, pp. 89–97, 2004.

[75] E. Jonathan and D. P. Bertsekas, "On the douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1, pp. 293–318, 1992.

[76] C. A. Bateson, G. P. Asner, and C. A. Wessman, "Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 1083–1094, March 2000.

[77] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse Unmixing of Hyperspectral Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011.

[78] N. Bhatia, V. A. Tolpekin, I. Reusen, S. Sterckx, J. Biesemans, and A. Stein, "Sensitivity of Reflectance to Water Vapor and Aerosol Optical Thickness," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3199–3208, 2015.

[79] R. S. Fraser and Y. J. Kaufman, "The Relative Importance of Aerosol Scattering and Absorption in Remote Sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-23, no. 5, pp. 625–633, 1985.

[80] B. N. Holben, T. F. Eck, I. Slutsker, D. Tanre, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov, "AERONET–A federated instrument network and data archive for aerosol characterization," *Remote Sensing of Environment*, vol. 66, no. 1, pp. 1 – 16, 1998.

[81] B. G. Gao and A. F. H. Goetz, "Column atmospheric water vapor and

vegetation liquid water retrievals from airborne imaging spectrometer data," *Journal of Geophysical Research: Atmospheres*, vol. 95(D4), pp. 3549–3564, 1990.

[82] Y. J. Kaufman and C. Gao, B., "Remote sensing of water vapor in the near IR from EOS/MODIS," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 5, pp. 871–884, 1992.

[83] C. Gao, B. and Y. J. Kaufman, "Water vapor retrievals using moderate resolution imaging spectroradiometer (MODIS) near-infrared channels," *Journal of Geophysical Research: Atmospheres*, vol. 108(D13), 2003.

[84] D. S. Fung, "Methods for the estimation of missing values in time series," Master's thesis, School Of Mathematics and Engineering, Edith Cowan University, Perth, 2006. [Online]. Available: Link

[85] P. Chylek, C. C. Borel, W. Clodius, P. A. Pope, and A. P. Rodger, "Satellite-based columnar water vapor retrieval with the multi-spectral thermal imager (MTI)," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 12, pp. 2767–2770, Dec 2003.

[86] M. S. Rahman, L. Di, R. Shrestha, E. G. Yu, L. Lin, L. Kang, and D. M., "Comparison of selected noise reduction techniques for modis daily ndvi: An empirical analysis on corn and soybean," in *2016 Fifth International Conference on Agro-Geoinformatics*, July 2016, pp. 1–5.

[87] N. Bhatia, M. D. Iordache, A. Stein, I. Reusen, and V. A. Tolpekin, "Propagation of uncertainty in atmospheric parameters to hyperspectral unmixing," *Remote Sensing of Environment*, vol. 204, pp. 472–484, 2018.

[88] N. Bhatia, A. Stein, I. Reusen, and V. A. Tolpekin, "An optimization approach to estimate and calibrate column water vapour for hyperspectral airborne data," *International Journal of Remote Sensing*, vol. 39, no. 8, pp. 2480–2505, 2018.

[89] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall/CRC, 1986.

[90] T. Duong, *ks: Kernel Smoothing*, 2016, R package version 1.10.3.

[91] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2008.

[92] B. W. Silverman, "Using Kernel Density Estimates to investigate multimodality," *Journal of the Royal Statistical Society*, vol. 43, no. 1, pp. 97–99, 1981.

[93] P. Hall and M. York, "On the calibration of Silvermans test for multimodality," *Statistica Sinica*, vol. 11, pp. 515–536, 2001.

[94] J. Li and A. D. Heap, "Spatial interpolation methods applied in the environmental sciences: A review," *Environmental Modelling & Software*, vol. 53, pp. 173–189, 2014.

[95] O. Babak and C. V. Deutsch, "Statistical approach to inverse distance interpolation," *Stochastic Environmental Research and Risk Assessment*, vol. 23, no. 5, pp. 543–553, July 2009.

Biography

Nitin Bhatia was on 10th October 1982 in New Delhi, India. In 2003, he received a diploma in electrical engineering from Aryabhat Polytechnic (Aryabhatt Institute Of Technology), Delhi. In 2006, he received his engineering degree in computer science (IT) from Indraprastha University (GGIPU), Delhi. He received his M.Sc. degree in Geo-information Science and Earth Observation in 2010 from Department of Earth Observation Science (ITC), Enschede, The Netherlands in collaboration with Indian Institute of Remote Sensing (ISRO), India. In the meantime, he worked as a software developer in the period of 2006–2008 and as a catastrophic modeller in the period of 2010-2011. He then began to pursue the present Ph.D research, which is a collaboration between Flemish Institute of Technological Research (VITO), Belgium, and Department of Earth Observation Science, Faculty of Geo-Information Science and Earth Observation of the University of Twente, The Netherlands. He is currently employed as a research officer in Australian research council (ARC), center of excellence for coral reef studies. His current research focuses on processing of the remote sensing data to analyse the spatial resilience in social-ecological systems. His research interests includes uncertainty and sensitivity analysis of remote sensing products developed in remote sensing processing chain, atmospheric correction, statistical analysis of remotely sensed images, and spatial resilience in social-ecological systems.

Author publications

7.4.1 ISI journals

1. N. Bhatia, V. A. Tolpekin, I. Reusen, S. Sterckx, J. Biesemans, and A. Stein.,"Sensitivity of reflectance to water vapour and aerosol optical thickness", in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8, no. 6, pp. 3199–3208, 2015.

2. N. Bhatia, M. D. Iordache, A. Stein, I. Reusen, and V. A. Tolpekin,"Propagation of uncertainty in atmospheric parameters to hyperspectral unmixing", in Remote Sensing of Environment, vol. 204, pp. 472–484, 2018

3. N. Bhatia, A. Stein, I. Reusen and V. A. Tolpekin, "An optimization approach to estimate and calibrate column water vapour for hyperspectral airborne data", in International Journal of Remote Sensing, vol. 39, no. 8, pp. 2480–2505, 2018.

4. N. Bhatia, A. Stein, I. Reusen and V. A. Tolpekin, "Estimation of AOD under uncertainty: an approach for hyperspectral airborne data", Submitted to special issue Remote Sensing, special issue in uncertainty in remote sensing image analysis.

7.4.2 Conference proceedings

1. M. D. Iordache, N. Bhatia, J. M. Bioucas-Dias, A. Plaza,"Uncertainty propagation from atmospheric parameters to sparse hyperspectral unmixing", in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 6133–6136, Beijing, 2016.

2. N. Bhatia, J. Biesemans, V. Tolpekin, I. Reusen, S. Sterckx and A. Stein,"Global sensitivity analysis of water vapour and visibility for atmospheric correction", in 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp. 1–4, Lausanne, 2014.

3. N. Bhatia, J. Biesemans, V. Tolpekin, I. Reusen, S. Sterckx and A. Stein,"Optimizing the range of atmospheric condition parameters to avoid over and under-estimation of uncertainty", in $6^t h$ Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp. 1–4, Lausanne, 2014.

7.4.3 Oral and posters presentations

1. N. Bhatia, V. A. Tolpekin, J. Biesemans, I. Reusen, S. Sterckx, and A. Stein,"Exploration of the correlation structure of the atmospheric condition parameters in support of the design of uncertainty propagation methods for hyperspectral image processing workflows", Poster presentation, EARSeL conference, Nantes, 2013.

2. N. Bhatia, J. Biesemans, V. A. Tolpekin, I. Reusen, S. Sterckx, and A. Stein,"Optimizing atmospheric condition parameters range to avoid over/under-estimation of uncertainty"Poster presentation in in 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lausanne, 2014.

3. Bhatia N., Biesemans J., Tolpekin V.A., Reusen I., Sterckx S., Stein A.,"Global Sensitivity Analysis of Water vapour concentration and Visibility for Atmospheric correction", Oral presentation in in 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lausanne, 2014.