UNIVERSITY OF SOUTHAMPTON

Developing a novel method to retrieve high spatial resolution Aerosol Optical Thickness (AOT) from satellite data

by

Robin Wilson

A thesis submitted for the
degree of Doctor of Philosophy (PhD)

in the
Faculty of Social and Human Sciences
Geography and Environment

Supervisors:
E. J. Milton and J. M. Nield and J. Noble

January 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES
GEOGRAPHY AND ENVIRONMENT

Doctor of Philosophy (PhD)

Developing a novel method to retrieve high spatial resolution Aerosol Optical Thickness (AOT) from satellite data

by Robin Wilson

Aerosol Optical Thickness (AOT) data have many important applications including atmospheric correction of satellite imagery and monitoring of particulate matter air pollution. Current data products are generally available at a kilometre-scale resolution, but many applications require far higher resolutions. For example, particulate matter concentrations vary on the scale of tens of metres, and thus data products at a similar scale are required to provide accurate assessments of particle densities and allow effective monitoring of air quality and analysis of local air quality effects on health.

This thesis describes the development of a novel method which retrieves per-pixel AOT values from high-resolution (30m) satellite data, and this method is the main novel contribution to scientific knowledge of this PhD. This method is designed to work over a wide range of land covers including both bright and dark surfaces - and requires only standard visible bands, making it applicable to a range of data from sensors such as Landsat, DMC, SPOT and Pleiades. The method is based upon an extension of the Haze Optimized Transform (HOT), which was originally designed for estimating the haziness of each pixel in a satellite image, based upon the distance from a 'Clear Line' in feature space. In this research, the HOT method is adapted and used to estimate AOT instead. Significant extensions include Monte Carlo estimation of the 'Clear Line', object-based correction for land cover, and modelling of the HOT-AOT relationship using radiative transfer models.

Validation against ground and satellite measurements, as well as simulated data, shows that 40–50% of the pixels have an error within ± 0.1, not much lower than many presently available low-resolution products, with further work likely to improve the accuracy. Two example applications show the potential of this method for per-pixel atmospheric correction and monitoring the spatial pattern of particulate matter pollution. This novel method will enable many new applications of AOT data that were impossible with low-resolution data.

Contents

List of Figures

List of Tables

Abbreviations & Nomenclature

6S	Second Simulation of the Satellite Signal in the Solar Spectrum
AATSR	Advanced Along-Track Scanning Radiometer
ACCA	Automatic Cloud Cover Assessment
AERONET	Aerosol Robotic Network
AHOT	Advanced Haze Optimized Transform
AISA	Airborne Imaging Spectrometer
AOT	Aerosol Optical Thickness
ARSF	Airborne Research and Survey Facility
ARVI	Atmospherically Resistant Vegetation Index
ASD	Analytical Spectral Devices
ASTER	Advanced Spaceborne Themal Emission and Reflection Radiometer
ATBD	Algorithm Theoretical Basis Document
ATSR	Along Track Scanning Radiometer
AURN	Automatic Urban-Rural Network
AVHRR	Advanced Very High Resolution Radiometer
BIGF	British Isles Continuous GNSS Facility
BOA	Bottom of Atmosphere
BRDF	Bidirectional Reflectance Distribution Function
BSHTI	Background Suppressed Haze Thickness Index
BUV	Backscattered Ultraviolet
CCN	Cloud Condensation Nuclei
CFARR	Chilbolton Facility for Atmospheric and Radio Research
DALY	Disability-adjusted Life Year
DDV	Dense Dark Vegetation
DMC	Disaster Monitoring Constellation
DN	Digital Number
ENVI	Environment for Visualising Images
ERTS	Earth Resources Technology Satellite
ESRI	Environmental Sciences Research Institute
ETM	Enhanced Thematic Mapper
GDAL	Geographic Data Abstraction Library
GEOBIA	Geographic Object-based Image Analysis
GEOS	Goddard Earth Observing System

GOES	Geostationary Operational Environmental Satellite
GNSS	Global Navigation Satellite System
GOCART	Global Ozone Chemistry Aerosol Radiation and Transport
GSFC	Goddard Space Flight Centre
HOT	Haze Optimized Transform
HRV	High Resolution Visible
HYSPLIT	Hybrid Single Particle Lagrangian Integrated Trajectory Model
HOTBAR	Haze Optimized Transform based Aerosol Retrieval
IPCC	Intergovernmental Panel on Climate Change
IR	Infrared
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LiDAR	Light Detection and Ranging
LUT	Lookup Table
MAIAC	Multi-angle Implementation of Atmospheric Correction
MERIS	Medium Resolution Imaging Spectrometer
MISR	Multi-angle Imaging SpectroRadiometer
MNF	Maximum Noise Fraction
MODIS	Moderate Resolution Imaging Spectroradiometer
MODTRAN	MODerate resolution atmospheric TRANsmission
MSS	Multispectral Scanner
MTF	Modulation Transfer Function
NaN	Not a Number
NCAVEO	Network for the Calibration and Validation of Earth Observation data
NCEP	National Centres for Environmental Prediction
NDHI	Normalized Difference Haze Index
NDVI	Normalized Difference Vegetation Index
NERC	Natural Environment Research Council
NIR	Near Infrared
NLCD	National Land Cover Dataset
OBIA	Object-based Image Analysis
ODR	Orthogonal Distance Regression
OLI	Operational Land Imager
OMI	Ozone Monitoring Instrument
OSIS	Observation of Shadows for aerosol Inversion over 3D Scenes
PBL	Planetary Boundary Layer
PCA	Principal Components Analysis
PM	Particulate Matter
$PM_{2.5}$	Particulate Matter with a diameter $< 2.5\mu m$
PM_{10}	Particulate Matter with a diameter $< 10\mu m$
PSF	Point Spread Function
PWC	Precipitable Water Content
RMSE	Root Mean Square Error

RTM	Radiative Transfer Model
SeaWIFS	Sea-Viewing Wide Field-of-View Sensor
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SLC	Scan Line Corrector
SNR	Signal to Noise Ratio
SPARC	Spectra Barrax Campaign
SPOT	Satellite Pour lObservation de la Terre
SRF	Spectral Response Function
SSA	Single Scattering Albedo
SWIR	Shortwave Infrared
SYNTAM	Synergy of Terra and Aqua MODIS
TARFOX	Tropospheric Aerosol Radiative Forcing Observational Experiment
TCHT	Tasselled Cap Haze Transform
TEOM	Tapered Element Oscillating Microbalance
TM	Thematic Mapper
TOA	Top of Atmosphere
VIIRS	Visible Infrared Imaging Radiometer Suite
WRS	Worldwide Reference System
α	Angstrom exponent
λ	Wavelength
ρ	Reflectance
ρ_g	Ground reflectance
ρ_p	Path reflectance
ρ_s	At-sensor (Top of Atmosphere) reflectance
τ	Aerosol Optical Thickness
L	Radiance
L_g	Ground-leaving radiance
L_p	Path radiance
L_s	At-sensor (Top of Atmosphere) radiance
m	Air mass
T	Transmittance
T_{high}	High threshold for edge detection
T_{low}	Low threshold for edge detection
x_a, x_b, x_c	6S atmospheric correction coefficients

Declaration of Authorship

I,

Robin Wilson

declare that the thesis entitled

Developing a novel method to retrieve high spatial resolution Aerosol Optical Thickness (AOT) from satellite data

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:
  - Wilson, R.T., Milton, E.J. and Nield, J.M. (2014). Spatial variability of the atmosphere over southern England, and its effect on scene-based atmospheric corrections. *International Journal of Remote Sensing 35*(13), 5198–5218.
  - Wilson, R.T. (2012). Py6S: a Python interface to the 6S radiative transfer model. *Computers and Geosciences 51*, 166–171.
  - Wilson, R.T., Milton, E.J. and Nield, J.M. (2012). Spatial variability of the atmosphere across southern England and the resulting error in assuming a uniform atmospheric correction. *Remote Sensing and Photogrammetry Society Conference 2012, Greenwich, UK* 12-14 Sept 2012.

Signed: .....

Date: .....

Acknowledgements

I couldn't have completed this PhD without the help and support of a huge number of people - far too many to list all of them. However, thanks must go first to my wife, Olivia, for all of her support throughout my PhD. She married me one year into my PhD, and so knew exactly what she was getting into. This PhD has been a long process with many difficult periods and through all of this she has been there to listen, support, suggest ideas and proof-read - whichever was necessary at the time!

As important to the success of this PhD as my wife (but in a very different way...) I would like to thank my supervisors Ted, Jo and Jason for all of their help during my PhD. Ted has supervised me through my undergraduate dissertation and my PhD, and has always believed in me. My PhD was Ted's last before retirement, and I hope that it was a fitting end to his career. Jo and Jason both had the pleasure of supervising a remote sensing PhD without being experts in remote sensing, but still contributed a huge amount, for which I am very thankful.

Others who were less directly involved, but whose support has meant a huge amount to me include Nicki Lewin (for all of the administrative support, encouragement and conspiring with Olivia to stop me doing too much), Alison Lawrence, Charlotte Upfold, Dan Barnes-Davies (if even he did pretend for years that my PhD was boring...), Dr Armstrong, colleagues within Geography and the ICSS, members of the RSPSoc Wavelength student community, and everyone else I've forgotten to mention. I will always be thankful for the support and encouragement that my parents and grandparents have always given me - even if they didn't necessarily understand everything I was talking about! Even though this thesis doesn't contain any conversation, it does contain many pictures - so I hope it satisfies the requirements of a good book[1].

On a more formal note, thanks are also due to:

- The Institute for Complex Systems Simulation, and all their staff, who funded and administered my PhD
- The Engineering and Physical Sciences Research Council for providing the ICSS funding under the EPSRC Doctoral Training Centre grant EP/G03690X/1
- The AERONET Principal Investigators and their staff for establishing and maintaining the many AERONET sites used in this thesis
- The staff at the Chilbolton Facility for Atmospheric and Radio Research for supporting the AERONET site there and the NCAVEO field campaign

[1] For those who don't recognise this, it refers to a comment from Alice at the start of Lewis Carroll's *Alice's Adventures in Wonderland*:
"what is the use of a book," thought Alice, "without pictures or conversations?"

- Participants in the NCAVEO 2006 Field Campaign, from which data were used in Chapter 2. These data were provided courtesy of NCAVEO via the NERC Earth Observation Data Centre (NEODC).
- The services of the Natural Environment Research Council (NERC) British Isles continuous GNSS Facility (BIGF), www.bigf.ac.uk, in providing archived GNSS data and products
- Shunlin Liang, Yuwei Li and the authors of Yan *et al.* (2003) for providing access to the Shunyi ground reflectance dataset
- Alexei Lyapustin and Yujie Wang for providing access to the entire US MAIAC data archive
- The IRIDIS High Performance Computing Facility, and associated support staff and services, at the University of Southampton
- The open-source software community, particularly OSGEO, for creating and maintaining the wide range of open-source software tools used in this thesis

Chapter 1

Introduction

Aerosol Optical Thickness (AOT) is a measure of the haziness of the atmosphere due to aerosol particles. AOT measurements are essential for a range of applications including satellite image atmospheric correction and monitoring of air pollution. Many methods have been developed to retrieve AOT from satellite data, but there are currently no methods operationally providing high spatial resolution AOT data (that is, fine-resolution data with a spatial resolution of 100m or better) even though there is a well-acknowledged need for these data. This thesis aims to fill this gap in the literature by developing a novel method to retrieve AOT from satellite data at a high resolution.

1.1 Aerosol Optical Thickness & its applications

Aerosol Optical Thickness (AOT) is a dimensionless measure of the reduction in the intensity of light passing through the atmosphere caused by scattering and absorption by particles in the atmosphere (aerosols). These aerosols come from a range of sources, and include dust, salt and soot, and the aerosol type determines the effect on light: for example, soot particles absorb light far more than most other aerosols (Kondratyev *et al.*, 2005). AOT can be measured by observing the effects that aerosols have on light: these methods are principally based on the scattering effect of aerosols, which has a strong wavelength dependence, with smaller wavelength light (such as blue light) being scattered far more than longer wavelength light (such as red or near-infrared light).

Ground measurements of AOT have been acquired for many years, by assessing the reduction in intensity of the light from the sun between the top of the atmosphere (TOA) and the ground surface (bottom of atmosphere; BOA). The TOA radiance can be accurately estimated from solar irradiance models, or measured by satellites, and the BOA radiance can be measured using a sun photometer (which has a very small field of view to allow it to measure the intensity of light coming from just the solar disc). Combining these two measurements, along with some ancillary data such as pressure and location, allows estimation of AOT with a very high accuracy (Morys *et al.*, 2001). Over three hundred of these instruments have been deployed into the Aerosol Robotic Network

(AERONET; Holben *et al.*, 1998), which acquire measurements of AOT approximately every fifteen minutes during daylight hours.

Satellite-based measurements of AOT provide a far greater spatial coverage than AERONET, and the majority of applications of AOT use satellite measurements due to the sparsity of AERONET measurement locations. Algorithms for estimating AOT from satellite data are based on the same theory as sun photometer measurements (that is, the wavelength-dependent scattering effect of aerosols), but in this case the light passes through the atmosphere twice (from the sun to the surface, and then reflected from the surface to the satellite), adding extra complexity. Satellite AOT retrieval algorithms apply complex methods to attempt to remove the noise caused by the land surface reflectance, and extract just the scattering signal from the data (for example, Levy *et al.*, 2010; He *et al.*, 2014). The resulting measurements have a lower accuracy than sun photometric measurements, but can cover the whole globe at resolutions ranging from 100km to 1km.

As well as affecting the passage of light through the atmosphere, aerosols have significant effects on human health – primarily through the breathing in of very small particles, known as $PM_{2.5}$ (Davidson *et al.*, 2005). These two primary effects are the basis for the main applications of AOT measurements: satellite image atmospheric correction and air quality monitoring - although AOT measurements are also used for climate monitoring (due to the aerosol effect on cloud formation, and thus on the Earth's energy balance; Le Treut, 2012), fire monitoring (by mapping smoke plumes; Hsu *et al.*, 1996), monitoring of volcanic ash distribution (for example, for flight safety; Christopher *et al.*, 2012), and monitoring of dust emission and transport (Ashpole and Washington, 2012).

Satellite images are affected by various atmospheric constituents including aerosols, water vapour, ozone and mixed gases – this is why satellite data can be used to estimate AOT (Petty, 2006). However, when satellite images are used to monitor the Earth's surface these atmospheric effects are noise which must be removed through a process of atmospheric correction – neglecting to remove these effects can lead to significant errors (Samanta *et al.*, 2010). To perform an atmospheric correction accurately, the concentrations of each atmospheric constituent must be known, which, in the case of aerosols, means knowing the AOT over the image. Most current atmospheric correction algorithms assume a constant AOT across the image, which is unlikely to be true in many cases: to get accurate results an atmospheric correction must be performed with a per-pixel parameterisation of atmospheric constituents, so that changes across the image can be taken into account (discussed further in Chapter 2). To parameterise these per-pixel atmospheric corrections, AOT must be known across the image, and the only realistic source for these data is satellite-derived AOT.

Aerosols, particularly those of an anthropogenic origin such as soot particles, are a major source of pollution with significant impacts on human health (Davidson *et al.*, 2005). The danger of this particulate matter pollution is related to the size of the particles: smaller particles can penetrate further into the lungs and thus cause more significant health

impacts (Seaton *et al.*, 1995). There has been a global focus on so-called $PM_{2.5}$ pollution, that is particulate matter with a diameter less than 2.5µm. These tiny particles - around a fortieth of the diameter of a human hair - cause diseases including lung cancer, asthma and cardiovascular disease, and were responsible for 3.2 million premature deaths worldwide in 2010 (Lim *et al.*, 2012). Governments across the world operate networks of ground stations to monitor $PM_{2.5}$ levels for compliance with standards, to monitor sources and sinks, and to allow research on the health effects of $PM_{2.5}$. However - like AERONET sites - these monitoring stations are relatively sparsely distributed and thus it is impossible to examine the spatial distribution of $PM_{2.5}$ pollution in detail. As particulate matter pollution consists, by definition, of aerosols, it is possible to develop robust relationships between AOT and $PM_{2.5}$ concentrations (for example, van Donkelaar *et al.*, 2010), and thus use satellite-derived AOT datasets to produce maps of $PM_{2.5}$ distribution.

1.2 The gap in the literature

Many methods have been developed to retrieve AOT from data collected from a range of satellites. AOT retrieval techniques can be categorised by the method through which they attempt to remove the 'noise' of land surface reflectance (as discussed in more detail in Chapter 3). This is usually implemented as a method to separate the at-sensor radiance (the light received at the satellite) into two components: the path radiance (caused by aerosol scattering, and thus the 'signal' that must be extracted from the data) and the ground radiance (caused by the surface reflectance, and considered in this case to be 'noise' that must be removed). Many of these methods are based upon one of the following:

- Empirical relationships between bands which are relatively unaffected by aerosols, such as short-wave infrared (SWIR) bands, and visible bands (such as Remer *et al.*, 2006)
- The assumption of temporal stability in ground reflectance (such as He *et al.*, 2014), even though this assumption is nearly always incorrect, as even supposedly invariant surfaces undergo significant changes in reflectance over a range of timescales (Anderson and Milton, 2006)
- The use of extra data (such as polarisation or multi-angular measurements; Mishchenko and Travis, 1997; Lyapustin *et al.*, 2011)

The majority of these algorithms produce data with a kilometer-scale spatial resolution; a few recent methods provide higher resolution data (although these methods are not yet providing data operationally). There is a distinct need for high-resolution data from a variety of application areas: for example, Hoff and Christopher (2009) state that air quality monitoring requires data at a resolution of 1–100m, and no operational methods currently provide data at this resolution. Furthermore, most current algorithms have strict data requirements (for example, the presence of SWIR bands, the availability of long time series of data, or the availability of a 'clear' image to use as a reference), significantly limiting their application, both in terms of the range of sensors which can be

used and the areas in the world in which the methods can be applied. Thus, there is a gap in the literature which can be filled by the development of an algorithm which can retrieve high-resolution (of 100m or finer) AOT from standard visible and near infra-red optical satellite imagery (that is, without requiring measurements of short-wave infra-red light, the polarisation of light, or the acquisition of data at multiple angles). The development of such an algorithm to fill this gap will be the focus of this thesis.

1.3 Approach

As discussed in Chapter 3, other researchers have attempted the development of a high-resolution AOT retrieval algorithm, but these algorithms suffer from the problems listed previously: strict data requirements limiting their applicability and lack of operational implementation (including broad scale validation). The issues with this previous work suggests that a change of algorithmic approach may prove fruitful. The most fundamental change of approach is to remove the focus on separating the at-sensor radiance into path and ground radiance components: this is at the centre of the majority of current algorithms, but this does not necessarily need to be the case. Other application areas within remote sensing, such as haze assessment algorithms, deal with atmospheric scattering without an explicit radiance separation stage, and there is a potential for these algorithms to be extended to retrieve AOT.

Haze assessment algorithms were originally developed in the 1970s to automatically classify hazy areas within satellite images, and to then attempt 'correction' of these hazy areas to make the data within them more usable. After significant work in the 1970s and 1980s, these algorithms were 're-discovered' in the early 2000s and brought up to date in the Haze Optimized Transform (HOT; Zhang *et al.*, 2002a). This method produces a single HOT value for each pixel in the image, based upon the distance of the pixel from a 'Clear Line' defined in blue-red feature space. This value quantifies the haziness of the pixel, and can then be used to perform a correction to attempt to remove the haze effects.

The American Meteorological Society Glossary of Meteorology defines haze as "particles suspended in air, reducing visibility by scattering light" (American Meteorological Society, 2013), which is fundamentally the same as the definition of aerosols, and thus the haziness of a pixel is fundamentally linked to AOT. The HOT algorithm was not designed to estimate AOT, but given that haze is just another name given to the visible effect of high AOT, it seems logical that there may be potential to extend the HOT to estimate AOT. This will remove the explicit radiance separation step from the algorithm, and the ability to use spatial context in correcting the HOT should allow reduction of signal-to-noise issues which have affected previous high-resolution algorithms (as noted by Remer *et al.*, 2006, 2013).

1.4 Aims & Objectives

The overall aim of this thesis is to develop a method to retrieve high-resolution AOT from satellite images. Specifically, the aim is to develop a method to produce a per-pixel AOT product from Landsat images, based upon an extension of the HOT, producing AOT with a spatial resolution of 30m. The objectives of this PhD are:

1. To assess the validity of the assumptions behind the Haze Optimized Transform
2. To develop an improved version of the Haze Optimized Transform, focusing particularly on the correction procedure
3. To develop a method to estimate AOT from the Haze Optimized Transform
4. To assess the accuracy of the resulting AOT values both in terms of absolute accuracy and 'spatial coherence' (the relative accuracy between adjacent pixels)
5. To apply the new method to some potential application areas where high-resolution AOT data are needed

1.5 Thesis structure

After **Chapter 1** (*Introduction*), the pilot study in **Chapter 2** (*Scoping the problem*) investigates the spatial variability of the atmosphere over southern England and its effect on atmospheric correction of satellite imagery. This self-contained study concludes that performing full spatially-variable atmospheric correction is essential, and that this requires per-pixel AOT products at high resolution.

Chapter 3 (*Literature Review*) explores the physical basis of atmospheric radiative transfer before detailing the established methods for monitoring AOT from both ground and satellite instruments, haze assessment methods for use with satellite imagery and potential applications of high-resolution AOT data. The chapter concludes with a synthesis which identifies a key need for a new algorithm to retrieve high-resolution AOT from satellite images, and suggests that an approach based upon the extension of haze assessment methods may prove fruitful.

Chapter 4 (*Investigating the Haze Optimized Transform*) investigates the conceptual basis of one of these haze assessment methods and validates the assumptions involved. This requires the development of specialised datasets, many of which are used throughout the rest of the thesis, and the methods used for generating these are discussed in detail in the early part of this chapter. By the end of this chapter, the fundamental basis of the Haze Optimized Transform has been shown to be correct, and a number of issues have been investigated in depth.

Chapter 5 (*Improving the Haze Optimized Transform*) discusses the extension of the Haze Optimized Transform to allow it to be used in the context of AOT estimation, with a focus on a new method for estimating the Clear Line and a Object-based Image Analysis approach to HOT correction. By the end of this chapter the improved HOT product is ready to be used to estimate AOT.

This estimation process is dealt with in **Chapter 6** (*Investigating the HOT-AOT relationship*), in which the actual link between HOT and AOT values is assessed, and then a method is developed to estimate this relationship for an arbitrary image. Again, specialised datasets are required, and their development is discussed early in the chapter. This chapter finishes by putting together all of the work done so far to produce an algorithm to retrieve high-resolution AOT from Landsat ETM+ images, and closes with example inputs and outputs.

Chapter 7 (*Validation*) assess the accuracy of the results from the new algorithm by comparison with ground-based and satellite-based measurements, as well as further use of simulated images, and concludes with an estimated AOT uncertainty which should be applied to all results from the algorithm.

Chapter 8 (*Example applications*) provides indicative examples to show how the new high-resolution AOT data could be used for per-pixel atmospheric correction and the assessment of $PM_{2.5}$ air pollution. A particular focus is on the novel applications which the availability of high-resolution data makes possible.

Finally, **Chapter 9** (*Conclusions & Further Work*) summarises the novel contributions made in this thesis, assesses the advantages and disadvantages of a haze-assessment approach to AOT estimation and discusses potential future extensions to the algorithm and its validation.

Chapter 2

Scoping the problem

This chapter consists of a report on a self-contained research project performed at the beginning of the PhD which used presently-available datasets to assess the spatial variability of the atmosphere (principally in terms of AOT) and its effects on satellite image atmospheric correction. Knowledge about the level of spatial variability in AOT is essential when deciding whether new high-resolution AOT products should be produced: if AOT has very low spatial variability then there is little need for new high-resolution products. Similarly, one of the main applications of a high-resolution AOT product is for atmospheric correction of satellite images: if satellite images can be accurately atmospherically-corrected by assuming a constant AOT across the image then there is little need for a new high-resolution AOT product.

The content of this chapter was published in the International Journal of Remote Sensing as Wilson *et al.* (2014) with the title *Spatial variability of the atmosphere over southern England, and its effect on scene-based atmospheric corrections.* This paper was jointly published by three authors: EJ Milton and JM Nield provided supervision during the research process, contributing ideas and assisting with editing of the manuscript, but the rest of the research and writing (including the development of the original idea, data acquisition, processing, radiative transfer simulations, development of statistical approaches and analysis of the reslts) was carried out by RT Wilson alone.

2.1 Introduction and Background

Remotely sensed data are typically used to generate quantitative products which require a high degree of accuracy, for example, satellite sensor data are applied to estimate the Global Climate Observing System Essential Climate Variables (GCOS, 2004) such as snow cover (Hall *et al.*, 1995), sea-surface temperature (Brown *et al.*, 1999), albedo (Wielicki *et al.*, 2005), water vapour (Gao and Kaufman, 2003) and net primary productivity (Running *et al.*, 2004). To produce these variables, remotely-sensed data must undergo atmospheric correction to remove the perturbing effects of the atmosphere from the data, and thus allow results to be determined accurately in physical units

(a) Overview (b) Study area

FIGURE 2.1: Location of the NCAVEO Field Campaign

(Slater, 1980). A range of atmospheric correction methods can be used with satellite sensor data but most methods assume that the atmosphere is spatially-uniform across the image. However, over large images such as those from Landsat (185km x 185km) or DMC (a swath width of 650km) the atmosphere is likely to vary and so uniform correction methods may introduce significant errors in the resulting data products.

This chapter investigates errors associated with uniform atmospheric correction over large images of southern England. First, the spatial variability of the atmosphere over southern England on a clear (cloud-free) day is quantified, and then the magnitude and range of errors associated with uniform atmospheric correction over this area is assessed, both in terms of radiance and NDVI.

2.1.1 Study area and period

The study used data over southern England (the grey area in Figure 2.4) from the 16th and 17th June 2006. These were typical mid-latitude clear days during the NCAVEO Field Campaign (Milton *et al.*, 2011). This unique field campaign was carried out in the area surrounding Chilbolton, near Andover, Hampshire, UK (Figure 2.1) and collected a multi-scale, multi-sensor dataset to enable research into the calibration and validation of earth observation data. The data included large quantities of satellite, airborne and ground data (Table 2.1), with as much of the data as possible collected on a single day to allow easy comparison.

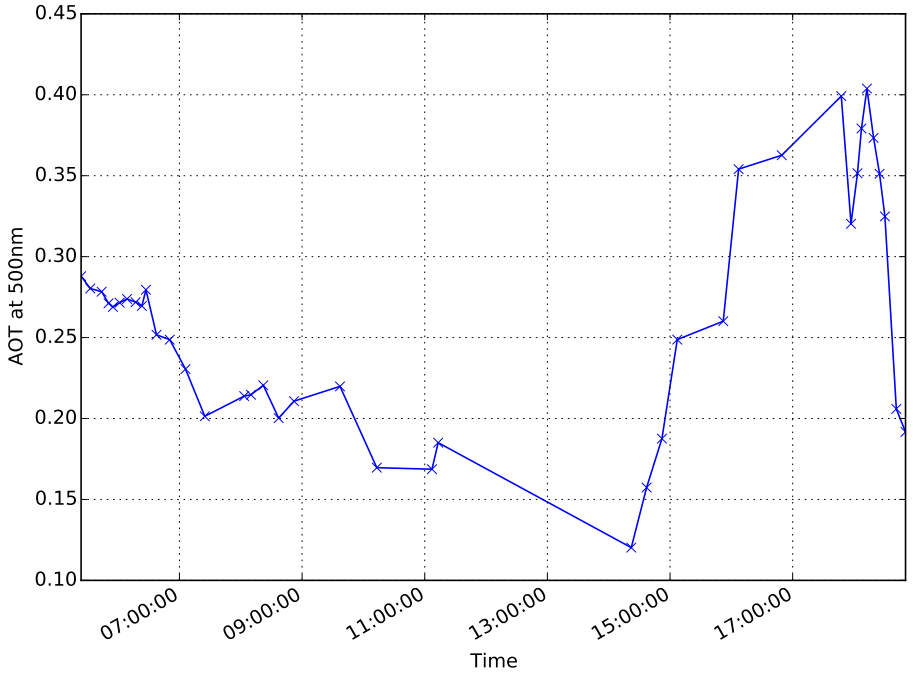
TABLE 2.1: Summary of the datasets collected by the NCAVEO Field Campaign

Platform	Datasets	Resolution
Satellite	CHRIS/Proba	34m
	UK-DMC	32m
	AlSat	32m
	Nigeria-Sat	32m
	SPOT-5 HRG	10/20m
Airborne	CASI-2	2.5m
	CASI-3	1m
	AISA Eagle	1m
	AIC digital camera	1m
	Digital Multispectral Camera	0.6m
	LiDAR	1m
Ground	Land cover	-
	LAI	-
	Vegetation canopy structure	-
	Leaf chlorophyll content	-
	River survey	-
	AOT and PWC from Cimel/Microtops	-
	Atmospheric water vapour profile	-
	Sky spectral irradiance distribution	-
	Diffuse:Global ratio	-
	Field spectroscopy of a range of sites	-

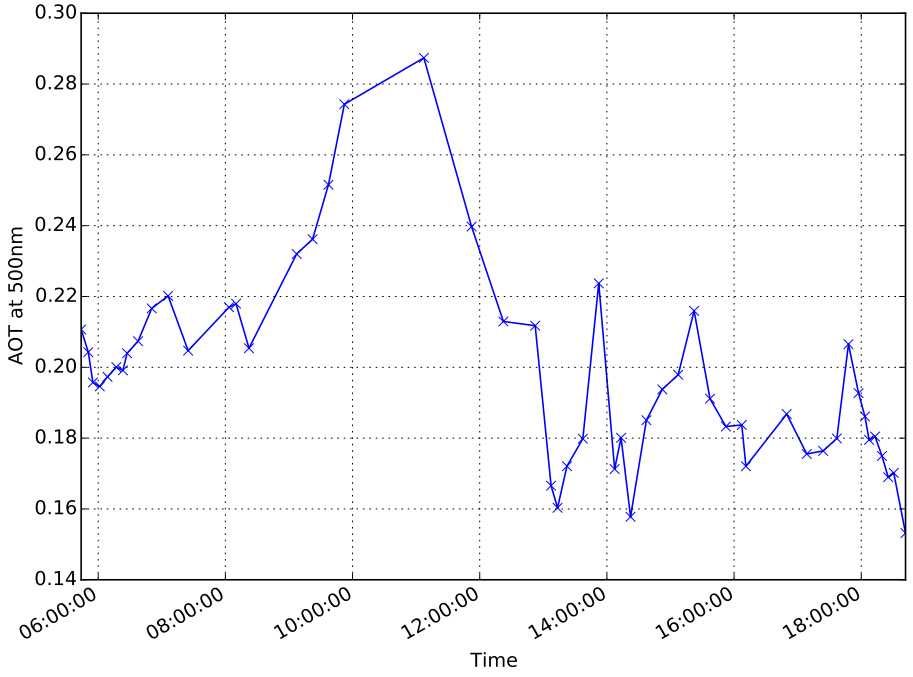
The meteorological situation changed significantly during these two days, as a high-pressure system migrated from the southern Atlantic Ocean, over southern England to Germany. The passage of this weather system caused significantly different wind directions on the two days (with average directions of 287° on the 16th and 192° on the 17th). Field observations confirm that conditions during the day on the 17th were more variable than the 16th (Milton *et al.*, 2011), with an increase in AOT (Figure **REF HERE**) associated with a reduction in sky clarity after 10:30 UTC. On the 16th there was also an increase in AOT, but this did not occur until late afternoon, by which time all data collection had been completed.

2.1.2 Background

In the early days of satellite remote sensing, simple scene-based atmospheric correction techniques such as Dark Object Subtraction were generally used. In the 1990s and 2000s there was significant development in per-pixel approaches designed for use with hyperspectral imagery. However, there is a lack of true pixel-based correction methods for multispectral imagery. Tools such as ATCOR (Richter, 2004) and FLAASH (Cooley *et al.*, 2002) can be used to perform a pseudo-pixel-based correction of multispectral images, but the methods that are used for extracting spatially-variable atmospheric information from multispectral imagery are limited. Thus, the majority of multispectral atmospheric corrections performed today are scene-based, using constant values of atmospheric



(a) 16th



(b) 17th

FIGURE 2.2: AOT measured at the Chilbolton AERONET site during the NCAVEO Field Campaign

parameters across the scene and not taking into account the spatial variability of the atmosphere over the image.

It is particularly important to quantify errors resulting from these scene-based atmospheric corrections due to three recent developments within remote sensing. First, the use of large images with a variety of spatial resolutions in environmental studies is becoming increasingly common, due both to the increased availability of large images and the policy-driven need for large area studies, particularly those relating to environmental change. Typical sensors include MODIS (500m resolution) and the Disaster Monitoring Constellation (30m resolution), which both produce images that cover very large areas (a single DMC image can cover approximately half the area of England). Second, with the incorporation of atmospheric correction tools into image processing software, atmospheric correction can now be performed by users who may have little knowledge of the possible uncertainties of the results. Third, data obtained from quantitative analysis of remotely-sensed images are now in widespread use for a variety of important scientific projects and errors in this data could have serious consequences. In climate modelling, significant errors in input data caused by incorrect atmospheric correction could result in misleading predictions being reported to policy makers. For example, Saleska *et al.* (2007) stated that the Amazon rainforest was more resilient to short-term climatic fluctuations than previously thought (as shown by a significant increase in Enhanced Vegetation Index), but Samanta *et al.* (2010) showed that these inferences were due to the use of cloud- and aerosol-contaminated satellite data in the original study.

2.1.3 Atmospheric parameters of interest

The primary atmospheric constituents which affect remotely-sensed measurements are mixed gases, ozone, aerosols and water vapour. Concentrations of atmospheric mixed gases are controlled by atmospheric pressure, and ozone concentrations can be modelled effectively by latitude and season (van Heuklon, 1979). However, aerosol and water vapour concentrations vary significantly both spatially and temporally, and thus contemporaneous data on these must be provided when atmospherically-correcting satellite images.

The *Aerosol Optical Thickness* (AOT), also known as the *Aerosol Optical Depth* (AOD), is a dimensionless measure of the degree to which aerosols restrict the transmission of light through the atmosphere, defined as the integrated extinction coefficient due to aerosols through a vertical column of unit area in the atmosphere (Iqbal, 1983).

Water vapour in the atmosphere can be quantified in two ways: Integrated Water Vapour (IWV), the vertically integrated mass of water per unit area ($\mathrm{kg\,m^{-2}}$), or Precipitable Water Content (PWC), the height of an equivalent column of liquid water (mm) (Iqbal, 1983).

2.1.4 Previous work

Previous studies that have assessed AOT variability have typically used i) low resolution data or ii) daily, weekly or monthly composites which are relevant for climate-related studies but not for assessing the spatial variability in AOT at the specific instant of satellite image acquisition.

González *et al.* (2003) and Koelemeijer *et al.* (2006) both examined the spatial variation of AOT across Europe using MODIS and ATSR-2 data respectively. These data were averaged to monthly or yearly periods, and so only provide estimates of an average variability of 0.2–0.5. González *et al.* (2003) found a wide range in AOT values across Europe, with values of 0.5–0.6 in industrialised areas of Germany and northern Italy, and values of 0.1 in rural areas of France, Spain and Norway, which suggests that local emissions are particularly important in determining AOT values. Koelemeijer *et al.* (2006) also found significant local effects, with many cities easily distinguishable as peaks in the data, and particularly low AOT in mountainous areas. The AOT values in southern England from the same study reflect this, with high values around London and the Thames Estuary and generally low values in rural Cornwall. González *et al.* (2003) also found that AOTs can increase by up to 300% over distances of around 50–10km (along a transect from Germany to Belgium), and similar gradients occurred in their data for the UK (eg. an increase of 275% from east Kent to mid Oxfordshire). AOTs also vary temporally, both diurnally (Smirnov *et al.*, 2002) and over weekly periods (Bäumer and Vogel, 2007). These anthropogenic variations, along with the prevailing meteorological situation, significantly impact the spatial variability of AOT, and thus yearly or monthly averages do not provide the information required for assessing the effect of uniform atmospheric correction procedures. For example, Figure 2.3 shows monthly average AOTs measured at the Chilbolton AERONET site: this shows a pattern of higher AOTs in spring and lower AOTs in winter, and an overall range of approximately 0.07–0.30, but does not tell us anything about the spatial variability of AOT at a specific point in time.

2.1.5 Uniformity assumptions in atmospheric correction methods

Very few implementations of atmospheric correction methods take into account the spatial variability of the atmosphere, even though these methods are conceptually able to work with a variable atmosphere. Typical relative or empirical methods, such as Dark Object Subtraction (Chavez, 1975; Moran *et al.*, 1992) and the Empirical Line Method (Smith and Milton, 1999), use averages of measurements taken across the image, thus ignoring the data on spatial variability which would be present in these measurements. Physical correction methods involve running a Radiative Transfer Model (for example, Berk *et al.*, 1999; Vermote *et al.*, 1997) on each pixel in the image, and thus could easily take into account spatial variability in atmospheric parameters.

The Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS; Masek *et al.*, 2006) partially accounts for spatial variability by estimating AOT over areas of dense dark vegetation (DDV) in the image using the Kaufman *et al.* (1997) method. The AOT data

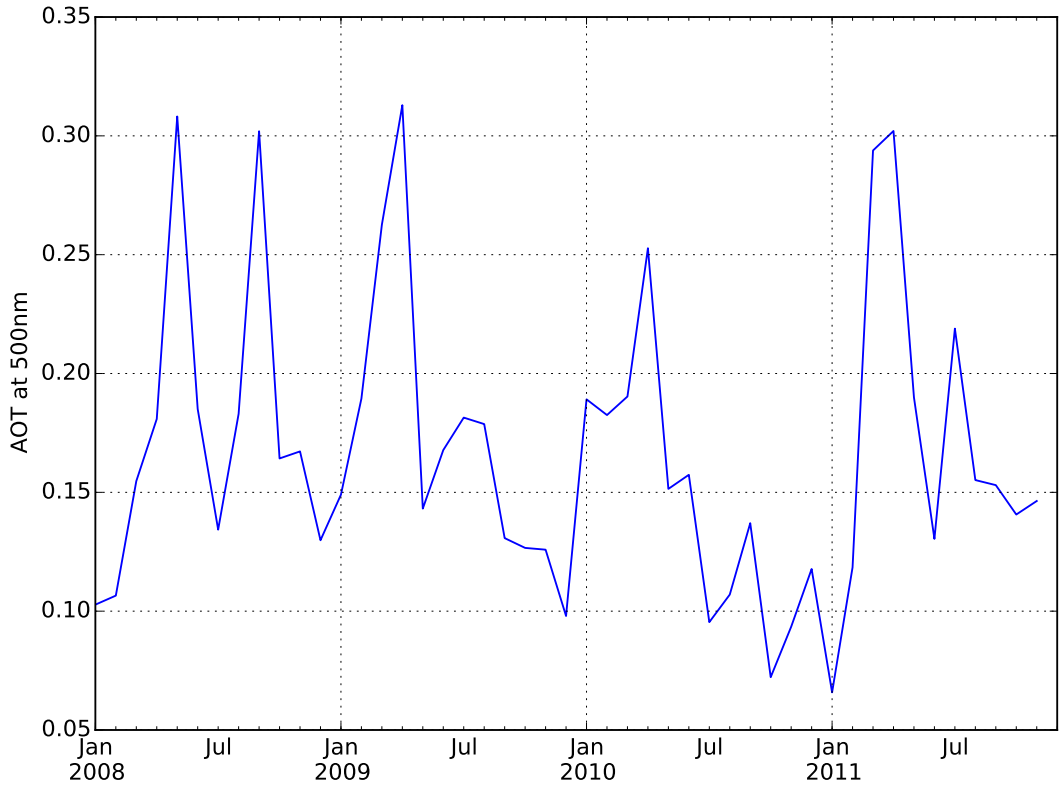


FIGURE 2.3: Monthly average AOT values measured at the Chilbolton AERONET site between 2008 and 2011 (the longest subset with no significant no data periods)

are then interpolated to 1km resolution and used to parameterise 6S (Vermote *et al.*, 1997) to perform the atmospheric correction. However, there are several problems with this method: i) it does not take into account fine-scale variability in AOT; ii) it can only estimate AOT over areas of DDV thus making it impossible to use over areas without DDV, such as deserts; and iii) it is only implemented for Landsat images as the AOT retrieval method requires the use of the Landsat short-wave infra-red bands.

There has been discussion within the community as to whether atmospheric correction is required in all situations. Song *et al.* (2001) state that atmospheric correction is not required for applications which require only a single image and do not need the data to be in physical units. For example, they argue that performing a maximum likelihood classification of a single Landsat image using training samples derived from the image data itself would give exactly the same result with and without atmospheric correction. This is because a uniform atmospheric correction would simply apply the same correction to each pixel in each band, thus changing the mean of each land cover class, but not altering the covariance of the classes. However, if a spatially-variable atmospheric correction were to be performed the correction applied each pixel would be different (based upon the atmospheric conditions over that pixel), and thus the covariances of the classes would change. Similarly, if the image was acquired through a spatially-variable atmosphere but a uniform atmospheric correction was performed there would be a

different error for each pixel, which would affect the covariances of the classes. Thus, if the atmosphere is spatially-variable, a full spatially-variable atmospheric correction is required even for the uses specified in Song *et al.* (2001).

2.2 Data Sources & Validations

AOT and PWC can be measured using a variety of ground instruments and satellite products with a range of spatial and temporal resolutions (Table 2.2). The uncertainty of these methods depends on the location they are used in, with factors such as land cover and aerosol type having a large effect, so it is important to perform validation for the study site.

Validation of satellite measurements against ground measurements is challenging for several reasons including: i) the lack of exactly coincident measurements, ii) differences in cloud screening, iii) different measurement variables, and iv) the fundamental difference between areally integrated measures from satellites and point-based measurements from ground instruments. The Ichoku *et al.* (2002) spatio-temporal subset approach for validation is used here, comparing a spatial subset from the satellite data (5 x 5 pixels) with a temporal subset from ground measurements (±30 minutes). Ichoku *et al.* (2002) justified the size of these subsets based upon an estimate of average aerosol front speed, the requirement to obtain a statistically-significant sample size, and the observation that larger window sizes could introduce errors from cloudy pixels and changing topography.

2.2.1 AERONET Sun Photometry

Sun photometers estimate AOT and PWC based upon measurements of solar irradiance in multiple wavelengths. Here, automatically cloud-screened data (Level 1.5) collected by the AERONET Cimel CE-318 sun photometer situated at the Chilbolton Facility for Atmospheric and Radio Research (CFARR) (Holben *et al.*, 1998) are used. A simple time-for-space substitution was used to obtain an estimate of the spatial variability over the whole area from this single point measurement, by taking the AERONET measurements over the entire daylight period and assuming that they are representative of the AOT across the whole study area.

Sun photometers are used as reference data within this study, as they are currently the most accurate method for measuring AOT (Wang *et al.*, 2009). Errors are low: approximately ±0.02 for AOT (Eck *et al.*, 1999), and with an PWC RMSE of 2.9 mm (Liu *et al.*, 2011).

The time-for-space assumption was examined by modelling the passage of aerosol particles across the UK during the study period using the Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT; Draxler and Rolph, 2003), using gridded one degree resolution meteorological data from the NCEP Global Data Assimilation System archive (Kalnay *et al.*, 1996). The model was parameterised to simulate an ensemble of possible back-trajectories for a single particle located above the AERONET site at 19:00

TABLE 2.2: Summary of characteristics and accuracy of all data sources considered

Source	**Type**	**Spatial Resolution**	**Temporal Resolution**	**Official validation**	**Study area validation**	**Closest acquisition time**
AERONET	Ground	One location	Every 15 minutes, in good weather	± 0.02	—	09:37
Met Office	Ground	36 stations across study area	Hourly	—	RMSE: 0.05–0.47 (for visibilities 40–10km)	10:00
MODIS AOT (MOD04)	Satellite	10km	Daily merged *or* once per orbit	$\pm 0.05 \pm 0.15\tau$	No sig difference	16th: 11:34 17th: 10:38
GlobAerosol	Multi-Satellite	10km	Daily	RMSE: 0.12	16th: No sig difference 17th: Sig difference	Daily
MODIS PWC (MOD05)	Satellite	1km	Daily merged *or* once per orbit	RMSE: 1.7mm	16th: No sig difference 17th: Sig difference	16th: 11:34 17th: 10:38
BIGF Water Vapour	Ground	25 stations across study area	Hourly	—	RMSE: 1.5mm	10:00

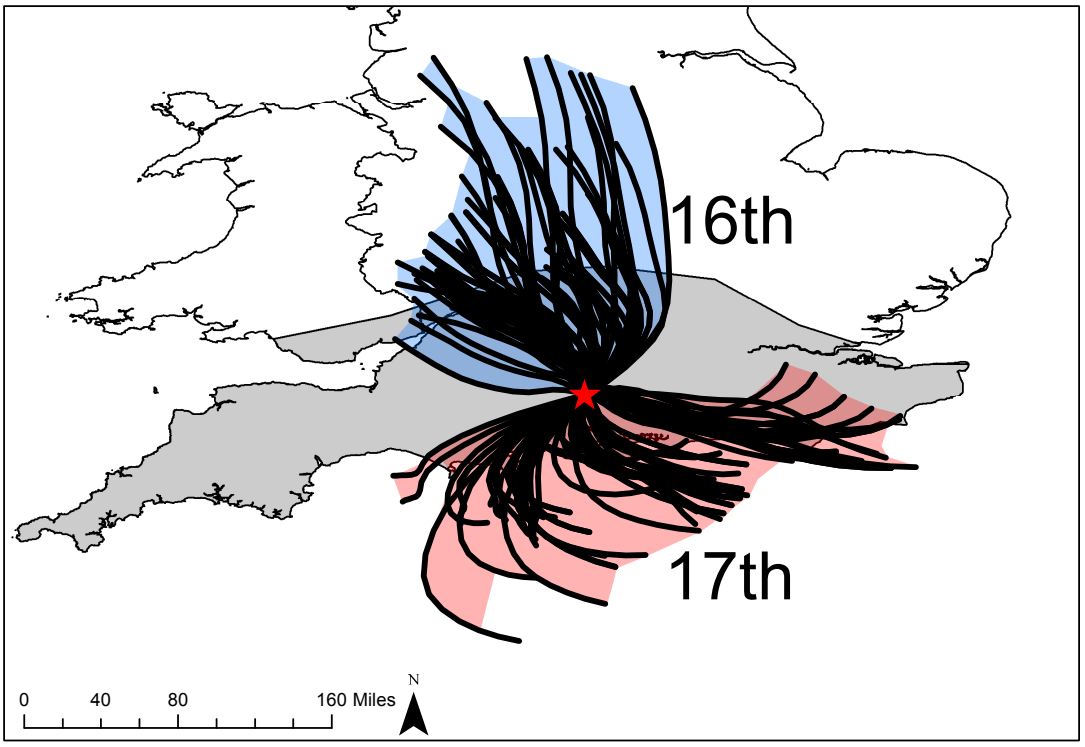


FIGURE 2.4: Estimates of 'contributing areas' for the Chilbolton AERONET site for the 16th and 17th June 2006 in blue and red respectively, with the study area shown in grey. The black lines show the individual trajectories towards Chilbolton, Hampshire (marked with a red star) computed by HYSPLIT using the ensemble mode with heights of 500m, 1000m and 1500m above ground level.

UTC back to its starting location at 05:00 UTC (the start and end times of the AERONET data during the study period). Simulations were run for particles at heights of 500m, 1000m and 1500m to capture the differing trajectories produced by height-varying winds. These heights were chosen based upon the finding of Matthias *et al.* (2004) that 80–90% of the AOT is produced by aerosols in the planetary boundary layer, which was found to be at a height of 1204 ± 481m at the Aberystwyth station, located approximately 100km outside the study area.

A simple 'contributing area' for the AERONET site was then calculated as the concave hull of the resulting trajectories. These estimated areas for the 16th and 17th June 2006 (Figure 2.4) show that the time-for-space substitution covered 23% and 19% of the study area on the 16th and 17th respectively. The contributing area for each day was very different due to contrasting meteorological conditions, and a large proportion of the contributing area was outside of the study area (62% and 63% respectively).

2.2.2 Met Office Visiometry

AOT was estimated from hourly measurements of horizontal visibility (accurate to $\pm 10\%$) acquired by a network of UK Met Office stations across the study area (UK Met Office, 2006) using Koschmieder's equation (Koschmieder, 1924; Horvath, 1981):

$$V = \frac{3.912}{\tau} \tag{2.1}$$

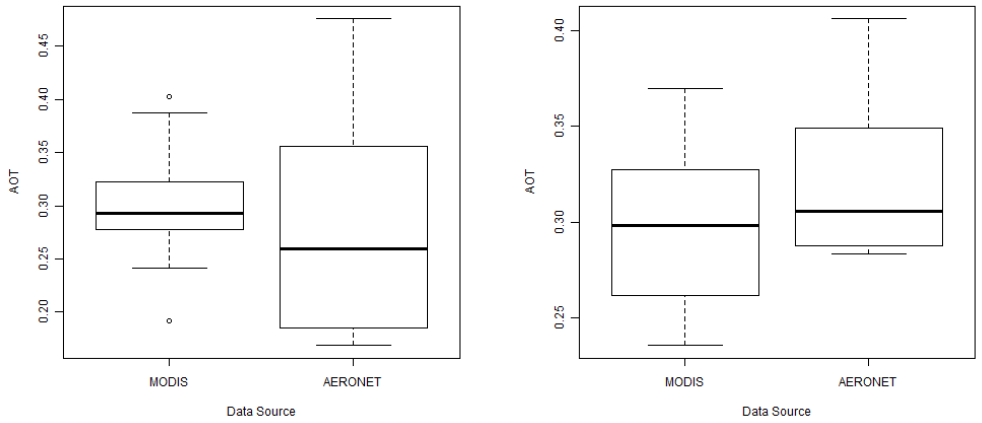
where V is the visibility in km and τ is the AOT.

Koschmieder's equation relates horizontal visibility and horizontal extinction coefficient measurements, but is now widely used for calculating vertical extinction coefficients (that is, AOT) from horizontal visibility. This mixing of horizontal and vertical measurements relies on many assumptions which are often invalid (Chan, 2009), and there are broader issues with the choice of coefficients in the equation (Middleton, 1952; Horvath, 1971, 1981). Previous studies comparing AOT and visibility-based estimates of AOT, have found correlations ranging from 0.38 (So *et al.*, 2005) to 0.89 (Chen *et al.*, 2009). However, despite these limitations, visibility-based estimates of AOT are still useful due to their high spatial and temporal resolutions as well as the wide availability of data collected according to World Meteorological Organisation standards.

2.2.3 MODIS AOT (MOD04)

The MOD04 product from the MODIS sensors on the Terra and Aqua satellites provides AOT estimates at 10km resolution using an algorithm based on short-wave infra-red measurements and the use of a Radiative Transfer Model lookup table (Remer *et al.*, 2006). The official MODIS validation report for the latest version of the algorithm (Collection 5.1) (Remer *et al.*, 2006) states that 67% of the retrievals were within the expected uncertainty ($\pm 0.05 \pm 0.15\tau$), which has been confirmed by independent validations (Levy *et al.*, 2010). Results improve when only pixels with the highest Quality Assurance Confidence were used, as in this study. Many assessments of MOD04 accuracy in the literature are based upon previous versions of the algorithm (Collection 4), but the current algorithm (Collection 5.1) has significantly improved the accuracy. The accuracy is seasonally-variable (El-Metwally *et al.*, 2010), likely due to the seasonal changes in aerosol types present over some of the sites used in their study.

Validation for the study area was performed between the MOD04 product and the AERONET site at Chilbolton using the Ichoku *et al.* (2002) method (Figure 2.5). The results on the 16th show that AERONET had a significantly larger range than the MOD04 product, due to the poor performance of the time-for-space assumption combined with potential cloud masking issues and the naturally higher variability found in point measurements as opposed to those averaged over large pixels. Although there were differences in range, results from t-tests showed that there was no significant difference between the samples obtained from MODIS and AERONET ($p = 0.83$ and $p = 0.28$ for the 16th and 17th respectively), and thus they are likely to have come from the same distribution of AOT values.



(a) 16th June 2006 (b) 17th June 2006

FIGURE 2.5: Boxplots showing the summary statistics for the validation of the MODIS AOT product against AERONET

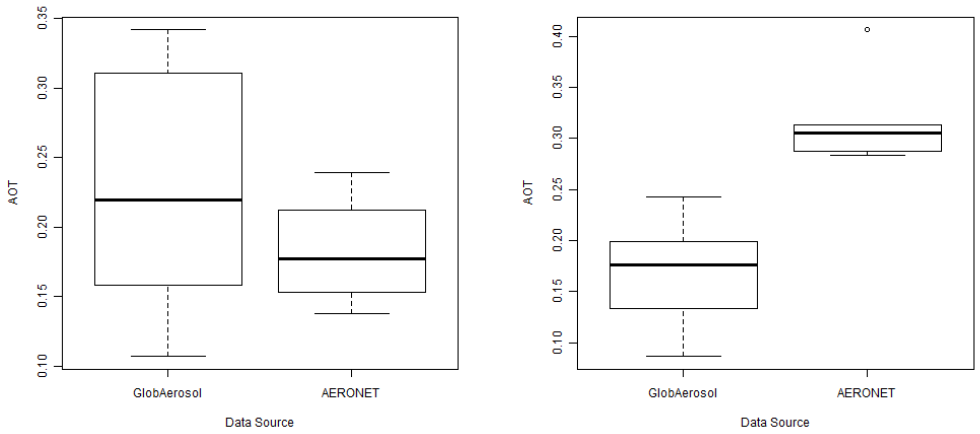
2.2.4 GlobAerosol

The GlobAerosol product is produced by merging AOT products from the ATSR-2, AATSR, MERIS and SEVIRI sensors (Thomas *et al.*, 2010), at a 10km resolution. Poor data are excluded based upon various checks, and merging is performed using temporal interpolation, with observations weighted by their error estimates (Siddans *et al.*, 2007). The official validation against AERONET measurements found that the AATSR-derived dataset was most accurate, with a RMSE of 0.07 (Poulsen *et al.*, 2009). Although the merged product has a lower accuracy (Poulsen *et al.*, 2009), the major advantage is that the merging process ensures higher spatial coverage of the area.

A comparison between the GlobAerosol merged product and the AERONET site at Chilbolton (Figure 2.6) show a smaller difference between AERONET and GlobAerosol measurements on the 16th than on the 17th June. Results from t-tests indicate that there was a significant difference between the AOT samples from MODIS and AERONET on the 17th ($p = 0.0008$), but no significant difference on the 16th ($p = 0.20$).

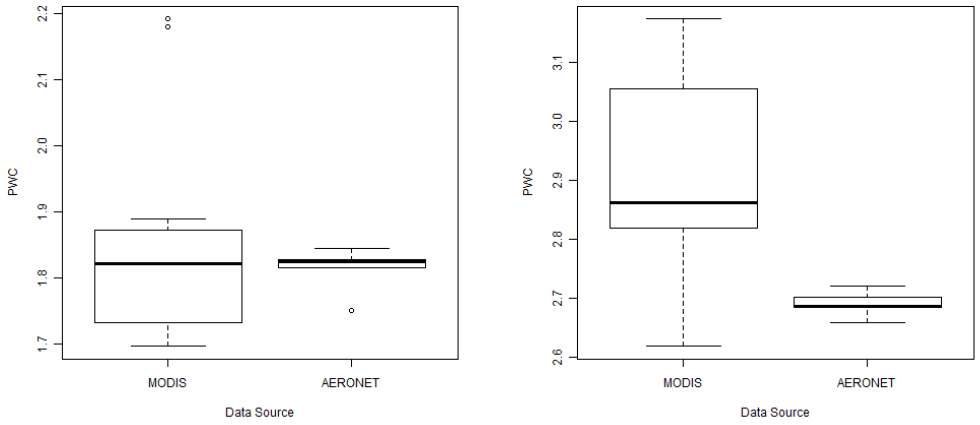
2.2.5 MODIS PWC (MOD05)

The MOD05 product provides PWC estimates at 1km resolution, based upon a ratio of adjacent bands with and without water absorption features. Official validation for the MODIS water vapour product is limited (Gao and Kaufman, 2003), with a RMSE based on a microwave radiometer dataset of 1.7mm, corresponding to an approximate 5–15% error for the PWC range found over southern England (10–40mm). Comparisons of the MOD05 PWC estimates to radiosonde and GPS-based measurements at Herstmonceux in southern England (50.889 N, 0.324 E) found a positive bias of 10% and 7% respectively (Li *et al.*, 2003), and comparisons in the Tibetan Plateau produced a similar result to the official validation (1.95mm RMSE).



(a) 16th June 2006 (b) 17th June 2006

FIGURE 2.6: Boxplots showing the summary statistics for the validation of the GlobAerosol product against AERONET



(a) 16th June 2006 (b) 17th June 2006

FIGURE 2.7: Boxplots showing the summary statistics for the validation of the MODIS Water Vapour product against AERONET

Validation between the satellite data and the AERONET PWC measurements at Chilbolton (Figure 2.7) indicated similar relationships between the spatial and temporal subsets on the 16th ($p = 0.437$), but not on the 17th ($p = 0.004$, as found with the other datasets. Again, the satellite data has a larger range, but on both days the AERONET data were encompassed within this range.

2.2.6 GPS Water Vapour

Measurements of delays in the GPS L-band radio signals passing through the atmosphere can be used to quantify the water vapour in the atmosphere above the GPS receiver (Bevis *et al.*, 1992). The British Isles Continuous GNSS Facility (BIGF; Natural Environment Research Council, 2012) uses these methods to provide estimates of integrated water

vapour at all BIGF stations on an hourly basis. Previous validations of GPS-derived water vapour estimates against radiosonde and satellite data have generally produced errors of 1–2mm (Becker *et al.*, 2003; Wang *et al.*, 2007; Wolfe and Gutman, 2000; Tregoning *et al.*, 1998). However, all the data used in these validations were processed from the GPS data by the authors, using models which were specifically parameterised for the area of study. The BIGF product is a global operational product, and thus uses a standard parameterisation across all sites, which may be expected to reduce the accuracy.

A BIGF measurement site is co-located with the AERONET site at Chilbolton, and validation was performed for all days with at least two matching measurements in the period August 2009–November 2010. This produced an average daily RMSE of 1.5mm, with a maximum of 6.3mm, and a Pearson correlation coefficient of 0.976, showing a good agreement between GPS-derived and AERONET-derived estimates.

2.3 Simulation of uniform atmospheric corrections

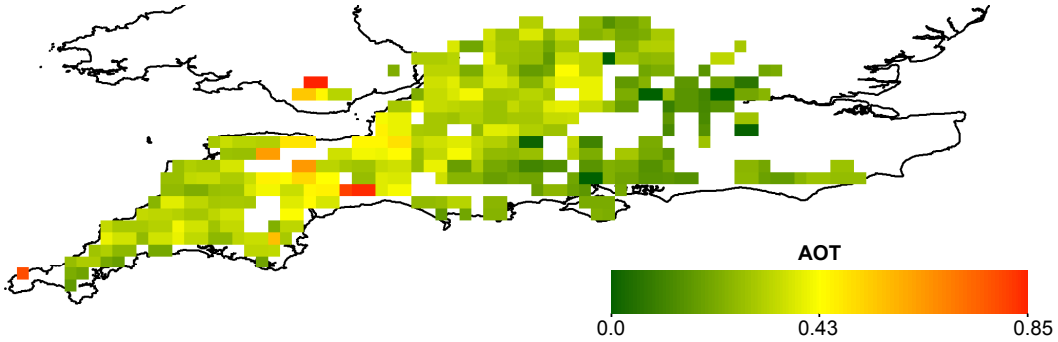
The 6S radiative transfer model (Vermote *et al.*, 1997) was used to simulate a uniform atmospheric correction over southern England, using the data on spatial variability described above. The Py6S (Wilson, 2012, available in Appendix A) interface to 6S was used to allow hundreds of individual simulations to be run in an automated manner. Simulations were run in two stages: first to generate a top-of-atmosphere (TOA) radiance from a representative vegetation spectrum under a given set of atmospheric parameters (P_{up}), and second to atmospherically correct the TOA radiance to a ground reflectance under a different set of atmospheric parameters (P_{down}). P_{up} was set to the 5% or 95% quantile of the AOT or PWC values and P_{down} was set to the mean of the AOT or PWC values (from Tables 2.3 and 2.4), thus simulating the uniform atmospheric correction of a pixel measurement which was actually acquired in extreme conditions. Simulations were performed for Landsat bands 1–4, and 6S parameters other than AOT and PWC were set to appropriate values for southern England. Results from the simulations were retrieved as reflectance values. To assess the effect on a standard remote-sensing product, NDVI was also calculated from these reflectances.

The errors resulting from a uniform atmospheric correction are conceptually the same as the errors resulting from uncertainty in the atmospheric parameters: both are caused by differences between the true parameter value and the value used for correction. Thus, a sensitivity analysis was also performed to assess the effects of the uncertainties of the data sources (as listed in Table 2.2) on remote sensing products.

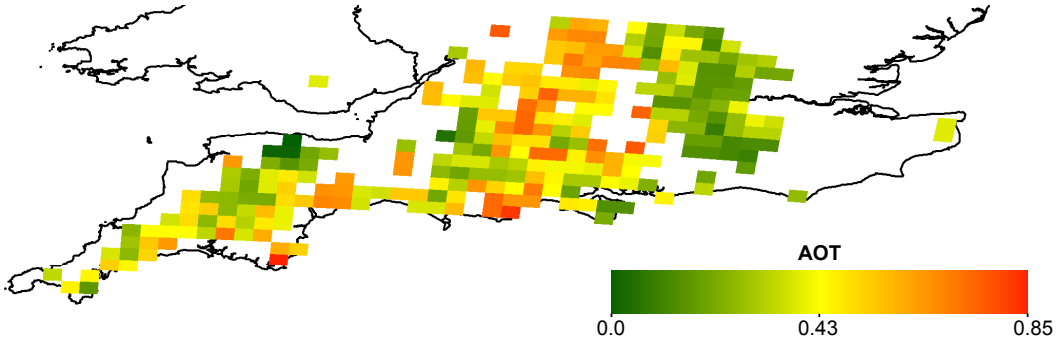
2.4 Spatial variability over the study area

2.4.1 Aerosol Optical Thickness

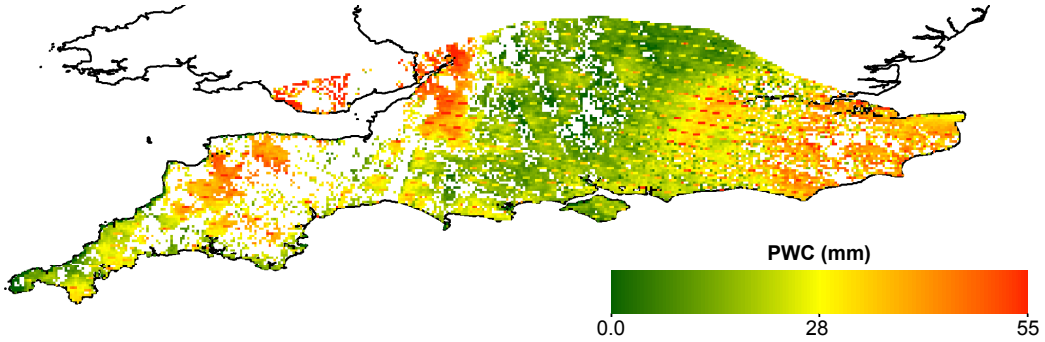
The AOT range over the study area on the 16th and 17th June 2006 was approximately 0.1–0.5 (Table 2.3). This is large, given that these measurements were acquired on days which had mostly clear skies across the study area, and shows that there is more spatial



(a) GlobAerosol



(b) MODIS AOT (MOD04)

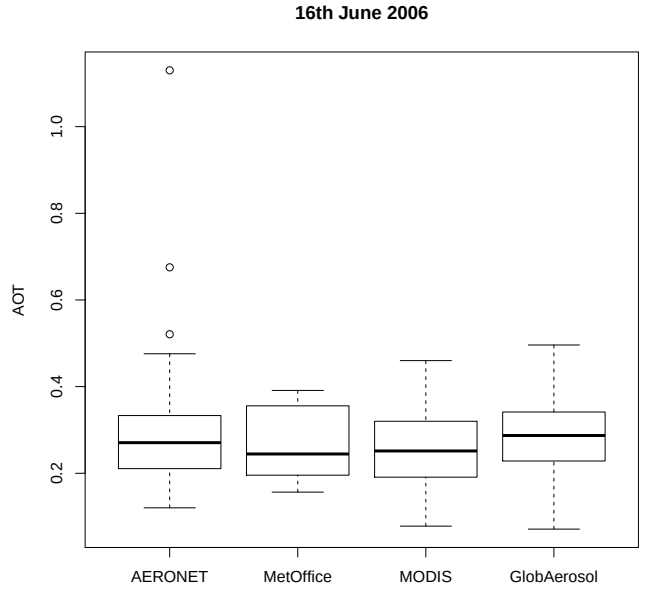


(c) MODIS PWC (MOD05)

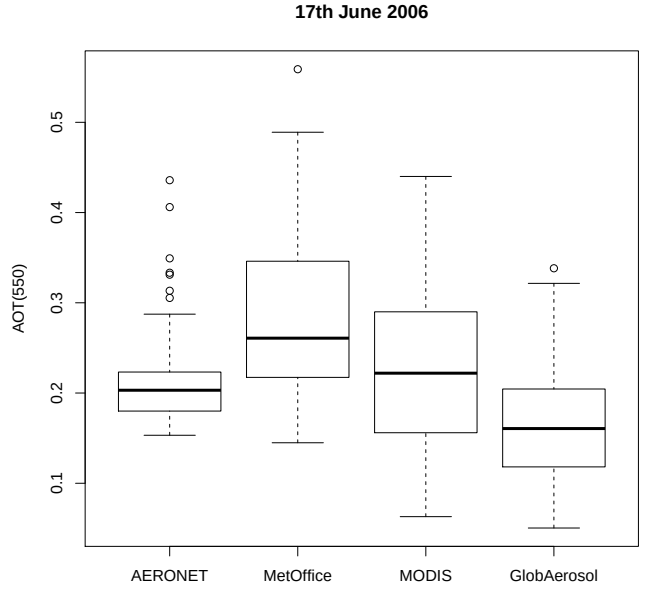
FIGURE 2.8: Examples of the three satellite data products used in this study. All images are from the 16th June 2006.

variability in AOT than visual examination of sky conditions suggests.

The overall range in AOT was similar each day, but all data sources have significantly higher variability on the 17th (Figure 2.9). This is consistent with the more changeable weather conditions on the 17th, as noted by the records from the NCAVEO Field Campaign (Milton *et al.*, 2011). Similarly, the median values for each dataset are very similar on the 16th, but not on the 17th. In all cases the satellite-based datasets (MODIS and GlobAerosol) have a lower minimum, which is likely to be due to errors in separating the at-sensor radiance into the ground reflectance and aerosol scattering components.

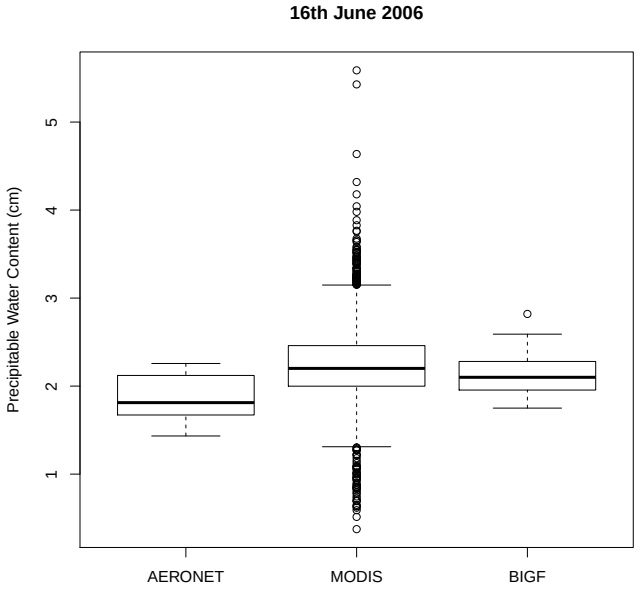


(a) 16th June

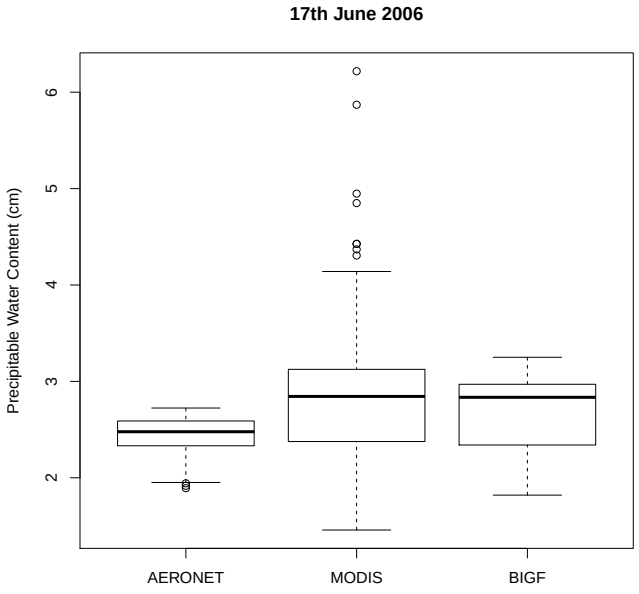


(b) 17th June

FIGURE 2.9: Boxplots showing the range of AOT values found over southern England during the study period according to each data source.



(a) 16th June



(b) 17th June

FIGURE 2.10: Boxplots showing the range of PWC values found over southern England during the study period according to each data source.

TABLE 2.3: Summary statistics showing the range of AOT values across southern England during the study period for each data source.Q05 and Q95 are the 5% and 95% quantiles respectively.

(a) 16th June 2006

Source	**Min**	**Max**	**Mean**	**Q05**	**Q95**
AERONET	0.120	1.130	0.291	0.156	0.464
Met Office	0.156	0.391	0.260	0.156	0.391
GlobAerosol	0.071	0.496	0.287	0.155	0.417
MODIS	0.078	0.460	0.258	0.139	0.398

(b) 17th June 2006

Source	**Min**	**Max**	**Mean**	**Q05**	**Q95**
AERONET	0.153	0.436	0.216	0.160	0.332
Met Office	0.145	0.559	0.296	0.175	0.489
GlobAerosol	0.050	0.338	0.164	0.082	0.258
MODIS	0.063	0.440	0.223	0.090	0.386

TABLE 2.4: Summary statistics showing the range of PWC (in cm) across southern England during the study period for each data source Q05 and Q95 are the 5% and 95% quantiles respectively.

(a) 16th June 2006

Source	**Min**	**Max**	**Mean**	**Q95**	**Q05**
AERONET	1.433	2.257	1.859	2.236	1.486
BIGF	1.750	2.820	2.138	2.565	1.813
MODIS	0.212	5.588	2.180	2.772	1.646

(b) 17th June 2006

Source	**Min**	**Max**	**Mean**	**Q95**	**Q05**
AERONET	1.892	2.724	2.433	2.688	1.975
BIGF	1.820	3.250	2.676	3.190	1.904
MODIS	1.458	6.218	2.763	3.379	1.995

The AOT data obtained from Met Office visibility measurements has a similar range to the other datasets on the 16th, but over-estimates the AOT on the 17th. This is likely caused by failure of the assumptions inherent in the visibility to AOT conversion due to the meteorological conditions. For example, local conditions could reduce horizontal visibility at ground-level, but not significantly affect the vertical extinction coefficient measured by AOT.

Generally the AERONET measurements have the lowest inter-quartile range (16th: 0.12, 17th: 0.04) due to the time-for-space substitution not capturing the variability across the entire study area, but have several high outliers (including a value of 1.13 on the 16th). These outliers are likely to be due to poor performance of the automated cloud screening algorithm used for the level 1.5 data, which performs relatively poorly for large areas of

TABLE 2.5: Effects of a uniform atmospheric correction performed over an area with the AOT variability from each data source. Values in the table are the reflectance differences for [95% perturbation; 5% perturbation], with reflectance values in percent. Note that increases in AOT (using the 95% percentile of the AOT data from the data source) cause increases in reflectances for all bands, but a decrease in NDVI.

(a) 16th June 2006

Source	$\rho_{\mathbf{B}}$	$\rho_{\mathbf{G}}$	$\rho_{\mathbf{R}}$	$\rho_{\mathbf{NIR}}$	**NDVI**
AERONET	+1.3; -1.0	+1.1; -0.8	+1.0; -0.8	+0.1; -0.1	-0.026; 0.027
Met Office	+1.0; -0.8	+0.8; -0.6	+0.7; -0.6	+0.1; -0.1	-0.017; 0.022
GlobAerosol	+1.0; -1.0	+0.8; -0.8	+0.7; -0.7	+0.1; -0.1	-0.018; 0.027
MODIS	+1.1; -0.9	+0.9; -0.7	+0.8; -0.7	+0.1; -0.1	-0.019; 0.024

(b) 17th June 2006

Source	$\rho_{\mathbf{B}}$	$\rho_{\mathbf{G}}$	$\rho_{\mathbf{R}}$	$\rho_{\mathbf{NIR}}$	**NDVI**
AERONET	+0.9; -0.4	+0.7; -0.4	+0.6; -0.3	+0.1; -0.1	-0.014; +0.014
Met Office	+1.5; -0.9	+1.2; -0.8	+1.1; -0.7	+0.1; -0.1	-0.030; +0.025
GlobAerosol	+0.7; -0.6	+0.6; -0.5	+0.5; -0.5	+0.1; -0.2	-0.009; +0.017
MODIS	+1.2; -1.0	+1.0; -0.8	+0.9; -0.7	+0.1; -0.2	-0.022; +0.025

TABLE 2.6: Effects of a uniform atmospheric correction performed over an area with the PWC variability from each data source. Values in the table are the differences from the true results for [95% perturbation; 5% perturbation], and reflectance values are in percent. ρ_{B}, ρ_{G}, ρ_{R} and ρ_{NIR} are the reflectances in the Landsat blue, green, red and NIR bands respectively.

(a) 16th June 2006

Source	$\rho_{\mathbf{B}}$	$\rho_{\mathbf{G}}$	$\rho_{\mathbf{R}}$	$\rho_{\mathbf{NIR}}$	**NDVI**
AERONET	0.00; 0.00	-0.02; +0.02	-0.02; +0.02	-0.37; +0.40	-0.002; +0.002
BIGF	0.00; 0.00	-0.02; +0.01	-0.02; +0.02	-0.39; +0.32	-0.002; +0.002
MODIS	0.00; 0.00	-0.03; +0.02	-0.03; +0.02	-0.49; +0.44	-0.003; +0.006

(b) 17th June 2006

Source	$\rho_{\mathbf{B}}$	$\rho_{\mathbf{G}}$	$\rho_{\mathbf{R}}$	$\rho_{\mathbf{NIR}}$	**NDVI**
AERONET	0.00; 0.00	-0.01; +0.02	-0.01; +0.02	-0.22; +0.43	-0.001; +0.002
BIGF	0.00; 0.00	-0.02; +0.04	-0.02; +0.04	-0.41; +0.72	-0.002; +0.004
MODIS	0.00; 0.00	-0.03; +0.03	-0.03; +0.04	-0.49; +0.70	-0.003; +0.007

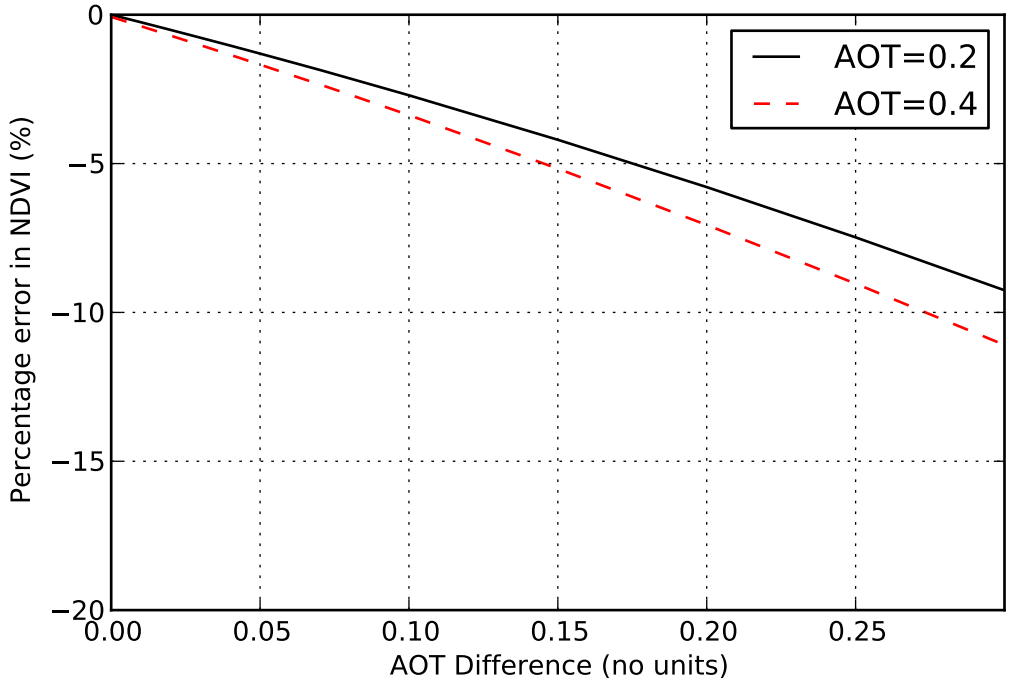


FIGURE 2.11: Sensitivity analysis showing percentage change in NDVI caused by correcting a standard green vegetation spectrum with a erroneous AOT value, for two standard AOT values (0.2, corresponding to a visibility of around 20km; and 0.4, corresponding to a visibility of around 10km). The x-axis shows the difference between the true AOT and the AOT used for correcting, and the y-axis shows the resulting error.

temporally and spatially homogeneous cloud (Smirnov *et al.*, 2000).

2.4.2 Precipitable Water Content

The range agreement between the data sources for PWC is weaker than for the AOT datasets (Figure 2.10), but an approximate range for PWC, taking into account expected values (Randel *et al.*, 1996) and obvious outliers on the 16th is approximately 1.5–3.0 cm on the 16th and around 2.0–3.5 cm on the 17th (Table 2.4). The larger values on the 17th were likely caused by the southerly/south-westerly winds bringing moist air-masses from the Atlantic Ocean over the study area. Again, the satellite data had the largest range, with many outliers for MODIS on the 16th June. High outliers may have been caused by incorrect cloud-screening (as clouds will have a significantly higher PWC) but the very low values (a minimum of 0.2 cm for MODIS) may be plausible in certain areas (a maximum PWC of 3 mm was found in Norway by Mook, 1978). The MODIS PWC dataset has a significantly higher resolution than the AOT datasets (1 km compared to 10 km) and is likely to record more small-scale variation that would be averaged out in a lower-resolution dataset, and thus have a larger range. The BIGF data compares well with the other datasets, with a smaller range than MODIS but similar inter-quartile ranges, showing the utility of this relatively-new measurement approach.

TABLE 2.7: Noise Equivalent Delta Radiance (NEΔL) and Noise Equivalent Delta Reflectance (NE$\Delta\rho$) for the visible and near infra-red Landsat bands under the simulation conditions.

	Blue	**Green**	**Red**	**NIR**
NEΔL	1.11	0.85	0.89	0.61
NE$\Delta\rho$	0.056	0.046	0.057	0.058
NE$\Delta\rho$/Error	26.6	26.1	19.1	3.4

2.4.3 Summary

All data sources confirm that the AOT and PWC over the study area was not uniform during the study period. As there were mainly clear skies during this time it is likely that the measured variability is a 'best case' scenario and that AOT variability will be greater in other situations. Thus, the assumption of atmospheric spatial uniformity made by atmospheric correction methods is not valid across a large area in southern England, and probably other mid-latitude areas.

2.5 Effects of uniform atmospheric correction

To estimate the implications of uniform atmospheric correction the 5% and 95% quantiles of AOT and PWC measured over the study period were used (as they represent the 'extreme' values that affect 10% of the pixels in the image) to simulate the effects of a uniform atmospheric correction for these 'extreme' pixels.

2.5.1 Aerosol Optical Thickness

The results of the Radiative Transfer Model simulations show that atmospheric correction of data acquired under a high AOT and corrected with a lower AOT produces erroneously high reflectances (Table 2.5). This is due to the increased scattering caused by the aerosols which was not corrected by the atmospheric correction. The error has a significant spectral dependence, with higher errors for lower wavelengths (blue) and very low errors for high wavelengths (NIR), and an overall range of 0.1-1.3 percentage points of reflectance. In this situation, the NDVI decreased, as the red reflectance increased relative to the NIR reflectance. The absolute NDVI difference was low, but the percentage error reached 5% for the Met Office dataset on the 17th June. Even relatively small errors in NDVI may affect derived products such as estimates of biomass production, for example, Kaufman (1993) found that a NDVI difference of 0.04 corresponded to biomass production errors of 11–30%.

To put these errors in context: the Noise Equivalent Delta Radiance ($NE\Delta L$) and equivalent Noise Equivalent Delta Reflectance ($NE\Delta\rho$) were calculated for each Landsat band under the simulation conditions, using the $NE\Delta\rho$ formula and data in Scaramuzza *et al.* (2004) (Table 2.7). The errors due to a uniform atmospheric (Table 2.5) are significant, at almost thirty times more than the NE$\Delta\rho$ for bands 1 and 2, approximately twenty times more for band 3 and three times more for band 4.

2.5.2 Precipitable Water Content

The reflectance differences caused by PWC perturbations are significantly smaller than those for AOT perturbations, with a maximum error of 0.5 percentage points in the NIR and 0.02 percentage points in the visible (Table 2.6). They have the opposite spectral dependence to the differences due to AOT, with low errors at short wavelengths, but high errors at longer wavelengths. This is because most multi-spectral satellite bands are deliberately located away from areas of the spectrum which experience significant water absorption, but Landsat band 4 (NIR, 0.76–0.90μm) covers a water absorption feature (Gao and Goetz, 1990). Although there is a differing effect between the NIR and red bands, it is not as significant as with the AOT perturbations, and thus the NDVI differences are much lower (with a maximum of 0.007).

2.5.3 Sensitivity Analysis

The sensitivity analysis (Figure 2.11) shows the how errors in AOT propagate to the resulting NDVI values for a range of errors. Comparing the results to the uncertainties of each dataset (Table 2.2) shows that there is a serious problem in using these datasets to provide AOT values for use in atmospheric correction procedures. The NDVI changes resulting from the official error estimates for each of the AOT data sources (Table 2.8), generate errors ranging from 2% to 7% for all data sources except AERONET. The effect of the AERONET error on NDVI is acceptable at less than 1%. This suggests that only AERONET data should be used for parameterising atmospheric correction models, but this research has also shown that a fully spatially variable correction is needed, and as AERONET sites are sparsely distributed this is not possible. The errors shown in this sensitivity analysis are a result of inaccuracies in the AOT data sources, and have no relationship to the uniform atmospheric correction issues that our main study focuses on. Thus, these errors could occur in any atmospheric correction that uses visibility, MODIS or GlobAerosol data to obtain the AOT input, regardless of whether the atmosphere is spatially uniform or variable.

2.5.4 Summary

Performing a uniform atmospheric correction for water vapour does not introduce unacceptable errors in reflectance or NDVI, with a maximum error of 0.7 percentage points and 0.6% respectively. Further simulations have shown that even in areas with very high water vapour amounts, such as tropical rainforests, NDVI values are unlikely to be significantly affected by variability in water vapour unless that variability approaches 80% of the mean value.

In contrast, performing a uniform atmospheric correction over a spatially-variable AOT distribution may cause errors in reflectance of up to 1.5 percentage points, and errors of 5% in NDVI values. Overall 5% of the pixels in the image may have a reflectance error $>$ +1.5 percentage points, and another 5% of the pixels may have an error $<$ -1.0 percentage points. To put this in context, 10% of the pixels in a Landsat image is approximately 3.8

million pixels, covering an area of approximately 3,500km^2.

TABLE 2.8: Resulting error in NDVI caused by AOT uncertainties (according to the official validation) for each data source, for AOTs of 0.2 and 0.4

	NDVI Error (%)	
Data Source	**AOT = 0.2**	**AOT = 0.4**
AERONET	0.51	0.70
MetOffice	3.37	7.06
GlobAerosol	3.29	4.08
MODIS	2.13	3.72

2.6 Conclusions

The spatial variability of the atmosphere over southern England was investigated by acquiring data on the Aerosol Optical Thickness (AOT) and Precipitable Water Content (PWC) from a wide range of ground- and satellite-based sources on two clear days. All data sources except the AERONET network of ground-based sun photometers had high uncertainty, but it was possible to extract a range of AOT and PWC over the study area for each day of 0.1–0.5 and 1.5–3.0 cm respectively. These ranges show that there is significant variation in these properties across this area.

The errors which would be caused by performing a uniform atmospheric correction over the study area were assessed through simulations using Py6S. These showed that ignoring the spatial variation in AOT when performing atmospheric corrections could cause errors in reflectance and NDVI of up 1.3 percentage points and 5% respectively, but that ignoring spatial variation in PWC caused maximum errors of 0.7 percentage points and 0.6% respectively (an acceptable error primarily due to the strategic location of multispectral sensor bands away from water absorption features).

In conclusion, the results from this study show that there is significant variation in AOT and PWC across southern England during clear days. The variation in PWC is not significant in terms of the errors resulting from a uniform atmospheric correction, but ignoring the variation in AOT by performing a uniform atmospheric correction could cause significant errors (reflectance errors of over twenty times the NE$\Delta\rho$, and NDVI changes > 0.03). (reflectance errors of over twenty times the NE$\Delta\rho$, and NDVI changes > 0.03). The widespread availability of scene-based atmospheric correction procedures in modern image processing systems invites users to disregard spatial variability in the atmosphere and risks introducing significant errors into key derived products.

Chapter 3

Literature Review

This chapter examines the relevant literature regarding AOT measurement, starting with an introduction to AOT and atmospheric radiative transfer and continuing with discussion of ground measurement of AOT. The majority of the chapter discusses methods used for estimating AOT from satellite measurements, and methods used for haze assessment of satellite imagery. Finally, a range of potential applications of AOT data - particularly high-resolution data - are discussed.

3.1 Atmospheric aerosols and AOT

3.1.1 What are aerosols?

Aerosols (from the Greek *aero* (air) *sol*ution) are fine solid or liquid particles suspended in a gas. In terms of atmospheric aerosols, the particles are suspended within the gases which comprise the atmosphere (primarily nitrogen and oxygen) and are affected by atmospheric dynamical processes, including diffusion, advection and deposition.

The size of atmospheric aerosol particles can range over four orders of magnitude, from a few nanometres to around 100 micrometres (Petty, 2006). Compared to many other atmospheric constituents, aerosols have a short residence time in the atmosphere, and therefore their concentrations vary significantly both spatially and temporally. Aerosols are formed by natural or anthropogenic processes from a range of geographically-distributed sources, undergo movement within the atmosphere, and are then deposited either through dry deposition (directly onto the Earth's surface) or wet deposition (where aerosol particles are incorporated into clouds as cloud condensation nuclei or washed out of the atmosphere by precipitation).

Aerosols can be classified by their composition, size and source, and atmospheric concentrations may range up to 10^7 or 10^8 particles per cubic centimetre (Kondratyev *et al.*, 2005). Any volume of atmosphere is likely to contain both fine and coarse particles (those with a diameter less than and greater than 2.5 µm respectively). Generally these two categories of particle sizes behave differently and thus undergo different dynamical

processes.

According to Kondratyev *et al.* (2005), the major natural sources of aerosols are :

- Soil and rock debris, producing dust aerosols (mineral dust)
- Volcanic eruptions, producing ash and dust aerosols
- Sea spray, producing salt aerosols
- Biomass burning, producing carbonaceous soot aerosols
- Reactions between naturally-occurring atmospheric gases, producing sulphate aerosols

Human activity has contributed significant to the aerosol load, and the major sources of anthropogenic aerosols are:

- Fuel combustion
- Industrial processes
- Domestic (non-industrial) sources
- Transportation
- Human-induced natural sources (for example, anthropogenically-driven erosion or intentional biomass burning)

The majority of these anthropogenic sources contribute carbonaceous, soot and sulphur aerosols to the atmosphere.

3.1.2 What effects do aerosols cause?

The presence of aerosols in the atmosphere causes several effects (Kondratyev *et al.*, 2005):

- **Light scattering and absorption:** Aerosols in the atmosphere cause increased scattering or absorption of light. This can have a significant impact on the planetary radiation balance by causing light to be scattered away from the Earth before reaching the ground, or by absorbing light and thus heating the atmosphere. This then has an effect on the Earth's climate, and uncertainties in modelling of aerosol effects on climate have driven much of the aerosol research in recent decades (Le Treut, 2012; IPCC, 2007). The scattering effects of aerosols also cause problems for satellite imaging, as the light received at the satellite will have been affected by scattering and absorption in the atmosphere.
- **Increased cloudiness:** Aerosol particles present in the atmosphere can provide cloud condensation nuclei (CCN) for the water vapour droplets in clouds to condense around, thus promoting cloud formation (Twomey, 1974; Albrecht, 1989) and affecting the Earth's climate.
- **Human health impacts:** Depending on their size, aerosols in the atmosphere may be able to penetrate deep inside the lungs of a human when they breathe and cause a wide range of health impacts. The most dangerous particles are those smaller than 2.5 µm, as these can reach the alveoli where oxygen exchange between the lungs and the blood occurs, and they can cause a range of problems including lung

cancer, asthma and heart disease (Seaton *et al.*, 1995). Particulate matter pollution is a serious health problem, particularly in rapidly industrialising countries, causing over 3.2 million premature deaths worldwide in 2010 (Lim *et al.*, 2012).

The magnitudes of each of these effects vary depending on the type, size and concentration of the aerosols.

3.1.3 How can we measure atmospheric aerosols?

3.1.3.1 Direct measurement

The most obvious way to obtain information about atmospheric aerosols is to collect a sample of aerosol particles and analyse them. This can be done at ground level or from aircraft using a range of different types of detectors which can measure many properties including total number concentration, cloud condensation nuclei concentration, size distribution, particle density, particle water content and aerosol chemical composition (McMurry, 2000).

This method can provide very accurate data, but is time-consuming and expensive, particularly if we require data over large spatial or temporal scales. It is particularly difficult to use this method to obtain data on atmospheric aerosol contents over a vertical path through the whole atmosphere, as this requires airborne measurements at many heights, ideally simultaneously.

Regular measurements of this type tend to be restricted to networks for measurement of aerosol properties near ground level, for use in studies about particulate matter pollution and human health. For example, the London Air Quality Network records a variety of aerosol data automatically at around fifty sites over the whole Greater London area (Fuller and Green, 2006), and similar networks exist in many other cities. Short periods of detailed physical measurements such as these are sometimes collected by field campaigns such as the Tropospheric Aerosol Radiative Forcing Observational Experiment (TARFOX; Smirnov *et al.*, 2000), the Southern African Regional Science Initiative (SAFARI-2000; Swap *et al.*, 2003; Haywood *et al.*, 2003) and the Fennec Saharan dust monitoring campaign (Ryder *et al.*, 2013).

3.1.3.2 Indirect optical measurements

An alternative way to find out about atmospheric aerosols is to measure some of the effects discussed earlier. Optical measurements of aerosols use the known effects that aerosols have on the passage of light through the atmosphere to infer information about the aerosols present. This can be done either by taking measurements at ground level of the sunlight transmitted through the atmosphere, or by taking measurements from satellites of light that has been transmitted through the atmosphere twice: once on the path from the sun to the ground, and one after it has been reflected from the ground towards the satellite.

The major effect of the atmosphere on light is to reduce its intensity, to a different extent

for different wavelengths of light, and this reduction can be used to calculate the Aerosol Optical Thickness (AOT; also known as Aerosol Optical Depth, or AOD), which is a measurement of the attenuation of light due to aerosol scattering and absorption. This can then be used to infer a number of the aerosol properties discussed above.

3.2 Atmospheric radiative transfer

3.2.1 Overview

As light passes through the atmosphere it interacts with atmospheric constituents. These interactions principally take two forms:

- **Absorption:** This is a reduction of the intensity of the light, normally caused by a collision with a molecule. The magnitude of the reduction is usually dependent on the wavelength, and the principal atmospheric absorbers are the major atmospheric constituent gases listed in Table 3.1.
- **Scattering:** This is a change in the direction of the light, which is caused by a collision with a molecule (such as those of the gases that make up the atmosphere) or aerosol particle.

3.2.2 Absorption

Absorption occurs when photons collide with molecules in the atmosphere and transfer some of their energy to the molecule, which absorbs the energy and re-radiates it. This re-radiated energy can generally be ignored as it is very weak and located in the microwave or far-infrared regions of the spectrum.

The amount of absorption depends on the type of molecule (for example, water vapour is more absorbing than nitrogen). Individual collision events are rarely studied when investigating radiative transfer; instead the collective effect of a large number of molecules is calculated by relating the absorption to the concentration of absorbing molecules in a section of the atmosphere.

3.2.3 Scattering

Scattering of light in the atmosphere can be split into two types based on the mathematics which best model the effects of the scattering: Rayleigh scattering and Mie scattering.

Rayleigh scattering occurs when the diameter of the scattering particle is significantly

TABLE 3.1: Composition of a dry atmosphere (plus water vapour) by percentage volume

Nitrogen	78%
Oxygen	21%
Argon	0.9%
Carbon Dioxide	0.04%
Other gases	0.06%
Water vapour	∼40%

smaller than the wavelength of the light. This relationship is calculated based on the scattering size, x,

$$x = \frac{2\pi r}{\lambda} \tag{3.1}$$

where r is the radius of the particle, λ is the wavelength of the light, and Rayleigh scattering occurs where $x \ll 1$. In the atmosphere, this generally means that Rayleigh scattering occurs when photons collide with gas molecules. The magnitude of Rayleigh scattering for a given wavelength is described by the proportionality below

$$S_R \propto \frac{1}{\lambda^4} \tag{3.2}$$

Where S_R is the magnitude of Rayleigh scattering and λ is the wavelength of light. Thus, Rayleigh scattering occurs significantly more at small wavelengths (blue light) than large wavelengths (red and infrared light).

Mie scattering occurs when the diameter of the scattering particle is of the same order of magnitude as the wavelength of the light. In the atmosphere, this means that all scattering by aerosols is Mie scattering. The mathematics behind Mie scattering is a full analytical solution of Maxwell's equations (albeit with many assumptions), which is too complex to describe in detail here (for more details see Petty, 2006).

3.2.4 Absorption and Scattering combined

Some particles in the atmosphere both absorb and scatter. For example, aerosols are known mainly for their scattering effect, but certain types of aerosols, such as soot particles, can be effective absorbers. The balance between scattering and absorption is represented by the *single scattering albedo* which is defined as

$$SSA = \frac{\text{ext}_s}{\text{ext}_s + \text{ext}_a} \tag{3.3}$$

Where SSA is the single scattering albedo, ext_s is the extinction of light due to scattering and ext_a is the extinction of light due to absorption. The single scattering albedo varies between 0 and 1, with 1 signifying purely scattering, and 0 signifying purely absorption.

3.2.5 Atmospheric transmittance and optical depth

The wavelength-dependence of absorption and scattering processes means that light at some wavelengths is significantly affected by its passage through the atmosphere, whereas light at other wavelengths is barely affected at all. This can be quantified by the atmospheric transmittance, defined as

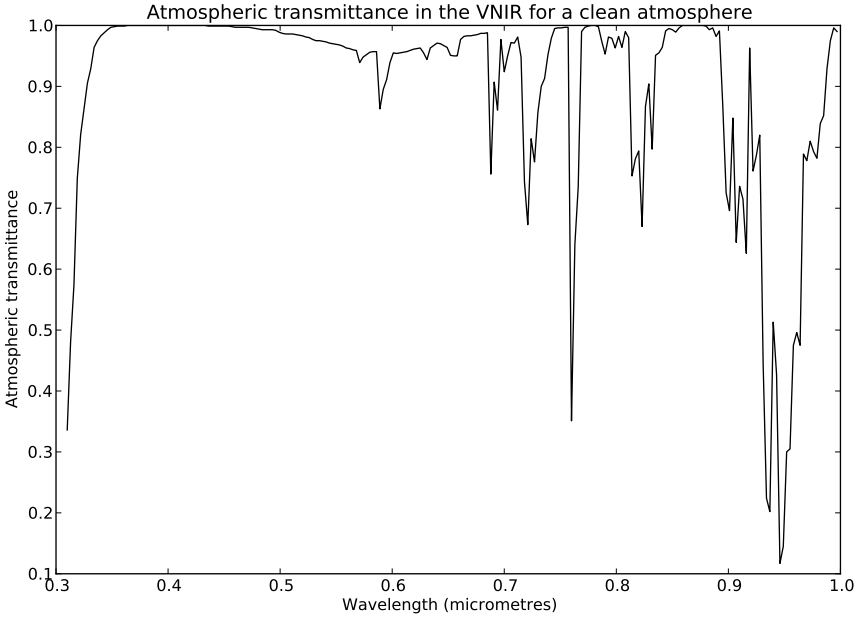


FIGURE 3.1: Transmittance through a clean atmosphere (no aerosols) for light in the visible and near infra-red spectrum. The higher the transmittance value, the more transparent the atmosphere. Data generated using the Py6S interface to the 6S Radiative Transfer Model (Wilson, 2012).

$$T_\lambda = \frac{I_\lambda}{I_{\lambda,0}} \tag{3.4}$$

where I_λ is the intensity of light at wavelength λ at the Earth's surface, and $I_{\lambda,0}$ is the extra-terrestrial (top-of-atmosphere) intensity of light at wavelength λ. Thus, the transmittance is simply the proportion of the top-of-atmosphere intensity which remains after passage through the atmosphere (see Figure 3.1 for a graph of transmittance through a clean atmosphere).

The Beer-Lambert law defines a logarithmic relationship between the transmittance and the optical thickness, τ, also known as the optical depth,

$$T_\lambda = e^{-\tau_\lambda} \tag{3.5}$$

Note that τ is dependent on the wavelength, λ, although these subscripts will be left out of the following equations for clarity. Measurements taken through the atmosphere at different angles will have different path lengths, and this must be corrected for by use of the optical air mass, m,

$$m = \frac{1}{\cos\theta}\frac{p}{p_0} \tag{3.6}$$

which calculates the length of the path at angle θ relative to the length of a vertical path for an atmospheric pressure p, where p_0 is the standard atmospheric pressure of 1000 hPa.

The transmittance can be decomposed into separate transmittances caused by different processes or different substances, for example, the intensity at the Earth's surface can be calculated as

$$I_\lambda = I_{\lambda,0}\, T_{\text{abs}}\, T_{\text{mol}}\, T_{\text{aero}} \tag{3.7}$$

where each of these transmittances is calculated from an optical depth,

$$I_\lambda = I_{\lambda,0} \exp(-m\tau_{\text{abs}} - m\tau_{\text{mol}} - m\tau_{\text{aero}}) \tag{3.8}$$

In these equations the total optical depth was decomposed into optical depths for absorption, molecular (Rayleigh) scattering and aerosol (Mie) scattering, but they could be decomposed further into ozone absorption, water absorption and so on. With the current level of decomposition, however, the equation above defines τ_{aero}, the Aerosol Optical Thickness (AOT) or Aerosol Optical Depth (AOD).

It should be remembered that each of these optical depths are wavelength dependent, and therefore the AOT will be different for each wavelength. Thus, whenever an AOT value is stated the wavelength it has been measured at must be stated as well. AOTs measured at 500 nm or 550 nm are often considered standard measurements. The Ångstrom exponent, α, describes the logarithmic dependency of the AOT on wavelength,

$$\frac{\tau_{\lambda_1}}{\tau_{\lambda_2}} = \left(\frac{\lambda_1}{\lambda_2}\right)^{-\alpha} \tag{3.9}$$

$$\alpha = -\frac{\log \frac{\tau_{\lambda_1}}{\tau_{\lambda_2}}}{\log \frac{\lambda_1}{\lambda_2}} \tag{3.10}$$

and has an inverse relationship with the average particle diameter. Thus, the Ångstrom exponent can be used for two separate purposes:

- To interpolate AOT values at wavelengths where they have not been directly measured, by calculating the Ångstrom exponent from the wavelengths that have been measured and using the formulae above for interpolation.
- To estimate the average particle size from measurements at various wavelengths, and therefore infer the particle size distribution

3.2.6 Satellite radiance

Light measured by a satellite at the top of the atmosphere is affected by atmospheric radiative transfer twice: once on its way from the sun to the surface, and again after the reflection from the surface towards the satellite. The radiance measured at the satellite sensor can be written as:

$$L_s(\lambda, \theta_s, \theta_v, \phi_s, \phi_v) = \rho(\lambda, \theta_s, \theta_v, \phi_s, \phi_v) L_{i,\lambda} + f(\omega, \Theta, \tau) L_{i,\lambda} + f(\rho(\lambda, \theta_s, \theta_v, \phi_s, \phi_v), \rho^{\text{adj}}(\lambda, \theta_s, \theta_v, \phi_s, \phi_v), L_{i,\lambda}) \quad (3.11)$$

where λ is the wavelength, θ_s and θ_v are the solar and viewing zenith angles, ϕ_s and ϕ_v are the solar and viewing azimuth angles, L_i is the incident radiation at the top of the atmosphere, ρ is the reflectance of the ground surface within the pixel, ρ^{adj} is the reflectance of the ground surface adjacent to the pixel and ω, Θ and τ are the aerosol single scattering albedo, phase function and optical depth respectively.

Taking each term separately and removing the dependencies on wavelength and angles for clarity, this equation can be written as

$$L_s = L_g + L_p + L_a \quad (3.12)$$

where L_g is the ground radiance, L_p is the path radiance and L_a is the radiance caused by the adjacency effect, as shown in Figure 3.2.

3.3 Ground measurement of AOT

3.3.1 Sun photometry

Calculating AOT from measurements of the direct solar irradiance at ground level involves a straightforward re-arrangement of the definition of AOT given in §3.2.5, based upon the Beer-Lambert law (Shaw, 1983; Schmid *et al.*, 1997),

$$\tau_{\text{aero}} = \frac{\ln I_{\lambda,0} - \ln I_\lambda}{m} - \tau_{\text{abs}} - \tau_{\text{mol}} \quad (3.13)$$

where $I_{\lambda,0}$ and I_λ are the radiances at the top-of-atmosphere and surface respectively, both at wavelength λ, m is the path length and τ_{abs} and τ_{mol} are the optical depths due to absorption and molecular scattering respectively.

m can be calculated as in equation 3.6, using the solar zenith angle (easily calculated from time, date and location) as θ, and $I_{\lambda,0}$ can be taken from any one of a number of top-of-atmosphere solar irradiance tables (for example Wehrli, 1985), ensuring that correction is performed for the varying Earth-Sun distance. τ_{mol} is easily calculated based on the wavelength under consideration and the atmospheric pressure (Kneizys *et al.*, 1981). τ_{abs} is composed of many components: absorption due to water vapour, oxygen,

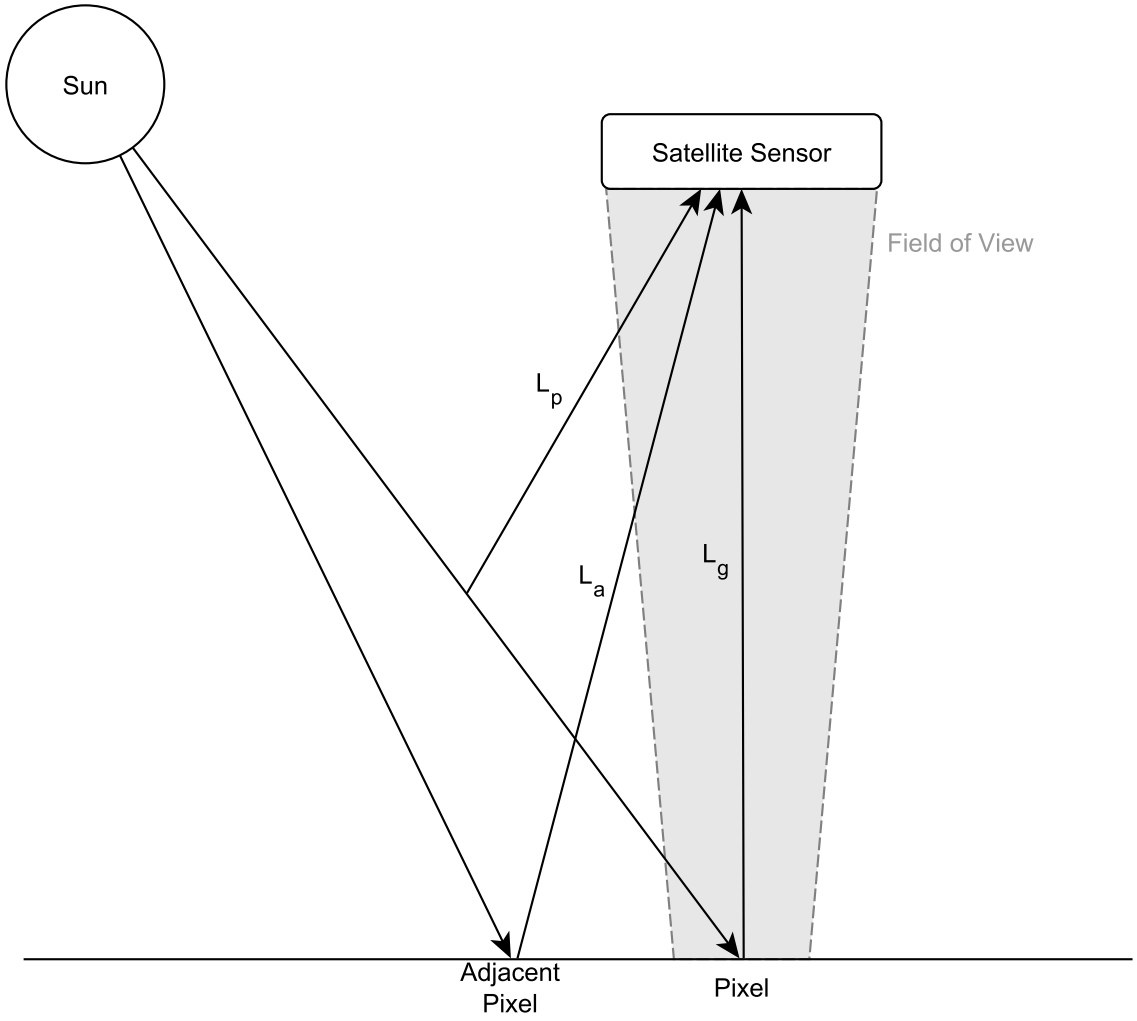


FIGURE 3.2: Diagram showing the components of the measured at-sensor radiance for a satellite sensor at the top of the atmosphere.

ozone and others, but by using wavelengths away from major absorption bands (for example, 0.94 µm for water and 0.76 µm for oxygen) we can ignore all of these except for ozone, as ozone absorbs radiation across a relatively wide wavelength range. Taking only Ozone into account, τ_{abs} can be calculated from standard ozone transmittance formulae such as those in Leckner (1978). Thus, calculating AOT from a measurement of I_λ is simply a matter of substituting these values into the equation above.

3.3.1.1 Instruments

Sun photometers are instruments that can measure I_λ accurately and for a very narrow wavelength range (ideally as close to monochromatic light as possible). To ensure that only direct solar radiation is measured the instrument must be pointed directly at the solar disc: this can be done manually or by an automatic sun-pointing device.

Instruments have developed significantly since the invention of the 'bolometer' (an early type of sun photometer) by Langley in the late 19th century (Langley, 1884). Most modern instruments use silicon photodiodes to detect the intensity of the light,

interference filters to restrict the light entering the instrument to a narrow wavelength range, and baffling and collimators to restrict the field-of-view of the instrument and ensure that scattered light is not measured. Due to problems with the stability of the silicon photodiodes, these instruments require regular calibration to ensure that the derived AOT measurements are accurate. Calibrations usually use the so-called Langley-method, which uses data collected at a high-altitude field site (such as Mauna Loa, Hawaii) over a range of different solar angles to compute the calibration constants (Schmid and Wehrli, 1995).

The two most commonly-used sun photometers are the Microtops II and the Cimel CE-318. The Microtops II (Morys *et al.*, 2001) is a hand-held instrument, shown in Figure 3.3(a), which measures the radiance in five narrow channels and calculates the AOT as well as water vapour and ozone absorption. The instrument has been designed to be portable and easy to use, while also collecting accurate measurements (to within approximately 1% accuracy under standard conditions). When taking a measurement the user must point the device directly at the solar disc, using a built-in sighting device: this allows use anywhere without necessarily using a tripod, but can lead to issues with the reliability of measurements, as the user must hold the device very still. The Microtops II instrument costs approximately £10,000.

The Cimel CE-318 instrument, shown in Figure 3.3(b), is designed for automated operation. It includes an automatic sun-tracking mount, which enables measurements to be taken without human intervention. Again, the instrument uses silicon photodiodes and interference filters and produces high-accuracy results. The sun-tracking mount is designed to very high tolerances, and therefore the pointing accuracy is far higher than the manual pointing by humans that is required for the Microtops II. According to Eck *et al.* (1999) the uncertainty of AOT data from the CE-318 is ± 0.02. The CE-318 does not have a sensor at 500 nm or 550 nm, so the AOT_{550} value cannot be directly computed, instead it is interpolated using the Ångstrom exponent calculated from measurements in two other wavelengths. The Cimel instrument is significantly more expensive than the Microtops, costing approximately £35,000.

3.3.1.2 AERONET

The Aerosol Robotic Network (AERONET; Holben *et al.*, 1998) is a global network of Cimel CE-318 sun photometers which automatically acquire measurements of atmospheric parameters including AOT. There are over 300 stations worldwide from which data are regularly acquired and made publicly available via the AERONET GSFC website. Measurements of radiance from the solar disc in multiple wavelengths are taken every 15 minutes during daylight hours in good weather, and these are converted to AOT as described above. The brightness of the area around the sun (the solar aureole) and its gradient to about 6 ° from the sun angle are dependent on the size distribution of the aerosol particles. Thus, measurements taken in these areas can be used to obtain information on the aerosol size distribution, and therefore infer the aerosol composition.

(a) Microtops II (b) Cimel CE-318

FIGURE 3.3: Photographs of the two most common commercially-available sun photometers. (Images courtesy of the NERC Field Spectroscopy Facility)

AERONET data are inverted against modelled predictions to produce this data using the Dubovik and King (2000) method.

All of these data are provided at three levels, with data volume decreasing significantly at higher levels as more data points are excluded:

- **Level 1:** Raw data from the instruments
- **Level 1.5:** As above, with automatic cloud-screening applied
- **Level 2:** As above, with a manual quality-control check applied

AERONET have also run a series of field campaigns, known as Distributed Regional Aerosol Gridded Observation Networks (DRAGON) since 2011. These involve the temporary creation of a relatively dense network of AERONET sites in a particular area (for example, the Houston Metropolitan Region or South Korea) for a period of a few months, providing dense, high-quality AOT data.

3.3.1.3 LED-based sun photometers

Within the last decade, there have been developments in creating cheap sun photometers based upon Light Emitting Diodes (LEDs). The development of these instruments were motivated by the high cost of competing devices (such as the Microtops and Cimel instruments described above) and the subsequent issues with setting up widespread AOT measurement networks. Under the auspices of the Global Learning and Observations to Benefit the Environment (GLOBE) programme (Finarelli, 1998), Brooks and Mims (2001) developed a sun photometer which uses LEDs as sensors. Although LEDs are usually used to produce light by passing a current through them, they can also work in reverse. That

is, shining light onto them will produce a current, and this means that they can be used as a simple photodiode (Acharya, 2005).

The light produced by LEDs has a very narrow wavelength range, and this also holds for the range of wavelengths that they are sensitive to when used as a photodiode. This is the major benefit of LEDs over standard photodiodes, as this eliminates the need for interference filters in front of the sensors. These tend to degrade over time, and the photodiodes themselves can also drift over time, whereas LEDs have been found to be very stable (Mims, 2003), and thus LED-based instruments are likely to produce more reliable results with fewer calibration issues. These instruments can be made significantly more cheaply than standard interference-filter-based instruments, and this has raised possibilities for the development of more widespread networks of sun photometers.

SkySci is a collaborative project between the academic units of Geography & Environment and Electronics & Computer Science at the University of Southampton, based upon an idea developed by the author. The project aims to develop a handheld sun photometer which would operate in a similar way to the Microtops (see §3.3.1.1). However, through the use of LEDs as sensors, the SkySci instruments can be produced for less than £50 each - in comparison to a cost of around £10,000 for a Microtops instrument. The instruments are designed to be linked to a smartphone, allowing measurements to be automatically sent to a central server in Southampton and combined into a detailed AOT dataset - a significant advantage over the GLOBE programme instruments. A grant obtained from the EPSRC Digital Economy 'IT as a Utility Network+' has been used to develop a prototype instrument (Figure 3.4), and plans are currently in place to acquire funding to produce a large number of instruments and collect measurements through a Citizen Science project.

3.3.2 Visibility-AOT relationship

Instead of measuring it directly, AOT can be estimated from horizontal visibility measurements using Koschmieder's equation (Koschmieder, 1924; Horvath, 1981)

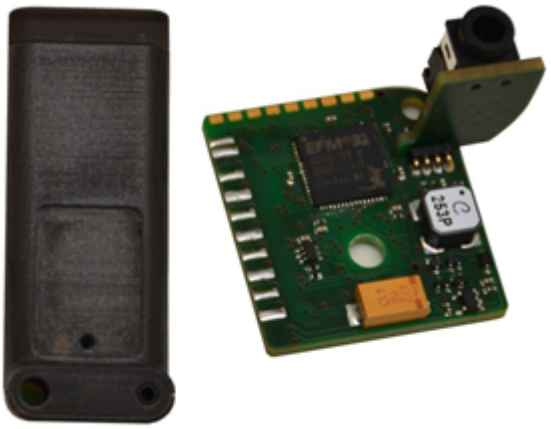
$$V = \frac{3.912}{\tau} \tag{3.14}$$

where V is the visibility and τ is the AOT.

This is derived from a formula for the visibility of objects in the atmosphere

$$V = \left(\frac{1}{\tau}\right) \ln \left|\frac{C_0}{\epsilon}\right| \tag{3.15}$$

where ϵ is the liminal contrast and C_0 is the inherent contrast of the object. Koschmieder suggested that under most viewing conditions the liminal contrast is 0.02, and proposed a



(a) Case and circuit board

(b) Instrument on test in Tenerife

(c) Instrument connected to phone

FIGURE 3.4: Prototype LED instruments

theoretical 'black object' whose inherent contrast is -1. Substituting these into this equation gives the final equation above.

This equation has been used extensively for estimating AOT, primarily for correction of satellite images (Retalis *et al.*, 2010), and could see wider operational use with visibility data from meteorological station networks. However, there are several issues with the use of this formula.

The equation was originally intended to relate horizontal visibility and the horizontal extinction coefficient, but is now widely used for calculating vertical extinction coefficients from horizontal visibility. This mixing of horizontal and vertical measurements assumes that

- The horizontal and vertical extinction coefficients are equal
- The relationship between the horizontal visibility and horizontal extinction coefficient also holds for horizontal visibility and the vertical extinction coefficient

Chan (2009) found that these assumptions did not hold for data collected over Hong Kong International Airport during winter 2008–9, where they observed several occasions when the vertical extinction coefficient increased significantly, but the horizontal extinction coefficient changed very little. Measurements with a ground-based Doppler LiDAR showed that this was caused by high aerosol concentrations at altitudes of 2–4km. These increased concentrations were not associated with any atmospheric phenomena that would be noticeable in a horizontal view such as clouds or dust storms.

Separately from the issues raised above, there are significant assumptions in the conversion formulae, discussed in more depth by Middleton (1952), and in the meteorological community Koschmieder's formula has been widely challenged (Horvath, 1971, 1981), both in terms of his choice of constants and the applicability of his formula to non-ideal conditions, but no widely-accepted replacement has been proposed.

3.4 Satellite measurement of AOT

Measurement of AOT from satellite sensors is very appealing, as it allows data acquisition across large areas without requiring large numbers of expensive ground measurement devices. However, measuring AOT accurately from satellites is challenging.

The formula for the radiance received at a satellite was given in equation 3.11, and the simplified version in 3.12. This simple formula defines the at-sensor radiance as the sum of the ground radiance, path radiance and radiance due to the adjacency effect:

$$L_s = L_g + L_p + L_a \tag{3.16}$$

The AOT can be estimated from the path radiance, and the adjacency effect is often ignored. Thus, to obtain a measurement of the path radiance we need to know the

ground-leaving radiance, that is, we need to solve the following equation,

$$L_p = L_s - L_g \tag{3.17}$$

where L_p is the path radiance, L_s is the at-sensor radiance, and L_g is the ground-leaving radiance.

All methods for retrieving AOT from satellites have to solve this problem, and they can be categorised by the methods through which they solve it. These categories are examined in more detail in §3.4.2

3.4.1 Estimation of AOT from path radiance

The radiance separation step will result in an estimation of the path radiance, which can be used to estimate AOT. The path radiance is the radiance reaching the satellite sensor which has been caused by scattering into the light path to the sensor, rather than by reflectance from the ground surface. A certain amount of path radiance is caused by the Rayleigh scattering processes which are present even in a perfectly clear atmosphere, but the majority is caused by the presence of aerosols. Thus,

$$L_p = L_{pr} + L_{pa} \tag{3.18}$$

where L_p is the path radiance, and L_{pr} and L_{pa} are the path radiances caused by Rayleigh scattering (which is easily calculated from the air mass and surface pressure; Kaufman, 1993) and aerosol scattering respectively, and

$$L_{pa} = L_p - L_{pr} \tag{3.19}$$

The path radiance due to aerosols is directly related to the AOT, through the formulae of radiative transfer. As these formulae are very complex, it is easiest to use a Radiative Transfer Model to produce a lookup table (LUT) giving the AOT for a range of path radiance measurements. Such a table can be created using a model such as 6S (Vermote *et al.*, 1997). Thus, once the path radiance has been calculated, the AOT can be retrieved by selecting the appropriate cell of this lookup table, interpolating if necessary.

So far, we have not taken into account differing aerosol types. The RTM will have had to have been parameterised using a specific aerosol model to produce the lookup table, and if this aerosol model is not a good representation of the real-world conditions for the measurement location then the resulting AOT will be erroneous. The simplest way to solve this problem is to have a defined aerosol model for each measurement location (each pixel, in the case of a satellite image) and use a LUT specific to that aerosol model to extract the AOT. Many algorithms use this approach, although some attempt to extract

the aerosol model and the AOT from the satellite data itself. These methods will be discussed more in §3.4.3.

3.4.2 Radiance separation methods

A range of methods for separating the path radiance and ground radiance components of the at-sensor radiance are described below, with examples of AOT retrieval algorithms which use these methods. Many of these methods are concerned with obtaining the ground reflectance, from which the ground-leaving radiance can be calculated as

$$L_{g\lambda} = \rho_\lambda L_{i_\lambda} \tag{3.20}$$

where ρ_λ and L_{i_λ} are the ground reflectance and the irradiance, both at a specific wavelength, λ. Once $L_{g\lambda}$ has been calculated, it can be substituted into equation 3.17, and the path radiance can be calculated.

3.4.2.1 Assume a known ground reflectance

The simplest way to separate the radiance into that caused by ground reflectance and by scattering into the light path is to assume a known ground reflectance, for example, a reflectance of zero. This was the approach that was taken for the first satellite measurements of AOT over the oceans by Griggs (1975) using ERTS-1 data. This method was also used for the early algorithms for retrieving AOT from the AVHRR sensor (Nagaraja Rao *et al.*, 1989), also over water. This is because water is a good absorber of infra-red radiation, so a legitimate assumption can be made that the reflectance in the NIR band is zero (or at least almost zero), and therefore the radiance measured at the satellite is the path radiance, making the solution very simple,

$$L_p = L_s - 0 \tag{3.21}$$

This also explains why the accuracy of AOT retrieval using any method is generally higher over water, as the reflectance is generally low, so any errors in estimating the reflectance (for example, over areas of the ocean with sun glint) will not have a large impact. These early measurements provided a good idea of the potential of satellite measurements for estimating AOT, and provided a good coverage of the globe (as water covers 71% of the Earth's surface). However, the lack of coverage over land was a major problem, particularly for certain applications such as assessing aerosol effects on public health, monitoring dust emission from deserts and the effects of volcanic dust, and they precipitated a major programme of work to develop better AOT retrieval techniques from satellites (Penner *et al.*, 1994).

Water isn't the only surface for which a certain reflectance can be assumed, there are certain land surfaces for which the reflectance can be assumed to be stable. Much

research has shown that Dense Dark Vegetation (DDV) tends to have reflectances of 1–2% in the red band and 2–3% in the green band (Kaufman and Sendra, 1988), and thus the radiance over DDV can be calculated, giving a solution of,

$$L_p = L_s - L_{DDV} \tag{3.22}$$

Of course, this limits the algorithm to retrievals over DDV, and thus a continuous map of AOT cannot be acquired over the whole globe - particularly for areas without dense dark vegetation (such as deserts, urban areas and ice caps). Thus, the AOT retrieval algorithm must assess whether each pixel contains DDV or not before deciding whether to proceed with the retrieval. This assessment is usually performed using thresholds for reflectance and a vegetation index: for example, an NDVI of ≥ 0.75 and a low reflectance $0.01 \leq \rho_{\mathrm{SWIR}} \leq 0.05$ as used in Lyapustin *et al.* (2004). Alternative methods which ensure that a reasonable number of pixels are selected involves choosing the pixels with the top x% of the NDVI values in the image, and then selecting the darkest y% of these (Kaufman and Sendra, 1988).

Early versions of the MERIS algorithm used the DDV approach with an extension to deal with angular effects on DDV reflectance, which are particularly important for wide-swath sensors such as MERIS (Santer *et al.*, 1999). AOT data from the POLDER sensor was used to select MERIS images with negligible aerosol content, and a RTM combined with a canopy model were used to create a detailed model of the BRDF of DDV surfaces in various world biomes. The reflectances from these BRDF models were used to calculate the ground radiance, rather than using the standard 1–2% reflectance values. These models were also used to generate a specific DDV selection threshold using Atmospherically Resistant Vegetation Index (ARVI; Kaufman and Tanré, 1996) for all biomes and viewing angles.

An alternative approach to create a situation where an assumption of a known reflectance is valid is to use a different set of wavelengths. In the ultraviolet (UV) spectrum (approximately 10–400 nm) the Earth has a very low ground-leaving radiance, which can therefore be assumed to be zero. Furthermore, the backscattered ultraviolet radiation (BUV) measured in the atmosphere is significantly affected by the aerosol content, principally by aerosol absorption. This effect was first noticed due to its confounding effect on the retrieval of ozone amounts from early instruments such as the Total Ozone Mapping Spectrometer (TOMS; Dave, 1978). This effect was then used in the development of algorithms for the measurement of a uncalibrated 'aerosol index' from TOMS by Hsu *et al.* (1996). The reason for the delay of almost twenty years between the original suggestion of these issues and their potential by Dave (1978) and an early algorithm implementation was due to the computational challenges in producing accurate radiative transfer models for the UV spectral region (Torres *et al.*, 1998). This is particularly difficult as multiple scattering effects are very significant in these wavelengths, and thus models are significantly more computationally intensive.

Torres *et al.* (1998) developed two methods to retrieve physically-meaningful aerosol properties, including AOT, from TOMS measurements. One approach was based on a standard inversion of radiances at 340 nm and 380 nm, which produced measurements of AOT and the single-scattering albedo, and was sensitive to all aerosol types. The other was based on a simulation of the expected radiances in a clean atmosphere (that is, an atmosphere with no aerosols, and purely Rayleigh scattering and gaseous absorption taking place) and an inversion of the differences between this and the measured radiances. Similar approaches (described in detail by Torres *et al.*, 2007) are still used for data from the Ozone Monitoring Instrument (OMI; the successor to TOMS).

3.4.2.2 Estimate the ground reflectance from another band

The reflectance of the majority of land surfaces cannot be assumed, so if AOT retrievals are required over surfaces other than water and DDV then they must be calculated using a different method. Some sensors have a Short-wave Infrared (SWIR) band at around 2.12 µm, and the atmosphere is almost entirely transparent to aerosols in this band due to the wavelength dependence of aerosol scattering. Thus, the radiance recorded in the SWIR can be assumed to be the same as the ground-leaving radiance, and is thus a true picture of the actual situation in the ground, uncontaminated by aerosols. If an empirical relationship between the SWIR radiance and radiances in other bands can be determined then radiances in other bands can be determined, and the path radiance in those bands can be calculated. Thus,

$$L_p = L_s - f(L_{SWIR}, \lambda) \tag{3.23}$$

where L_{SWIR} is the radiance in the SWIR band, and f uses empirical relationships to determine the radiance in the band for which AOT calculation is being performed. This is currently the most common approach used by operational AOT retrieval algorithms. The details of the empirical relationships vary between sensors, and are discussed below for some important products.

Early versions of the MODIS AOT algorithm (Kaufman *et al.*, 1998) up to Collection 004 assumed that aerosols were entirely transparent in the 2.12 µm wavelength, and then used two 'generally applicable' ratios (Kaufman *et al.*, 1997)

$$\rho_{0.47} = \frac{\rho_{2.12}}{4} \tag{3.24}$$

$$\rho_{0.66} = \frac{\rho_{2.12}}{2} \tag{3.25}$$

to derive the reflectance in the other two bands (blue and red) from this reflectance.

However, results showed that these ratios performed poorly, as surface reflectance values

were often over-estimated, leading to errors in retrieved AOTs. For the updated Collection 005 version of the algorithm (Remer *et al.*, 2006), surface reflectance in all three visible bands was estimated from the 2.12 µm band as a function of viewing angle and 'greenness' (calculated using a short-wave-infrared-based version of the NDVI). These give more accurate results for surface reflectance values, and therefore more accurate AOT retrievals (Levy *et al.*, 2010). However, there are still issues with the estimation methods, as Jethva *et al.* (2009) found significant differences between estimated and measured surface reflectances. The original MODIS algorithm produced results with a 10km spatial resolution, but Collection 006 has recently been released with an experimental high-resolution AOT retrieval method producing 3km resolution data (Remer *et al.*, 2013; Munchak *et al.*, 2013). This uses a very similar algorithm to the 10km data, but produces data with a higher uncertainty due to signal-to-noise ratio issues.

Once the surface reflectance has been estimated, the ground-leaving radiance can be calculated and subtracted from the satellite signal to leave the path radiance. This is then inverted using LUTs derived from the MIEV (Wiscombe, 1980) and RT3 (Evans and Stephens, 1991) radiative transfer models. For efficiency reasons, the table values were computed for AOTs of 0.0, 0.25, 0.5, 1.0, 2.0, 3.0 and 5.0, nine solar zenith angles, sixteen sensor zenith angles and sixteen relative azimuth angles, with interpolation performed between these values as necessary.

A slightly different approach was taken by later versions of the MERIS AOT retrieval algorithm, which used a method based upon using ARVI values to estimate ground reflectance. This is based on work by Santer *et al.* (2007) who found a strong linear relationship between ARVI value and the reflectance in the red band which held for DDV and many other surfaces (for all ARVI values ≥ -0.5). This allows surface reflectance to be estimated for a wide range of surfaces, thus extending aerosol measurements over areas with no DDV.

3.4.2.3 Use multi-angular measurements

The fundamental problem with separating the at-sensor radiance into path radiance and ground-leaving radiance components is that the problem is ill-posed: there are two unknowns but only one equation. Thus, all previous approaches have used estimates for one term to make the problem well-posed.

However, instead of reducing the number of unknowns, the number of equations can be increased. This can be done by taking multiple observations of the ground surface under slightly different conditions. In this case, multiple observations at different angles are used. Each of the angular measurements will consist of the path radiance plus the ground-leaving radiance, and each will also have a different air mass (the length of the path through the atmosphere that the light takes). If we assume that the ground surface is Lambertian (that is, it has uniform reflectance in all directions) then the differences in path radiances at different angles will caused entirely by the air mass, and the

measurements can then be combined into a system of simultaneous equations which will be well-posed, and are therefore solvable. For example,

$$m_1 L_p = L_s^1 - L_g \tag{3.26}$$
$$m_2 L_p = L_s^2 - L_g \tag{3.27}$$
$$m_3 L_p = L_s^3 - L_g \tag{3.28}$$

where m_1 to m_3 are the air masses for each measurement and L_s^n and L_g are the at-sensor and ground radiances respectively for angle n.

However, it is well-established that the majority of land surfaces are not Lambertian, so the bi-directional reflectance distribution function (BRDF) of the ground surface must be taken into account. Unless we know the shape of the BRDF *a priori*, which is unlikely as few surface BRDFs are stable, then there are now more unknowns, and the problem becomes ill-posed again. Flowerdew and Haigh (1996) solved this problem during the development of the ATSR-2 aerosol retrieval algorithm by assuming that the shape of the BRDF function is independent of wavelength (although, of course, the magnitude of the function varies significantly with wavelength). Thus the BRDF at wavelength λ_2 can be calculated simply as a function of the BRDF at another wavelength, λ_1

$$\text{BRDF}(\lambda_2) = k\ \text{BRDF}(\lambda_1) \tag{3.29}$$

where k is the ratio linking the magnitude of the BRDF function at wavelengths λ_1 and λ_2. This is known as the *k-approximation*, and has been shown to be reasonable for a variety of surface types (Flowerdew and Haigh, 1995). A very similar algorithm is used for retrievals from the AATSR sensor.

The ATSR-2 and AATSR sensors acquire measurements at two angles (forward-viewing and nadir), whereas the MISR sensor acquires measurements at nine angles (nadir, plus four forward and aft viewing angles). This provides significantly more data, thus making the problem well-posed, and allowing an accurate estimation of the aerosol optical thickness along with derivation of a suitable aerosol model. The Diner *et al.* (2001) algorithm consists of two separate methods for determining AOT, which are chosen depending whether DDV areas are available in the image. If they are, then retrieval is based upon the standard DDV method, with the multi-angle measurements used to derive the appropriate aerosol model to use. If DDV areas are not available then the algorithm proceeds similarly to the ATSR-2 algorithm described above. This allows the highest accuracy retrievals to be performed when DDV is available, with data of a moderate quality still available in other situations.

Multi-angle Implementation of Atmospheric Correction (MAIAC; Lyapustin *et al.*, 2011)

is an algorithm which simultaneously retrieves aerosol information and surface bidirectional reflectance from MODIS data, thus removing the need for the *k-approximation*. The aerosol retrieval part of the algorithm can operate over a range of surface types and has been used to retrieve AOT at 1km resolution (as opposed to the standard 10km resolution of the standard MODIS AOT product). The algorithm uses a time-series of MODIS retrievals, which cover a range of angles due to the changing MODIS viewing geometry, and operates in a iterative manner where the aerosol retrieval for a pixel in a particular MODIS scene depends on the BRDF calculated for that pixel from the previous MODIS image in the time series. The BRDFs are estimated as part of the aerosol retrieval procedure using a minimisation of the Ross-Thick-Li-Sparse BRDF model (Wanner *et al.*, 1995), with the multi-angular observations required to fit the BRDF model provided by the differing observation and illumination geometries between subsequent MODIS acquisitions, or estimated from the standard MODIS dark target algorithm for the first image in the time series. Once the BRDF has been retrieved for a particular MODIS scene, the AOT can be retrieved easily using the equations above, as due to the BRDF being known these are now well-posed. The cyclical nature of the algorithm means that the quality of the retrievals improve along the time series as the definition of the BRDF of each pixel improves, and the overall accuracy is very close to that of the MODIS 10km AOT product (Emili *et al.*, 2011).

3.4.2.4 Use measurements of polarisation

As well as wavelength, direction and intensity, light can also be described in terms of its polarisation. This refers to the direction of oscillation of the electric field which is part of the light wave, and is usually described by the four components of the Stokes vector (I, Q, U and V). The polarisation of light can be changed by interactions of the light, which include reflection by the ground surface and scattering by aerosols. Thus, using polarisation measurements, as well as measurements of wavelength, direction and intensity, can help retrieve aerosol contents. Generally, measurements of the polarisation of the measured light provide more measurements for the system of equations, thus making the problem better-posed. Furthermore, when a sensor acquires measurements of polarisation and total irradiance for each band, there are enough parameters to invert the measurements directly to AOT and aerosol model using a LUT. This removes the requirement to calculate path radiance and then convert this to AOT. Mishchenko and Travis (1997) developed a retrieval algorithm which is as straightforward as choosing the entry from the LUT that satisfies two equations,

$$\frac{I_{meas} - I_{calc}}{I_{calc}} \leq 0.04 \tag{3.30}$$

$$\frac{1}{2}(|Q_{meas} - Q_{calc}| + |U_{meas} - U_{calc}|) \leq 0.002 \tag{3.31}$$

where the *meas* and *calc* subscripts refer to the measured polarisation and the

polarisation from the lookup table respectively. This method has a theoretical accuracy in AOT retrieval of ±0.015.

3.4.2.5 Assume temporal stability of ground reflectances

This is another approach that provides more data, and therefore makes the inversion problem well-posed. In this case, a number of images are acquired over time, and the differences between the radiance values in them are used to estimated the AOT for each image, after assuming that the ground reflectance has been stable over all of the image acquisition times.

The earliest uses of this approach were the contrast-based methods developed by Tanré *et al.* (1988) and Holben *et al.* (1992). These were based upon the observation that as AOT increased, the visible contrast between pixels with different reflectances decreased, that is, for two close pixels i and j, which have a difference in reflectance

$$\Delta\rho_{ij} = \rho_i - \rho_j \tag{3.32}$$

the difference in at-sensor radiance

$$\Delta L_{ij} = L_i - L_j \tag{3.33}$$

is inversely-related to the atmospheric transmittance

$$\Delta L_{ij} \propto \frac{1}{T} \tag{3.34}$$

Thus, given a time-series of images, a knowledge of the real value of $\Delta\rho_{ij}$ and an assumption that this is stable over time, the change in contrast between the images can be used to estimate the atmospheric transmittance, and therefore AOT.

There are several limitations to this technique:

- It can only be performed over land surfaces where the reflectance can legitimately be assumed to be stable over the time period under question. True invariance in reflectance is almost impossible to find, but several techniques have been developed for finding pseudo-invariant sites for other procedures such as instrument calibration and empirical atmospheric correction (Schott *et al.*, 1988; Canty *et al.*, 2004; de Vries *et al.*, 2007), and these can be applied to select the appropriate pixels to use with this method.
- The true (ground) value of $\Delta\rho_{ij}$ must be determined accurately: this can either be done from a clear image (although finding an image with an AOT of zero is almost impossible) or from ground measurements.

- Due to issues with reflectance anisotropy, the images used in the time series must be collected under similar angular conditions or a robust BRDF modelling procedure must be included in the method.
- The contrast reduction due to AOT is about 40% less for low-resolution imagery (such as AVHRR or MODIS) when compared to high-resolution images (Holben *et al.*, 1992). This has made it hard to apply this technique for modern low-resolution sensors such as MODIS. Data can still be retrieved, but is only useful when large contrasts in AOT are expected.

Early application of these methods took place largely over arid regions, where the high reflectance made retrieval with the other methods available at the time (mainly based on DDV reflectance) impossible. More advanced work has taken place recently, for example Liu *et al.* (2002) measured the contrast between pixels in a SPOT image using a multi-directional structure function and produced a significant improvement in AOT accuracy.

A newer approach called Synergy of Terra and Aqua MODIS (SYNTAM; Tang *et al.*, 2005) also makes use of a ground reflectance stability assumption, but is not based upon contrast differences. Instead, the problem is approached in a similar way to the multi-angular and polarization retrievals discussed above, as the additional data allows the ground reflectance to be accurately estimated, and makes the problem well-posed. The algorithm takes advantage of the two MODIS sensors in orbit, on the Terra and Aqua satellites, whose overpass times are normally separated by less than three hours. Over this short time interval the ground reflectance can be assumed not to change (apart from angular effects due to changing illumination and viewing geometry) and the aerosol type and properties (such as size distribution) can also be assumed to be stable. Therefore, we now have two equations, and have fixed some of the unknowns to be the same in both equations, thus producing a well-posed system of equations.

$$L_p = L_s^1 - f(\text{BRDF}, L_g) \tag{3.35}$$

$$L_p = L_s^2 - f(\text{BRDF}, L_g) \tag{3.36}$$

$$\tag{3.37}$$

However, under real-world conditions, BRDF effects mean that the measured reflectances cannot be assumed to be equal, as both the solar illumination geometry and the viewing geometry will have changed between the observations. The SYNTAM algorithm uses the *k-approximation* (Flowerdew and Haigh, 1995) as used by the multi-angular observations from MISR and AATSR above to deal with this issue. As the equations are non-linear and very difficult to solve, the Newton iteration algorithm is used to search for the best result. A limited accuracy assessment showed a maximum error of ± 0.1, although this may not be representative of the algorithm performance over a broader area.

3.4.3 Aerosol model choice

Some of the methods described in §3.4.2, such as those based upon polarised and multi-angular observations, are able to retrieve more than just AOT: they can retrieve details about the aerosol type, its size distribution, and other parameters. However, the majority of methods only calculate the path radiance and thus further details on the aerosol properties are required to invert this path radiance and estimate AOT. The key aerosol properties are the size distribution, single scattering albedo and scattering phase function, and the choice of these properties is known as the aerosol model.

An aerosol model must be provided for each pixel, and these values are normally taken from an aerosol climatology. Several of these exist, and they generally provide information on aerosol properties based upon latitude, longitude and season. Originally many algorithms used the climatology from d'Almeida *et al.* (1991), which was derived from a wide range of ground and airborne measurements. However, more recent work has either used data from AERONET (Holben *et al.*, 2001) or from other satellite sensors (such as MODIS; Remer *et al.*, 2008). The accuracy of the climatology has been identified as one of the major causes of uncertainty in satellite-derived measurements of AOT (King *et al.*, 1999).

3.4.4 Cloud screening

Cloud screening is a very important pre-processing procedure which must take place before the estimation of AOT. However, it is often difficult to develop an appropriate cloud-screening algorithm, as both errors of commission and omission can cause serious problems. Some cloud-masking algorithms may classify strong aerosol emission episodes (such as a large amount of dust aerosol) as cloud - thus removing this episode from the resulting AOT dataset - but others may be too liberal, and allow cloud-contaminated pixels to remain, thus falsely inflating AOT values. This is a particular issue for sensors with large pixel sizes (such as POLDER or OMI), as there is a greater likelihood of sub-pixel cloud contamination, which is notoriously difficult to detect but which may still have a significant impact on the resulting AOT value.

Krijger *et al.* (2007) analysed the number of cloud-free pixels available at a number of resolutions and found a log-linear relationship between pixel area and percentage of pixels that were cloud free. Considering images over the whole globe, only 16% of 10km pixels are entirely cloud free, significantly limiting the volume of data available for AOT retrieval. However, extrapolating the data in Krijger *et al.* (2007) to estimate the proportion of 30m pixels which are cloud free (in this case, extrapolation seems reasonable as the data in the paper behaves almost perfectly linearly over five orders of magnitude of pixel size) produces an estimate of 56% - providing far more data for AOT retrieval.

AOT retrieval algorithms vary as to how they approach cloud screening: some use the standard cloud mask product for the relevant sensor, but some develop their own cloud masking algorithms (such as the MODIS AOT product; Martins *et al.*, 2002). Using the

standard mask has the advantage that these masks have normally been well validated, but they may not be good enough at picking up sub-pixel clouds for use with AOT retrievals.

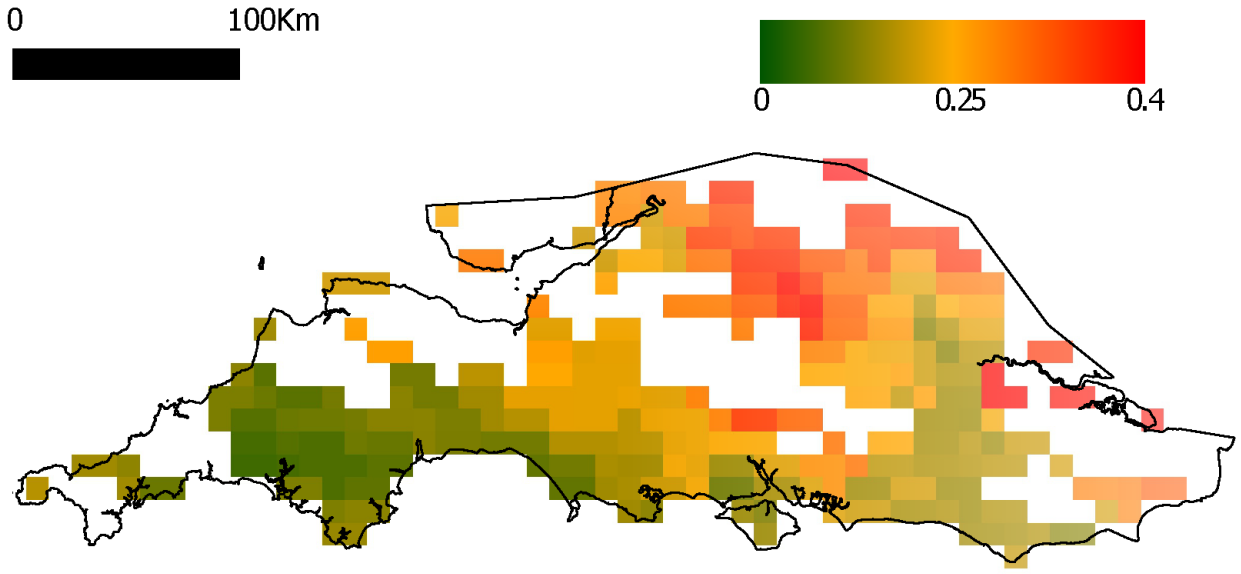
3.4.5 High resolution methods

The spatial resolution of the AOT products produced by the methods described above vary from around 20km to 20m. The majority of operational products are produced at a resolution of around 10km. Many operational products, such as the MODIS AOT product, are produced from data at a far higher resolution than 10km, but they then aggregated to improve the signal to noise ratio. For example, the MODIS signal to noise ratio was improved from 0.66 to 13 by aggregating 1km pixels to 10km pixels (Remer *et al.*, 2006) and the lower signal to noise ratio produced when the data were aggregated to 3km pixels is responsible for the increased error in the 3km product (Remer *et al.*, 2013).

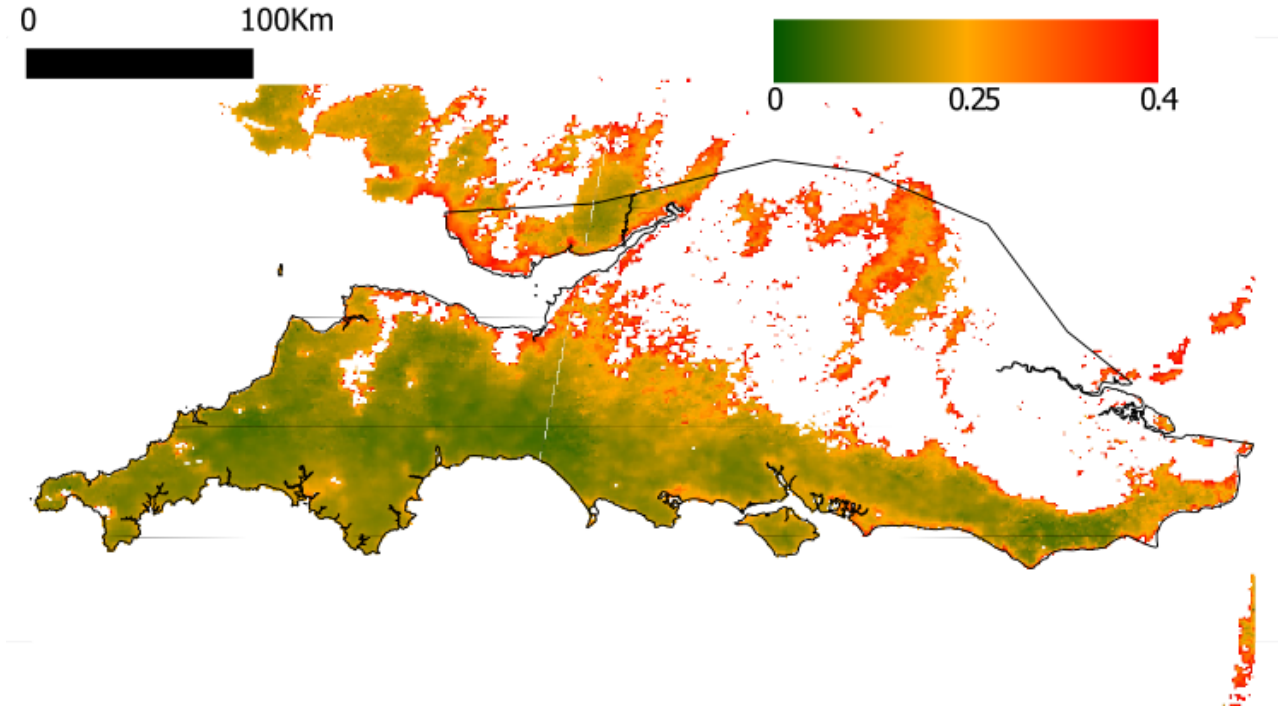
SYNTAM and MAIAC (discussed in §3.4.2.5 and §3.4.2.3 respectively) both retrieve AOT at 1km resolution from MODIS data, although they are not currently implemented as operational products. Figure 3.5 shows a comparison between 10km MOD04 and 1km MAIAC data, displaying the significant increase in the detail - and therefore the utility - of the data at a higher resolution.

Some of the methods described above have been applied to high-resolution imagery, such as that from Landsat TM and SPOT, giving AOTs at resolutions of 10–30m; significantly higher than the kilometer-scale products described above, although none of these products are operationally produced as yet. However, it is difficult or impossible to use many of the radiance separation methods described above with high-resolution data. For example, no instruments currently acquire multi-angular or polarisation data at high resolution, and many high-resolution sensors do not have the bands required to allow estimation of ground reflectances from SWIR measurements. Furthermore, most high-resolution sensors, with the notable exception of the Disaster Monitoring Constellation (DMC), do not have a short enough revisit time to legitimately assume that the ground reflectances are temporally stable between images. Thus, most high-resolution retrieval methods are limited to using DDV or other well-characterised surfaces, thus limiting their applicability in many environments (including urban areas and deserts).

For example, Liu and Liu (2009) used SPOT data to estimate AOT using the DDV method, producing a R^2 value of 0.76 when compared with AERONET data. However, this method does not perform a full spatial retrieval of AOT, mainly due to the fact that not all pixels contain DDV. Thus, the AOT value is retrieved for all DDV pixels and all other pixels in the image are assigned the AOT value of the DDV pixel nearest to them, using one of the early methods used by the ATCOR model (Richter, 1996a). Hadjimitsis and Clayton (2009) use a very similar method to obtain point (as opposed to fully spatially-distributed) AOT data near major airports in the UK and Cyprus using lakes as dark targets. Several other methods have been developed which are also based on retrieval for DDV pixels, but which use a better interpolation method. For example,



(a) MOD04 data at 10km resolution



(b) MAIAC data at 1km resolution

FIGURE 3.5: Maps of AOT over southern England on the 17th June 2006. White areas were cloudy during acquisition. Note how much extra detail is available in the 1km data, including higher AOT over urban areas, and around much of the coast.

Zhong (2011) calculate AOT over water and DDV pixels and then linearly interpolate these AOTs across the image.

A few other authors have claimed to have developed algorithms to retrieve per-pixel AOT from high-resolution images, but on closer examination the methods have fundamental issues which mean that the algorithms do not perform a true per-pixel retrieval. For example, Themistocleous and Hadjimitsis (2013), Chen *et al.* (2011) and Themistocleous *et al.* (2012) retrieve AOT over DDV, or a similarly limited range of land covers, and then interpolate across the rest of the image. Sifakis and Iossifidis (2014) claim to retrieve AOT at 30m resolution but actually aggregate pixels, producing AOT at 500m instead.

Approaches which do not involve the separation of the at-sensor radiance into the ground radiance and path radiance components have proved fruitful for use with high-resolution imagery. For example, methods based upon the 'blurring' effect caused by aerosols have also been used with Landsat and SPOT imagery, with early work by Tanré *et al.* (1988) and Sifakis and Deschamps (1992) using a reference image (acquired under very low AOT conditions) and then comparing the structure function (which measures the blurring effect between adjacent pixels) between this image and the image from which AOT is to be retrieved. More recent work by Liu *et al.* (2002) has improved this method by taking into account surface anisotropy, and application to SPOT images has retrieved AOT with a RMSE of 0.122 and a mean absolute error of 0.09. However, a fundamental limitation of these approaches is that they require a reference image which has uniform low AOT, which is difficult as a) there is significant spatial variation in AOT (as examined in Chapter 2) and b) many areas of the world (such as highly polluted cities) rarely, if ever, have a low AOT.

The Observation of Shadows for aerosol Inversion over 3D Scenes (OSIS; Thomas *et al.*, 2011) method estimates AOT from very high-resolution (sub-metre) imagery by observing the changes in radiance in areas of shadow/non-shadow transition. This is only applicable in areas with large areas of both deep shadow and bright sun, where a 3D model of surface features is available, and is thus most suited to urban areas and imagery collected at large solar angles.

Landsat is one of the few high-resolution sensors to have a SWIR band, and thus aerosol retrieval using Landsat is not limited to well-characterised areas. Lyapustin *et al.* (2004) developed a method based on 3D radiative transfer theory for aerosol retrieval from Landsat over a wide range of surfaces. Validation showed a high accuracy ($\pm 0.02 - -0.03$), although the authors admit that this may not be a realistic assessment of the quality of the method as the volume of validation data was very low, and validation was only performed over one area where the atmospheric conditions were very stable. Although the accuracy may be high, the revisit period of Landsat is a minimum of 16 days - and the cloud-free revisit period is likely to be significantly higher than this - which severely limits the utility of the resulting AOT data. As this method relies on the Landsat SWIR band it is not applicable to other high-resolution sensors with shorter revisit times

(such as the DMC sensors, SPOT, or the soon-to-be-launched Sentinel-2).

3.4.6 Interpolation

Many of the methods described above only produce AOT over certain land cover types, such as DDV or other dark surfaces, therefore producing areas with no data and requiring the use of interpolation methods to produce an AOT dataset covering the whole of the study area. This type of dataset is not always required, but it is essential for certain applications. For example, to use AOT measurements for a spatially-variable satellite image atmospheric correction a measurement is required for every pixel (Richter, 2004). A range of approaches have been used, ranging from a simple linear interpolation (Richter, 1996a) through data assimilation procedures such as Optimal Interpolation (Nirala, 2008) to geostatistical techniques such as kriging (Kanaroglou *et al.*, 2002). Many of these methods were developed in eras of lower computing power, and thus the interpolation methods used may have been simplified for computational efficiency (for example, Richter, 1996b), at the expense of statistical validity.

3.4.7 Merged products

An alternative approach to fill in the gaps in AOT measurement left by some sensors - and also improve their accuracy - is to merge the AOT products from multiple sensors. Figure 3.6, from Mélin *et al.* (2007) shows the benefits of merging data from two sensors (in this case SeaWIFS and MODIS) to increase data availability over an area.

Often merging is performed in conjunction with interpolation, with the aim of producing an accurate and complete AOT dataset for a certain area, even for pixels in which no measurements are available. It is important to decide which order to carry out interpolation and merging: whether all data are interpolated to create a set of no-gaps datasets and then merged, or whether data are merged and any remaining gaps are filled (Zubko *et al.*, 2010). The decision should be based on the confidence in the merging and interpolation procedures used.

All merging algorithms require a way of co-locating measurements from multiple instruments in both space and time before the data can be merged. This in itself can be challenging as different satellites have different resolutions, revisit periods and orbital paths, and is often performed by projecting all of the data into one co-ordinate system and simply merging data which are overlain (Nirala, 2008), although some algorithms use more complex methods based upon sensor Point Spread Functions (Gupta *et al.*, 2008).

Once a set of data co-located in time and space have been selected, the merging can be performed using methods ranging from simple arithmetic averaging (Zubko *et al.*, 2010) and weighted averaging (for example, by pixel counts) to statistical methods such as Maximum Likelihood Estimation (Zubko *et al.*, 2010), universal kriging (Chatterjee *et al.*, 2010) and mutual information (Li *et al.*, 2012) to machine learning techniques such as neural networks (Xu *et al.*, 2005).

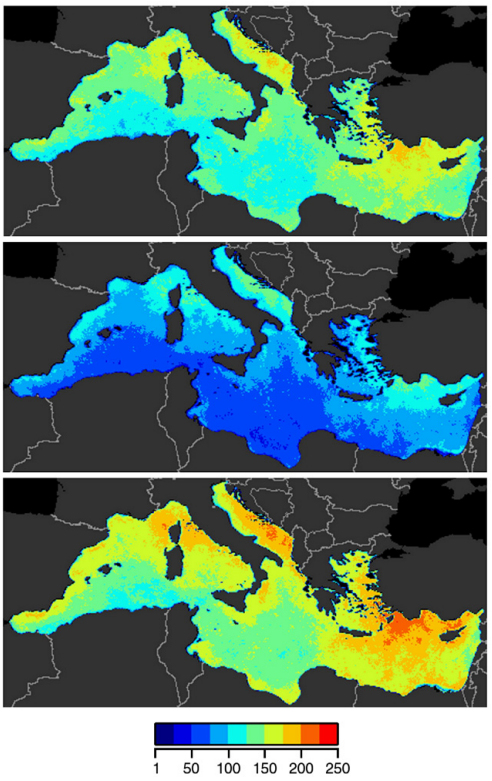


FIGURE 3.6: Number of days of AOT data available over the Mediterranean Sea. Top: SeaWIFS, Middle: MODIS, Bottom: Merged. From Mélin *et al.* (2007), Figure 13.

Many studies have shown increases in accuracy when using merged products (Xu *et al.*, 2005; Chatterjee *et al.*, 2010) but this is not always the case. For example, the GlobAerosol project was a Data User Element project run by the European Space Agency to produce a coherent, gap-filled and accurate merged AOT product for the whole globe (Thomas *et al.*, 2010). However, validation showed that the merged product had a lower accuracy than some of its constituents, and suggested that this lower accuracy may have been caused by one poor constituent dataset, or by the temporal interpolation that was required in the merging process (Poulsen *et al.*, 2009). Thus, it should be noted that merging does not always lead to higher quality products.

3.4.8 Overview of satellite approaches

Table 3.2 lists the advantages and disadvantages of each AOT retrieval approach discussed in the previous sections and Table 3.3 lists each sensor currently providing operational AOT data, along with the resolutions and accuracies of that data. Note that generally the methods that provide highest accuracy (such as polarisation-based and multi-angular retrievals) have the lowest spatial and temporal resolution. Figure 3.7 shows the development of the AOT retrieval methods described above over time, with operational products highlighted. The diagram shows how many methods developed with very little influence from other methods - for example the use of measurements of polarization for measuring AOT - and how there are operational products based upon all methods except

those which assume stable reflectances.

TABLE 3.2: Advantages and disadvantages of AOT retrieval approaches

Method	**Advantages**	**Disadvantages**
Assume a known ground reflectance For example Nagaraja Rao *et al.* (1989)	• Simple • Computationally-efficient	• Limits retrieval to oceans, DDV and other stable and characterised areas • Reflectance may not be as well-known as thought (for example, sun glint and the range in reflectance of DDV)
Estimate the ground reflectance from another band For example Remer *et al.* (2006)	• Allows retrievals over a variety of surfaces (but not all)	• Empirical relationships between bands do not work for all surfaces types - thus causing errors for certain surfaces
Use multi-angular measurements For example Lyapustin *et al.* (2011)	• Retrievals can be very accurate • Retrievals can take place over most surface types • Can retrieve physical properties as well, thus better characterising AOT	• Measurements are only available from a few satellites • Measurements tend to be at low spatial resolutions •
Use measurements of polarisation For example Mishchenko and Travis (1997)	• Retrievals can be very accurate • Can retrieve physical properties as well, thus better characterising AOT •	• Measurements are only available from a few satellites • Measurements tend to be at low spatial resolutions •
Assume ground reflectances are temporally stable For example Tang *et al.* (2005)	• Allows retrieval over a wide range of surfaces •	• Ground reflectances are never truly stable - this causes errors • Retrieval limited to areas where an assumption of stability holds

Table 3.3: Sensors currently providing operational AOT data, and associated methods, resolutions and accuracy.

Sensor	**Method**	**Spatial Resolution (km)**	**Revisit Time (days)**	**Accuracy**	**References**
AVHRR	DDV	100	1	RMSE: 0.03–0.04	Veefkind *et al.* (1999) Stowe *et al.* (1997)
MODIS	SWIR-based estimation	10	1	$\pm 0.05 \pm 0.15\tau$	Remer *et al.* (2006)
MODIS	SWIR-based estimation	3	1	$\pm 0.05 \pm 0.25\tau$	Remer *et al.* (2013)
VIIRS	SWIR-based estimation	6	1	Mean error: 0.02 StDev error: 0.11	Jackson *et al.* (2013)
MISR	Multi-angle	17.6	9	$\pm 0.04 \pm 0.18\tau$ 70% within $\pm 0.2\tau$	Liu *et al.* (2004) Kahn (2005)
OMI	UV	13 x 24	1	RMSE: 0.15 RMSE: 0.05	Ahn *et al.* (2008) Livingston *et al.* (2009)
POLDER	Polarisation	20	1	Regression: $0.96\tau_{sp} + 0.028$	Goloub *et al.* (1999)

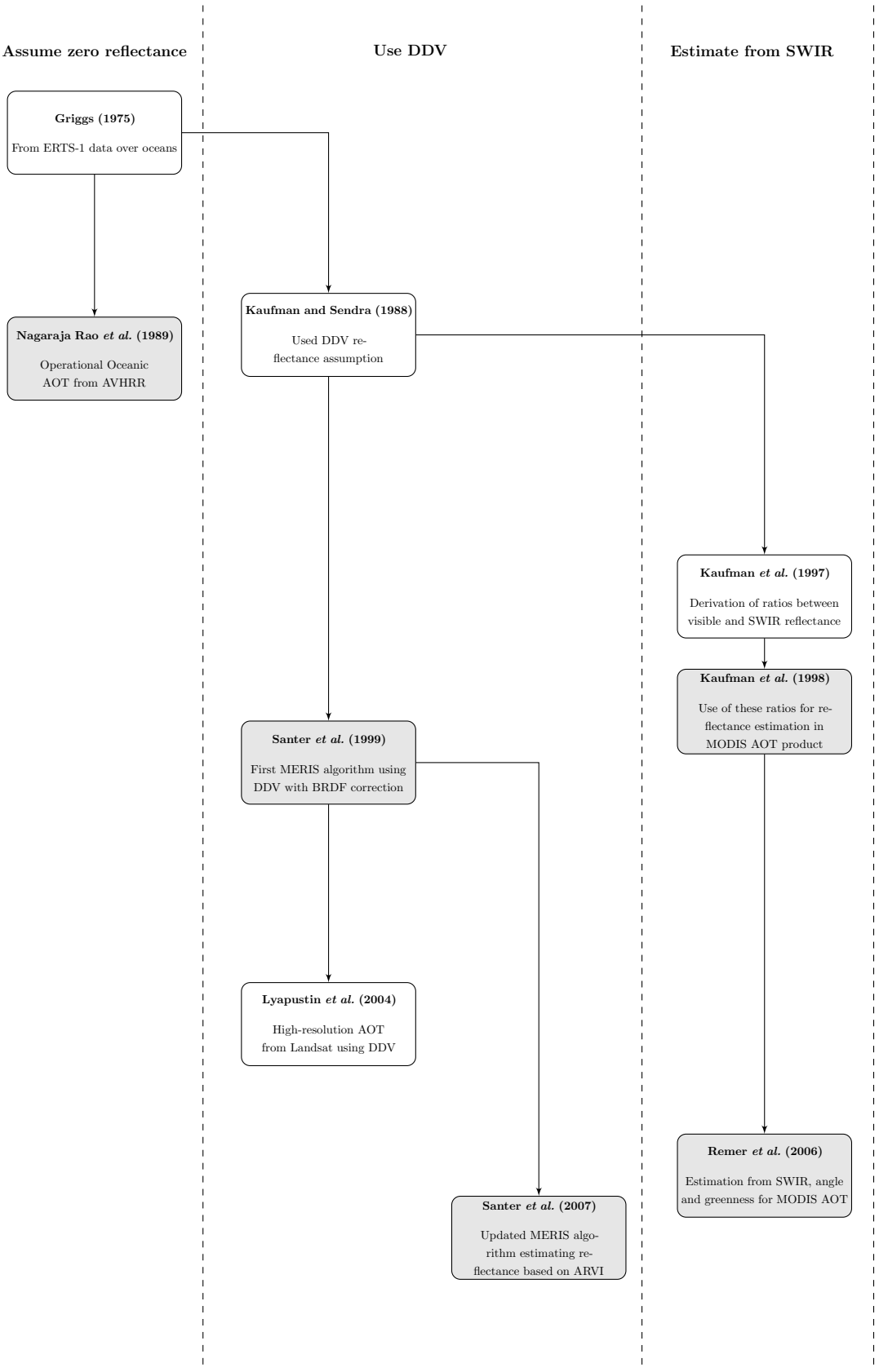
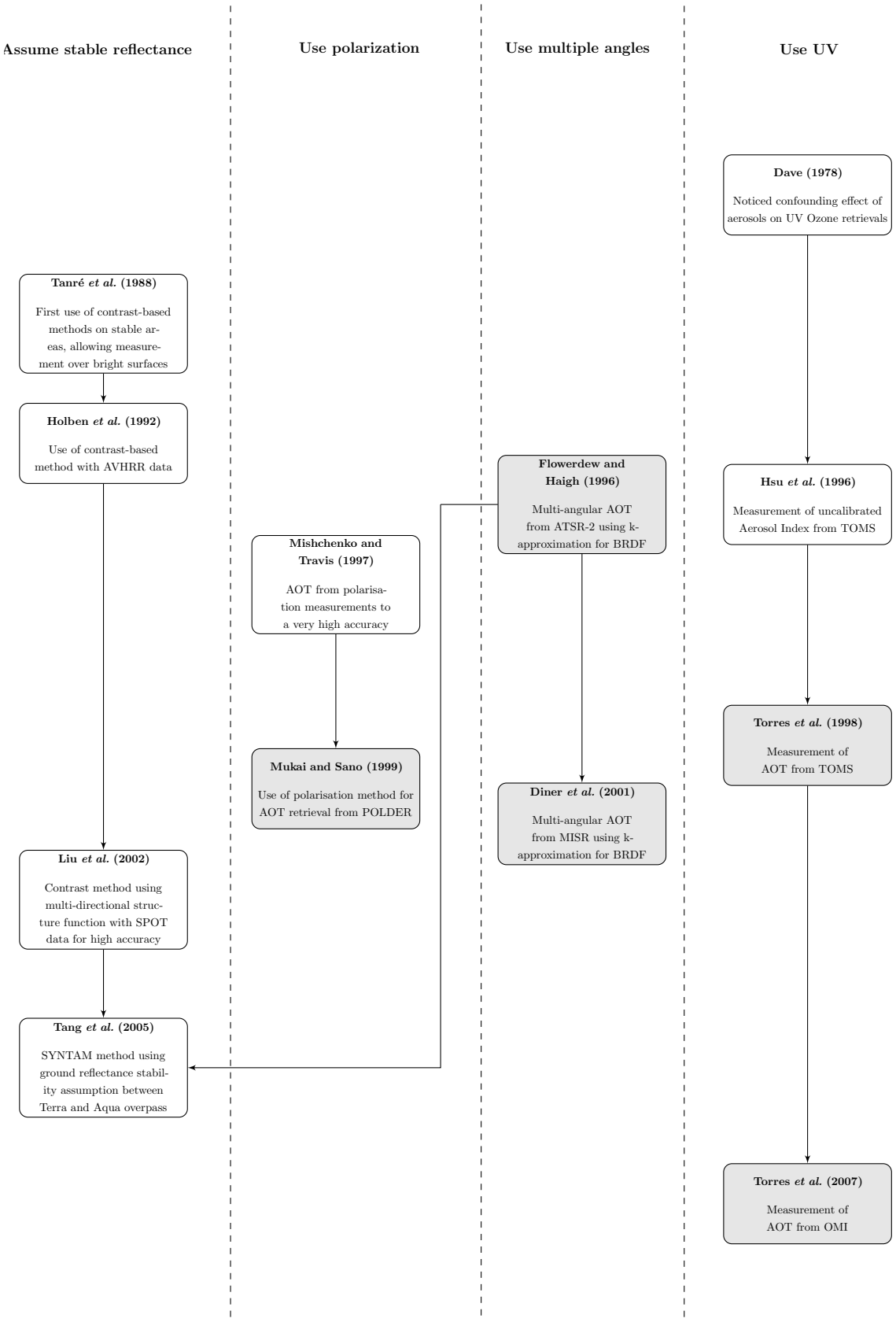


FIGURE 3.7: Diagram showing the development of satellite imagery AOT retrieval methods, with arrows showing influences. Time progress downwards from the top (1970s) to the bottom (2000s). Shaded boxes are papers describing operational algorithms.
(Continues on next page)


Assume stable reflectance
Use polarization
Use multiple angles
Use UV
Dave (1978)
Noticed confounding effect of aerosols on UV Ozone retrievals
Tanré et al. (1988)
First use of contrast-based methods on stable areas, allowing measurement over bright surfaces
Holben et al. (1992)
Use of contrast-based method with AVHRR data
Flowerdew and Haigh (1996)
Multi-angular AOT from ATSR-2 using k-approximation for BRDF
Hsu et al. (1996)
Measurement of uncalibrated Aerosol Index from TOMS
Mishchenko and Travis (1997)
AOT from polarisation measurements to a very high accuracy
Mukai and Sano (1999)
Use of polarisation method for AOT retrieval from POLDER
Torres et al. (1998)
Measurement of AOT from TOMS
Diner et al. (2001)
Multi-angular AOT from MISR using k-approximation for BRDF
Liu et al. (2002)
Contrast method using multi-directional structure function with SPOT data for high accuracy
Tang et al. (2005)
SYNTAM method using ground reflectance stability assumption between Terra and Aqua overpass
Torres et al. (2007)
Measurement of AOT from OMI

3.5 Haze assessment of satellite imagery

A separate set of literature has developed focusing on the issue of haze in satellite images and methods to classify haze-contaminated areas of an image and then remove the effects of the haze. In practice, this normally means producing a 'haze image' where each pixel contains a number representing the haziness of that pixel, and then using this image to perform some sort of 'haze correction' or 'haze reduction' operation. Haze is the noticeable visual effect of a high AOT and is caused by the increase in path radiance. Therefore the first part of these methods are basically alternative ways to assess the AOT for each pixel in an image, although they are generally not referred to as such.

Similarly to the standard AOT retrieval algorithms detailed in §3.4, the key issue when generating these 'haze images' is to ensure that the values used to represent the haziness of a pixel are not affected by the ground reflectance. That is, a change in ground reflectance with the same haze amount should lead to the same result, making haze information comparable across pixels with different land covers.

The haze removal methods discussed below, in approximately chronological order, are all image-based methods: that is, they do not require any data to function apart from that inherent within the image itself. This is important as it allows use of these methods on a wide range of images which may not all have ground data associated with them. These techniques do not perform a full atmospheric correction of the image; they aim to produce a map of the haze over the image and then use this to normalise the image to ensure a uniform amount of haze in each pixel (ideally uniformly-clear). Here we will focus on the methods used to map the spatially-variable haze amounts across the image, rather than the image normalisation methds, as this is most relevant to our goal of producing a high-resolution AOT dataset.

3.5.1 Tasselled Cap Transformation

The Tasselled Cap Transformation was developed by Kauth and Thomas (1976) from observations of the changing spectral response of an agricultural vegetation pixel over the growing year, as the surface transitions from bare soil, to a full canopy, and back to bare soil again. They plotted these trajectories in a four-dimensional feature-space consisting of the Landsat MSS bands, and found that different surface types appeared at different places on the graph, producing an image that looked like a line drawing of a cap with tassels on it (see Figure 3.8).

As exploring a four-dimensional space is challenging, they developed the Tasselled Cap Transformation, which extracted four separate features from measurements in the Landsat MSS bands. Each of these features were orthogonal to each other, thus allowing independent examination of the separate components of the spectral response of each pixel, and were calculated using simple affine transformations of the raw Digital Number data. These features were,

- **Brightness** (TC1)

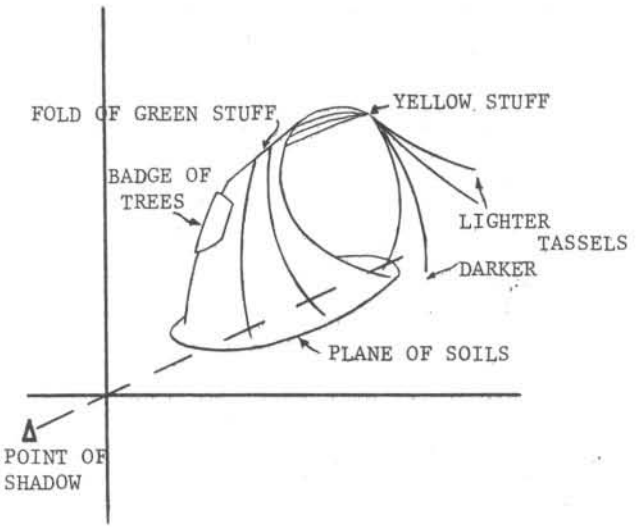


FIGURE 3.8: The 'Tasselled Cap' in Landsat MSS feature space. Taken from the original description in Kauth and Thomas (1976), ©1976, 2014 IEEE

- **Green-stuff** (TC2)
- **Yellow-stuff** (TC3)
- **Nonsuch** (TC4)

As these features are orthogonal, the *green-stuff* feature could be used for assessing the phenological development of vegetation, without contamination by the overall *brightness* of that pixel, which may be affected by atmospheric or geometric conditions (although in some situations there can be some 'leakage' of response from one feature to another). In a similar way to a principal components analysis, the first few tasselled cap components contain the majority of the variability of the data (TC1 and TC2 together account for nearly 95% of the variability according to Crist, 1985) suggesting that certain tasselled cap components can be ignored when processing data so as to reduce processing time. In the years after its original development, Tasselled Cap Transformations were developed for many other sensors including Landsat TM (Crist and Cicone, 1984), SPOT (Silva, 1992) and MODIS (Lobser and Cohen, 2007).

In the first paper on the Tasselled Cap Transformation (Kauth and Thomas, 1976), there was a suggestion that one or more of the features may be useful for assessing the hazyness of each pixel, and this idea was further investigated by several other authors. The original suggestion was that atmospheric haze would cause a shift of the data points in the 'yellow-stuff' and 'non-such' directions, and that by measuring the magnitudes of these shifts some information on the haziness of the atmosphere could be determined (Kauth and Thomas, 1976), although they did not implement a haze diagnostic based on these ideas. It is feasible that the 'yellow-stuff' feature may contain information on haze, as it primarily consists of the difference of the first two Landsat MSS bands (blue and green) which are highly-correlated for standard surface retrievals but affected differently by haze. Thus, movement away from the high correlation (measured as larger than expected

differences between the bands) suggest haze contamination.

As part of the development of a full Landsat data pre-processing system a haze diagnosis and correction method called XSTAR was implemented by Lambeck (1977). This was also based upon the 'yellow-stuff' vector, in this case measuring a noticed translation of the brightness-greenness hyperplane in the direction of the 'yellow-stuff' vector for haze-contaminated pixels. The translational offset was calculated from a standard 'yellow-stuff' value derived from analysis of a range of "average Landsat scenes" (Lambeck, 1977), and a multiplicative transformation was then performed to translate each pixel back to this average 'yellow-stuff' value.

Crist (1985) developed the Tasselled Cap Haze Transformation, also known as TC4, although it is not simply the fourth of the Tasselled Cap transformed bands. To reduce confusion we will refer to it here as the TCHT. The development was done using the Tasselled Cap Transformation for the Landsat TM sensor, which has more bands than Landsat MSS, and therefore produces more Tasselled Cap bands. The first three Tasselled Cap bands seem to be strongly related to surface features ('Brightness', 'Greenness' and 'Wetness') and therefore the haze transformation was derived from the rest of the Tasselled Cap bands. The derivation used a Principal Components Analysis of simulated images produced under clear and hazy conditions, and produced a haze measure calculated as

$$\text{TCHT} = 0.88\text{TM}_1 - 0.08\text{TM}_2 - 0.46\text{TM}_3 - 0.003\text{TM}_4 - 0.6\text{TM}_5 + 0.1\text{TM}_7 \qquad (3.38)$$

TCHT is conceptually similar to the use of the 'yellow-stuff' feature by Lambeck (1977), as it too primarily consists of the difference between two highly correlated bands. In this case the blue and red bands are used, as they are very highly correlated but are further apart on the spectrum than the blue and green bands used by the 'yellow-stuff feature', and therefore the differences in the influence of haze on the bands is larger.

The TCHT has been applied by several researchers to normalise the haze in images. For example, Lavreau (1991) applied this to normalise spatially-variable haze in Landsat TM images by relating the haze diagnostic values to the DNs in each band, and applying this relationship to correct all pixels to an equivalent of a haze diagnostic of zero.

The widely-used ATCOR atmospheric correction tool (Richter, 2004) uses the TCHT as part of its optional haze-reduction procedure. It is important to note that although the ATCOR software performs atmospheric correction, the haze-reduction procedure is an optional pre-processing step in the software, and none of the information about haze is used in the atmospheric correction procedure itself. The algorithm originally used a simplification of the TCHT created by reducing it to only the two most significant terms, that is:

$$\text{Haze Diagnostic} = 0.88\text{TM}_1 - 0.46\text{TM}_3 \qquad (3.39)$$

A threshold for this diagnostic was used to mask hazy regions, and further thresholds were used to exclude urban areas (which often receive abnormally high haze diagnostic values) and thick cloud. A histogram matching procedure was then used to produce a corrected image with significantly less influence from haze (Richter, 1996a).

Since its original development, the algorithm has seen several improvements:

- Richter (1996b) implemented a transition region between hazy and non-hazy areas, as the previous algorithm resulted in sharp transitions that were very noticeable in the resulting images. Kriging and other geostatistical techniques for interpolating haze diagnostic values were not used as they were considered too computationally-intensive, and instead a simple linear interpolation over a range of around ten Landsat pixels was implemented, with a separate correction performed for each 'contour level' of the interpolation.
- More recent versions of ATCOR (no exact date when the change occurred is available) only use the TCHT-derived haze diagnostic for the original masking of the haze, and instead use the Haze Optimised Transform (HOT; Zhang *et al.*, 2002a, discussed below) to estimate the haziness of each pixel (Richter, 2012).

An essential quality of any haze assessment procedure is that the haze measure responds only to haze and not to ground reflectances. Unfortunately all of the Tasselled Cap-based haze assessment methods described above are sensitive to ground reflectances in certain situations. Crist (1985) noted that the TCHT measure responded particularly to certain land covers: for example, roads and water bodies, and thus required some filtering before use. More concerning is that there appears to be a differential response of the haze diagnostic over green and senescent vegetation, significantly limiting the application of the haze diagnostic over areas with a range of vegetation types.

Richter (1996a) dealt with some of these issues by filtering out urban areas and areas of thick cloud, but did not suggest any solutions for the effects caused by green/senescent vegetation. Furthermore, Zhang *et al.* (2002a) found a significant positive relationship between the TCHT value and the surface reflectance, thus suggesting that an increase in the 'haze diagnostic' could be caused by an increase in ground reflectance rather than haziness. Indeed, these issues with ground reflectance contamination of the TCHT may have driven the switch to the Haze Optimised Transform for correction within the ATCOR software.

3.5.2 Haze Optimised Transform (HOT)

The Haze Optimised Transform (HOT; Zhang *et al.*, 2002a) was designed to address some of the limitations in Tasseled Cap-based haze assessment. The transform itself is designed specifically to extract the 'haze signal' from an image, without any contamination from the ground reflectance, and has been developed entirely for this purpose, in contrast to the Tasselled Cap Transformation for which haze assessment was not the primary aim.

The HOT is, however, based upon the same fundamental basis as the Tasselled Cap derived haze measures: the high correlation between visible bands which is disturbed by haze. The Tasselled Cap approaches calculate the 'haze' diagnostic as the difference between the two bands, whereas the HOT takes a more robust approach based on vectors in feature space. There has been some criticism of the correlation-based approach, but as far as the author is aware, no detailed studies have been performed to assess the correlation between bands in a range of multispectral images. Liu *et al.* (2011) criticised the HOT because of its reliance on image correlation, stating that for 23 hazy Landsat TM images analysed, the correlation coefficients between the visible bands were generally lower than 0.9, and sometimes lower than 0.8. However, these assessments were performed on hazy imagery which hadn't been corrected, and the assumption underlying the HOT is that haze in an image will reduce the correlation between bands, thus these results are of no real significance.

When plotting any two of the blue, green and red reflectances for pixels in feature space a straight line in a positive direction can be seen, caused by the high correlation between these bands. This line is defined as the *Clear Line*, or *Clear-sky vector*. Increases in AOT values due to aerosols, thin cloud or other contamination, cause pixels to move away from this line in a predictable manner. Figure 3.9, taken from the original HOT paper (Zhang *et al.*, 2002a), shows the relationship between DNs in Landsat TM bands 1 and 3 for a range of surface types (A–K) under clear conditions (with no aerosols). The small numbers show how the original points migrate away from the Clear Line as the AOT increases from 0 to 6.7 (the large maximum here taking into account the optical thickness of clouds). The original definition of the HOT used Landsat TM bands 1 (blue) and 3 (red), but the authors comment that it should be equally applicable using any visible bands, as they are all highly correlated.

It is important to note that equally-numbered points for all surface types form a line almost parallel to the Clear Line: for example, the Clear Line in this graph has a slope of 1.03, and the slope of the line at the highest AOT (labelled `19` in the graph) is 0.98. It can be seen that the trajectories of each original point are different, but they all result in a similar displacement perpendicular to the Clear Line for the same increase in AOT, suggesting that this displacement reflects AOT only and is not affected by surface type. Thus, the Haze Optimised Transform is defined as the perpendicular distance from the Clear Line to each point. The formula used by Zhang *et al.* (2002a) is

$$\mathrm{HOT} = B_1 \sin\Phi - B_3 \cos\Phi \tag{3.40}$$

where Φ is the slope angle of the Clear Line.

Detailed simulations performed by Zhang *et al.* (2002a) show that the HOT is significantly less affected by ground reflectance than the Tasselled Cap based methods, which was also confirmed using real-world image data. Furthermore, the HOT was shown

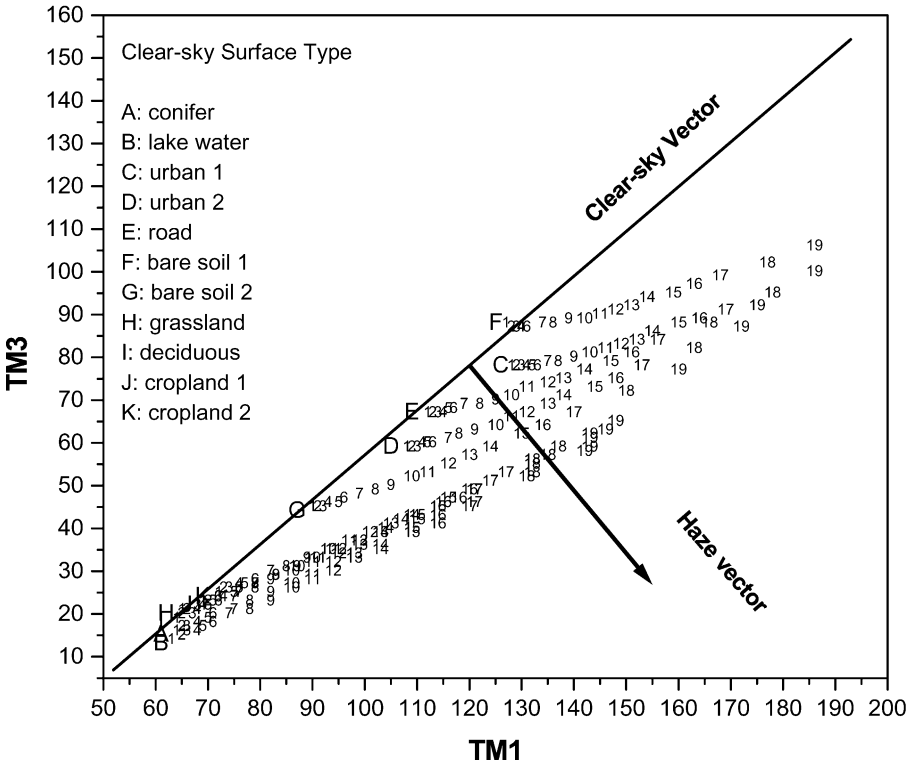


FIGURE 3.9: The fundamental basis of the Haze Optimised Transform, taken from Zhang *et al.* (2002a). The graph shows Landsat TM band 1 (Blue) against Landsat TM band 3 (Red), with the Clear-sky Vector (the Clear Line) showing the high correlation between the two. The numbered points show the DN values as the AOTs increase from 0 to 7.6 (in thirteen equal-sized steps), with a gradual migration away from the Clear Line visible as the Haze vector.

to be insensitive to aerosol or cloud types, with very similar trajectories away from the clear line for different types of haze contamination. This has advantages and disadvantages: it means that the HOT can be used without knowledge of the underlying aerosol type, but it also means that the HOT can only provide an estimate of the haze of a pixel but not the aerosol/cloud type.

Wang *et al.* (2013) found a strong quadratic relationship ($R^2 = 0.9999$) between AOT and HOT, suggesting that HOT could be used to estimate AOT accurately, which is unsurprising as the HOT value can be thought of as a proxy for the path radiance. The HOT-AOT relationship found varied slightly depending on the underlying surface, as Figure 3.10 shows. Furthermore, as the HOT was developed on a range of AOTs up to 6.7, it may not be sensitive enough to low AOTs to be used as an AOT mapping technique - this will need further investigation if this method is to be used to estimate AOT.

Although the HOT is a significant improvement on the Tasselled Cap-based methods, there are still some effects from ground reflectance. Examining Figure 3.9 in detail shows that the clear-sky values for three land-cover classes do not lie exactly on the Clear Line: these are lake water (B), urban 1 (C) and urban 2 (D). As these classes are offset from the Clear Line in the direction of the haze vector, they will have higher HOT values than

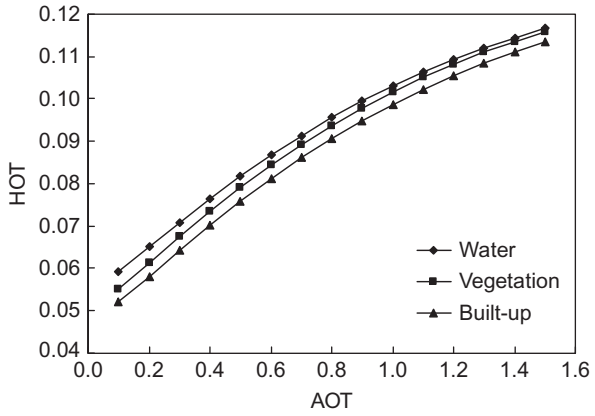


FIGURE 3.10: Relationship between AOT and HOT for three different surface types. All relationships are quadratic and have a R^2 value over 0.99. Taken from Wang *et al.* (2013)

they should, thus giving the false appearance of more haze over these areas. Similarly, bare soil (F) is offset from the Clear Line in the opposite direction, thus giving lower than expected HOT values. Although snow was not shown in Figure 3.9, its extreme brightness - combined with large variation in reflectance - means that it is likely to also produce anomalous HOT values.

3.5.3 Extensions of the HOT

Further work on the HOT by the original authors has mainly focused on analysing the effect that haze reduction based on the HOT has on images (Zhang *et al.*, 2002b; Zhang and Guindon, 2003). However, several authors have attempted to extend the Haze Optimised Transform itself to deal with some of its limitations.

3.5.3.1 Haze removal for high-resolution satellite data

Moro and Halounova (2007) applied the HOT to assess the haze in IKONOS images with a spatial resolution of 4m. They confirmed the results of Zhang *et al.* (2002a), finding that both urban areas and water received abnormally-high HOT values, thus causing errors when applying the haze reduction methodology. Attempts were made to structurally decompose the image into haze features and ground features using methods such as wavelet decomposition and image morphological operators, but this was found to be impossible as haze variation and ground features have similar dimensions and scales of variability. Smoothing using standard low-pass filters was also attempted, but this was found to result in a loss of useful information about the small-scale variability of the haze, although it was noted that smoothing may be an acceptable technique for lower-resolution imagery.

The final chosen method was to mask water and urban areas from the analysis entirely (using a simple NIR threshold for water, and a general brightness threshold for urban areas) and then interpolate HOT values for these areas using either splines or simple linear interpolation. This was found to give reasonable output maps of haze as well as

appropriate final images after corrections were applied.

3.5.3.2 Advanced Haze Optimised Transform

The Advanced Haze Optimised Transform (AHOT; He *et al.*, 2010) was designed to deal with some of the issues with the original HOT procedure. This involved two main changes to the algorithm:

- An empirically-derived offset was applied to the HOT formula to ensure that the HOT of areas in the manually-defined clear region were zero.
- A post-processing procedure was applied to remove the effects of land cover on the HOT values

The post-processing procedure was first attempted based on the method from Liang *et al.* (2001, 2002), which involves generating a land cover map using a clustering method on the infra-red bands and then subtracting the mean HOT for a pixel's land cover from that pixel. However, this did not produce acceptable results, and a method based on thresholding was also found to be unworkable, so a separate procedure was developed based upon the methods of 'filling sinks' and 'flattening peaks' often used in Digital Elevation Model processing. In this case the sinks and peaks were spuriously low and high HOT values respectively. The sinks were found and filled using the method of Planchon and Darboux (2002). But, extracting spurious HOT peaks was more difficult as it is important not to accidentally select genuinely high HOT values. The peak detection algorithm used here was based upon selecting regions with sharp HOT value transitions at their edges, as sharp transitions in haze occur rarely in nature, and thus these regions are likely to have spuriously high values caused by land cover effects. These methods gave high-quality results, with the resulting mean and standard deviation of HOT similar across all land cover classes.

3.5.3.3 Background Suppressed Haze Thickness Index

The Background Suppressed Haze Thickness Index (BSHTI; Liu *et al.*, 2011) is an alternative haze measure which is designed to suppress the background noise (for example, from land cover changes) as much as possible. Like the Tasselled Cap transformation, it is a linear combination of bands

$$\text{BSHTI} = k_1 b_1 + k_2 b_2 + k_3 b_3 + k_4 \tag{3.41}$$

where b_1, b_2 and b_3 are the blue, green and red bands respectively and k_1 to k_4 are a set of coefficients. These coefficients are chosen so as to maximise the score function

$$\text{score function} = \frac{m_{\text{TR}}}{\sigma_{\text{CR}}} \tag{3.42}$$

subject to the constraint

$$m_{\mathrm{CR}} = 0 \tag{3.43}$$

where TR is a manually delineated thick haze region, CR is a manually delineated clear region and m and σ are the mean and standard deviation respectively. Thus, the score function is also the function for the signal-to-noise ratio of the image, and the constraint ensures that the mean value in the clear region is zero. The coefficient values (k_1 to k_4) can be calculated easily using a set of simultaneous equations based upon the covariance of the bands.

By choosing separate coefficient values for each image, rather than using standard values like the Tasselled Cap haze diagnostic, BSHTI ensures that the haze measurement has the best signal to noise ratio possible for the image, but this also means that the resulting haze measurements are not comparable across different images.

After haze detection has taken place using BSHTI, a 'haze perfection' step is performed to ensure that no effects remain from land cover variations. This step assumes that real haze varies slowly and continuously across the image, and therefore sudden increases or decrease in haze values are taken to be evidence of incorrect values which are removed using sink filling and peak flattening procedures very similar to those used in He *et al.* (2010).

3.5.4 Alternative methods

3.5.4.1 Normalised Difference Haze Index

The Normalised Difference Haze Index (NDHI; Zha *et al.*, 2012) is similar in design to the well-known NDVI and was developed using MODIS bands 1 and 4 (red and green respectively),

$$\mathrm{NDHI} = \frac{M_1 - M_4}{M_1 + M_4} \tag{3.44}$$

The denominator of this fraction is higher for pixels covered by haze and the numerator will normally be negative, and of a larger magnitude for hazy pixels, thus giving a higher NDHI for these pixels. Zha *et al.* (2012) found higher correlations between NDHI and ground-level particulate matter than those found using AOT, but there are some significant limitations to this method, which make it almost unusable for estimating AOT:

- The NDHI values can be significantly contaminated by the ground reflectance for pixels which are only covered by thin haze, thus this method is best suited to areas where the haze is almost opaque. Thus, the method is more suited to finding areas of thin cloud and thick haze, rather than mapping AOT values.
- The method cannot be used over urban areas as the relationship between NDHI values and haze is different over urban and non-urban areas: over urban areas high NDHI values signify less haze, whereas elsewhere high values signify more haze.

3.5.4.2 Multiple image methods

Several methods have been proposed for assessing haze in an image by comparing a hazy image to a clear image, for example Du *et al.* (2002) who use wavelet analysis to compare the images. The major limitation of these is that they require an entirely clear image, which is often difficult or impossible to acquire. Furthermore, it may be possible to collect an image with no obvious haze present within it, but it is impossible to obtain an image with an AOT of zero across it, as there are always aerosols present in the atmosphere. Thus, these methods cannot be used to assess AOT, just thicker haze.

3.5.5 Summary of haze-based methods

Figure 3.11 shows a summary of the haze assessment methods discussed in this section, with influences from previous work shown by arrows. It should be noted how much of the work is based upon the haze diagnostic of Crist and Cicone (1984).

3.6 Modelling aerosol dynamics

Aerosol concentrations in the atmosphere are not static: they change constantly over time through three major processes:

- **Emission/Formation:** The creation of aerosols in the atmosphere results either from their emission from a source (such as the entrainment of dust from a desert surface or the emission of soot aerosols from a car exhaust) or their formation within the atmosphere (for example, as part of a chemical reaction between other particles in the atmosphere).
- **Transport:** The movement of aerosol particles both horizontally and vertically, caused by atmospheric flow processes, primarily driven by pressure gradients.
- **Deposition:** The removal of aerosols from the atmosphere, normally resulting in the particles being deposited onto the Earth's surface. This can be caused by a range of processes including gravitational deposition (which particularly affects large particles) and precipitation washout (where aerosols become incorporated into rain drops and removed from the atmosphere when it rains).

Many models have been developed which incorporate all three of the dynamical processes above. The Global Ozone Chemistry Aerosol Radiation and Transport (GOCART) model was originally developed to model atmospheric sulphur dynamics (Chin *et al.*, 2000), but has since been extended to simulate a wide variety of aerosols. Emission is dealt with on a per-species basis: for example, Ginoux *et al.* (2001) parameterised dust uplifting into the atmosphere based upon surface type, wind speed and a threshold velocity, and deposition (including gravitational settling and precipitation washout) is parameterised based upon particle size and standard physical formulae. Modelling of the transport of aerosol particles is based upon meteorological fields of pressure and vertical motion, from which the x, y and z wind vectors can be determined. The advection (movement) of a particle is then simulated as a simple addition of these vectors to the initial position of the particle.

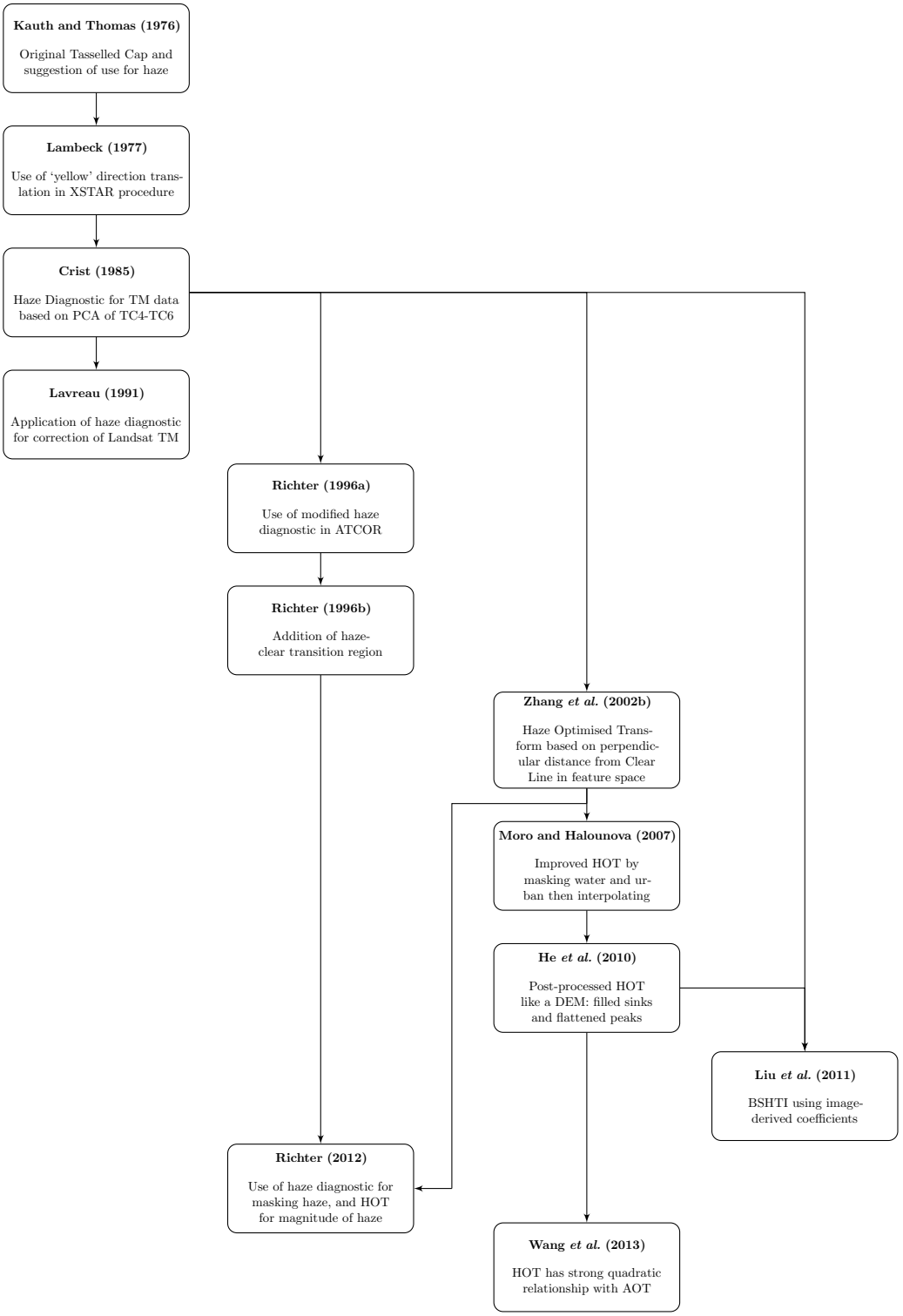


FIGURE 3.11: Diagram showing the development of haze assessment methods for satellite imagery, with arrows showing influences. Time progress downwards from the top (1970s) to the bottom (2010s)

Several similar models exist, including an extension to the community-driven Weather Research and Forecasting (WRF) numerical weather prediction model (Skamarock *et al.*, 2005) called WRF-CHEM (Grell *et al.*, 2005). This extension is fully integrated with the main WRF meteorological model, thus allowing meteorological processes to affect aerosols and aerosols to also affect meteorological processes.

Running any numerical weather prediction model is very computationally expensive and requires significant effort to be put into parameterisation. The Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Rolph, 2003) is a simpler model which can be configured to only simulate aerosol transport. Given the location of a single particle (which can be used to represent a group of particles), HYSPLIT can calculate a trajectory for that particle forward or backward in time, based upon forecast or archive data provided by meteorological organisations including the NCEP Global Data Assimilation System (Rodell *et al.*, 2004). Ensembles of simulations can also be run to take into account the chaotic nature of the atmosphere and the uncertainties in the model input data. An example of HYSPLIT output is shown in Figure 3.12.

Simple trajectory analysis with HYSPLIT has been used by several authors to characterise aerosol source regions. For example, McGowan and Clark (2008) identified seasonal dust transport routes from Lake Eyre, Australia by using HYSPLIT model runs at multiple altitudes, finding that dust emitted over Lake Eyre can be transported thousands of miles to locations including Borneo and Antarctica. Segura *et al.* (2013) developed an procedure for using HYSPLIT to determine the main aerosol type for a specific location and time. Their method used HYSPLIT to establish back-trajectories from the location, and then a method to assign aerosol emission types to each point along the trajectory, defining the aerosol type at the final location as the most common aerosol type along the trajectories. Wang *et al.* (2009) developed TrajStat, a GIS-based tool to analyse multiple HYSPLIT trajectories (from running HYSPLIT in the ensemble mode) and produce a statistical assessment of the contribution of each possible source area to the resulting aerosol load at the simulated location.

In a separate mode, HYSPLIT can calculate the dispersion of a pollutant released from an emission site. The processes driving the emission (such as entrainment of desert dust) are not simulated; the model is simply parameterised to release a certain amount of pollution (a 'puff') at regular intervals over time. This puff is then advected in the same way as a single particle, but also undergoes dispersion - that is, the puff gets larger over time. In this configuration, HYSPLIT can also simulate the deposition of particles through gravitational settling and precipitation washout. An example of the output of this mode of HYSPLIT is shown in Figure 3.13. This mode of HYSPLIT has been used to model diverse atmospheric constituents including particulate mercury pollution (Chand *et al.*, 2008) and nuclear fallout (Moroz *et al.*, 2010).

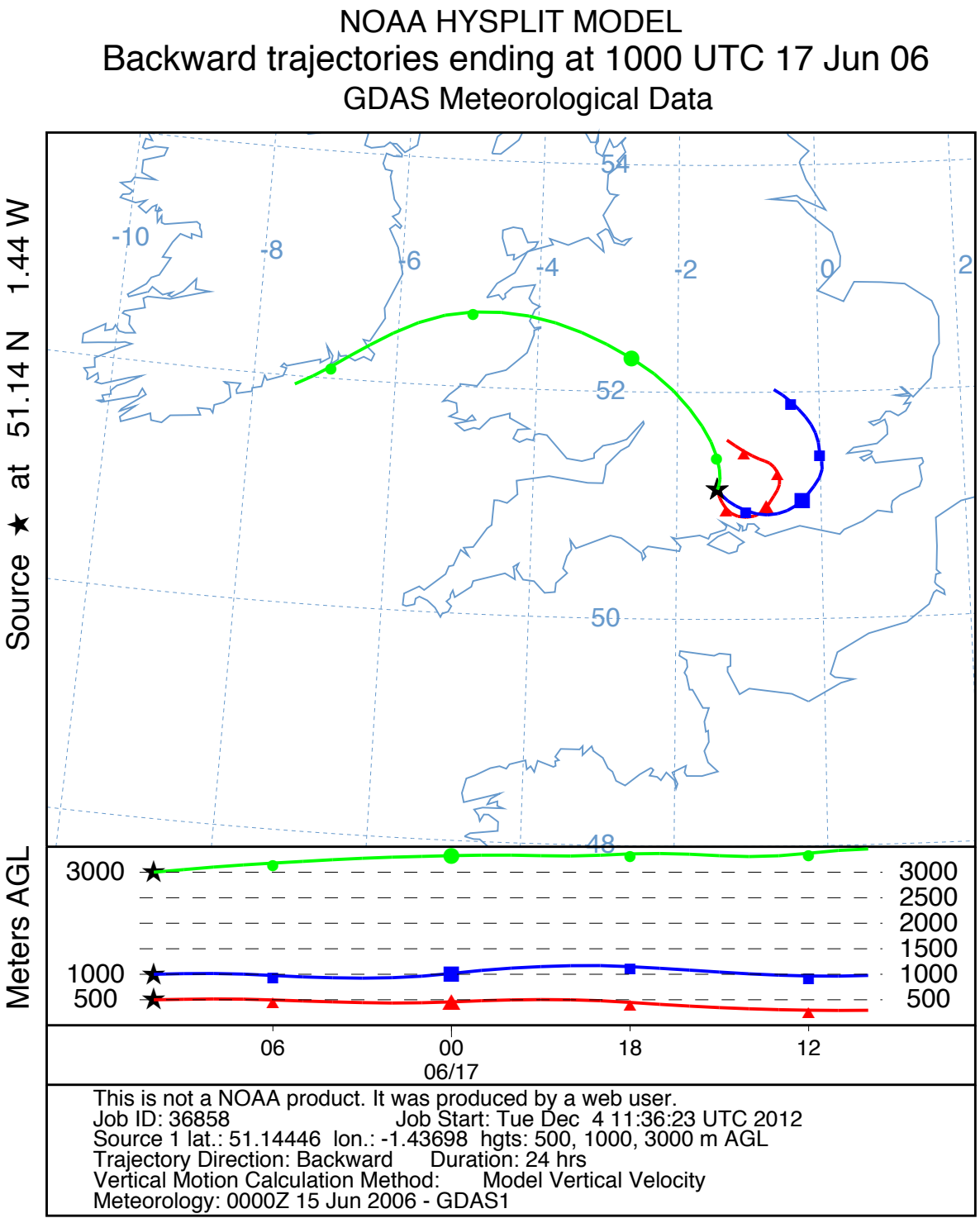


FIGURE 3.12: An example of a HYSPLIT model output for a backwards trajectory for a particle at Chilbolton, near Andover starting at 10am on the 17th June 2006. Note the very different trajectories for particles at different heights (500m, 1000m and 3000m above ground level).

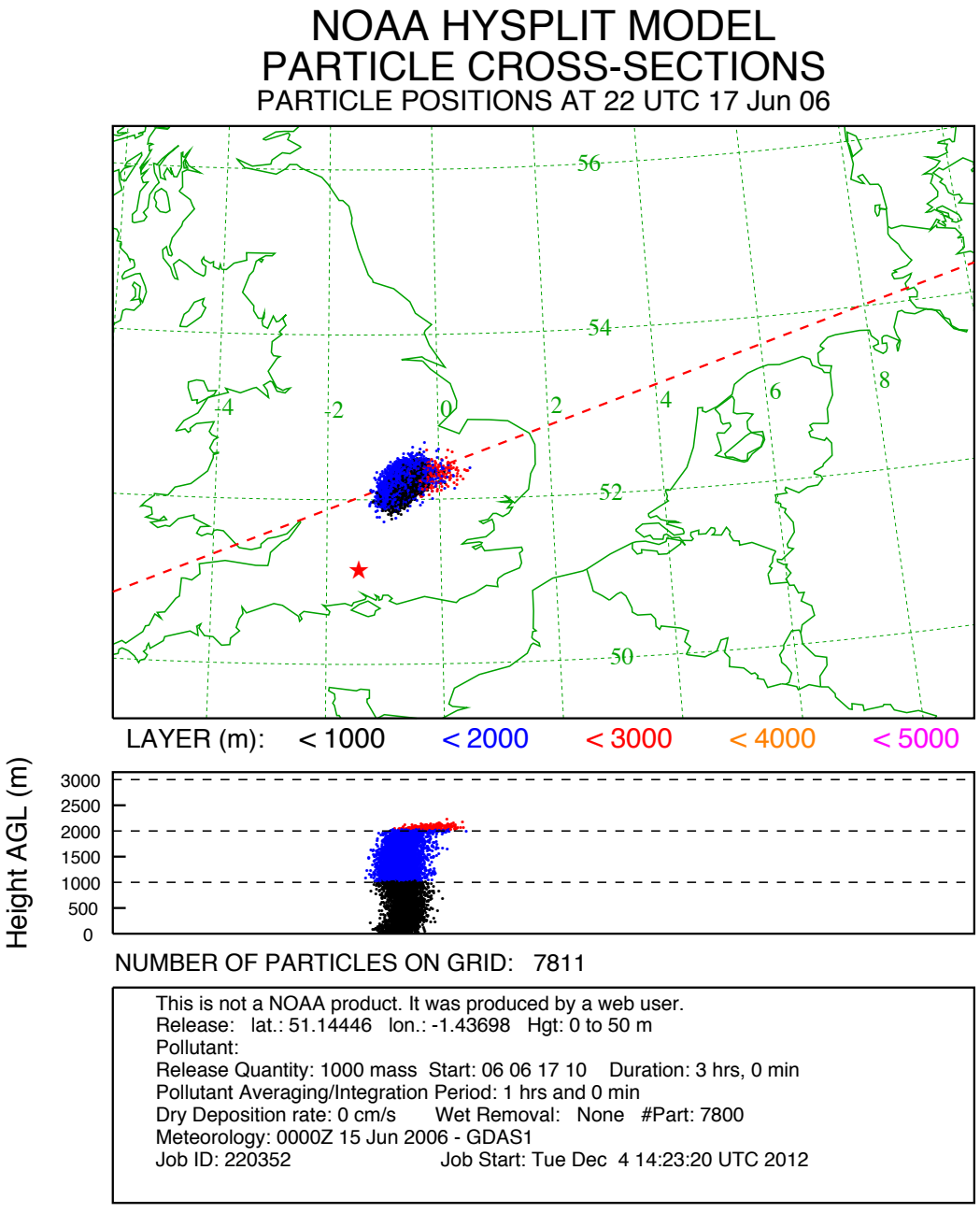


FIGURE 3.13: Example of the distribution of pollution released at Chilbolton, near Andover after 12 hours, with the release starting at 10am on the 17th June 2006. Note the height profile of the pollution.

3.7 Applications of AOT measurements

3.7.1 Air quality and human health

As outlined previously, particulate matter pollution is a significant issue worldwide, particularly in urban areas. The 2010 Global Burden of Disease report (Lim *et al.*, 2012) estimates that ambient particulate matter pollution was responsible for 3.2 million premature deaths worldwide in 2010, and 76,000 Disability Adjusted Life Years (DALYs; a measure of the number of years of life lost due to serious illness, disability or death). This is 3.1% of the total global DALY burden - a significant increase from the figures estimated in previous risk assessments (Ezzati *et al.*, 2002), due to the use of satellite data and global atmospheric models to produce a more accurate estimate of global $PM_{2.5}$ concentrations.

There is significant variation in the ranking of particulate matter pollution as a contributor of DALYs, ranging from the fourth largest contributor in East Asia (eclipsed only by high blood pressure, smoking and poor diet) to the 32nd largest contributor in Oceania. East Asia is an area with particularly high levels of particulate matter pollution and, with limited ground monitoring networks, satellite measurements are one of the few sources of data. An extreme pollution episode occurred over Beijing in January 2013, with Uno *et al.* (2014) reporting $PM_{2.5}$ concentrations of over 1000µg/m^3 (for comparison the World Health Organization's upper limit for $PM_{2.5}$ concentrations is 35µg/m^3), with significant spatial variability that could not be captured by ground-based instruments (Tao *et al.*, 2014).

Particulate matter concentrations are usually measured as PM_{10} or $PM_{2.5}$, which represent the concentrations (usually in µg/m^3) of particles with a diameter less than 10µm and 2.5µm respectively. Generally the focus for human health studies is $PM_{2.5}$ as these are the only particles that can penetrate deep inside the lungs, and thus cause the most significant health effects (Davidson *et al.*, 2005), and many regulatory bodies require monitoring of $PM_{2.5}$ by local authorities.

Many studies have attempted to derive a relationship between AOT and $PM_{2.5}$, but this is challenging as they measure fundamentally different things. AOT is a columnar measurement, which can be defined as the integral of the vertical extinction coefficient at all altitudes through the atmosphere, whereas PM is the concentration of particulate matter in the atmosphere at ground level. Thus, even if we can assume a direct relationship between concentration of particulate matter and extinction coefficient, the measurements are still performed over different height ranges. Hoff and Christopher (2009) define the relationship between AOT and $PM_{2.5}$ as

$$\mathrm{AOT} = \mathrm{PM}_{2.5} H S \tag{3.45}$$

where H is the height of the well-mixed planetary boundary layer (PBL) and S is the specific extinction efficiency of the aerosol at the ambient relative humidity. Thus,

according to this definition the relationship will depend on the planetary boundary layer height as well as the type and size distribution of the particles, as these control the extinction efficiency.

Hoff and Christopher (2009) also provide a review of a wide range of attempts to link AOT and PM, from studies using over a thousand ground measurements of PM over the continental United States (Engel-Cox *et al.*, 2004), to those based on one measurement site (Schaap *et al.*, 2009). The majority of these studies are empirical in nature: they obtain measurements of both AOT and $PM_{2.5}$ and try to estimate a relationship. The accuracy (in $PM_{2.5}$ units) of the relationship is rarely reported; instead, correlation coefficients are given. Overall, the maximum correlation coefficient reported was 0.98, and the minimum was 0.4, with an average of 0.69. This suggests a relatively strong relationship, although some studies perform significantly better than others. These studies tend to be those which incorporate other variables such as the PBL height and the humidity into the relationship. For example, Tao *et al.* (2012) found that including PBL height and relative humidity in a model relating AOT to $PM_{2.5}$ increased the R^2 from 0.32 to 0.62. Furthermore the variability of aerosol types was shown to be important by Engel-Cox *et al.* (2004) found significantly higher correlation coefficients in areas with a single dominant aerosol type when compared to areas with a wide range of aerosol types, sizes and compositions.

Recently there have been developments in more robust ways of estimating $PM_{2.5}$ from AOT measurements. van Donkelaar *et al.* (2010) developed a method to calculate a conversion factor, η, which can be used to estimate $PM_{2.5}$ as,

$$PM_{2.5} = \eta\tau \tag{3.46}$$

They used a three-dimensional chemical transport model (GEOS-Chem; Bey *et al.*, 2001) to estimate aerosol properties and planetary boundary layer height, taking into account meteorological climatologies and aerosol types. The model was used to estimate a monthly climatology of η at a spatial resolution of 0.1 x 0.1 degrees. The conversion factor was applied to estimated $PM_{2.5}$ from global MODIS and MISR AOT retrievals (filtered to remove retrievals which were likely to have a high error), and comparisons with ground measurements showed that approximately 67% of the retrievals were within $\pm 1 \pm 15\%$, with a global population-weighted uncertainty of 6.7 $\mu g/m^3$) - with the majority of this uncertainty coming from uncertainty in the satellite AOT measurement rather than the AOT-$PM_{2.5}$ relationship. These conversion factors have already been used by some researchers with currently-available AOT datasets such as those from MODIS and MISR. For example, Evans *et al.* (2013) assessed the health effects of $PM_{2.5}$ pollution estimated from MODIS AOT using the van Donkelaar *et al.* (2010) conversion factors, and found that 7.1% of total global mortality was attributable to anthropogenically-derived $PM_{2.5}$ – a higher estimate than that derived through the Global Burden of Disease report (Lim *et al.*, 2012) – and Dey *et al.* (2012) assessed the

spatial variability in particulate matter concentration over the Indian subcontinent, finding that 51% of the population were routinely exposed to particulate matter pollution over the World Health Organisation upper limit of 35 µg/m^3.

Several researchers using AOT data to monitor PM have raised issues with the low resolutions of the data available. Hoff and Christopher (2009) state that "Health exposure needs tend to focus on spatial scales ranging from 1 to 100m...neither surface monitoring or satellite measurements can deal with that spatial scale", Martin (2008) lists the development of higher-spatial resolution products, particularly for use in urban areas, as a key recommendation for future work, and Gupta *et al.* (2006) state that their work in cities was limited by the low resolution of the MODIS data that was available. A higher resolution of AOT data is required to allow widespread use of satellite-derived air quality and particulate matter data for human health applications. This is particularly the case in cities, where very high resolutions are required to monitor the dynamics of the particulate matter distribution (Loughner *et al.*, 2007)

Higher resolution will bring other benefits as well as the ability to resolve patterns of pollution in greater detail. Chudnovsky *et al.* (2013) found that there was a significant positive relationship between the AOT data resolution and the fit of the linear regression between AOT and $PM_{2.5}$, with a R^2 of 0.46 with 1km AOT data, compared to 0.18 with 10km AOT data. This was thought to be caused by the lower-resolution data 'hiding the variability' in the AOT data that was captured by the network of ground-based PM monitoring stations. The highest resolution AOT data used in the study was 1km, but given the reasoning for the increase in R^2 values, it would seem logical to suggest that this relationship may continue to hold for higher resolutions of AOT data. Furthermore, as discussed in §3.4.4, there are less problems with cloud contamination for high resolution data (Henderson and Chylek, 2005), both increasing the coverage (Chudnovsky *et al.*, 2013, found an increase in data availability of 50–70% due to fewer pixels being masked as cloud) and reducing the error (due to false negatives in the cloud screening process).

3.7.2 Atmospheric correction of satellite images

The majority of the scientific use of satellite data is now quantitative - that is, it is based upon the real physical values of light intensity measured by the satellite (usually in units of $\mathrm{W\,m^{-2}\,sr^{-1}\,nm^{-1}}$). However, the light reaching the satellite undergoes atmospheric radiative transfer twice (on the path from the sun to the ground and again from the ground to the sensor; see §3.2) and this can significantly affect the values recorded at the satellite. Atmospheric correction is the process of removing these atmospheric effects from the data. There are multiple methods for doing this including image-based methods (such as Dark Object Subtraction; Chavez, 1988) and empirical methods (such as the Empirical Line Method; Smith and Milton, 1999), but the most popular atmospheric correction methods use Radiative Transfer Models (RTMs).

RTMs simulate the physics of light passing through the atmosphere to model the

atmospheric effect on the satellite image, and then remove it. RTMs which are designed for working with satellite data include 6S (Vermote *et al.*, 1997) and MODTRAN (Berk *et al.*, 1987), which have been used in end-user-focused atmospheric correction software such as ATCOR (Richter, 2004). To accurately correct an image the RTM must be parameterised to represent the atmosphere at the time of image acquisition, and this requires atmospheric data including AOT, aerosol type and precipitable water vapour content. Assuming these parameters are accurately specified, the atmospheric correction performed by RTMs is more accurate than those performed by other methods.

Most current atmospheric correction software assumes a constant value of AOT across the whole image, but this is unlikely to be true. AOT can vary on very small scales, and thus significant differences in AOT are likely to be found between pixels in the same image, particularly if the image contains a variety of land uses (for example, urban/industrial compared to rural areas). Furthermore, the increases in swath sizes found with modern sensors (such as MODIS, MERIS and the high-resolution sensors on the satellites within the Disaster Monitoring Constellation) make it more likely that significant changes in AOT will be observed across an image, even if land uses do not change significantly. If a constant AOT cannot be assumed, then an AOT value must be provided separately for each pixel in the image. This could be produced from any high-resolution AOT dataset, but ideally should be produced from the image data itself (thus ensuring that the AOTs used for correction are the AOTs at exactly the time of image acquisition). For example, the MODIS atmospheric correction algorithm solves this problem by using the AOT data produced from the same MODIS image using the MODIS AOT algorithm described in §3.4. However, this approach is impossible to use with images from high-resolution sensors, as these sensors cannot produce fully spatially-explicit AOT data and therefore the only AOT data which is available is of a lower resolution than the imagery to be corrected.

As discussed in Chapter 2, satellite data are now used to estimate a range of physical measurements such as Leaf Area Index and albedo, and poor atmospheric correction can lead to significant errors in such products. For example, Saleska *et al.* (2007) stated that the Amazon rainforest was more resilient to short-term climatic fluctuations than previously thought (as shown by a significant increase in Enhanced Vegetation Index during a period of drought), but Samanta *et al.* (2010) showed that these inferences were due to the use of cloud- and aerosol-contaminated satellite data in the original study, and therefore were not genuine ground-level effects.

3.7.3 Other applications

Satellite-derived AOT measurements have a wide range of other applications, some of which are listed below. The discussion focuses on the utility of high-resolution AOT data.

- **Climate modelling:** As discussed in §3.1.2, the presence of aerosols in the atmosphere changes the Earth's energy balance by both direct effects, such as the scattering and absorption of light by aerosols, and indirect effects such as increased

cloudiness caused by aerosols promoting the formation of clouds. However, our understanding of the feedbacks between aerosol content, cloud cover and, ultimately, the energy balance is limited, and uncertainties surrounding the effect of atmospheric aerosols are considered to be some of the largest uncertainties in current climate modelling (Stevens and Boucher, 2012; Lohmann and Feichter, 2005; Le Treut, 2012). However, improvements in satellite remote sensing of aerosols have helped reduce this uncertainty: the most recent IPCC report stated "[aerosol] forcings are now better understood than at the time of the Third Assessment Report due to improved in situ, satellite and ground-based measurements and more comprehensive modelling, but remain the dominant uncertainty in radiative forcing." (IPCC, 2007). Further improvements in aerosol data from satellites are required to reduce these uncertainties further, and although Mishchenko *et al.* (2004) state that the spatial resolution requirement for aerosol measurements to quantify aerosol forcing in the climate system is 6km, higher resolution measurements will still be very useful. For example, higher resolution measurements are less likely to be contaminated by cloud cover, are easier to validate with ground measurements and can be aggregated for comparison with lower resolution products.

- **Validating aerosol transport models:** Many of the aerosol transport models described in §3.6 are designed to produce raster images showing the distribution of aerosols at the end of the model run. These models need validating, but validating spatially-distributed outputs from models against point ground measurements is difficult; satellite-derived AOT measurements provide a similar dataset allowing comparison (Chin *et al.*, 2002).
- **Fire/Smoke monitoring:** Satellite sensors are widely used for monitoring wildfires on a global basis (Lentile *et al.*, 2006) for a variety of purposes, including emergency planning and assessment of the impact of biomass burning on climate change. However, fires do not just damage the land surface local to the fire; they also produce smoke which can travel significant distances, spreading the impact of the fire. Smoke consists of many particles (mostly of soot) suspended in the atmosphere which can be seen in AOT images, and these can be used to monitor the distribution of smoke. For example, Damoah *et al.* (2004) monitored the spread of smoke plumes from forest fires in southeast Russia using MODIS, SeaWIFS and TOMS data and found that the smoke spread widely across the Northern hemisphere over a period of around 17 days. High-resolution data allow observation of smoke plumes from smaller fires which would not be visible on lower-resolution imagery (see Figure 3.14).
- **Volcanic ash distribution monitoring:** Several volcanic eruptions recently have released large volumes of ash into the atmosphere which have become distributed across wide areas and caused significant problems for air transport, often resulting in the closure of airspace (Wainwright, 2010). It is important to monitor the distribution of this ash at a high-resolution to ensure that public safety is ensured by closing airspace when needed, but that inconvenience is minimised by not closing

airspace that has insignificant levels of ash within it. Ash transport models are often used for this purpose, but they require input data regarding current airborne ash concentrations, and also require validation. AOT measurements from various satellites including MODIS and MISR have been widely used for these purposes, for example by the Met Office in the UK (Christopher *et al.*, 2012) as well as by organisations in Germany (Langmann *et al.*, 2012) and Spain (Toledano *et al.*, 2012).

- **Dust monitoring/modelling:** Measurements of AOT are of great importance to those studying dust - both in terms of the general distribution of dust from the world's source regions, and for modelling and predicting extreme dust events. Similarly to the other uses described above, AOT data are used to provide information on dust events (Kaufman *et al.*, 2005; Prasad and Singh, 2007; Gautam *et al.*, 2009), validate models (Tegen and Fung, 1994) and examine the effects of dust in the atmosphere (Zhang and Christopher, 2003; Han *et al.*, 2008). Schepanski *et al.* (2007) created a map of dust activation frequency in the Sahara using a combination of thermal data and AOT data from SEVIRI, but their results were at a low resolution (one degree). Similarly, Ashpole and Washington (2012) used SEVIRI data to track dust plumes across the Sahara, to investigate the spatial and temporal patterns of dust emission and transport High-resolution data will allow the validation of similarly high-resolution models, and provide greater detail on dust events - for example, allowing far more detailed examination of dust sources in the Sahara to be performed, and allowing identification of the part of a salt pan from which dust is being generated, rather than identifying the salt pan as a whole.

3.8 Synthesis and identification of research gaps

The wide range of uses of AOT data, discussed in §3.7, show its importance in many fields of research including air quality monitoring, satellite atmospheric correction and climate change modelling. Currently, AOT data are provided by measurements from sun photometers (§3.3), and retrievals from a range of satellites (§3.4 and Table 3.3). However, these measurements are not suitable for a number of potential applications due to their low spatial resolution. For example, Martin (2008) states that higher resolution AOT data are needed to allow useful analysis of air quality due to particulate matter pollution, particularly in urban areas, and Hoff and Christopher (2009) state that AOT measurements with a spatial resolution of 1–100m are needed to properly characterise the health effects of particulate matter exposure. Furthermore, high-resolution AOT data are needed to enable accurate per-pixel atmospheric corrections of the wide range of high-resolution satellite data which is now widely available from satellites such as Landsat, SPOT, IKONOS and DMC, as well as data from future satellites such as the Pléiades constellation and the Sentinel series of satellites.

The data that is currently available cannot fulfil these needs. AERONET has only 300 sites worldwide, and thus cannot provide high-resolution data, and most satellite AOT retrievals are at a resolution of around 10km. Recently some algorithms have been

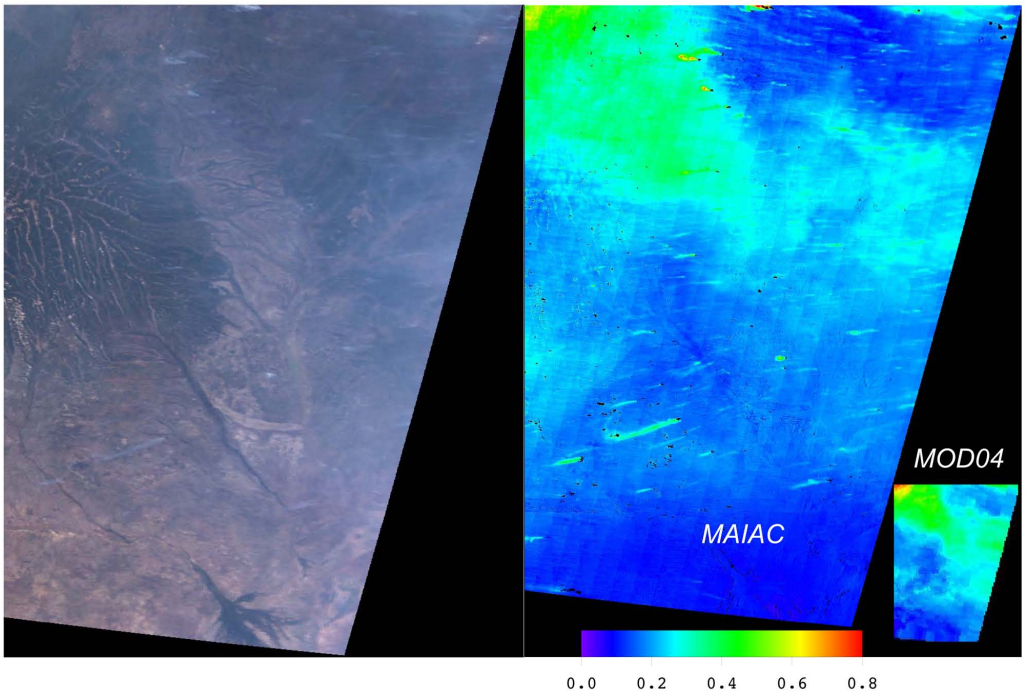


FIGURE 3.14: Left: MODIS RGB composite over South Africa during the biomass burning season. Right inset: MOD04 AOT product at 10km resolution, Right main: MAIAC AOT product at 1km resolution. Note how individual smoke plumes are visible at 1km resolution but not at 10km resolution. Figure from Lyapustin *et al.* (2011)

developed to allow higher resolution AOT retrieval (such as SYNAM and MAIAC which have been used to produce 1km data), but these products are not yet operationally produced and are time-consuming, computationally-expensive and difficult for end-users to produce.

Thus, there is a requirement for a high-resolution AOT product, and there is a gap in the research literature regarding methods which could be used to provide this product. Many of the methods described in §3.4 cannot be extended to produce high-resolution data as no satellites currently produce high-resolution multi-angular or polarisation data, and Landsat is the only high-resolution satellite which acquires measurements in shortwave infra-red bands. Building and launching high-resolution satellites to acquire these measurements would probably be feasible, but they are likely to be very expensive and have other limitations (for example, long revisit periods). Instead, developing a method which could retrieve AOT by using the visible and near-infrared bands would allow the method to be used on a wide-range of satellite images, including Landsat, SPOT and IKONOS, as well future satellites such as the Pléiades constellation and the Sentinel series.

Although the retrieval of high-resolution AOT data was described as a research gap above, there are several methods which retrieve relative measures of the haziness of the atmosphere from a wide range of satellite images. These haze assessment methods (§3.5) are generally used in the pre-processing of satellite images, but the relative measures of haziness they produce have not yet been linked to absolute measurements such as AOT.

Thus, there is a opportunity to fill this research gap by developing an algorithm to produce high-resolution AOT data by linking haze assessment methods (§3.5) with standard AOT retrieval methods (§3.4). The development and validation of such an algorithm will be the focus of the rest of this thesis.

Chapter 4

Investigating the Haze Optimized Transform

4.1 Introduction

In the literature review, a gap in the literature was identified, with an opportunity to fill this gap by developing an algorithm to produce high-resolution AOT data by linking haze assessment methods with standard AOT retrieval methods. The first stage in this process is to perform a detailed investigation of a haze assessment method to test that it performs as expected, check its assumptions, and identify any major issues.

The literature review showed clearly that the Haze Optimized Transform (HOT; Zhang *et al.*, 2002a) is the most advanced of the haze assessment methods: it is based on the same underlying theory as the Tasselled Cap Transformation and the NDHI, but is more robust to land-cover changes. The HOT was developed for the purpose of identifying, quantifying and removing haze and thin cloud contamination from satellite images and is based on the high correlation between the visible bands of satellite sensor data which defines a 'clear line' in feature-space (as shown in Figure 4.1). Pixels that are contaminated by haze will be offset from this line, and the HOT calculates the distance of each pixel from the 'clear line' and uses this as an estimation of the haziness of this pixel. A full discussion of the theoretical basis of the HOT, and related techniques, is given in §3.5.2.

This chapter discusses the datasets required to investigate the HOT, tests the key findings of Zhang *et al.* (2002a) and assesses the accuracy of the assumptions underlying the HOT.

4.2 Data

4.2.1 What data do we need?

The spectral points on the graph from the original HOT paper, shown in Figure 4.1, were generated from 4m-resolution PROBE-1 measurements of representative Canadian land covers, averaged to a resolution of 30m and passed through the MODTRAN Radiative

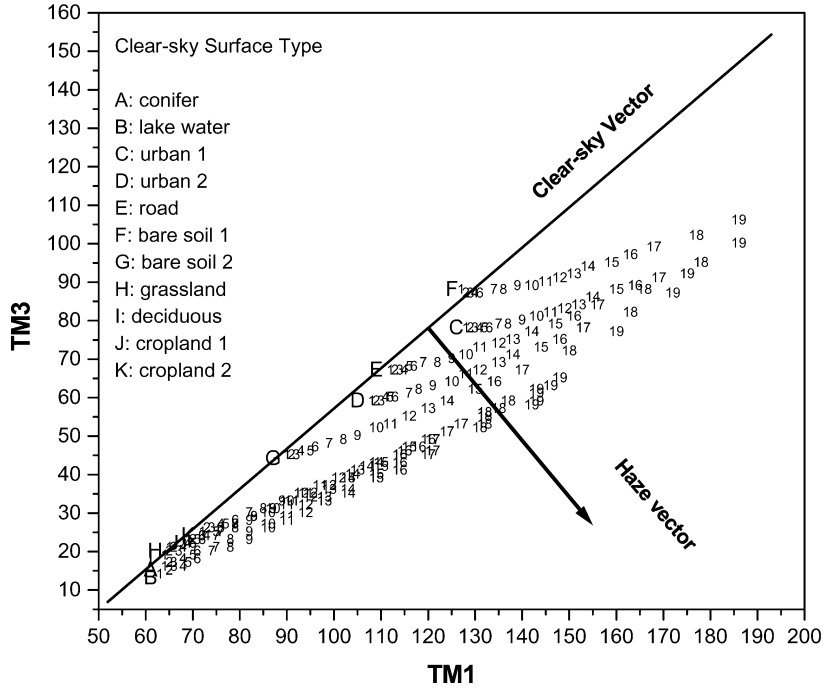


FIGURE 4.1: The basis of the Haze Optimized Transform (HOT), from Zhang *et al.* (2002a), using Landsat TM bands 1 (Blue) and 3 (Red)

Transfer Model to simulate the top of atmosphere (TOA) radiance in a clean atmosphere. However, to investigate the HOT in detail we need a set of realistic reflectance measurements at medium spatial resolution (20–30m) and high spectral resolution, with no atmospheric effects present. This will allow us to simulate the TOA radiance for a range of satellite sensors under various atmospheric conditions to test various aspects of the HOT. It is important to note that the any real-world data used must have been atmospherically-corrected on a per-pixel basis, or there will be residual atmospheric effects of varying magnitudes present in the data.

Furthermore, reflectance can be measured in many ways, producing measurements such as the Hemispherical-Conical Reflectance Factor (HCRF), the Bi-directional Reflectance Factor (BRF) and the Bi-hemispherical Reflectance Factor (BHRF), depending on the illumination and measurement geometries (Schaepman-Strub *et al.*, 2006). Almost all real-world measurements are of the Hemispherical-Conical Reflectance Factor (HCRF) so these measurements will be assumed throughout the rest of the thesis. It is possible that differences in the geometry of the conical field of view of the measurement instrument could cause problems - as these can vary significantly - but the majority of the work in this thesis uses Landsat which (due to its relatively narrow swath width) has a relatively constant field of view.

It was found to be impossible to create any dataset that fulfilled all of the requirements listed above, so three datasets have been created, which are described in the sections below and compared in Table 4.1.

TABLE 4.1: Comparison of datasets

Dataset	Realistic	Resolution	Atmospheric Correction	Spectral resolution	Range of land covers
Resampled airborne	Yes, real-world data	Any resolution > 1m	Uniform atmospheric correction performed over small image, possible per-pixel residual atmospheric effects.	Hyperspectral	Limited: Agriculture, Low-density housing.
LEDAPS-corrected Landsat images	Yes, real Landsat images	30m	LEDAPS atmospheric correction procedure (Wolfe *et al.*, 2004), pseudo-per-pixel	Landsat TM or ETM+ Bands	Good: Any land cover available
Landsat pixels over AERONET sites	Yes, real Landsat pixels	30m	Accurately performed using AOT from AERONET site within pixel	Landsat TM or ETM+ bands	Good: Most land cover categories covered

4.2.2 Resampled airborne data

A subset of an airborne hyperspectral image acquired by the Airborne Research and Survey Facility (ARSF) AISA Eagle sensor during the NCAVEO Field Campaign (Milton *et al.*, 2011) was resampled to a range of spatial resolutions from an initial resolution of 1m to create a simulated Landsat image. This image was chosen due to the range of coincident ancilliary data acquired during the field campaign, as well as the range of land-covers present within a small image subset.

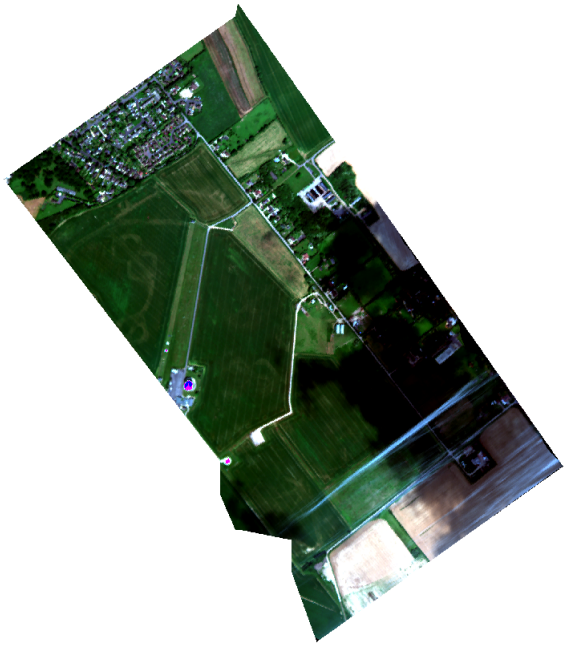
The airborne image was atmospherically corrected with ATCOR-4 (Richter, 2004) using a single parameterisation across the whole image, and thus there will be spatially-variable residual atmospheric effects in the corrected image. These effects will be present, but should be of low magnitude as the image subset is very small. The original 1m image and the resampled 30m image are shown in Figure 4.2. This dataset will be referred to henceforth as the *ARSF dataset*.

4.2.3 LEDAPS-corrected Landsat images

The Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS; Wolfe *et al.*, 2004) provides the closest to a per-pixel atmospheric correction of Landsat imagery that is available at present (August 2014). AOT is estimated above areas of dense dark vegetation (DDV) in the image, and then interpolated at a 1km resolution to cover the rest of the image, and each pixel is then corrected using its own interpolated AOT. DDV is selected based upon a simple threshold of pixels with a top of atmosphere reflectance in Landsat band 7 (approximately 2.08–2.35 μm) < 0.15 (Kaufman *et al.*, 1997; Ju *et al.*, 2012). This method relies on the presence of DDV in the image: thus it provides far better results over the Amazon rainforest than over the Sahara desert.

Assessment of LEDAPS surface reflectance products against similar MODIS products and field spectrometer measurements, as well as comparisons of LEDAPS DDV-derived AOT against AERONET measurements, show that LEDAPS does not always produce high quality corrections (Maiersperger *et al.*, 2013). In general, LEDAPS-estimated AOT were consistent with AERONET observations (correlation between 0.63 and 0.87), although in areas with limited DDV the AOT was significantly over-estimated (with errors up to 0.6). Comparisons between LEDAPS surface reflectance and field spectrometer measurements showed good results over a vegetation site surrounded by DDV (with errors within ± 0.03), but very poor results in arid regions (with errors up to 0.12). In general, results were better in the longer wavelengths, where the aerosol effects are less significant.

The use of interpolation to estimate AOT over non-DDV pixels can cause significant errors in certain situations. For example, there is often limited DDV within urban areas, so AOT values from DDV in the surrounding countryside are interpolated to produce values for the urban region, thus producing erroneously high or low values. A simple examination of this problem was carried out by comparing ground-based and LEDAPS-derived AOT values at a location outside of the Geography building on



(a) Original ARSF

(b) ARSF resampled to 30m spatial resolution

FIGURE 4.2: ARSF images

FIGURE 4.3: Example Landsat 7 image (LE72020252014184ASN00) over Southampton, showing the measurement location (red cross), and the closest areas of DDV (blue crosses), between which interpolation is performed. DDV was detected using the LEDAPS threshold described in the text, and the yellow cross shows Southampton Common which, although it is vegetated, was not detected as DDV.

Highfield Campus at the University of Southampton (50.925°, -1.414°) during the Landsat ETM+ overpass at 10:56 on 3rd July 2014. As shown in Figure 4.3, there are significant areas of DDV outside the urban area, but none within the city itself. Thus, the LEDAPS AOT is interpolated, producing an AOT value at the measurement site of 0.147, which is almost three times the ground-based measurement of 0.0596, taken contemporaneously with the Landsat overpass using a Microtops sun photometer.

The main advantage of this dataset is that it consists of genuine Landsat images, with a real-world mixture of spatial variability and spatial structure, and that these images are available both uncorrected and corrected with LEDAPS. There are significant issues - including potentially low accuracy over desert areas, urban areas, and other areas lacking DDV, and higher errors in the wavelengths which are key to this work (blue and red). However, this is the best spatially-variable atmospherically-corrected Landsat dataset available, so it will be used with its known issues taken into account.

LEDAPS images, along with the uncorrected original images, were acquired as listed in Table 4.2, with the aim of covering a broad range of global land cover types. This dataset will be referred to henceforth as the *LEDAPS dataset*.

TABLE 4.2: Landsat scenes used in this project. Both uncorrected and LEDAPS-corrected copies of the images were used.

Scene ID	**Name**	**Land covers**
LE72270602001259EDC00	Amazon	Forest
LE71720741999285EDC00	Botswana	Scrub & Salt Pan
LE72010242000171EDC00	London	Urban, Agriculture & Mixed Vegetation
LE72020252002087EDC00	SotonNewForest	Urban, Agriculture & Mixed Vegetation
LE72020241999287AGS00	Chilbolton	Urban, Agriculture & Mixed Vegetation
LE71790772000097EDC01	Namibia	Desert
LE71620462000186SGS00	Saudi Arabia	Desert
LE71990461999298EDC00	Sahara	Desert
LE70240321999240EDC00	Midwest	Agriculture & Urban

4.2.4 Atmospherically-corrected Landsat pixels

It is impossible to atmospherically-correct a whole Landsat image on a per-pixel basis, as no 30m resolution AOT products currently exist, so this dataset was created from Landsat pixels directly over many AERONET measurement sites where the atmospheric conditions are accurately known. The procedure for generating this dataset, known as *LandsatAERONET* is shown in Figure 4.4.

Landsat TM acquisitions from 2009 were used, as this was the most recent year during which the TM sensor had no significant problems, and pixels in a window centred on each AERONET site location were selected. The atmospheric correction was performed as accurately as possible, using the 6S Radiative Transfer Model (Vermote *et al.*, 1997) through the Py6S interface (Wilson, 2012) with AOT and Precipitable Water Content (PWC) data from AERONET measurements (Holben *et al.*, 1998), the aerosol profile from AERONET inversions (Dubovik and King, 2000) and ozone contents estimated using van Heuklon (1979) model.

The most accurately corrected pixel will be the pixel directly over each AERONET site, but using just these measurements significantly restricts the available data volume. Thus a configurable $n \times n$ window was used to select pixels surrounding the AERONET site, with $n = 9$ used as a good compromise between data volume and distance from the AERONET site (providing 81 pixels, including the central pixel, with a maximum distance from the site of 150m). The locations of each pixel are kept in the output dataset, allowing the dataset to be easily reduced to the equivalent dataset at any lower n value (for example $n = 1$ for just those pixels directly over each AERONET site).

In 2009 there were 291 AERONET sites operating worldwide, and of these 165 had a Landsat image over them within 15 minutes of AERONET measurement. Of these, 139 had all of the ancillary data needed for atmospheric correction, and after manual cloud screening, fully-corrected pixels over 124 separate sites were left (Figure 4.5). Taking into

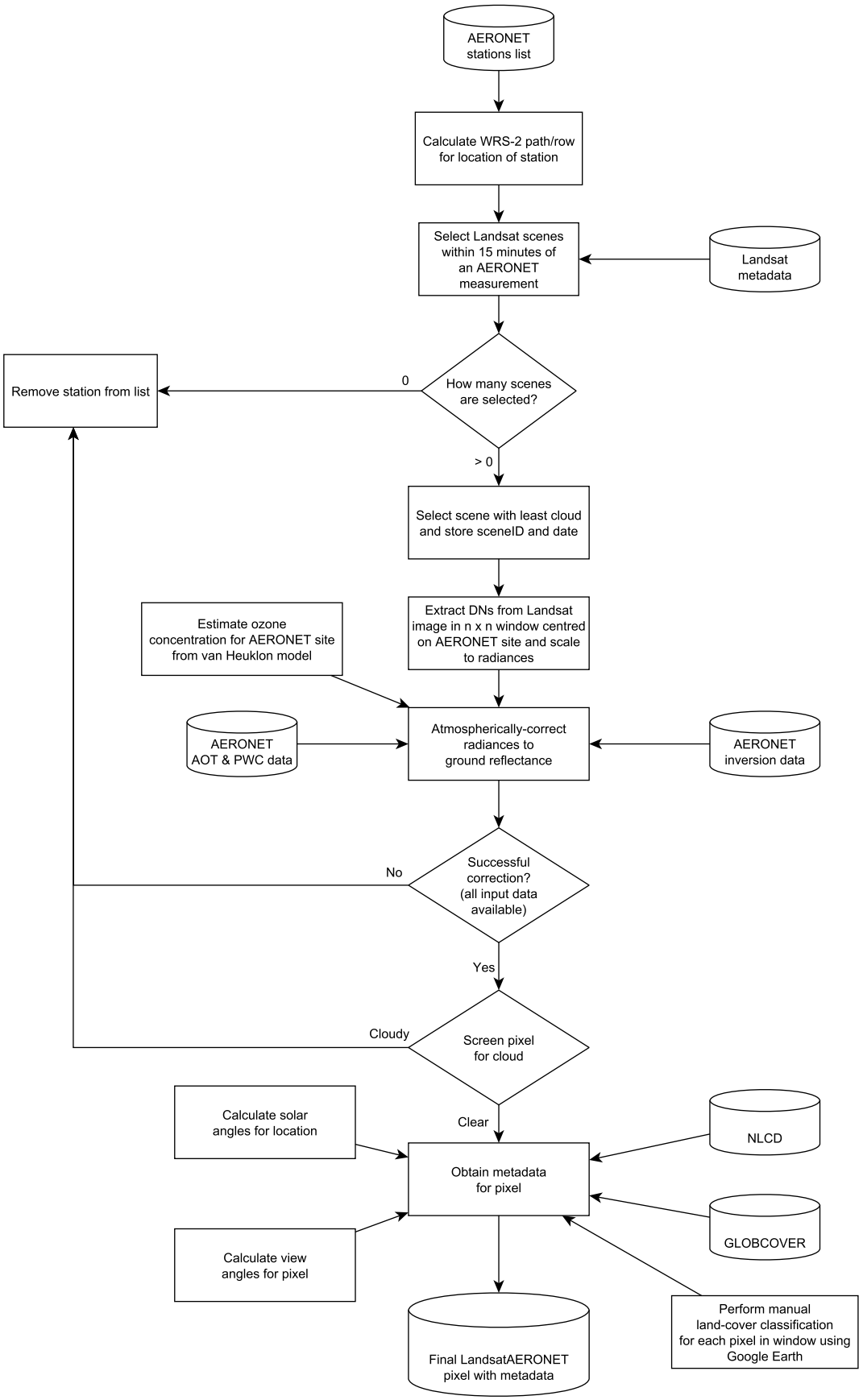


FIGURE 4.4: Flowchart showing the procedure for generating the LandsatAERONET dataset

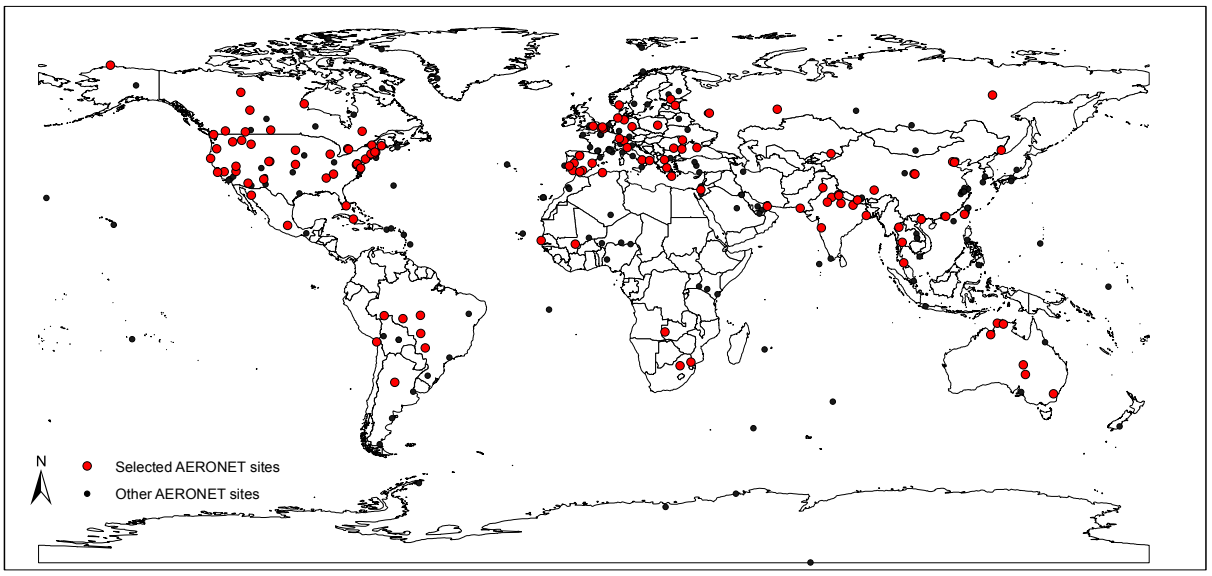


FIGURE 4.5: Locations of AERONET sites: the 127 selected sites are shown in red, other sites in grey.

account 80 surrounding pixels (for the $n = 9$ case), this gives a total of 10,044 pixels.

Land cover information was acquired for each of the corrected pixels from two land cover maps: GLOBCOVER (global at 300m spatial resolution, 127 AERONET sites covered; Arino *et al.*, 2007) and the US National Land Cover Database (NLCD, USA at 30m spatial resolution, 30 AERONET sites covered; Fry *et al.*, 2011). The GLOBCOVER data at 300m resolution was found not to be representative of the 30m subpixel land cover, and so the NLCD data at 30m resolution was more suitable. However, only 30 sites had NLCD data: too few to assess the effect of land cover on the HOT. Thus the land cover over all sites was manually classified using images from Google Earth acquired as close as possible to the relevant Landsat scenes. The NLCD legend (Table 4.3) was used as the classes are very well described, and this allows a validation of manually classified land cover against the NLCD data.

The percentage of NLCD pixels which were classified into the same class by the manual classification procedure was low, at only 50%. However, it is very difficult to distinguish some of the NLCD classes: for example, the difference between *Developed (Medium Intensity)* and *Developed (High Intensity)* is difficult to assess by eye, and *Deciduous Forest* and *Evergreen Forest* can be hard to distinguish in certain situations. The pixels were then classified into four broad land cover classes: Water/Snow/Ice, Bare, Urban and Vegetation, and 84% of the NLCD pixels were manually classified into the same class, a significant improvement from the individual NLCD classes.

The full list of data collected for each selected pixel is shown in Table 4.4.

The same process was carried out for Landsat ETM+ data, using data acquired before the Scan Line Corrector (SLC) failure (1999–2003), and, for AERONET sites located in the central portion of images, data acquired since the failure (as the central part of the image

TABLE 4.3: National Land Cover Database classes, and equivalent broad land cover classes. Further details, including full class descriptions, are available from Fry *et al.* (2011)

NLCD Class	**Broad Class**
Open Water	Water, Snow & Ice
Perennial Ice/Snow	Water, Snow & Ice
Developed, Open Space	Urban
Developed, Low Intensity	Urban
Developed, Medium Intensity	Urban
Developed, High Intensity	Urban
Barren Land (Rock/Sand/Clay)	Bare
Deciduous Forest	Vegetation
Evergreen Forest	Vegetation
Mixed Forest	Vegetation
Shrub/Scrub	Vegetation
Grassland/Herbaceous	Vegetation
Pasture/Hay	Vegetation
Cultivated Crops	Vegetation
Woody Wetlands	Vegetation
Emergent Herbaceous Wetlands	Vegetation

is not affected by the SLC failure). These datasets will be referred to henceforth as the *LandsatAERONET TM dataset* and *LandsatAERONET ETM+ dataset*.

4.2.5 Data comparison

TOA radiance data for Landsat TM bands 1 (Blue) and 3 (Red) were simulated in a clean atmosphere (that is, without aerosols) from the reflectance data in the Landsat-AERONET TM and resampled ARSF datasets as well as a LEDAPS-corrected Landsat image. The results (Figure 4.6) show that the data fall into the same region within the blue-red feature space, suggesting that all of the datasets are representing the same phenomena - in this case, atmospherically-corrected Landsat pixels. The resampled ARSF data are contained within the boundaries of the distribution of the Landsat-AERONET data, and the latter has a significantly larger range in both bands. This is because the Landsat-AERONET data cover a far wider range of land covers than the small ARSF image. The LEDAPS image has an even wider range, as each of the 38 million pixels in the LEDAPS image will be a unique mix of subtly different land covers.

TABLE 4.4: Data collected for each pixel in the LandsatAERONET dataset

Data
AERONET site name
Latitude
Longitude
Elevation
Landsat WRS-2 Path/Row
Landsat Scene ID
Image acquisition time
Aerosol Optical Thickness at acquisition time
Precipitable Water Content at acquisition time
Ozone amount (estimated from the van Heuklon (1979) model)
Atmospherically-corrected reflectances in bands 1–7 (excluding band 6)
Solar zenith angle
Solar azimuth angle
View zenith angle (calculated according to the procedure in Roy *et al.*, 2011)
View azimuth angle (calculated according to the procedure in Roy *et al.*, 2011)
Land Cover classification from NLCD
Land Cover classified manually
Simulated TOA radiance in a clean atmosphere for bands 1–7 (excluding band 6)

4.3 Validating HOT assumptions

4.3.1 Correlation

One of the main assumptions of the Haze Optimized Transform is that the visible bands are highly correlated, and thus the pixels in an image will cluster around a line in the visible feature-space. The strong positive correlation coefficients between the visible bands in the LandsatAERONET and resampled ARSF datasets (Table 4.5) show that this is the case.

Liu *et al.* (2011) state that the assumption of high correlation that underlies the Haze Optimized Transform method is incorrect, as their test of 23 hazy TM images found that the majority were < 0.9, a significant number were < 0.8 and some were as low as 0.4. However, these tests were performed on hazy images and the Haze Optimized Transform is fundamentally based upon the fact that as pixels become hazy they migrate away from the Clear Line and thus reduce the correlation. Thus, correlations on hazy images would be expected to be low, but correlations on atmospherically-corrected images would be higher. This relationship is shown in Table 4.6, where the majority of the correlations for LEDAPS-corrected images are higher than their part-corrected counterparts. The lowest correlations tend to be for images with a mixed land cover including significant amounts

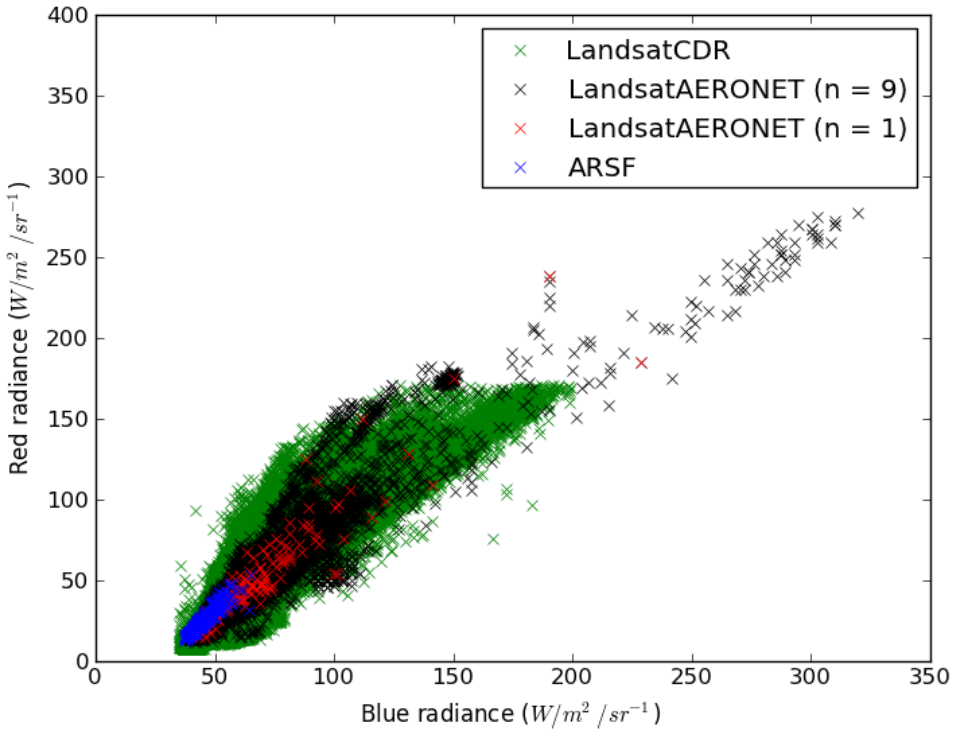


FIGURE 4.6: Comparison of Landsat at-sensor radiance simulated from the resampled ARSF and Landsat-AERONET TM datasets, shown with at-sensor radiance simulated from a LEDAPS-corrected Landsat image.

TABLE 4.5: Pearson correlation coefficients calculated between the visible bands

(a) LandsatAERONET

	Blue	**Green**	**Red**
Blue	1.00	0.98	0.94
Green		1.00	0.97
Red			1.00

(b) Resampled ARSF

	Blue	**Green**	**Red**
Blue	1.00	0.94	0.94
Green		1.00	0.97
Red			1.00

of bare land: for example, the Saudi Arabia and Namibia images have low correlations, but the Sahara image has a high correlation, as the entire image consists of a single land cover type. Some of the lower correlations are likely to be due to residual atmospheric contamination in the LEDAPS images - caused by poor estimation of the AOT by LEDAPS - which would result in points moving away from the line, and thus a lower correlation.

TABLE 4.6: Blue-Red correlations for all of the Landsat images used in this study, for the LEDAPS-corrected image and a part-corrected image (corrected for Rayleigh scattering, ozone and water vapour absorption, but not aerosols)

Image	**LEDAPS**	**Part-corrected**
Amazon	0.853	0.876
Botswana	0.886	0.887
Chilbolton	0.884	0.853
London	0.930	0.923
Midwest	0.903	0.883
Namibia	0.576	0.612
Sahara	0.958	0.948
Saudi Arabia	0.649	0.567
SotonNewForest	0.931	0.802

4.3.2 What is the trajectory of an AOT increase in feature space?

The Haze Optimized Transform is a measure of the distance of a point from the clear line which represents the correlation between all 'clear points'. For the HOT to have any relationship to AOT, the trajectory of a point which is increasing in AOT must have a different gradient to the clear line.

Figure 4.7 shows how a point (marked in blue) moves in feature space as its AOT is increased from 0 (no aerosols; the blue point) to 0.1, 0.2, 0.3, 0.4 and 0.5. This is as expected according to Zhang *et al.* (2002a), with the points approximately equally spaced for equal increments in AOT, and the trajectory of the points at a shallower angle than the clear line. There is a larger gap between the points for an AOT of 0 (a clean atmosphere) and an AOT of 0.1, as scattering increases significantly when introducing even a small amount of aerosols into a clean atmosphere.

4.4 Issues with the HOT

The HOT is not a perfect measure of the haziness of a pixel, and its value is affected by several other perturbing factors. This section examines some of these factors, and then examines possible methods to reduce the impact of these issues.

4.4.1 Breadth of the point cloud

The breadth of the point cloud of clear pixels in feature space is key to the accuracy and sensitivity of the Haze Optimized Transform. The breadth, in this case, refers to the average offset from the clear line, which is a representation of the range of the data in the axis perpendicular to the major axis of correlation in the data (shown as PC2 in Figure 4.9). The greater the breadth of the distribution of clear points (that is, points with no atmospheric contamination) around the clear line, the larger the average HOT value of a clear point and thus the greater the uncertainty in the AOT estimate from the HOT value. The majority of the issues discussed below are a problem because they cause an increase in the breadth of the point cloud.

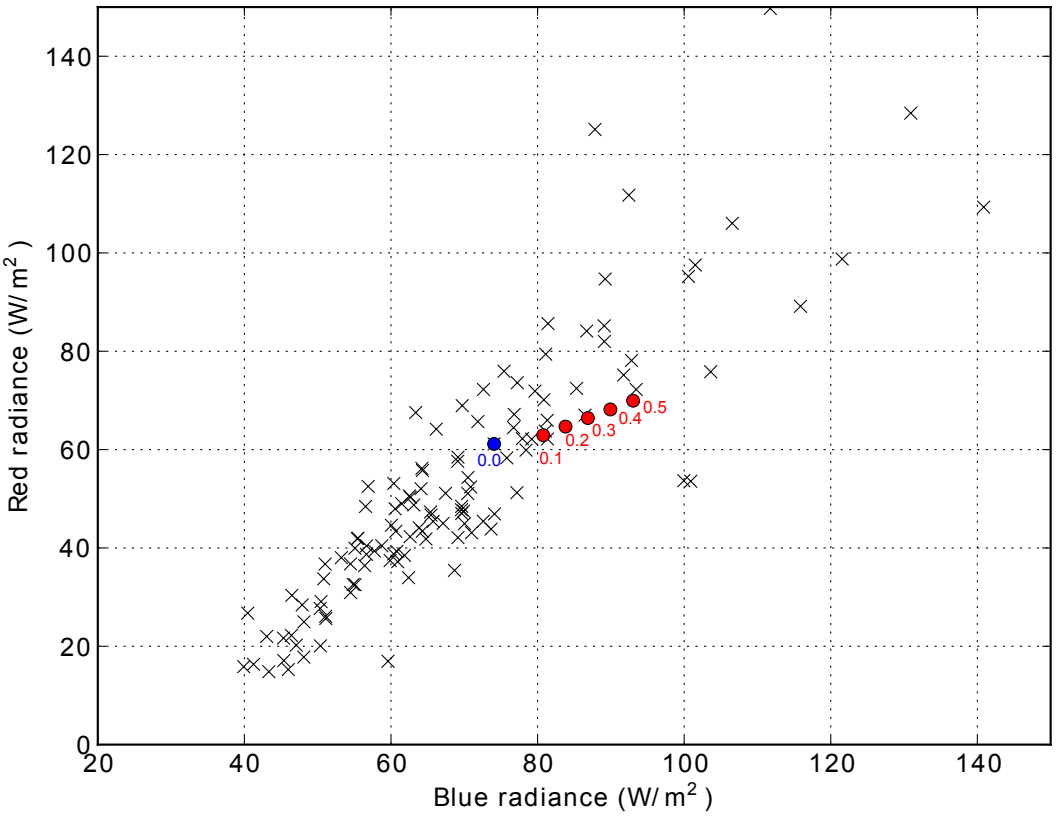
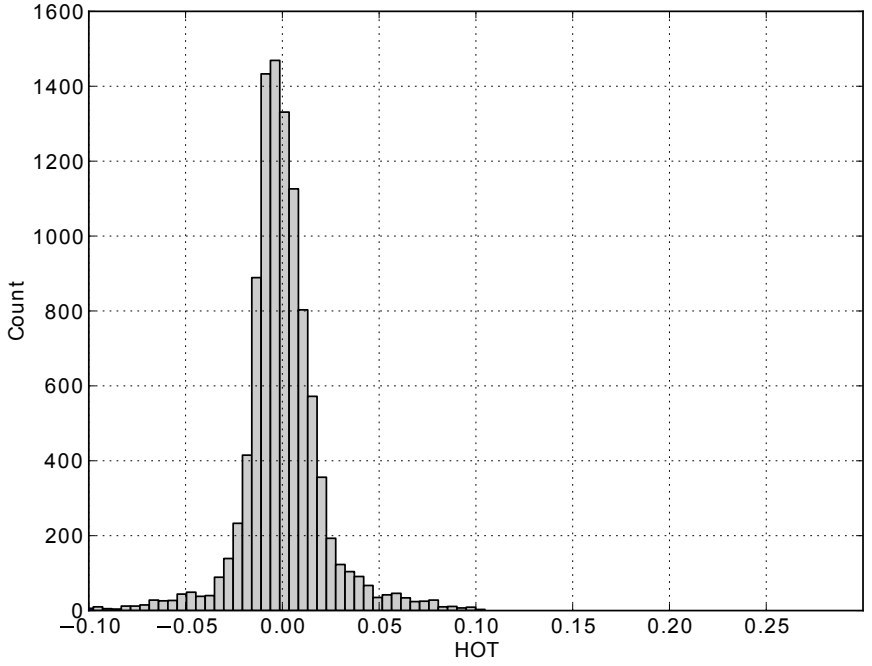


FIGURE 4.7: Trajectory of AOT-perturbed points in feature space. Black crosses show the $n = 1$ subset of the LandsatAERONET TM dataset. The selected point is in blue, and its movement in feature space for AOT increases in steps of 0.1 is shown in red

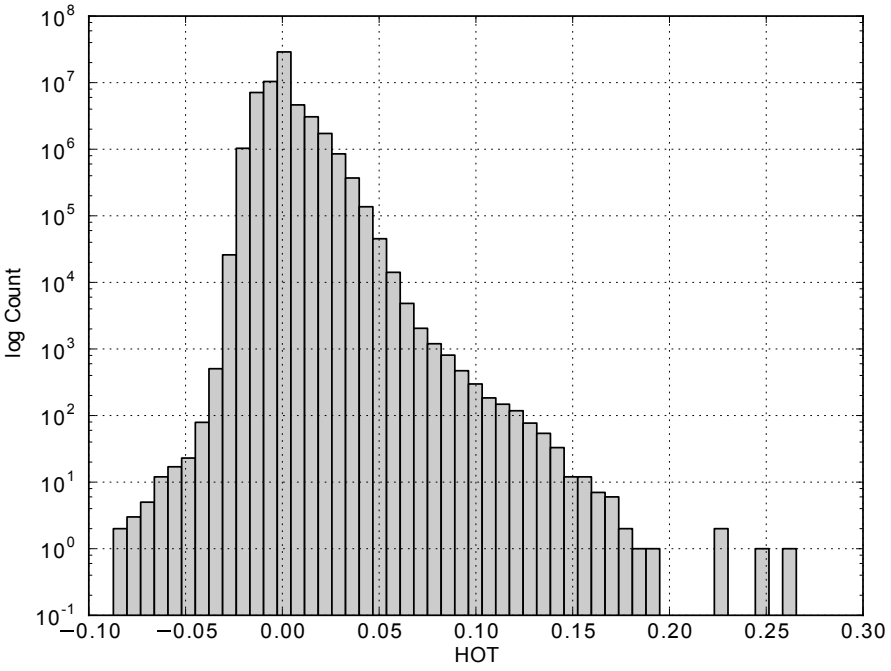
TABLE 4.7: Statistics of the HOT values for each of the two datasets, calculated using a Clear Line calculated from a standard regression of each dataset.

	n	**mean**	**min**	**max**
LandsatAERONET	10 000	0.000	-0.136	0.105
LEDAPS	5.83×10^7	0.000	-0.087	0.266

HOT values for the LandsatAERONET ($n = 9$) and Chilbolton LEDAPS image (chosen as an image with typical mixed urban/rural land cover) were calculated using a linear regression of all points as the Clear Line. Results are shown in Figure 4.8 and Table 4.7. Ideally, all of the HOT values should be zero, as these are all clear pixels. However, this is not the case - with a maximum negative offset of -0.136 and a maximum positive offset of 0.266. The Chilbolton LEDAPS image has significantly larger HOT values than the LandsatAERONET dataset, and this is likely to be due to LEDAPS over- or under-estimating the AOT over parts of the image (as with the Southampton example described in §4.2.3), thus leaving some atmospheric contamination in the images.



(a) LandsatAERONET TM ($n = 9$)



(b) Chilbolton LEDAPS image

FIGURE 4.8: Histograms showing the distribution of HOT values, calculated using a Clear Line calculated from a standard regression of each dataset. Note that the y axis of the LEDAPS histogram shows log values.

4.4.2 What factors control the location of the points in feature space?

To understand what controls the breadth of the point cloud, the factors that control the location of these points in feature space must be understood, so that some of these factors can be controlled to reduce the breadth of the distribution. Due to the range of metadata available, the LandsatAERONET TM dataset will be used to analyse these possible factors. Here the simulated top-of-atmosphere radiance in a clean atmosphere is used, calculated from the atmospherically-corrected reflectance of Landsat pixels directly above AERONET sites, and thus there should not be any per-pixel atmospheric effects left in the data.

There are two key dimensions of the location in feature space in this context: not the primary dimensions of blue and red reflectance, but 1) the location along the clear line ('brightness') and 2) the location perpendicular to this ('deviation'), as shown in Figure 4.9. Values for each pixel on these two new dimensions were calculated using Principal Components Analysis. Linear regression analysis was then carried out to assess the relationship between each principal component and various possible causal factors:

- Land cover according to the manual land cover classification (the variable which, if it were detailed enough and the atmospheric correction were perfect, should account for all of the variation in the dataset)
- AOT, PWC and Ozone amount (in case residual atmospheric effects were still present in the data)
- Solar zenith and azimuth, and view zenith and azimuth (in case of angular effects from either the atmosphere or anisotropic surface reflectance)

Due to the size of the full LandsatAERONET TM dataset ($n = 10044$) almost all of these variables show a significant relationship with the two principal components: with a large data volume, almost any statistical test will produce a significant p-value. Thus, instead of using p-values, the analysis below focuses on the proportion of the variance in the principal components that has been explained by each variable. A threshold has been set, and variables which explain less than 1% of the observed variance are not discussed below.

The following variables explained $> 1\%$ of the observed variance in the first principal component ('brightness'), thus affecting the position of a pixel along the clear line:

- **Land cover:** As expected, this has a significant relationship with the brightness, explaining 57% of the observed variance in brightness.
- **Precipitable Water Content:** This explains 2.8% of the variance in pixel brightness: the lower the precipitable water content, the brighter the pixel. Although poor atmospheric correction may be the cause of a relationship between PWC and brightness, it seems that this is actually caused by the relationship between land cover and climate. The brightest pixels in the LandsatAERONET dataset are those in desert regions, where there tends to be a low precipitable water content, and thus this falsely appears as a relationship between precipitable water

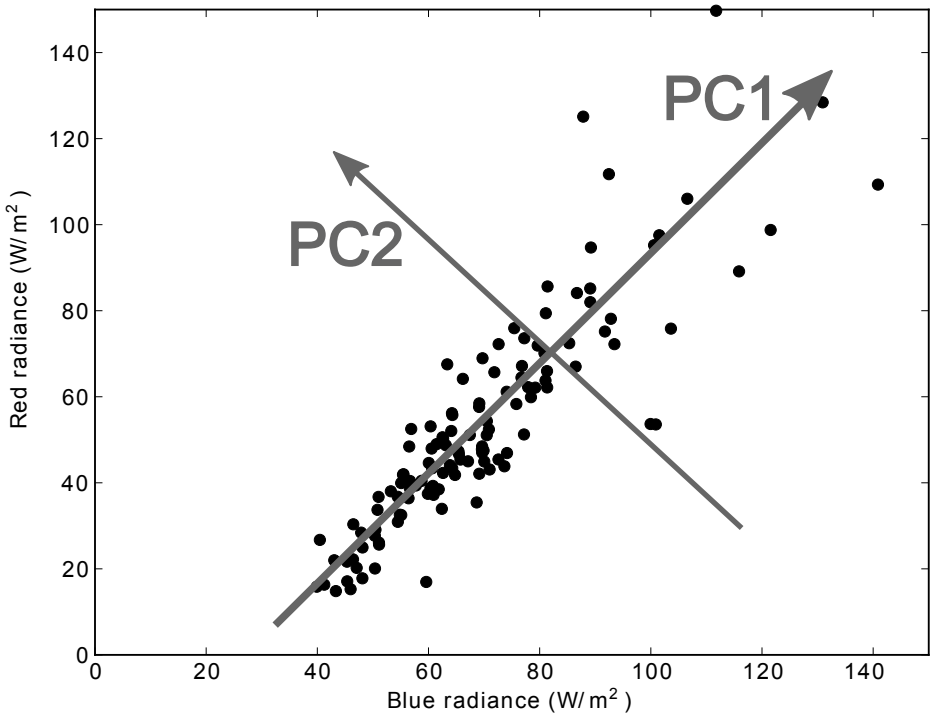


FIGURE 4.9: Diagram showing the two principal components, and their relationship to the LandsatAERONET dataset. PC1 represents brightness, and PC2 represents deviation. Note the directionality of the components, with the arrows pointing in the direction of increasing principal component values.

content and brightness.

- **Solar zenith angle:** This explains 1.7% of the variance in the pixel brightness, but this relationship is entirely dependent on a few outliers, and the relationship loses significance once the outliers are removed.

The following variables have a significant relationship with the second principal component ('deviation'), which - due to the directionality of the PCA transform - can also be thought of as 'excessive redishness'. This is the more important of the two principal components as it is highly correlated with the Haze Optimized Transform values and controls the breadth of the point cloud:

- **Land cover:** Again, land cover has a very significant relationship with the deviation from the clear line, explaining 36% of the variance of the deviation. In general, this means that some land cover types are located further away from the line than others, as suggested by Zhang *et al.* (2002a).
- **Solar zenith angle:** There is a negative relationship between the solar zenith and the 'excessive redishness' of a point, which explains 8.1% of the observed variance in the deviation from the clear line. That is, as the sun gets lower in the sky (an increasing solar zenith) the simulated top-of-atmosphere radiance becomes 'more blue'. This seems to be a residual atmospheric effect, as the path length of light in

the atmosphere becomes longer as the sun gets lower in the sky, and thus the light undergoes more scattering. This scattering is wavelength dependent, so blue light is scattered more than red light, and the pixel becomes 'more blue'. This atmospheric effect should not still be present in the data, as the atmospheric correction of each of the original LandsatAERONET pixels took into account the solar and view angles of the pixel, but it is unlikely to be present in a real-world image as every pixel will have been acquired at a very similar solar zenith angle.

- **Solar azimuth angle:** This explains 5.5% of the observed variance in deviations from the line, and the relationship is almost certainly caused by angular reflectance effects from the ground surface. For example, vegetation can vary significantly in reflectance depending on the azimuth angle it is viewed at, and this will affect the location of the point for that pixel in feature-space. As with the relationship above, it is unlikely to be present in a real-world image, as the solar azimuth angle for each pixel is likely to be similar.
- **Others:** Ozone concentrations and AOT explain 3.7% and 2.5% of the observed variance respectively, but there does not appear to be any significant relationship between these variables and deviation from the line: the R^2 values result from a few outliers and the distribution of the data (for example, most AOT values are relatively low, so the position of a point with high AOT has a significant influence on the relationship).

4.4.3 How does land cover affect the location of the points in feature space?

Figure 4.10 shows of the data points from the LandsatAERONET ($n = 9$) dataset, coloured according to the manual land cover classification results. As described in §4.2.4, this dataset consists of atmospherically-corrected Landsat pixels from sites across the globe. As the data comes from many different images, it will have been collected across a range of seasons and thus will contain the reflectance of vegetation at a range of phenological stages.

The graph shows groups of outliers in the plot of LandsatAERONET TM data, and the majority of these outliers are specific land cover types. The limited number of pixels covered with snow or ice are significantly brighter than the rest of the data (region H on the graph), and bright white building roofs (particularly those which are large enough to cover an entire 30m pixel) and very bright fields (such as oil seed rape) also produce very bright outliers (regions G and F respectively). These regions are outliers in terms of brightness, but are also slightly offset from the Clear Line, which would produce spurious HOT values.

Bare surfaces (region C) and areas of shrub/scrub (region A) are offset from the Clear Line in the red direction, and would therefore produce spuriously low HOT values, and water (region B) and some pasture/hay fields (region E) are offset in the blue direction, and would therefore produce spuriously high HOT values. Region D contains a mix of

urban data acquired under very polluted conditions (AOT = 1.7), which weren't fully corrected by the LandsatAERONET correction routine, and therefore - due to the relationships that the HOT relies on - the points are offset in the blue direction.

These offsets are very similar to those found by Zhang *et al.* (2002a) and shown in Figure 4.1, with water and urban areas producing spuriously high HOT values and grassland/shrubland and bare surfaces producing spuriously low HOT values. This land cover effect is the most significant factor affecting the deviation of points from the Clear Line, and thus some sort of 'land cover correction' procedure will be needed to remove this effect.

4.4.4 How does the breadth of the point cloud change based upon sensor resolution?

The image resolution is likely to significantly affect the breadth of the point cloud, as the measurements acquired for large pixels are in essence an average of the range of spectra present within the pixel. In the terminology of Strahler *et al.* (1986), an H-res image would produce a very broad point cloud as each pixel within the image would have the unique spectrum of an individual part of a separate object, whereas the pixel values in an L-res image would be averages over a wide range of individual spectra, and thus the point cloud would be narrower.

Figure 4.11 shows that this is the case, with mean HOT values decreasing rapidly as resolution increases. The trend for maximum HOT is more complex, with several sharp jumps followed by plateaus. The jumps are likely to be related to the spatial structure of the image. For example, averaging a number of pixels over one object, such as a building roof, is likely to give a fairly similar result to using a single pixel, but as soon as the area of pixels being averaged extends onto another land cover type then the result will change significantly.

This suggests a potential method to improve the resulting HOT values: resample the data to a coarser resolution. This will have the negative effect of removing some of the detail that we are trying to observe, so it should be not used unless absolutely necessary.

4.5 Conclusion

This chapter has examined, in detail, the fundamental basis and assumptions of the HOT. It has shown that the assumptions of the HOT, such as a high correlation between visible bands, are generally valid, and that the analysis of Zhang *et al.* (2002a) is correct. This is an essential stage in the development of a new algorithm based upon the HOT: it is now safe to extend the HOT as the basic method has been shown to be valid.

Issues with the HOT have been examined, and the breadth of the point cloud has been found to be a significant problem. The majority of this breadth is caused by the variation in land cover across a real-world image. Correction for these land cover effects would

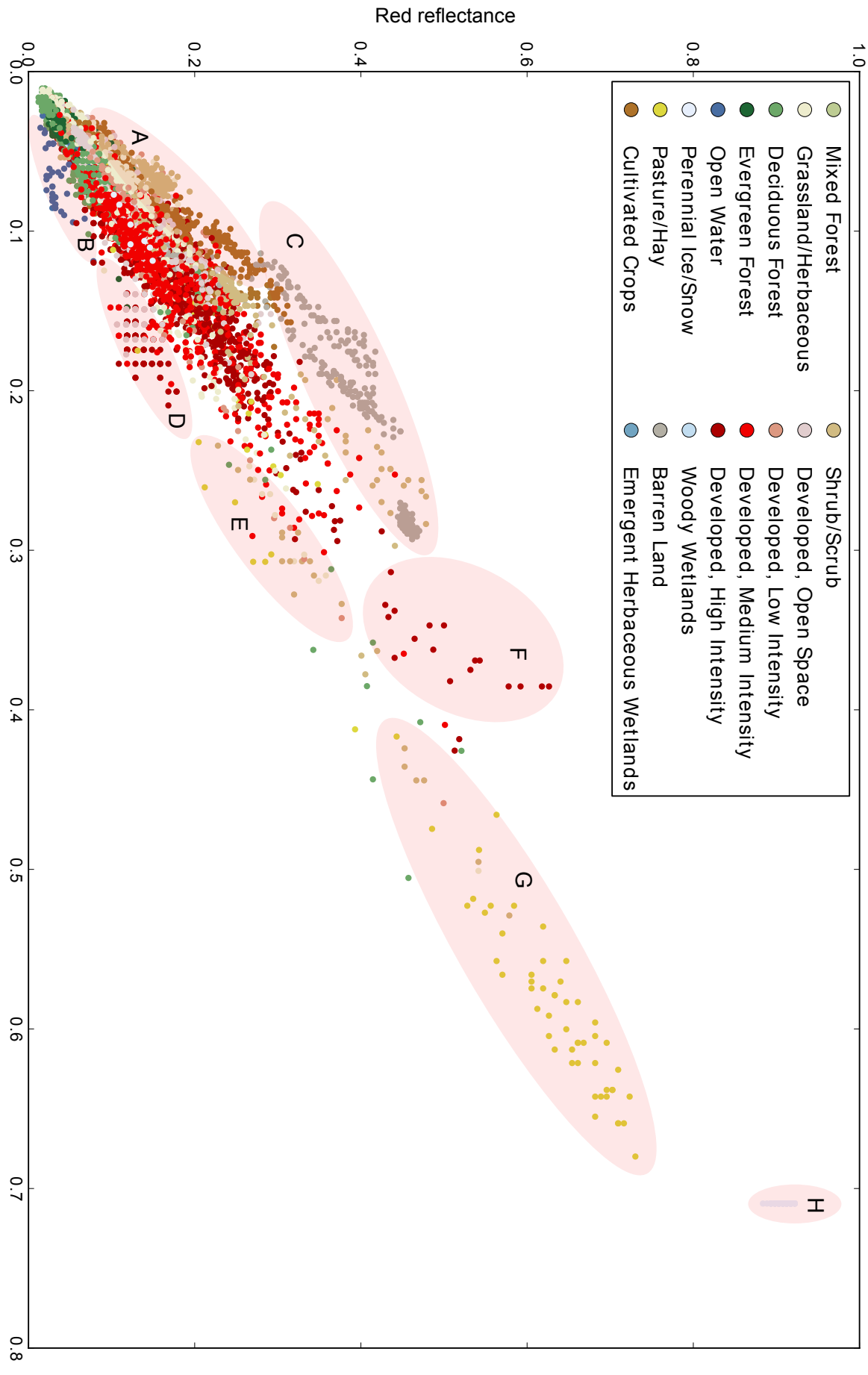
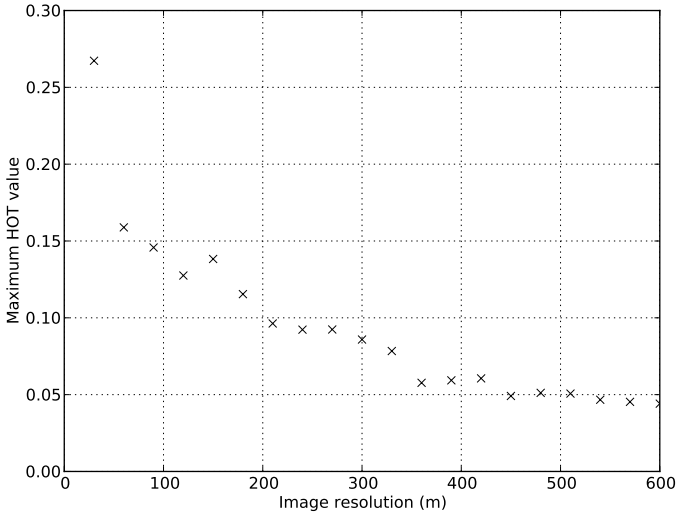
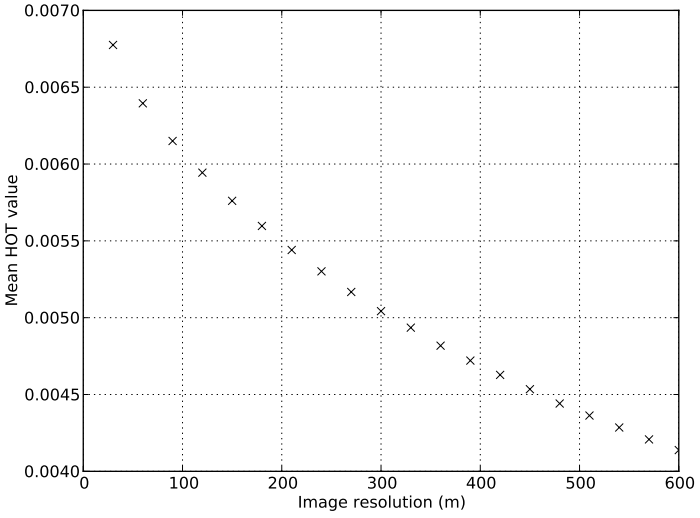


FIGURE 4.10: All data points from the LandsatAERONET ($n = 9$) dataset, coloured according to the manual land cover classification. Highlighted regions are referred to in the text.



(a) Maximum HOT



(b) Mean HOT

FIGURE 4.11: Maximum and Mean HOT values from the Chilbolton LEDAPS image resampled to a range of resolutions

significantly improve the resulting HOT data. A coarsening of sensor resolution has been found to significantly decrease this breadth, although this also removes some of the detail from the images and so is not a desirable solution.

To perform the 'validation' of the original HOT, three data sources were used: an atmospherically-corrected airborne image, LEDAPS-corrected Landsat images, and an entirely new dataset consisting of 'perfectly' atmospherically-corrected Landsat pixels over AERONET sites around the world. This latter dataset - which plays a significant role in the rest of the thesis - is the major novel contribution of this chapter. Although the HOT has been found to be fundamentally valid, further development will be required to allow it

to be used to estimate AOT: this is discussed in the next chapter.

Chapter 5

Improving the Haze Optimized Transform

5.1 Introduction

The previous chapter investigated the original HOT method, and found that its assumptions were valid and it performed as expected. However, the original Haze Optimized Transform is not suitable to be used for estimating AOT, due to its focus on haze as opposed to aerosols. This chapter focuses on developing a version focused on aerosols, as well as implementing a number of other improvements to the HOT, namely:

1. Use of calibrated satellite sensor data, rather than raw Digital Number values
2. Use of a method to estimate the Clear Line without requiring manual selection of a clear area of the image
3. Use of an Object-Based Image Analysis (OBIA) approach to perform a robust HOT correction

These will be discussed in detail in the rest of this chapter.

5.2 Pre-processing satellite sensor data before use

The original HOT used raw digital numbers (DNs) to define the location of the points in feature-space. The advantage of this approach is that it does not require any pre-processing of data, but there are significant disadvantages: DNs have no physical meaning, and are not comparable between different image acquisition conditions or different sensors.

Two options were investigated: using radiance values, and using partially-corrected reflectance values.

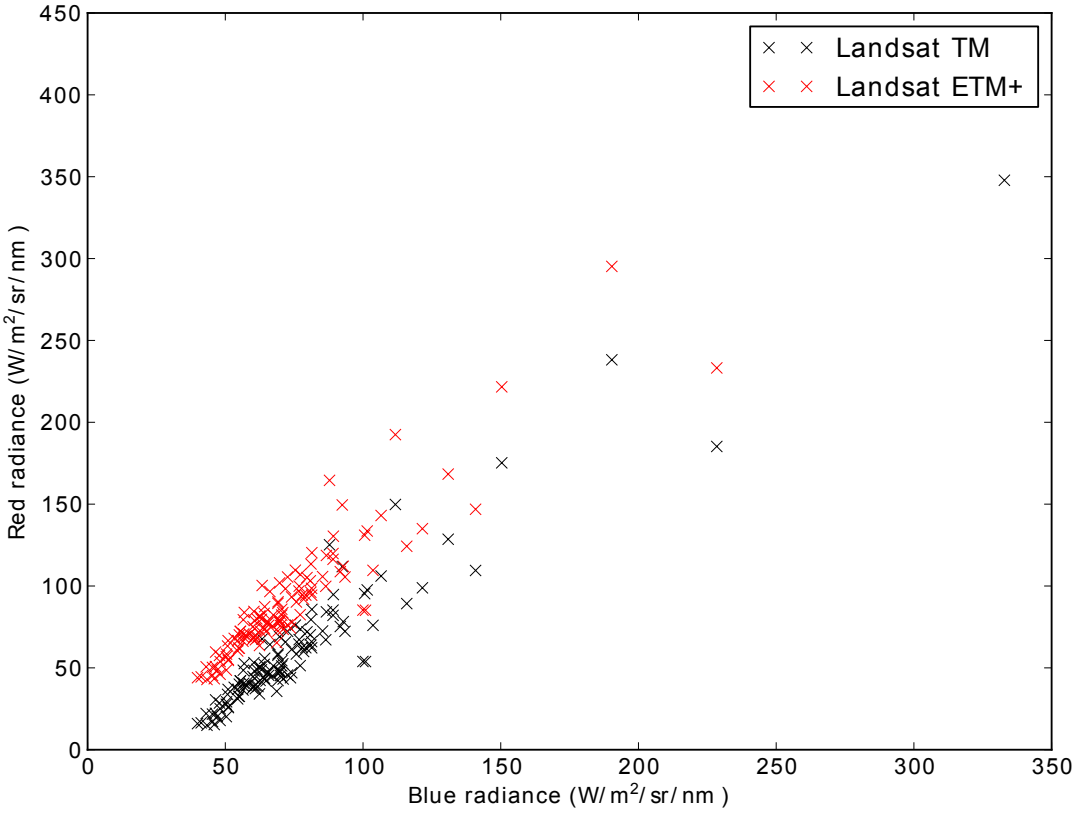


FIGURE 5.1: Comparison of the simulated radiances in a clean atmosphere produced from the $n = 1$ subset of the LandsatAERONET TM dataset using the spectral response functions of the Landsat TM and Landsat ETM+ instruments (which are built-in to 6S Vermote *et al.*, 1997). Both instruments have the same range for each spectral band, but different response functions.

5.2.1 Radiance

One potential improvement is to convert the raw DN values to radiance, using the scaling factors provided in the image metadata, before calculating the HOT value. Use of real physical units (radiance), rather than sensor-specific measurements (DN values) enables the comparison of HOT values between a range of sensors, making it easier to develop a method which can be applied to satellite data from a range of sensors.

Although using radiance enables the use of the HOT for a range of different sensors, care must still be taken as the differing spectral response functions between sensors can cause significant changes in the location of points within feature space (Steven *et al.*, 2003). For example, Figure 5.1 shows that there is a significant difference between simulated radiances of two very similar sensors: Landsat TM and Landsat ETM+. Similarly, radiance values are not comparable between different image acquisition conditions, as radiance is a measure of the light intensity reaching the sensor.

5.2.2 Partially-corrected reflectance

An alternative is to use reflectance, which is a measure of the proportion of light reflected by an object, defined as the ratio of the reflected light to the incident light. Two main categories of reflectance values are used in remote sensing:

- **Top of Atmosphere (TOA)** reflectance is the ratio of the irradiance at the top of the atmosphere to the radiance received at the satellite, including correction for the solar angle and Earth-sun distance.
- **Bottom of Atmosphere (BOA)** reflectance is the ratio of the irradiance at the ground surface to the ground-leaving radiance. Thus, it includes all corrections necessary to obtain irradiance at the ground surface and ground-leaving radiance, such as corrections for solar and viewing angles, Earth-sun distance, mixed gases, ozone, aerosols and water vapour.

As discussed in the previous chapter, reflectance can be measured through a range of illumination and observation geometries producing measurements such as the Hemispherical-Conical Reflectance Factor (HCRF) and the Bi-hemispherical Reflectance Factor (BHRF). Any measurement of reflectance from a satellite will, by definition, be a measurement of HCRF. As discussed in the previous chapter, Landsat pixels have a relatively constant field of view, and thus in this context, HCRF measurements from Landsat should be comparable.

As part of the development of an improved version of the HOT, an alteration of the BOA reflectance correction method was devised, which corrects for all atmospheric effects apart from aerosols: referred to as a 'pseudo-BOA reflectance'. This was implemented using Py6S, by configuring the atmospheric conditions based upon the image metadata and setting the aerosol parameterisation to 'No Aerosols' before using the built-in 6S 'atmospheric correction from radiance' function (described in detail in the original 6S paper; Vermote *et al.*, 1997)

In the context of the Haze Optimized Transform, we want to use data which are comparable between sensors, and which correct as much of the atmospheric variability as possible, while ensuring aerosol effects are still present in the data. This should make it easier to apply the AOT estimation procedure to different sensors, and will ensure that HOT values are only affect by aerosols, not any other atmospheric effects. Therefore, the correction to 'pseudo-BOA reflectance' was chosen, and has been applied to all images used from this point onwards.

5.2.3 Masking

A key component of all aerosol retrieval algorithms is the cloud masking procedure. Even very small amounts of thin cloud over a pixel can artificially increase the AOT value (Kaufman *et al.*, 2005), so the cloud-masking algorithm should be conservative: preferring false positives to false negatives.

Several aerosol retrieval algorithms have their own specially-designed cloud masking procedures (for example Martins *et al.*, 2002). In this case the FMask algorithm (Zhu and Woodcock, 2012) has been used, as it has a lower false negative rate than other Landsat cloud masking algorithms such as ACCA (Irish, 2000). FMask is an object-based masking algorithm which marks each pixel as one of the following:

- Clear land pixel
- Clear water pixel
- Cloud shadow
- Snow
- Cloud
- No observation (fill data)

In this case, any pixels which are not marked as *Clear land* are masked out. As well as masking clouds, this removes all other cloud-related phenomena which may affect the accuracy of the algorithm (shadows, possible clouds and so on), and also pixels containing snow or water, two landcovers which are likely to have erroneous HOT values. As shown in Figure 4.10, water pixels have an erroneously high HOT, and snow pixels have an extremely high brightness. The removal of these two landcovers is unlikely to lead to a significant reduction in the utility of this algorithm, as there are several existing methods for AOT retrieval over water, and many areas of remote sensing require special methods to be used over snow/ice. All pixels which are masked are set to NaN, as this value will carry through all future calculations, ensuring that masked pixels will stay masked during all steps of the algorithm.

5.3 Clear Line estimation

The original Haze Optimized Transform method involved estimating the parameters of the clear line (the gradient, m, and the y-intercept, c) by performing a regression of the blue and red bands in an area of the image with no haze. However, there is no part of any satellite image that will have an AOT of zero, so this approach is impossible here: the regression line must be estimated using some other approach.

Several methods for generating an appropriate regression line were developed and then compared. A key decision when comparing multiple methods is the criteria for assessing the quality of each method: how do we decide whether a method is 'better' than another method? This is particularly difficult in this situation, as we don't have a 'perfect answer' to which we can compare the results. The LEDAPS images are the best per-pixel corrected images that are available, but they have multiple issues (described in §4.2.3), so we cannot necessarily assume that the regression line that best fits through the LEDAPS point cloud will be the same as the Clear Line. However, there is no better way to quantitatively assess the quality of the lines produced by each method, so lines will be assessed by the sum of the squared residuals from a LEDAPS-corrected image, as well as a qualitative assessment of the position of the line.

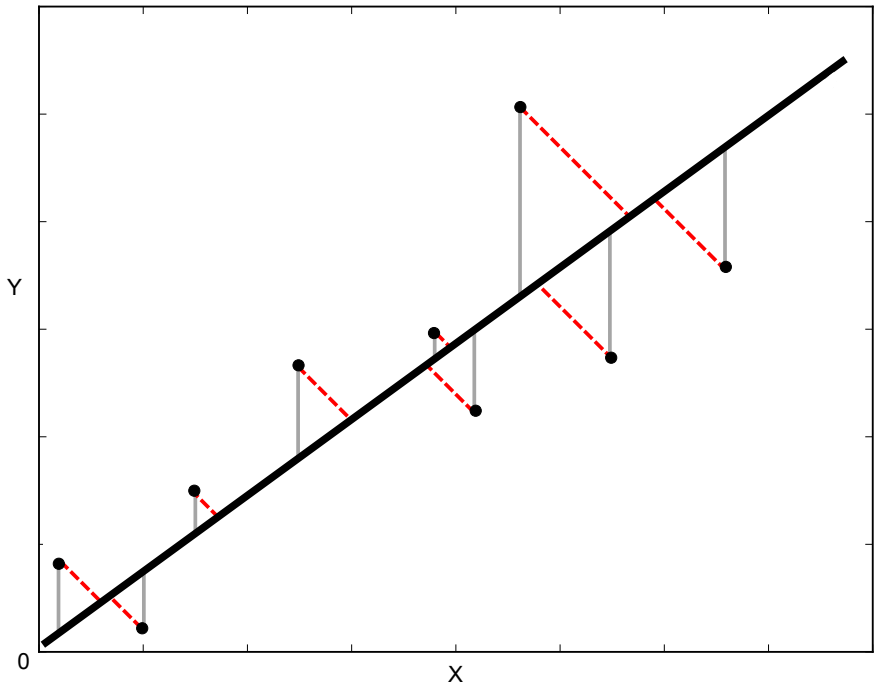


FIGURE 5.2: Graphical depiction of the calculation of residuals in standard regression (grey lines) and ODR (red dotted lines).

5.3.1 Orthogonal Distance Regression

Standard linear regression based upon a least-squares minimisation of the residuals is not appropriate because residuals in standard linear regression are calculated as the distance from the line to each point in the y direction; the error in x is not taken into account. This stems from the common use of linear regression to predict one variable from another where it is only the error in the predicted variable that is relevant. However, in this case we are not trying to predict the red reflectances of pixels from their blue reflectances, we are trying to establish a relationship between these two variables: so both x and y errors are important. To take this into account, all regressions were carried out using Orthogonal Distance Regression (ODR; Boggs and Rogers, 1990), which calculates residuals as the distance to each point along a line perpendicular to the regression line (Figure 5.2). The Scipy interface to the ODRPACK software (Boggs *et al.*, 1989) was used to perform ODR, with starting estimates for the parameters given by standard least-squares linear regression.

A comparison of standard regression and ODR is shown in Figure 5.3 for the Namibia and SotonNewForest LEDAPS simulated radiance images. In both cases, the ODR regression line would be a significantly better Clear Line than the standard regression line, as it is almost parallel to the primary axis of variation in the data (shown by the darkness of the point cloud), meaning that the residuals (that is, the HOT values) would not be affected significantly by the brightness of the point. The effect is greater for the Namibia image,

but is still present in SotonNewForest, and similar results are found for all of the other LEDAPS images.

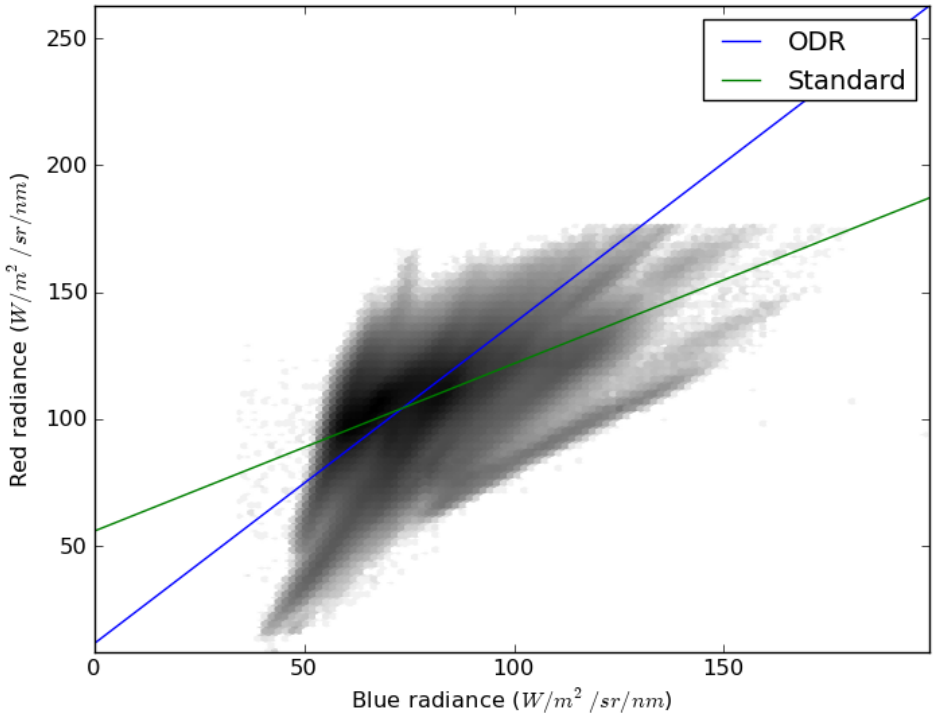
5.3.2 Estimation of Clear Line parameters

The parameters of the clear line should be constant for a set of images with constant land cover, acquired by the same sensor (with no degradation) under the same conditions (for example, viewing and irradiance angles, solar irradiance and ozone amount), but with different aerosol amounts (as, by definition, the aerosol content of pixels in the image should not affect the definition of the clear line). All of these criteria are unlikely to be met, as even images of the same location will have different conditions depending on the time and date of image acquisition. However, if we assume that the conditions of image acquisition do not significantly affect the parameters of the clear line - given the pseudo-BOA reflectance correction - then the only variable which will significantly affect the clear line parameters is the land cover.

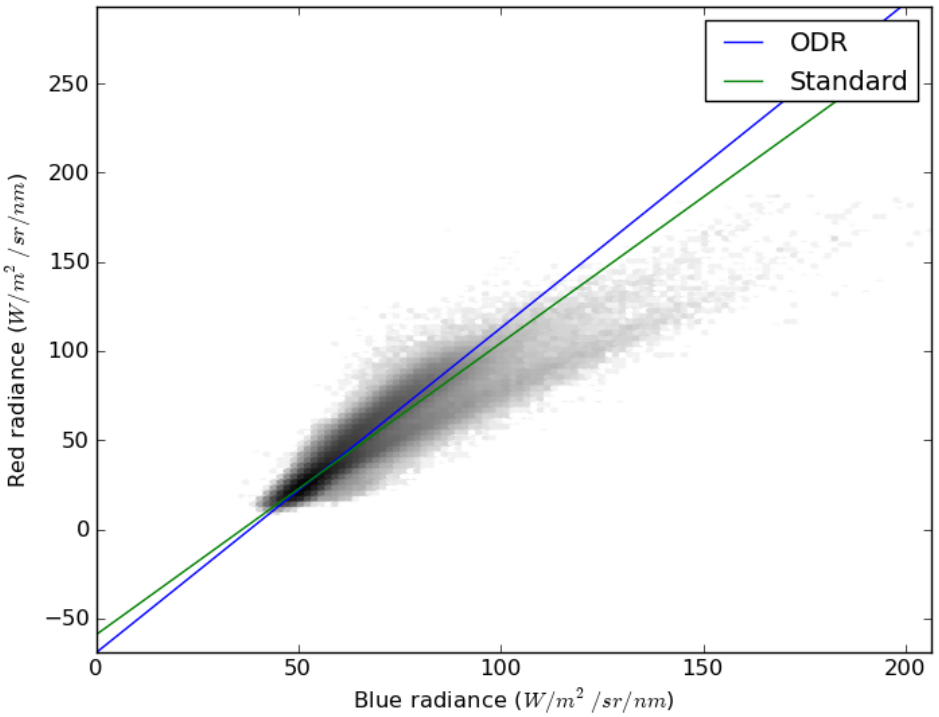
He *et al.* (2010) argued that the location of the clear line in most images was primarily controlled by two landcovers which are present in the majority of images: vegetation and asphalt. If we assume this is the case for all images we will want to process, then we can simply use the regression line calculated using all points in the LandsatAERONET dataset as a regression line for all images.

Some of the points in the LandsatAERONET dataset have unusually high reflectances which are unlikely to be present in many images, and may have an undue influence upon the regression line. Thus, the regression line based upon LandsatAERONET points where both the red and blue reflectances are low - in this case under a threshold of 0.5 - could prove to be better.

The assumption by He *et al.* (2010) that the land cover of an image does not affect the parameters of the Clear Line may not hold for images with very different land covers. Therefore, an alternative method was developed to take into account the land cover of the image when estimating the Clear Line parameters. The GlobCover dataset (Arino *et al.*, 2007) provides 300m land cover data across the whole world, and this was used to calculate the percentage of pixels in each of the GlobCover classes over the area covered by the satellite image. These percentages for each class were then mapped to the NLCD classes which were used for classifying the LandsatAERONET dataset, using the class correspondences in Table 5.1. A Monte Carlo sampling procedure was then applied, sampling 5,000 points from the LandsatAERONET dataset with the same proportion of land cover classes as found in the GlobCover subset. Regression (using the ODR method) was then carried out on this sample of LandsatAERONET points. This sample is representative sample of the ground reflectances of pixels within the image, and thus the regression line calculated from this sample should produce a good Clear Line. This novel approach can be traced back to the work of Lambeck (1977), who used 'average Landsat scene reflectances' as part of their image processing procedures: here a more advanced,



(a) Namibia



(b) SotonNewForest

FIGURE 5.3: Comparison of standard Ordinary Least Squares linear regression with Orthogonal Distance Regression for two example LEDAPS images. The results for the other LEDAPS images were very similar.

TABLE 5.1: GlobCover to NLCD class mapping, used in the Monte Carlo regression

GlobCover Class	**NLCD Class**
Post-flooding or irrigated croplands (or aquatic)	Open Water
Rainfed croplands	Cultivated Crops
Mosaic cropland / vegetation	Cultivated Crops
Mosaic vegetation / cropland	Pasture/Hay
Closed to open broadleaved/evergreen forest	Mixed Forest
Closed broadleaved deciduous forest	Deciduous Forest
Open broadleaved deciduous forest	Deciduous Forest
Closed needleleaved evergreen forest	Evergreen Forest
Open needleleaved deciduous or evergreen forest	Evergreen Forest
Closed to open mixed forest	Mixed Forest
Mosaic forest or shrubland / grassland	Shrub/Scrub
Mosaic grassland / forest or shrubland	Grassland/Herbaceous
Closed to open shrubland	Shrub/Scrub
Closed to open herbaceous vegetation	Grassland/Herbaceous
Sparse vegetation	Shrub/Scrub
Closed to open broadleaved forest (fresh flooded)	Deciduous Forest
Closed broadleaved forest or shrubland (salt flooded)	Deciduous Forest
Closed to open grassland (regularly flooded)	Grassland/Herbaceous
Artificial surfaces and associated areas	Developed, Medium Intensity
Bare areas	Barren Land (Rock/Sand/Clay)
Water bodies	Open Water
Permanent snow and ice	Perennial Ice/Snow
No data (burnt areas, clouds)	NaN

and more representative, 'landcover-weighted' average is used.

5.3.3 Comparison of methods

The sum of squared HOT values (ΣHOT^2; which is equal to the sum of the squared residuals calculated by the ODR) for each of the LEDAPS images and each of the three methods described above are shown in Table 5.2. The lower the ΣHOT^2, the better the method fits the LandsatAERONET dataset, and therefore the closer to the assumed 'best' Clear Line the estimated line is.

In all but the very bright images (Namibia, Sahara and Saudi Arabia), calculating the regression using only the LandsatAERONET data points with reflectances < 0.5 improved the fit significantly, and in approximately half of the images using the LandsatAERONET Monte Carlo approach reduced the ΣHOT^2 significantly - for two images the decrease was by an order of magnitude. In the other half of the images the LandsatAERONET Monte Carlo approach has a slightly poorer fit than the LandsatAERONET < 0.5 line, but the magnitude of the increase in ΣHOT^2 is significantly smaller than the magnitude of the decrease for the other half of the images. The LandsatAERONET Monte Carlo method also has the lowest standard deviation, showing that it performs more uniformly across a wide range of images: an important factor in making this method as widely applicable as

possible. Given these results, it was decided to use the LandsatAERONET Monte Carlo method, and this has been used to estimate the Clear Line from this point onwards.

Once the position of the Clear Line has been estimated, the HOT image is created by calculating the orthogonal distance from the clear line for each pixel in the image as

$$\text{HOT} = \frac{B_3 - (mB_1) - c}{\sqrt{m^2 + 1}} \tag{5.1}$$

where B_1 is the pixel value in band 1 (blue), B_3 is the pixel value in band 3 (red), m is the gradient of the clear line and c is the y-intercept of the clear line.

5.4 Object-based HOT correction

The main issue with the HOT is that its values are affected by land cover as well as aerosol concentrations; by attempting to correct for the effect of land cover, the uncertainty involved in estimating AOT from HOT should be reduced.

Previous work on the Advanced Haze Optimized Transform (AHOT; He *et al.*, 2010) and the Background Suppressed Haze Thickness Index (BSHTI; Liu *et al.*, 2011) has involved performing a correction procedure on the raw HOT values to remove this land cover effect. However, this is difficult as it is not just a signal-noise correction, as land cover is both a source of signal (as it is a driver of localised change in AOT - for example, the higher values generally found over urban areas) and of noise (anomalous HOT values caused by land cover effects). A good HOT correction method will remove anomalous HOT values caused by land cover effects, while not altering genuine HOT values: a challenging task.

Methods have generally focused on looking for sudden changes in HOT value, as these are unlikely to occur in genuine HOT data (due to atmospheric mixing causing dispersion of aerosol particles), but are often caused by sharp borders between different land covers. The correction method used by AHOT and BSHTI is the same, and involves treating the HOT image as a Digital Elevation Model (DEM). The correction task then involves 'filling sinks' and 'flattening peaks' to remove negative and positive anomalies respectively. These are implemented using mathematical morphology functions. However, a mathematical morphology-based approach, although computationally-efficient, is not flexible enough to allow the ideal correction to be implemented. For example, the peak flattening routine used by AHOT and BSHTI flattens a peak to a constant value, thus removing the spatial pattern of HOT values within the anomalous area.

An improved HOT correction method has been developed within the context of Geographic Object-based Image Analysis (GEOBIA). GEOBIA (Hay and Castilla, 2008; Blaschke, 2010) is a general approach to image processing which involves segmenting the image into a number of real-world objects and then processing each of these objects separately. The formal mathematical definition of segmentation states that every pixel in the image should be a member of one - and only one - segment: that is, the entire image

TABLE 5.2: Sum of squared HOT values (ΣHOT^2) for the three Clear Line estimation methods: using the regression line from the LandsatAERONET dataset, using the regression line from all points in the LandsatAERONET dataset which have a reflectance < 0.5 and calculating the regression line using the LandsatAERONET Monte Carlo method.

Site	**All LandsatAERONET**	**All LandsatAERONET** < 0.5	**LandsatAERONET Monte Carlo**
Amazon	7.0×10^4	2.3×10^2	3.3×10^2
Botswana	1.8×10^4	1.6×10^4	1.4×10^4
Chilbolton	6.6×10^4	2.2×10^3	3.8×10^3
London	5.4×10^4	1.3×10^3	2.0×10^3
Midwest	8.0×10^4	1.8×10^3	3.0×10^3
Namibia	5.1×10^4	1.2×10^5	7.0×10^4
Sahara	3.5×10^5	4.4×10^5	1.8×10^5
Saudi Arabia	1.2×10^5	1.8×10^5	4.1×10^4
Southampton	2.1×10^4	6.2×10^2	1.4×10^3
Sum	8.29×10^5	7.62×10^5	3.16×10^5
Mean	9.2×10^4	8.5×10^4	3.5×10^4
StDev	9.5×10^4	1.4×10^5	5.6×10^4
Max	3.5×10^5	4.4×10^5	1.8×10^5
Min	1.8×10^4	2.3×10^2	3.3×10^2

is segmented. In many applications, however, this is inefficient, as the majority of image objects are not used. For example, previous work by the author developed a method to only segment areas which had the potential to be Ground Calibration Targets for use in empirical atmospheric correction (Wilson and Milton, 2010).

In this application, image objects are only created where there are sharp boundaries in the HOT image, thus extracting a set of objects which have HOT anomalies which are likely to have been caused by land cover effects. Once the objects are created for each of the anomalous regions, the correction process can be implemented very flexibly on a per-region basis, taking into account all of the contextual information which the use of image objects provides.

The development and testing of the algorithm was carried out using LEDAPS images, which - theoretically - should have spatially-constant HOT values, due to the LEDAPS spatially-variable atmospheric correction removing all aerosol effects. Thus, theoretically, all anomalies seem in these images will be due to land-cover effects which the correction method should remove. However, issues with the LEDAPS correction mean that this is not necessarily the case, and that some areas will have genuinely higher or lower HOT values, due to a poor LEDAPS correction in that area. This will be analysed further when assessing the quality of the correction, in §5.4.8.

5.4.1 Segmentation algorithm

The segmentation algorithm takes a HOT image and creates image objects over sharply-defined areas with anomalous values, and consists of three main stages:

1. **Edge detection:** The Canny algorithm (Canny, 1986) was used to find sharp edges in the image. The algorithm uses hysteresis thresholding to mark pixels which have an edge magnitude $> T_{\text{high}}$ as edges, and then recursively mark any pixels which have an edge magnitude $> T_{\text{low}}$ and are connected to a previously-found edge pixel as edges too. In most Canny implementations, the thresholds are given as absolute edge magnitude values, but the range of magnitudes can vary significantly between images. Thus, a wrapper for the Canny algorithm was created which allows thresholds to be specified as a percentile of the histogram of magnitude values, given as a fraction between 0–1. T_{high} is generally set very high to ensure that only very strong edges are picked up, but T_{low} is set lower to ease the edge completion step.
2. **Edge Completion:** The output from edge detection often has gaps in the borders of objects, and these require filling in an intelligent manner. The Shih and Cheng (2004) algorithm for adaptive morphological edge linking, which fills gaps in edges using locally-aligned ellipsoidal morphological dilation, was implemented to do this. This method sometimes produces isolated small edges which are not linked to any other edges: these are automatically removed in the next step.
3. **Filling and labelling edges:** Now that the edges have been completed so that each object has a fully-defined border, the edges were filled to create image objects,

and each object was given a unique integer label. Small edges that have not been used in this process (for example, if they never got fully linked into a polygon) were then removed.

Full details of the segmentation process, and the relevant parameters used, are shown in Figure 5.4, and example images at each stage of the process are shown in Figure 5.5.

5.4.2 Correction algorithm

Through the use of a object-based approach, each anomaly will be represented by a separate image object, thus allowing correction algorithm to perform individual corrections for each anomaly, rather than performing a uniform correction across the whole image, as in previous methods (such as He *et al.*, 2010). The basic correction method, upon which several extensions have been implemented, is fast to execute and does not restrict the corrected values inside an object to be constant, as is the case with the AHOT and BSHTI approaches.

The algorithm is conceptually simple (although it is implemented in a somewhat more complex manner to obtain the best computational efficiency): the difference between the mean HOT value of the pixels within an positive HOT anomaly image object and those bordering it is subtracted from all of the pixels within the image object. This preserves the spatial pattern of HOT values within the object, but reduces the magnitude of the values to make them closer to the other pixels in the local area, thus removing the large HOT anomaly. The same method is implemented for negative HOT anomaly objects, but the difference is added rather than subtracted. Full details are provided in Figure 5.6.

5.4.3 Improvement of the segmentation algorithm

The results of this segmentation process were examined across all of the test LEDAPS images, and the final segmentation was observed to be poor in some situations, with both missing segments where they should have been present (under-segmentation), and segments in areas that did not have sharp boundaries (over-segmentation). This then led to a lack of correction, or over-correction, in these areas. Several refinements were made to the segmentation algorithm to take these into account. Whether a particular change improved the results was determined visually, based on both the generated segments image and the removal of sharp-boundaried anomalies in the LEDAPS images corrected using these segments.

- **Segmentation around regions with no data:** As discussed in §5.2.3, various regions in the input images were masked out, and the values set to NaN. However, the edge detection ignores all masked pixels and therefore the edges of objects that bordered masked areas were not being detected. This was resolved by performing two separate Canny edge detection procedures which were then combined using a binary OR operation: one to detect the standard edges in the image, and one in which a binary image consisting of NaN/non-NaN values was used to detect the

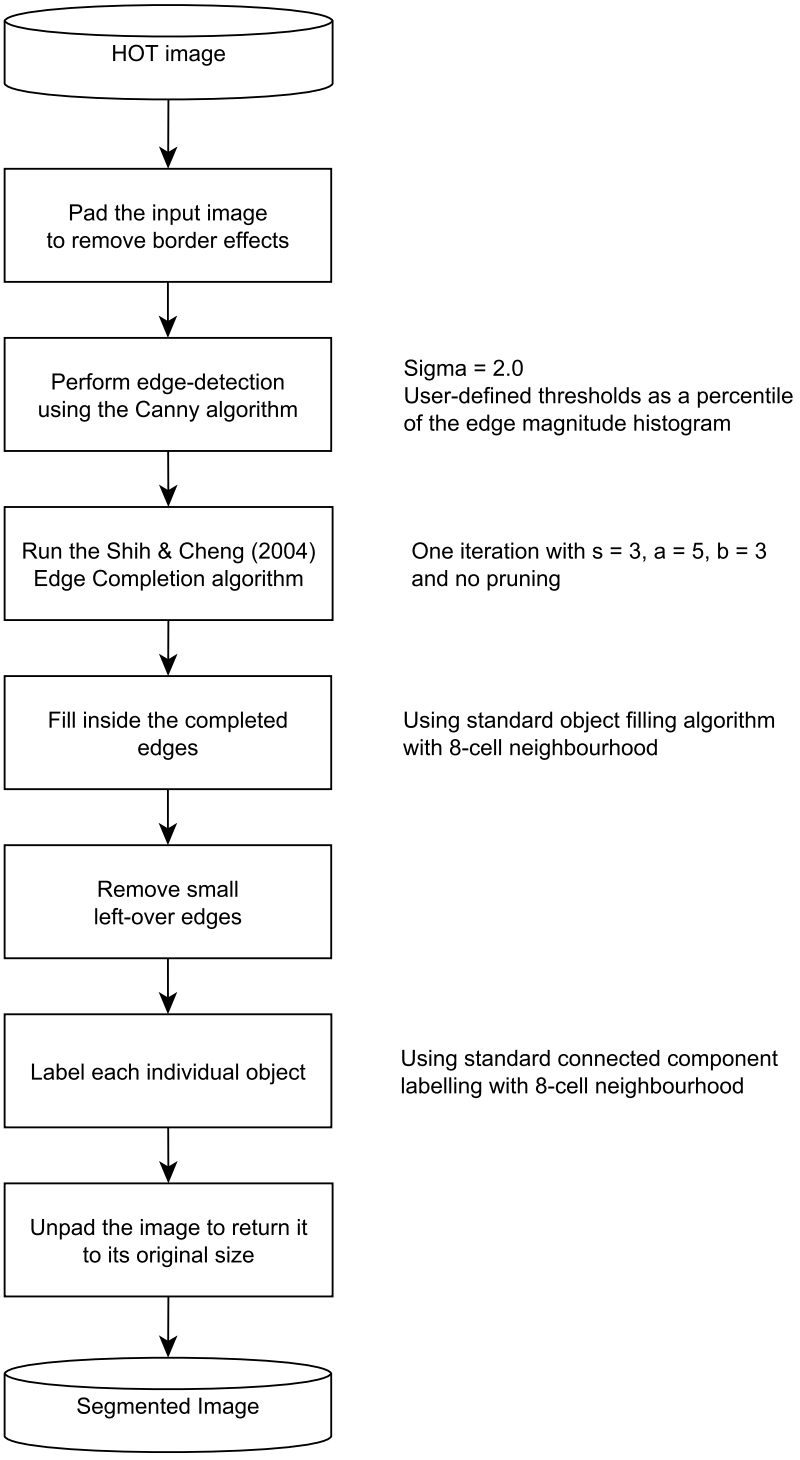
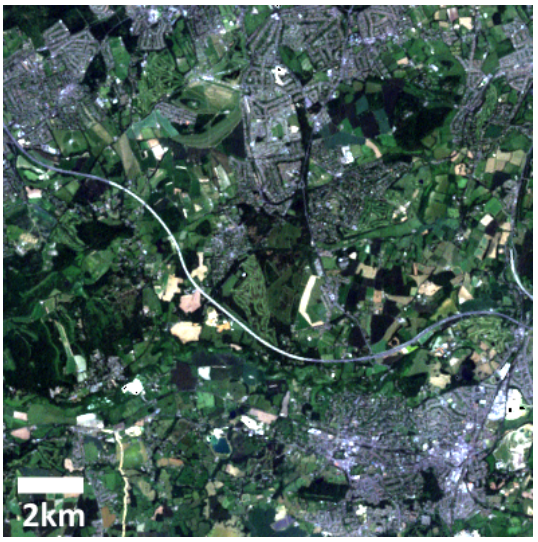
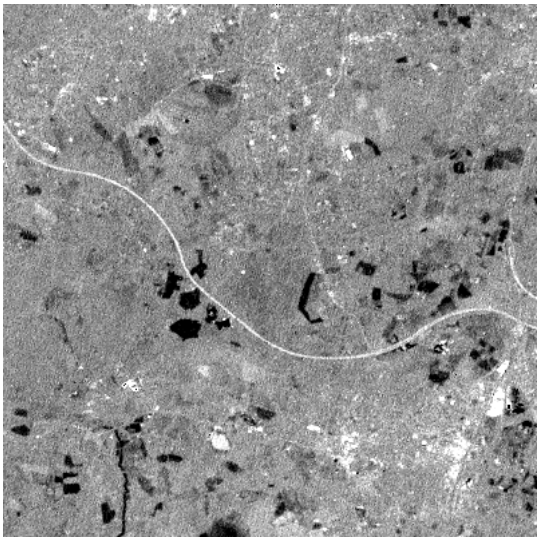


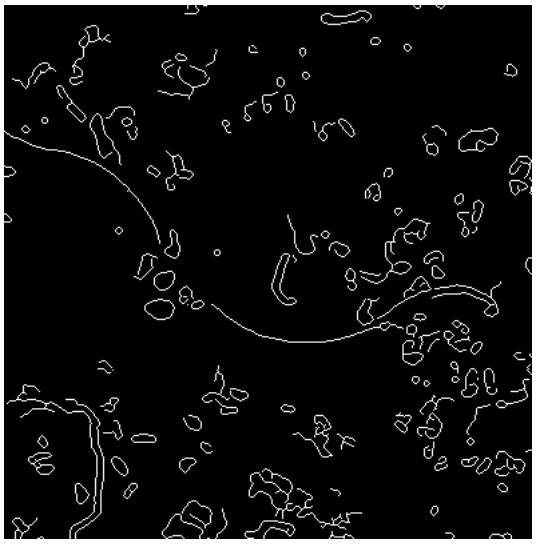
FIGURE 5.4: Flowchart showing the segmentation procedure, along with the parameters used.



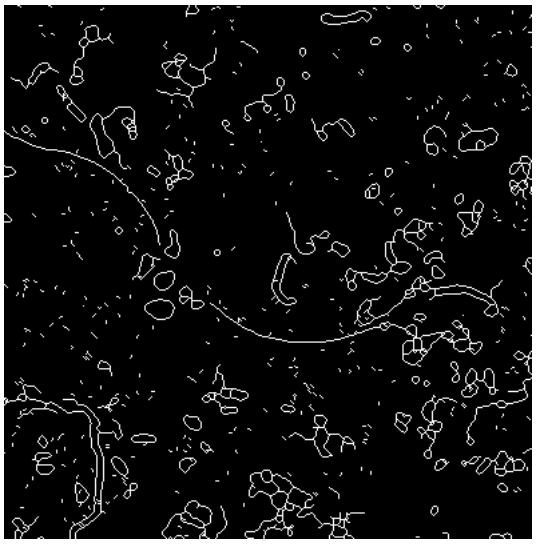
(a) RGB composite



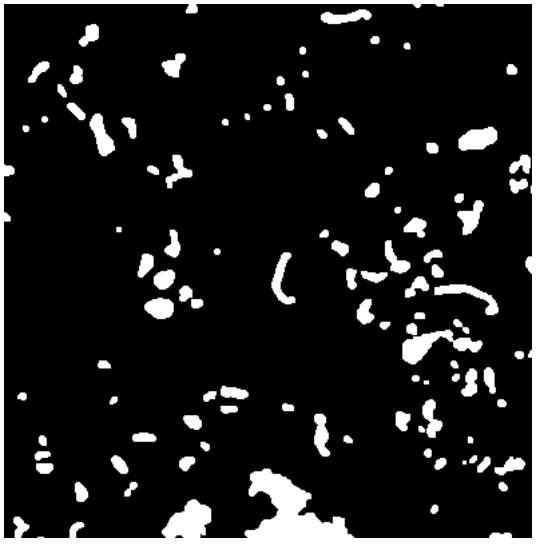
(b) HOT image



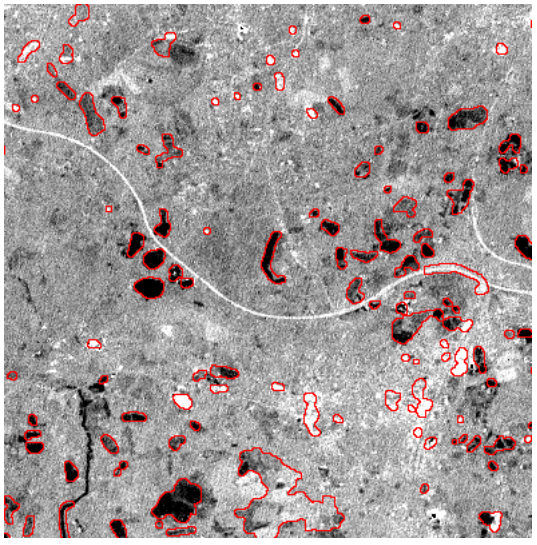
(c) Raw edges



(d) Completed edges



(e) Segmented areas



(f) HOT image with segmented areas overlayed

FIGURE 5.5: Example outputs at each stage of the segmentation process of a subset of the London Landsat image over Redhill, Surrey.

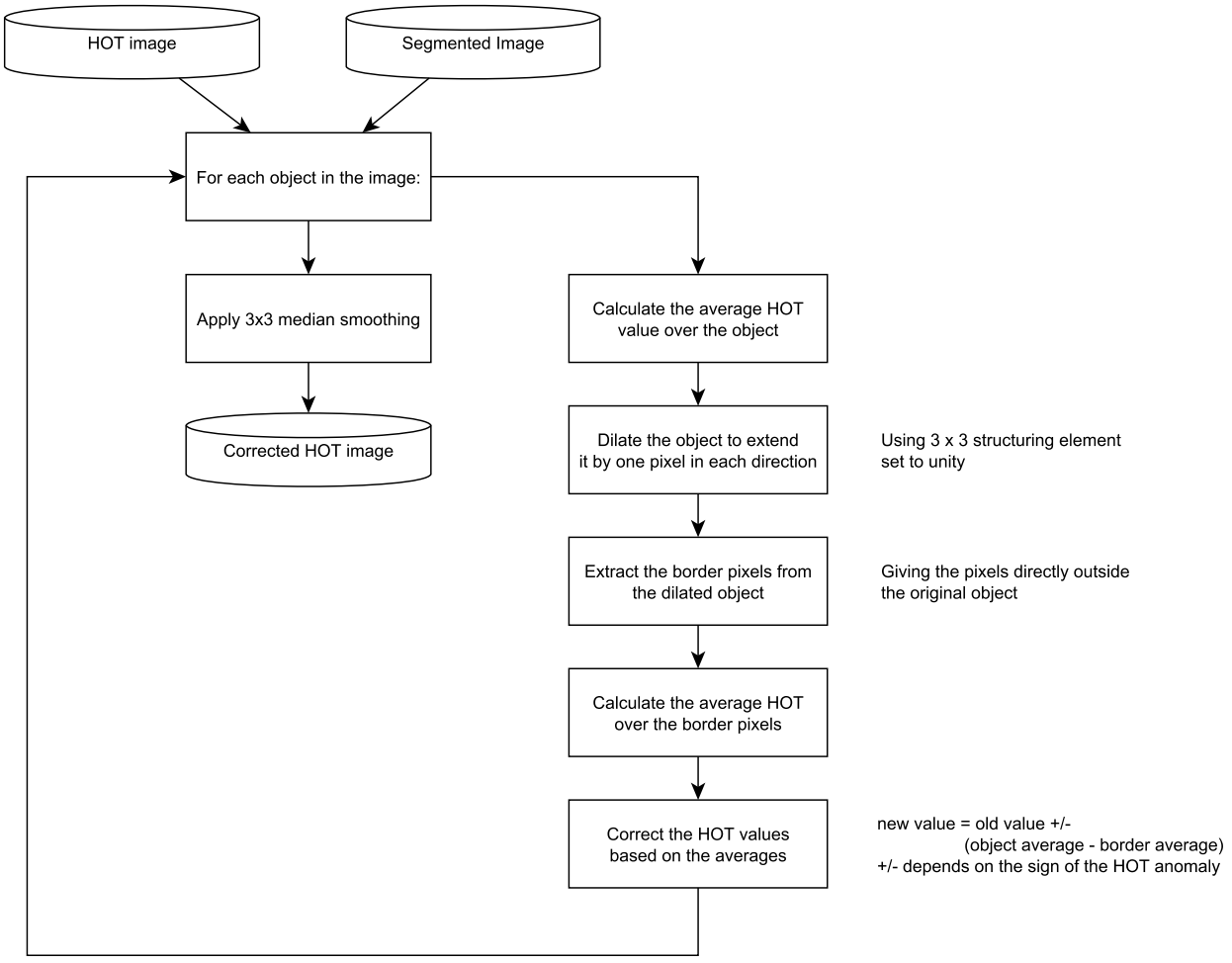
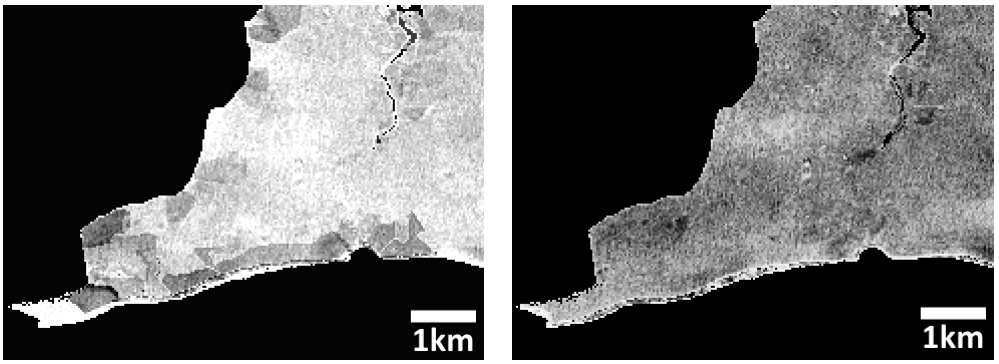


FIGURE 5.6: Flowchart showing the base HOT correction procedure.



(a) Before (b) After

FIGURE 5.7: Example of the improvement in segmentation around regions with no data. This extract from the SotonNewForest image shows the westernmost part of the Isle of Wight. In (a), correction artefacts around the coast are clearly noticeable. These are caused by segments not being created in areas adjoining masked areas - in this case, the sea. In (b) the correction is improved by the extension of the edge detection process to deal with masked areas, thus allowing segments to be created next to masked areas. The colour scale is the same for both images, with high HOT values shown in white, and low values in black.

edges of NaN regions. The improvement in the correction of part of the Isle of Wight coast is shown in Figure 5.7.

- **Perform full adaptive iteration for edge completion:** The Shih and Cheng (2004) algorithm can be run in a single step or as an iterative process with pruning of 'dangling' edges between each iteration. The former was originally used as it is quicker, but it was found that implementing the full adaptive iteration procedure reduced over-segmentation as 'dangling' edges were removed by the pruning procedure and thus did not form their own spurious objects.
- **Allow sub-segments:** Due to the mode of operation of the standard region labelling approach, the original segmentation algorithm simply filled inside the edges produced by the edge completion procedure. This means that if an object with a sharp boundary had another sharp-boundaried object within it, for example, a lake in the middle of an area of woodland, the whole area was segmented as one object. This effect was most noticeable in London which has relatively a sharp boundary around the main urban area, but also has various sharp-boundaried objects within it which have significantly different land cover, such as Hyde Park and other areas of green space. In the original version of the algorithm the whole of London was segmented as one object, leaving sharp boundaried objects inside the main urban area uncorrected (Figure 5.8(a)).

  The ability to create sub-objects was implemented by replacing the simple filling and labelling step in the segmentation algorithm by the following:

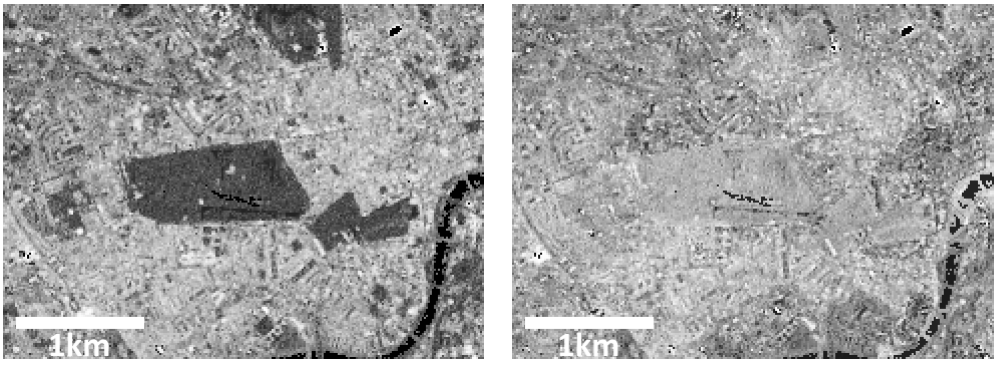
  1. Filling the completed edges to create the large objects (as in the previous version)
  2. Masking these filled objects with the completed edges, thus separating sub-objects with a line of blank pixels
  3. Labelling the result, which will result in separate labels for each sub-object due to the blank pixel separation

Adding the creation of sub-objects using this method improved the correction of areas in central London (Figure 5.8(b)), and elsewhere.

5.4.4 Improvement of the correction algorithm

Even with an improved segmentation, the correction procedure produced poor results in some circumstances. An iterative process of testing possible changes led to the following improvements in the algorithm. Whether a particular change improved the results was determined visually, based on the removal of sharp-boundaried anomalies in the LEDAPS images.

- **Offset borders:** The bordering pixels used to calculate the average HOT value around the image object are offset by one pixel, so rather than covering the immediate border of the object, there is one pixel gap first (see Figure 5.9). This reduces the effect of an erroneous segmentation of the object, and deals with adjacency effects in the HOT image, resulting in a reduction in correction artefacts



(a) Before (b) After

FIGURE 5.8: Example of the effect of allowing sub-segments to be created during the segmentation process. This extract from the London image shows Hyde Park, London and the surrounding area. In (a) the park - which is a significantly different land cover to the surrounding area, and has a sharp boundary - is not corrected at all, as it has not been individually segmented. In (b) it has been extracted as a sub-segment and is well corrected. The colour scale is the same for both images, with high HOT values shown in white, and low values in black.

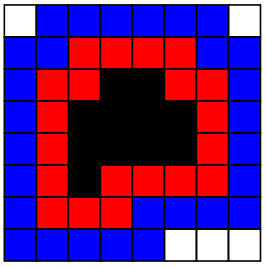
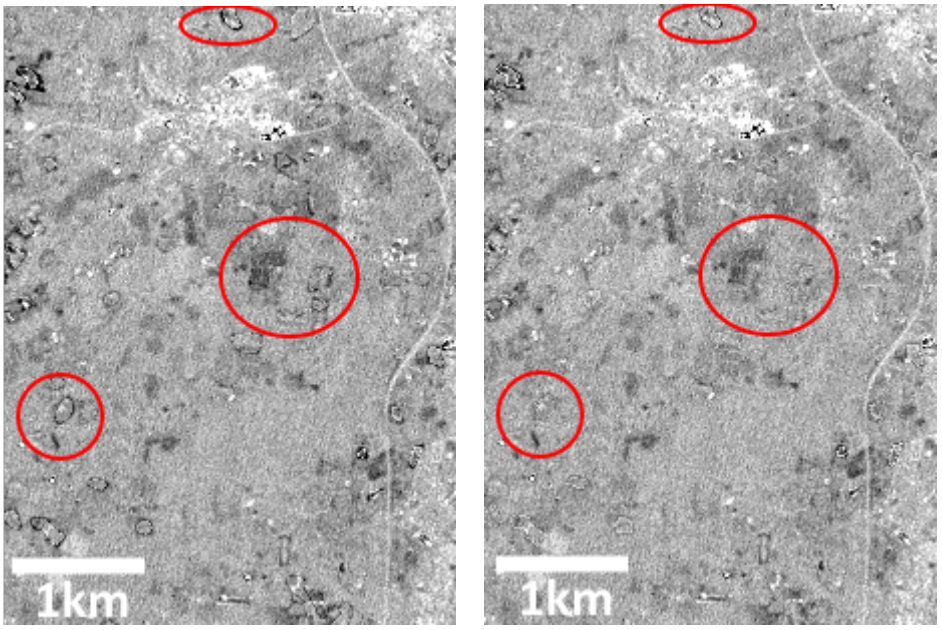


FIGURE 5.9: The two borders used by the correction algorithm. The black pixels are the image object, red pixels show the standard border, and blue pixels show the offset border.

around the edges of objects (Figure 5.10).

- **Correcting for edge effects:** The HOT values of pixels around the edge of dark objects are often brighter than the rest of the dark pixels (due to mixed pixels occurring here), and thus if they are corrected in the same way as the dark pixels they will be 'over-corrected' and appear as bright edges around the object (Figure 5.11). These edge effects are removed by selectively correcting the pixels surrounding the object, so that those which are of a similar magnitude to the object are corrected, but those which are significantly brighter do not have their values changed. The threshold used to decide which pixels to include is chosen using the Otsu (1975) algorithm, which finds the optimal threshold for separating two classes (in this case dark pixels and the brighter pixels at the edge of the dark object) based on an analysis of the histogram of values (Figure 5.12). The effect is a reduction in the magnitude - and in many cases the removal - of edge effects around bright and dark objects, particularly in areas with multiple objects close to each other. For example, Figure 5.13 shows a small landscaped business park next to a major road.



(a) Before (b) After

FIGURE 5.10: Example of the effect of offsetting the borders of the image objects when calculating statistics, taken from the SotonNewForest image, and showing the reduction in correction artefacts in (b) compared to (a). The colour scale is the same for both images, with high HOT values shown in white, and low values in black.

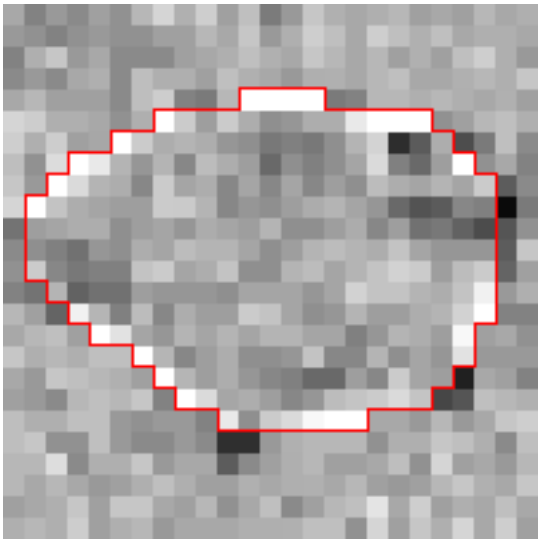


FIGURE 5.11: Example of the edge effects produced by the HOT correction algorithm when correcting a small area of dark forest pixels. The bright edge effects cover 1–2 30m pixels.

Correcting the multiple segments created for the various sharply-defined land covers within the business park (vegetation, buildings and roads) originally caused edge effects, which are removed when using the Otsu threshold.

- **Using more robust statistics to characterise the object-border difference:** Over- or under-correction of objects in the HOT image was observed in some situations. After investigation it was found that these were caused by

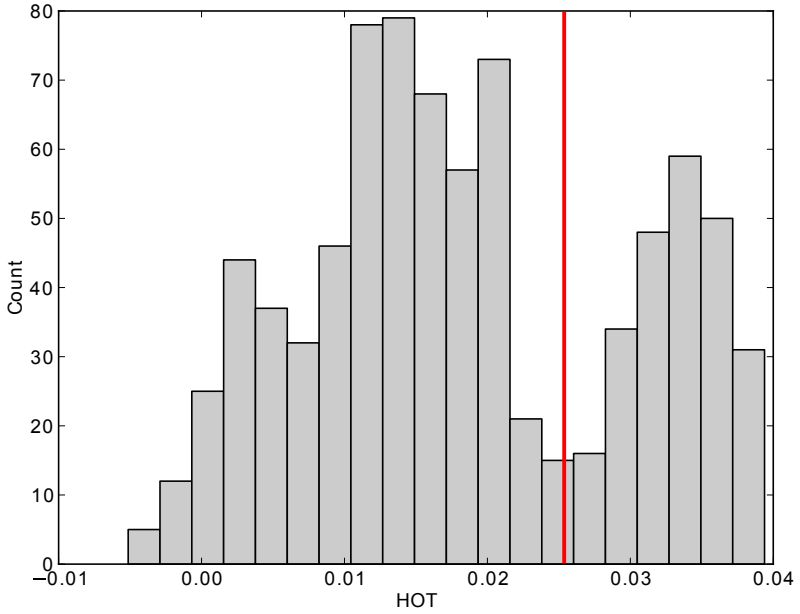
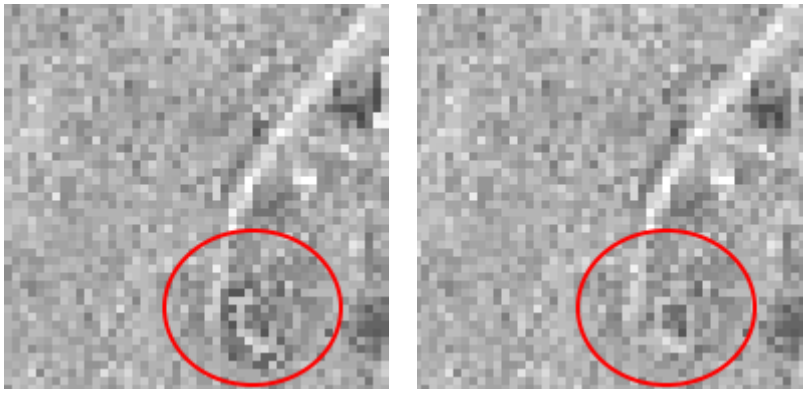
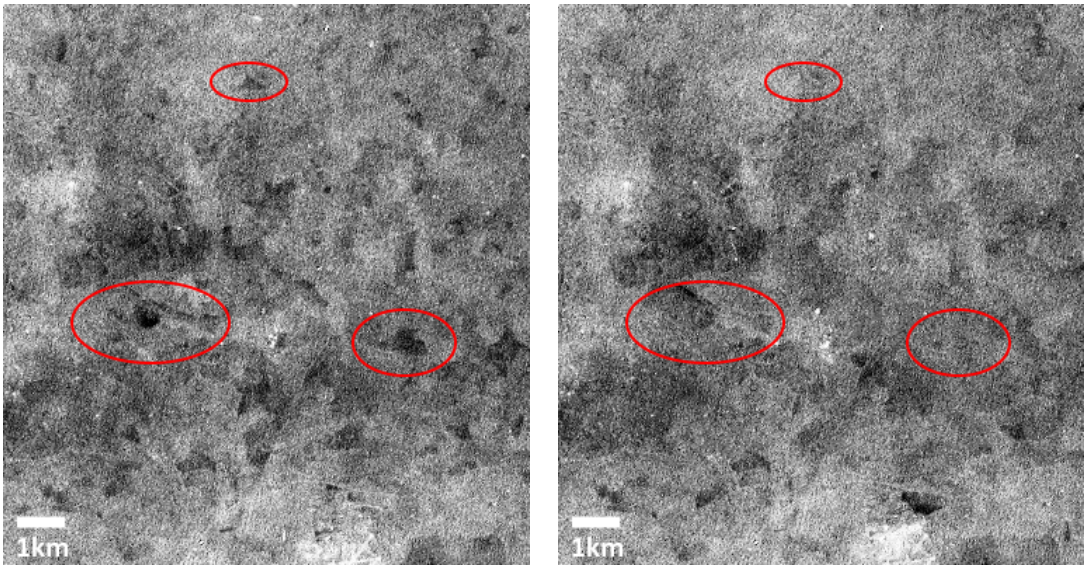


FIGURE 5.12: Example of the histogram of HOT values within a dark object. The histogram is noticeably bi-modal, and the Otsu algorithm chooses the threshold, shown by the red line, which best separates these two separate data distributions: the normal dark pixels (left) and the brighter edge pixels (right).



(a) Before (b) After

FIGURE 5.13: Example of the effect of correcting for edge effects by using the Otsu threshold, taken from the SotonNewForest image, and showing the reduction in correction artefacts in (b) compared to (a). The colour scale is the same for both images, with high HOT values shown in white, and low values in black.



(a) Mean (b) Median

FIGURE 5.14: Example of the effect of calculating the image object and border statistics using the median (b) instead of the mean (a), showing the improvement in the correction from using the median. The colour scale is the same for both images, with high HOT values shown in white, and low values in black.

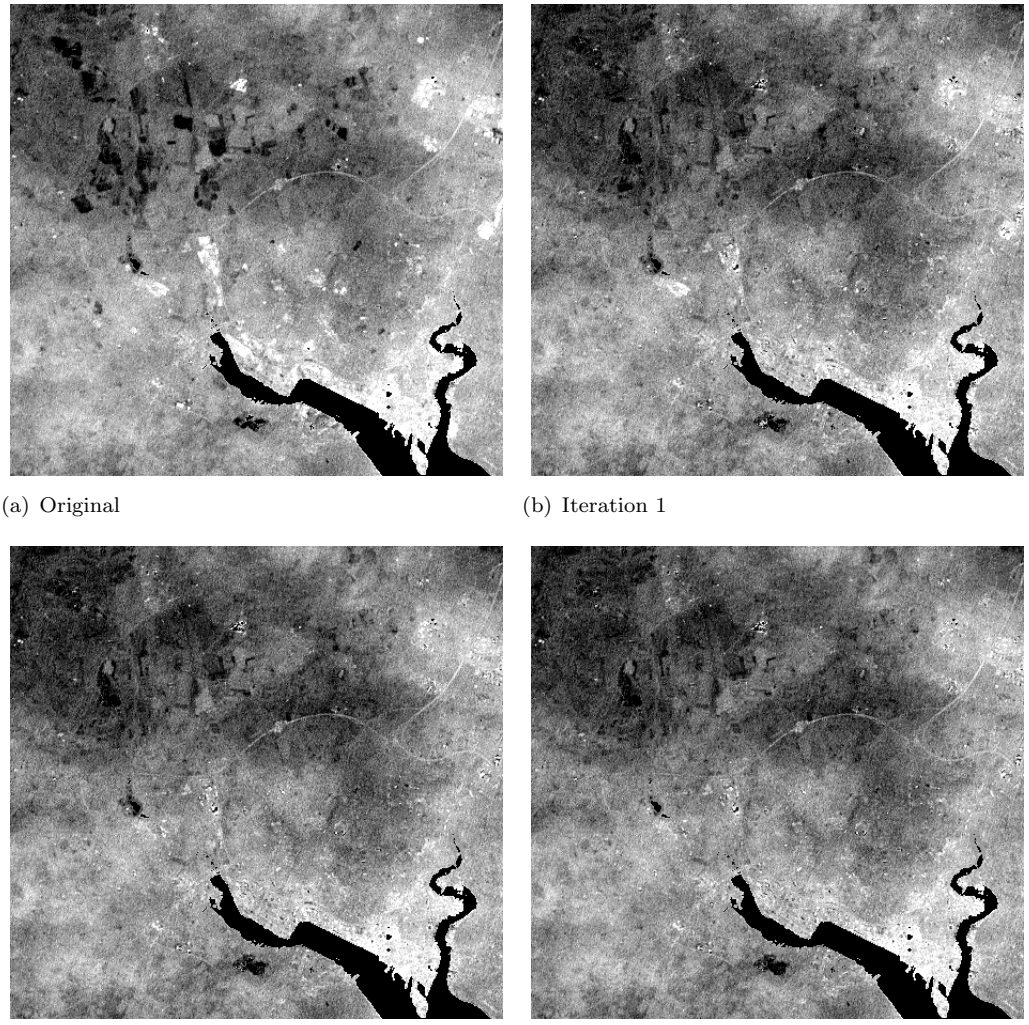
anomalous values in the object or border which were affecting the mean value used to calculate the correction. The mean is particularly sensitive to extreme values, and several more robust, but more complex, methods based upon the histograms of each object were tested. However, the best results were found by simply replacing the mean with the median, which is far less sensitive to extreme values. For example, Figure 5.14 shows the improvement in the correction of anomalies caused by vegetation in a rural area of the SotonNewForest image when the median is used.

5.4.5 Iteration

With the alterations described above, the combined segmentation and correction procedure removed a large number of the anomalies present in the original HOT images, and produced far better results than the previous version. However, there were still areas where it performed poorly - particularly in urban areas where a range of complex, mixed land covers gives rise to many small, but significant, HOT anomalies.

The best way to correct these was found to be by iterating the segmentation-correction process, using the corrected HOT image from the previous iteration as the input to the next segmentation and correction steps. Even over only a few iterations, the quality of the correction over urban areas improved (Figure 5.15).

Tests were performed to select the best method for deciding when to stop iterating: either 1) after a fixed number of iterations, or 2) based upon some sort of convergence in the correction process. Statistics were extracted to look for convergence, each calculated between two consecutive iterations: the number of pixels changed, the maximum pixel



(a) Original (b) Iteration 1

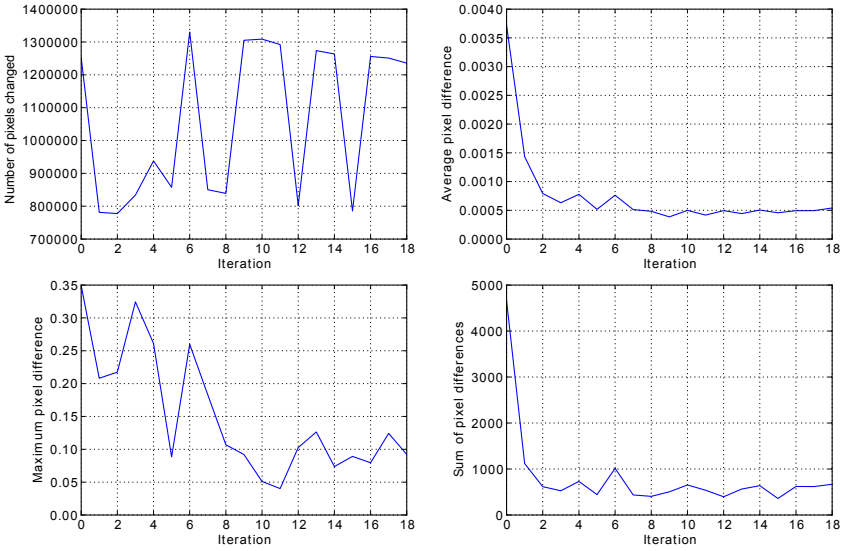
(c) Iteration 2 (d) Iteration 3

FIGURE 5.15: Example outputs over Southampton for the first three iterations of the segmentation-correction procedure, from the SotonNewForest image. The colour scale is the same for all images, with high HOT values shown in white, and low values in black.

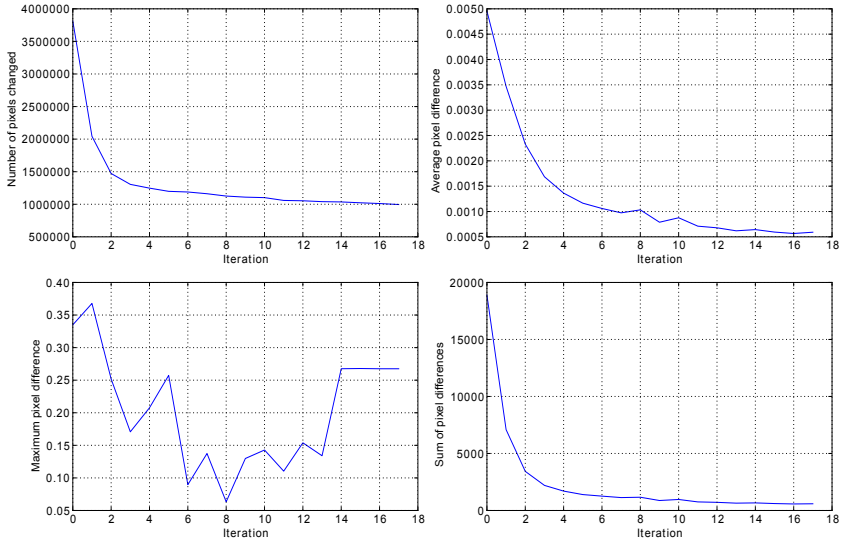
difference, the sum of all pixel differences, and the mean pixel difference (calculated with the number of pixels changed as the denominator to make the measure more robust).

Plots of these statistics over twenty iterations for several test images (Figure 5.16) showed that both the number of pixels corrected and the maximum pixel difference were very noisy (often oscillating between high and low values) and did not converge at all, even after forty iterations. The latter is due to later iterations finding a few extreme anomalies hidden within other objects, such as a bright building within a large area of green space, thus increasing the maximum pixel difference. In contrast, both the total pixel difference and the mean pixel difference showed an approximately exponential decline, gradually converging on a relatively constant value after a varying number of iterations, depending on the image, but always within twenty iterations.

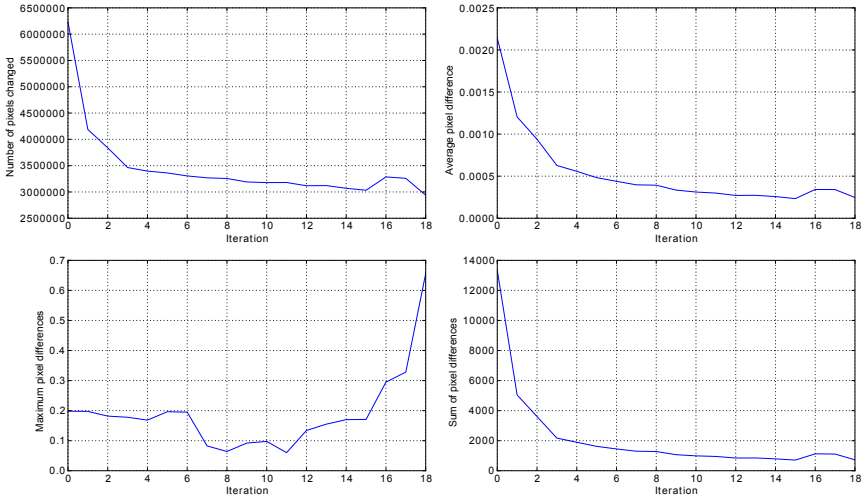
The mean pixel difference was chosen to be used as the threshold, as normalising by the



(a) Amazon

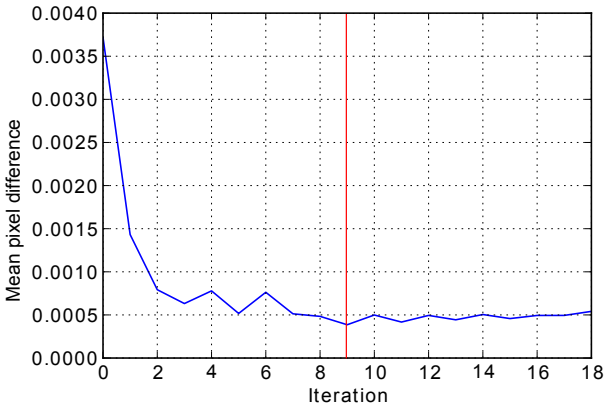


(b) Chilbolton

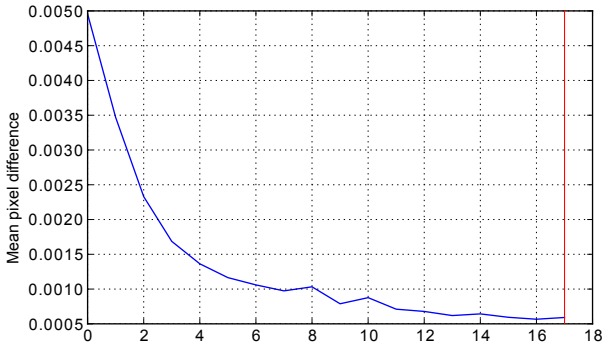


(c) Midwest

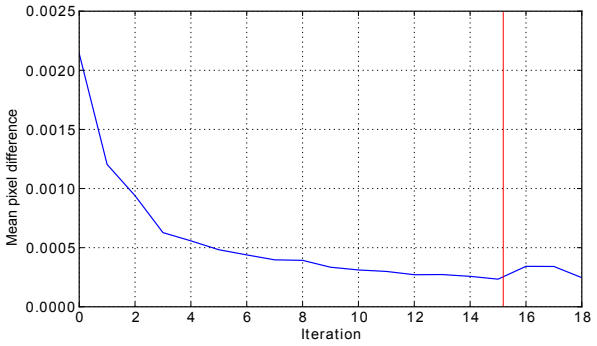
FIGURE 5.16: Statistics (number of pixels changed, mean pixel difference, maximum pixel difference and sum of pixel differences) calculated over 20 iterations of the segmentation-correction cycle.



(a) Amazon



(b) Chilbolton



(c) Midwest

FIGURE 5.17: Average pixel difference calculated over 20 iterations of the segmentation-correction cycle, with the termination iteration, according to the 10% threshold, marked in red.

total number of pixels changed to calculate the mean is likely to provide a more robust measure than using a simple sum. As the magnitude of this statistic varies between images a fixed threshold could not be used, so an individual threshold was chosen for each image, based upon the mean pixel difference between the original HOT image and the result of the first iteration. After experimentation, the threshold was set at 10% of this value, and the termination iterations are shown in Figure 5.17 as a red line.

This approach to terminating the iterative process does not always perform well. In some situations it can lead to an 'over-iteration', which can exacerbate problems with the

correction (see §5.4.8). Conversely, 'under-correction' can leave HOT anomalies in the image, but this seems to be very rare: the current configuration of termination conditions tends towards 'over-correction'.

5.4.6 Parameterisation

The segmentation/correction algorithm has two configurable parameters: T_{high} and T_{low}, the high and low thresholds for the Canny edge-detection method. These must be chosen by the user at the start of the algorithm run - normally after examining the results of segmentation with a range of parameters (Figure 5.18) - and the rest of the algorithm can be run without user intervention. During development, many other parameters were altered - such configuration for the edge-linking and pruning steps - but constant values producing good results for all images were discovered.

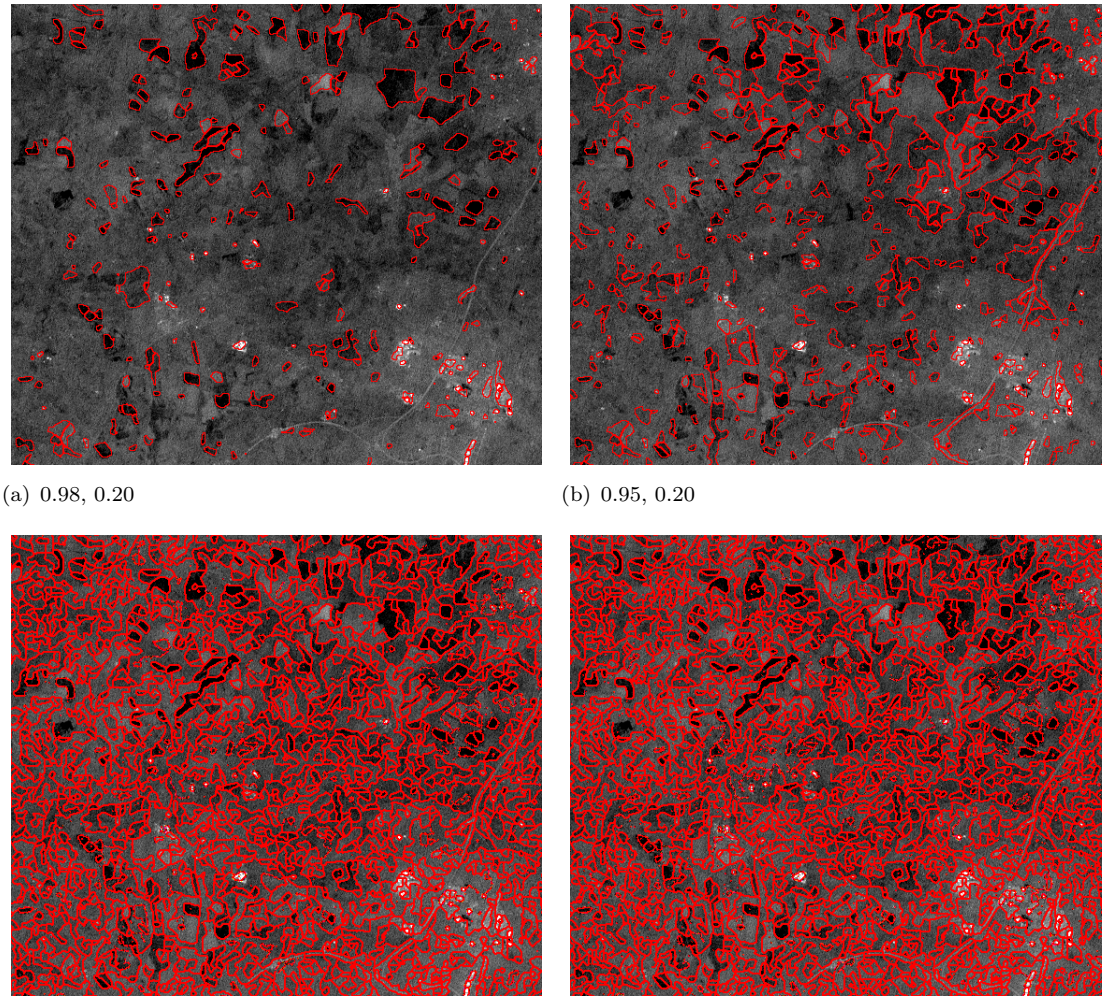
The segmentation parameters should be chosen so that segments cover all HOT anomalies, but are not created where HOT anomalies do not exist. Choosing appropriate segmentation parameters is important, as under-segmentation will cause uncorrected HOT anomalies to be present in the final result, and over-segmentation will have a broader smoothing effect on the image, potentially removing or adjusting genuine HOT differences. The exact choice of parameters can be difficult, as subtle changes in parameters can have large effects. For example, the parameters for Figure 5.18(a) and 5.18(b) are very similar (0.98, 0.20 and 0.95, 0.20), but 5.18(b) has significantly more segments. Conversely, there is a big difference in the parameters for 5.18(c) and 5.18(d) (0.85, 0.75 and 0.85, 0.50), but there is very little difference in the segmentations - as they both produce segments across the whole image. In the context of this example, parameters of 0.98 and 0.20 (Figure 5.18(b)) were chosen, as this segmentation picks up some of the HOT anomalies that are missed with a stricter parameterisation, but avoids over-segmentation.

After testing various choices of parameters, the parameters shown in Table 5.3 were chosen for each of the LEDAPS images, and these have been used from here onwards.

5.4.7 Cross-sensor applicability

The segmentation and correction algorithms were tested on both Landsat TM and ETM+ data, with the original HOT images created from a clear line based upon the LandsatAERONET TM and ETM+ datasets respectively. The development of the segmentation and correction procedure was performed using Landsat ETM+ images, and when the procedure was tested on Landsat TM images it was found to perform very poorly (Figure 5.19).

A range of potential reasons for this were investigated, principally focusing on the differences between the TM and ETM+ sensors, in terms of radiometric and geometric quality. A large Point Spread Function (PSF) effectively reduces the resolution of the image, smoothes the pixel data and, in the terminology of Strahler *et al.* (1986), makes the image 'more L-res'. Segmentation algorithms are known to perform more poorly with



(a) 0.98, 0.20

(b) 0.95, 0.20

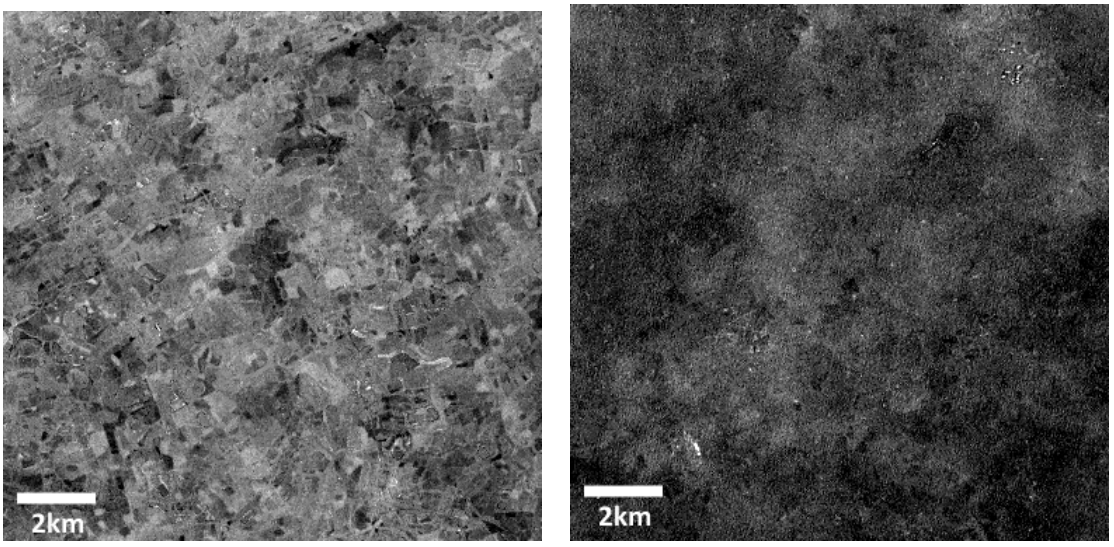
(c) 0.85, 0.75

(d) 0.85, 0.50

FIGURE 5.18: Example segmentations of a mixed urban/rural region within the SotonNewForest image, with a range of segmentation parameters. The parameterisation shown in (b) was chosen for this image.

TABLE 5.3: Segmentation parameters for each of the test images

Image	$T_{\mathbf{high}}$	$T_{\mathbf{low}}$
Amazon	0.98	0.95
Botswana	0.85	0.75
Chilbolton	0.85	0.70
London	0.85	0.70
Midwest	0.85	0.70
Namibia	0.87	0.70
Sahara	0.98	0.2
SaudiArabia	0.85	0.70
SotonNewForest	0.95	0.2



(a) Landsat TM (b) Landsat ETM+

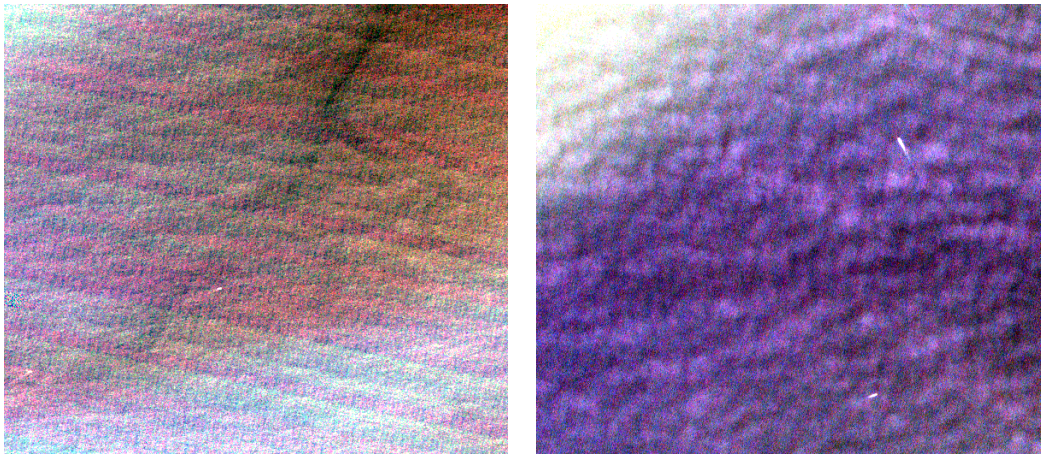
FIGURE 5.19: Example of correction results over an area in Dorset, from the SotonNewForest image, based on Landsat TM and Landsat ETM+ LEDAPS images. The colour scale is the same for both images, with high HOT values shown in white, and low values in black.

L-res data (Blaschke *et al.*, 2004). However, the TM and ETM+ sensors have very similar PSFs, with full-width half-maximum values of 36m and 34m respectively (Storey, 2001; Markham, 1985). Furthermore, the segmentation procedure seemed to perform relatively well on Landsat TM images, suggesting that it is the correction procedure which was performing poorly with these images.

Issues with noise in Landsat TM images are well-reported in the literature, particularly periodic noise such as banding, caused by differences in the optical quality between consecutive mirror scans in different directions (Masek *et al.*, 2001; Nichol and Vohora, 2004). This is visible in Figure 5.20, showing linearly-stretched RGB composites taken from an area of the English Channel in the SotonNewForest image. Applying the Maximum Noise Fraction transform (MNF; Green *et al.*, 1988) to the image emphasises the noise (Figure 5.21), showing significant banding with a period of 16-rows.

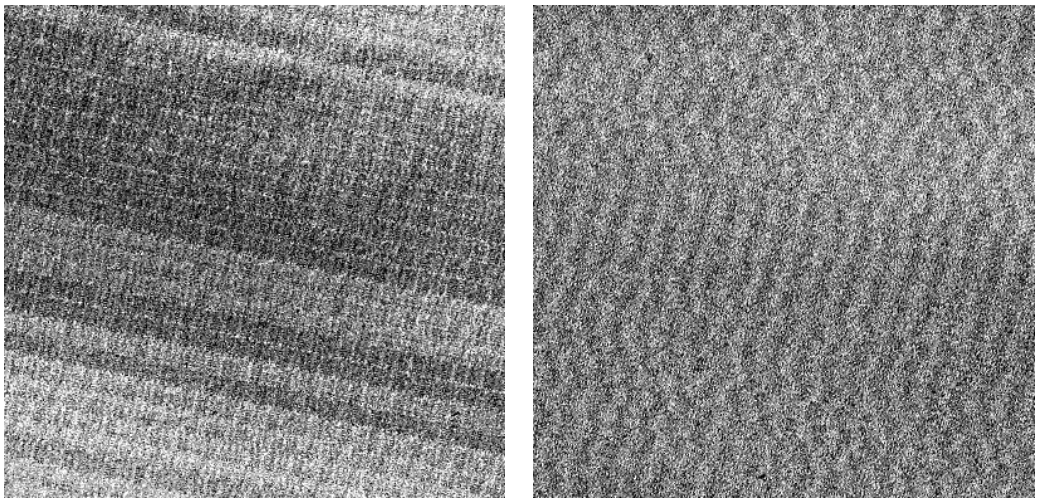
This type of coherent, periodic noise is likely to significantly affect the quality of the correction procedure. By perturbing the values of the pixels within and around the borders of the objects, it results in a significant increase in edge effects, as visible in Figure 5.19(a). Appendix B shows an example of this on a very small image, allowing manual calculation and explanation of the issue.

In modern terms, both Landsat TM and Landsat ETM+ are very old sensors, originally designed in the early 1980s and early 1990s respectively. There was a significant improvement in data quality between TM and ETM+, and that improvement has continued with the Operational Land Imager (OLI) instrument launched on Landsat 8 in February 2013, which has a better radiometric resolution (12-bit as opposed to 8-bit) and



(a) Landsat TM (b) Landsat ETM+

FIGURE 5.20: Linearly-stretched RGB composites from Landsat TM and Landsat ETM+ over part of the English Channel, showing the banding present in Landsat TM data.



(a) Landsat TM (b) Landsat ETM+

FIGURE 5.21: The fifth Maximum Noise Fraction band from MNF transforms performed on TM and ETM+ images of the English Channel. The MNF transforms were calculated from a uniform area of deep water in the images.

a significantly improved signal-to-noise ratio (Irons *et al.*, 2012). Most other modern optical sensors, including the High-Resolution Visible (HRV) instrument on the SPOT series of satellites, and the instruments planned for Sentinel-2, produce data of similarly high quality.

At the beginning of this research (2011) only Landsat 5 and 7 were operational, but Landsat 8 was launched in February 2013. The algorithm development was not extended to work with Landsat 8 at this stage due to limited availability of data in the early stages of the mission, the lack of a Landsat 8 LEDAPS product which was needed to create the simulated images, and the potential issues that increasing the scope of the project could have on timely completion of the research.

Landsat 8 has now been fully operational for over 18 months and is producing high-quality data. LEDAPS products are still unavailable for Landsat 8 (as of September 2014 they were estimated to be released in January 2015; personal communication, USGS Customer Support Team), but as the first version of the algorithm has now been developed there is less of a need to produce simulated images from LEDAPS products. Landsat 8 has many advantages which make it likely that the algorithm will perform better on Landsat 8 imagery than that from Landsat 7:

- **Signal-to-Noise ratio:** Landsat 8 has a significantly improved signal-to-noise ratio (SNR) when compared to Landsat 7 and previous sensors. For example, the SNR of the Landsat 7 blue band is 40, the requirement for Landsat 8 is 130, and the actual performance is 350 - over eight times better (Irons *et al.*, 2012). This will reduce the presence of noise in images, and thus limit the potential for noise (particularly in the blue band) to be interpreted as changes in AOT, and thus improve the accuracy of the AOT retrieval.
- **Radiometric resolution:** Landsat 7 used 8-bit quantisation, providing only 256 different levels of brightness for each pixel in a scene. Landsat 8 improves this to 12-bit, allowing 4096 brightness levels (Irons *et al.*, 2012). This will allow more accurate capture of subtle differences in brightness between pixels, thus improving the accuracy of the AOT retrieval.
- **'Deep Blue' band:** Landsat 8 adds two extra bands, one of which is a 'Deep Blue' band at the shorter-wavelength end of the blue region of the spectrum (approximately 430–450nm). The HOT is based upon both the wavelength dependence of aerosol scattering and the high correlation between blue and red reflectance. Using blue and red bands, but increasing the wavelength difference between them - as would be the case when using the 'Deep Blue' band - should increase the magnitude of the aerosol scattering effects, and thus improve the accuracy of AOT retrievals.

Given these improvements, it is very unlikely that the poor correction seen with Landsat TM data will be found with Landsat 8 data - or any other modern data - and the method should be able to be adapted to new sensors as they are developed and launched. The sensor requirements for the new algorithm are far less restrictive than most AOT retrieval algorithms, and the algorithm simply requires:

- A blue band (with wavelengths $\approx$ 400–500nm, with lower wavelengths likely to perform better)
- A red band (with wavelengths $\approx$ 600–700nm)
- Sufficient past data to produce a comprehensive version of the LandsatAERONET dataset for the new sensor (the original LandsatAERONET datasets were generated from a year of data, but more data may be needed for satellites with long revisit periods)
- A known Spectral Response Function to use for Py6S simulations

These requirements are met by a wide range of sensors with a variety of resolutions: here the focus is on high-resolution sensors, as a wide range of methods already exist to retrieve AOT estimates from lower resolution sensors such as MODIS and MISR. Potential sensors are listed in Table 5.4: their resolutions range from 2m to 30m and revisit periods from 1 day to 16 days. Although the algorithm could conceptually work with very high resolution data (for example, Pleiades data at 2m resolution) it would probably have significantly higher error (as shown in §4.4.4, where the resolution significantly affects the spread of points around the Clear Line). Furthermore, it is unlikely that AOT measurements need to be produced at such a high resolution, so work should focus on coarser resolution data such as that from Landsat 8, SPOT and Sentinel-2. Both Landsat 8 and Sentinel-2 have an extra blue band centred at a lower wavelength which is designed for aerosol remote sensing: using this band in place of the standard blue band should improve the results of the algorithm. Sentinel-2 also provides freely-available data with only a five day revisit period, thus allowing retrieval of AOT with both high spatial and high temporal resolution.

TABLE 5.4: Details on potential sensors with which the new algorithm could be developed to work. Data from ITC (2014); Drusch *et al.* (2012); Powell *et al.* (2007).

Sensor	**Spatial Resolution (m)**	**Revisit period**	**Data freely available?**
Landsat 8	30	16 days	Yes
ALI	30	16 days	Yes
SPOT-5	10	2–3 days	No
SPOT-6 & 7	6	1 day	No
Pleiades	2	1 day	No
RapidEye	5	1 day	No
Sentinel-2	10	5 days	Yes

To extend the algorithm to work with a new sensor, the following work must be carried out:

1. **Develop a new sensor version of the LandsatAERONET dataset:** This requires repeating the process described in §4.2.4 for the new sensor, including acquiring a wide range of images, extracting pixels over AERONET sites, atmospherically-correcting them, collecting metadata and checking for any remaining atmospheric contamination. The manual land-cover classification from Google Earth imagery cannot be automated and will be the most time-consuming part of the process, but the majority of the rest of the process is automated and can be easily altered to work with any sensor.
2. **Implement an automated cloud-screening process:** Cloud screening is an essential part of the pre-processing (see §3.4.4), but the FMask algorithm used in the new algorithm will only work with Landsat data. A broad range of other cloud screening algorithms exist (Irish, 2000), and a suitable algorithm must be implemented for the new sensor. Alternatively, manual cloud-screening can be

performed, but this is both time-consuming and generally of lower accuracy.

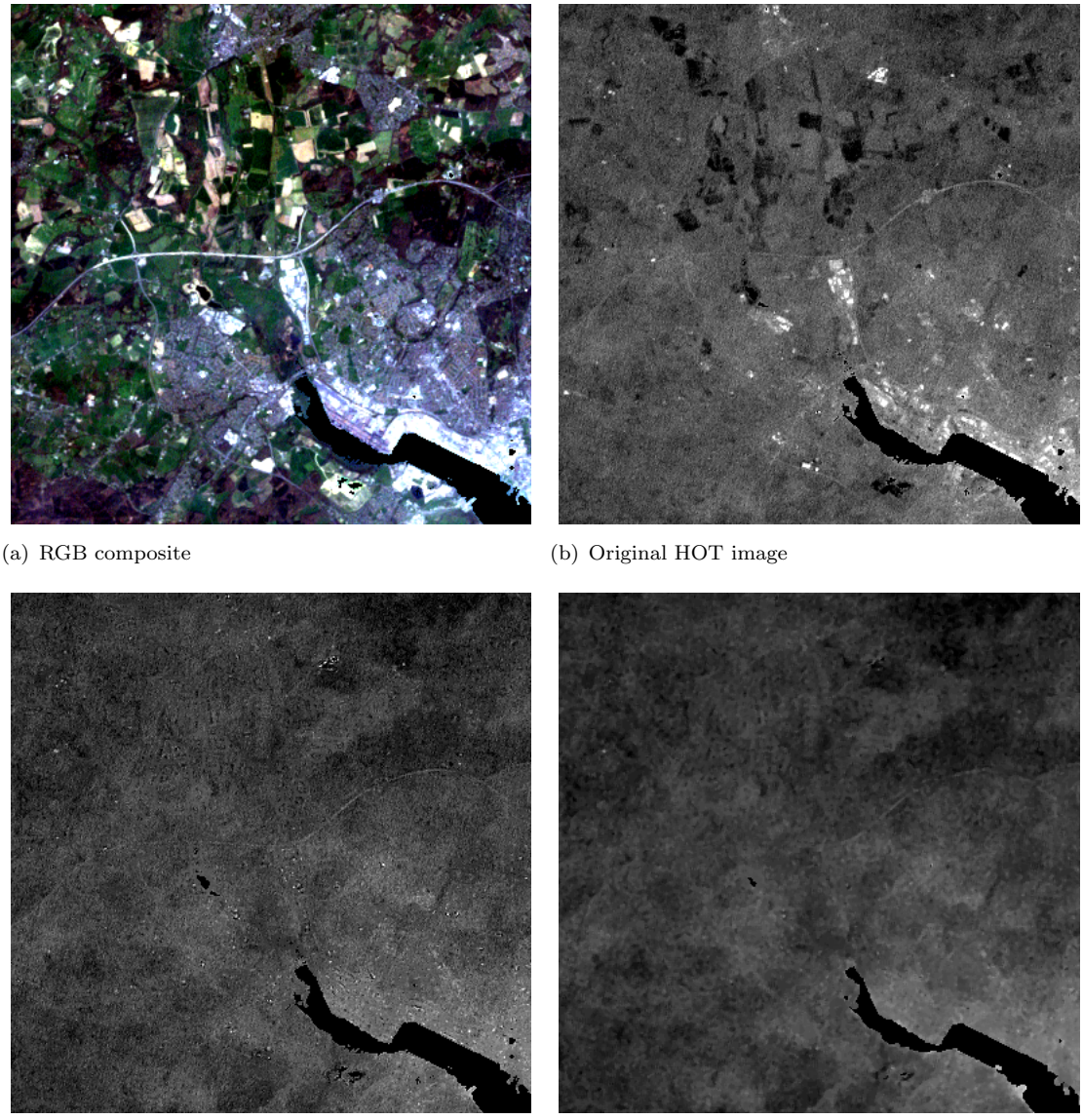
3. **Configure Py6S simulations for new bands:** Py6S simulations are used in various stages of the AOT retrieval process, and these simulations must be performed with the Spectral Response Function (SRF) for each of the bands used. A wide range of SRFs are built-in to the underlying 6S model, or have been added to Py6S, and it is easy to add further SRFs if necessary.
4. **Test the rest of the algorithm, altering where necessary:** Depending on the differences between Landsat ETM+ and the new sensor, changes may be necessary in the rest of the algorithm. This is likely to be limited to changes in the internal parameterisation of the various stages (such as changes in the parameters for edge linking or the Monte Carlo Regression), but for radically different sensors it could involve the addition of extra processing stages.

5.4.8 Evaluation of corrected HOT images

The examples images in Figure 5.22, and the profiles in Figure 5.23 show that the correction significantly improves the land-cover issues within the HOT image, but that the corrected images still have some individual pixel anomalies. These are remnants of objects which were not entirely identified or corrected by the algorithm. These single pixels anomalies cannot be corrected by the algorithm, as they will not be detected by the edge-detection procedure. A simple median smoothing function applied using a 3x3 pixel moving window removes these small anomalies, improving the correction, and so this has been added to the correction procedure as a final post-processing step.

He *et al.* (2010) produced boxplots showing the variability of HOT values over a wide range of different land covers (classified from the image using an unsupervised classification algorithm) to show the effectiveness of their HOT correction procedure. This is less appropriate to apply here as the average HOT value for any land cover should not be zero (as every pixel in an image will have some aerosols over it) and the average HOT should not be the same for each land cover (as it is likely that different land covers will have different amounts of pollution above them, and thus the HOT values should be different). Thus, we would not expect to see a large difference between pre-correction and post-correction boxplots, resulting in an almost 'flat' series of boxplots, as found by He *et al.* (2010, figures 4 and 10 respectively). However, the HOT correction procedure should result in a small amount of 'flattening' of the series of boxplots, and a reduction in the variance of the HOT over each land cover. This is shown to be the case in Figure 5.24.

The correction algorithm is a significant improvement on previous efforts to correct land-cover issues in the HOT, but its performance depends on the situation in which it is applied. Figure 5.25 shows examples of the original HOT and the corrected HOT for some situations where the correction method performs well. Even though these are the correction results from LEDAPS images, they do not have an entirely uniform HOT value after correction. This does not necessarily mean that the algorithm has not performed well: it is likely that the LEDAPS per-pixel atmospheric correction did not remove all of



(a) RGB composite

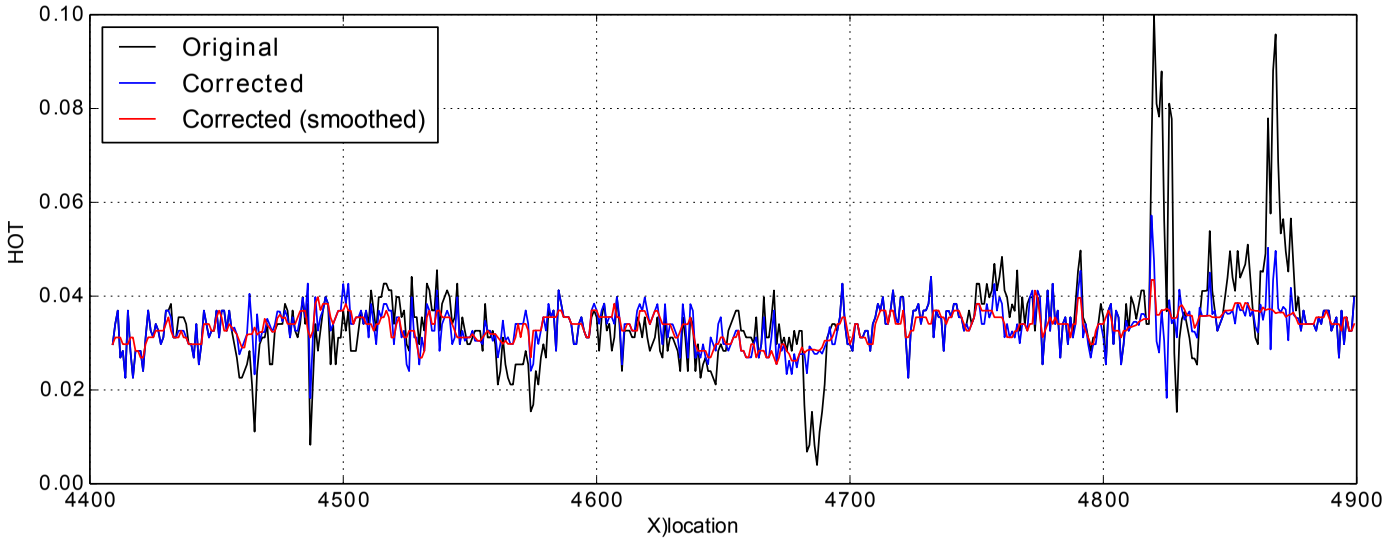
(b) Original HOT image

(c) Corrected HOT

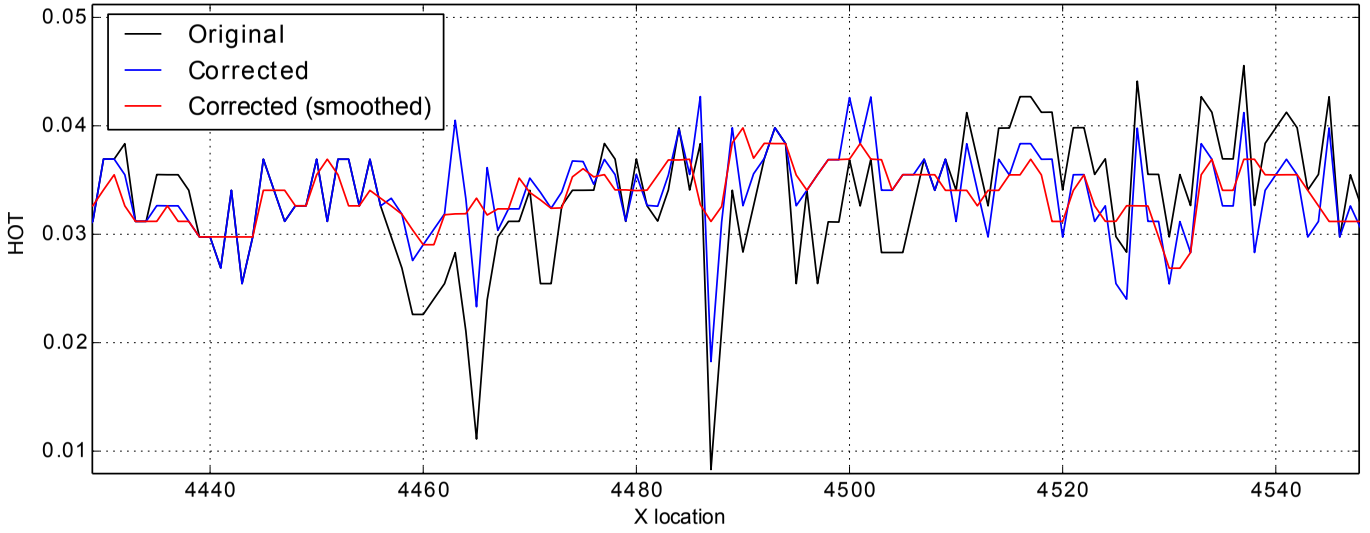
(d) Corrected HOT with 3x3 median smoothing

FIGURE 5.22: Example outputs of the HOT correction process on a subset of the SotonNewForest LEDAPS image covering the western side of Southampton. The colour scale is the same for all HOT images, with high HOT values shown in white, and low values in black

the spatially-variable atmospheric effects. For example, higher HOT values around central Southampton are visible in Figure 5.25(b), but this is not due to poor correction of land cover anomalies - comparing with Figure 5.25(a) shows that all of these have been removed - it is due to a genuine higher AOT over central Southampton which was not corrected for by LEDAPS, probably due to the lack of DDV over this area. Similar effects are visible in the other examples in Figure 5.25, most of which are likely to be due to uncorrected atmospheric effects, rather than poor correction. Taking these genuine AOT effects in to consideration, the correction procedure does well at removing all obvious land-cover effects, particularly in the SotonNewForest example (Figure 5.25(b)). The correction complex urban-rural mosaic in Figure 5.25(d) is corrected relatively well, with

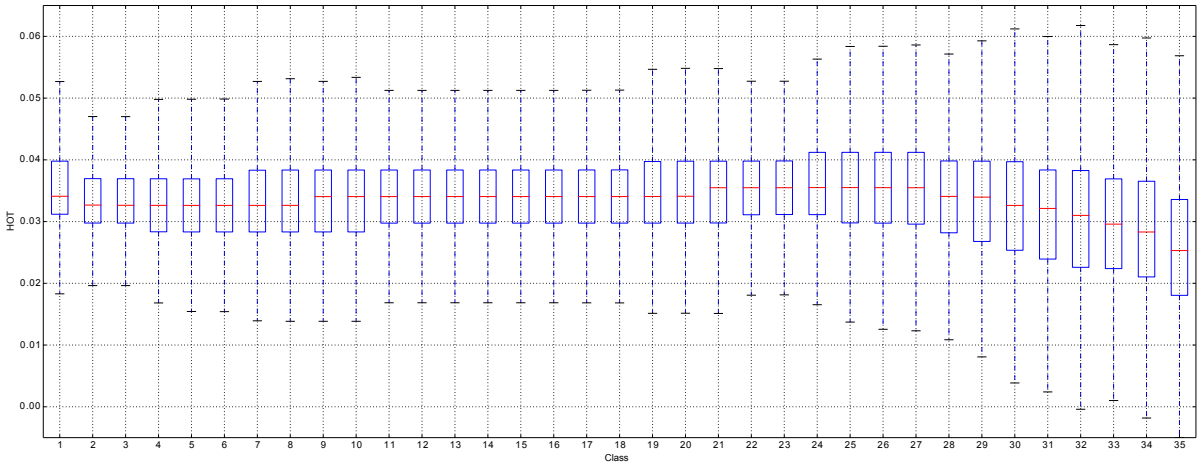


(a) Overview

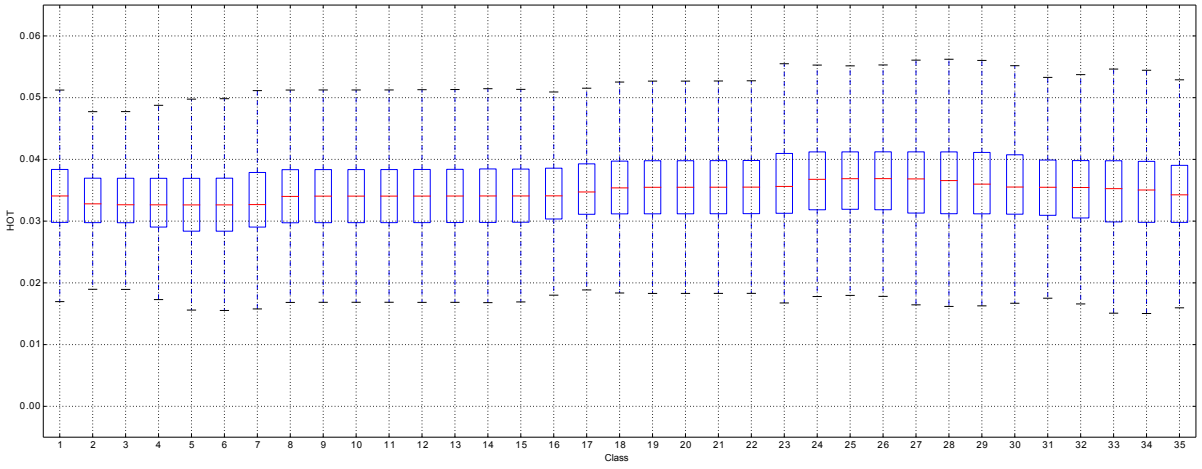


(b) Zoomed section

FIGURE 5.23: Horizontal profiles of the original HOT, corrected HOT and corrected-smoothed HOT from the SotonNewForest image. (a) shows an overview, with (b) showing a section in greater detail.



(a) Original HOT



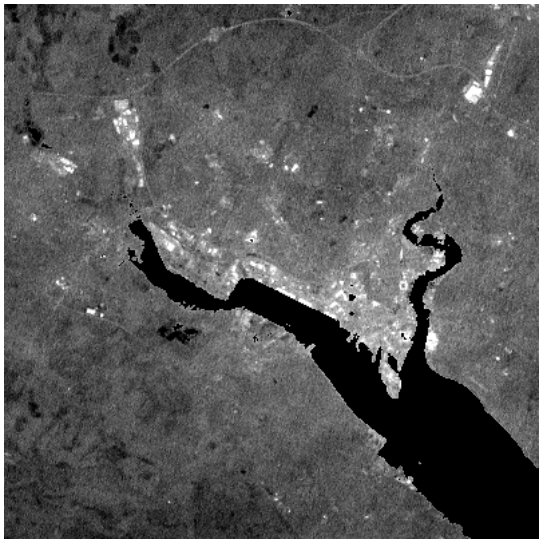
(b) Corrected HOT

FIGURE 5.24: Boxplots showing HOT values for 35 different land-cover classes within the SotonNewForest image, shown for (a) the original HOT and (b) the corrected HOT. Note the reduction in range, and alignment of all classes around a central value in (b), when compared to (a).

some residual effects visible from negative HOT anomalies over fields. It is difficult to assess the correction of Central London (Figure 5.25(f)) as there are likely to be significant changes in AOT over small spatial scales (the 'urban canyon' effect Vardoulakis *et al.*, 2003), which LEDAPS is unlikely to have corrected well for, due to a lack of DDV. However, obvious HOT anomalies such as Hyde Park have been corrected well, suggesting that the correction is performing well overall.

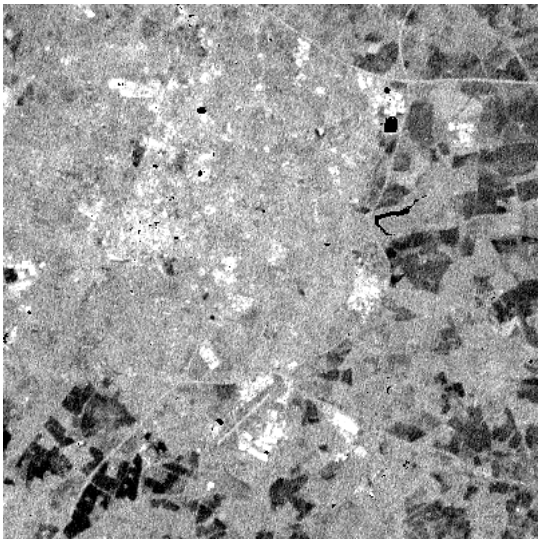
Figure 5.26 shows examples of issues with the HOT correction, caused by limitations in the current implementation of the algorithm. The majority of these can be improved in future versions of the algorithm. The original HOT images have not been shown here: when the correction is poor, the issue is normally easily noticeable from the corrected image.

Figures 5.26(a) and 5.26(b) show corrections that show contamination by edge-effects.

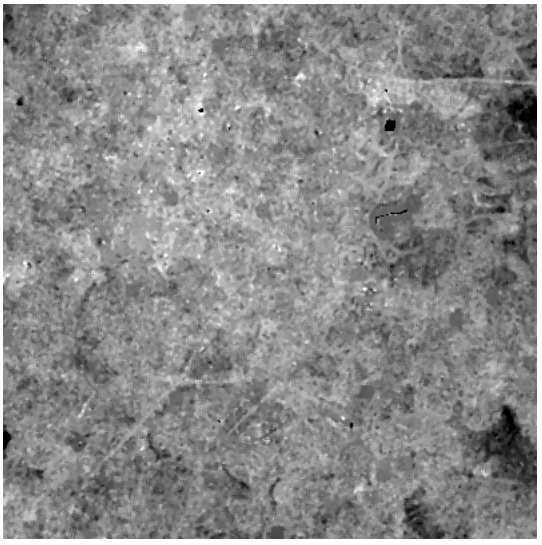


(a) Southampton, with surrounding rural area, from the SotonNewForest image (Original)

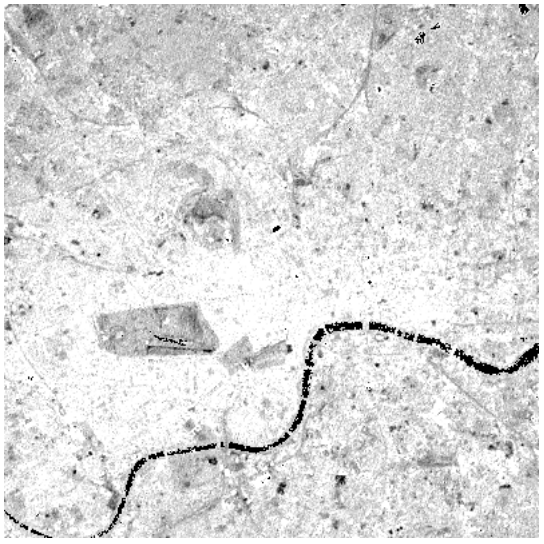
(b) Southampton, with surrounding rural area, from the SotonNewForest image (Corrected)



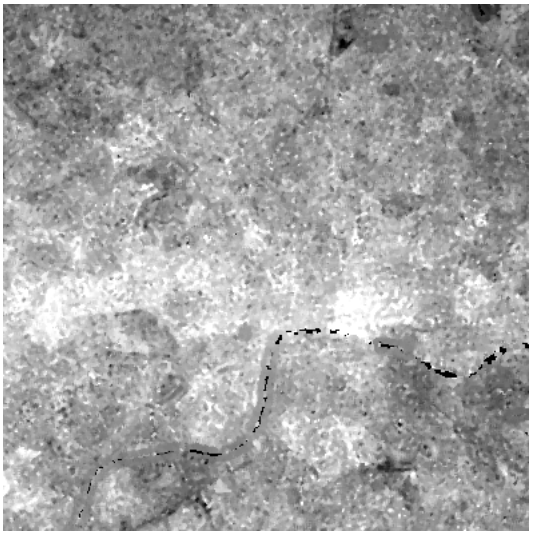
(c) Complex urban-rural area from the Chilbolton image (Original)



(d) Complex urban-rural area from the Chilbolton image (Corrected)

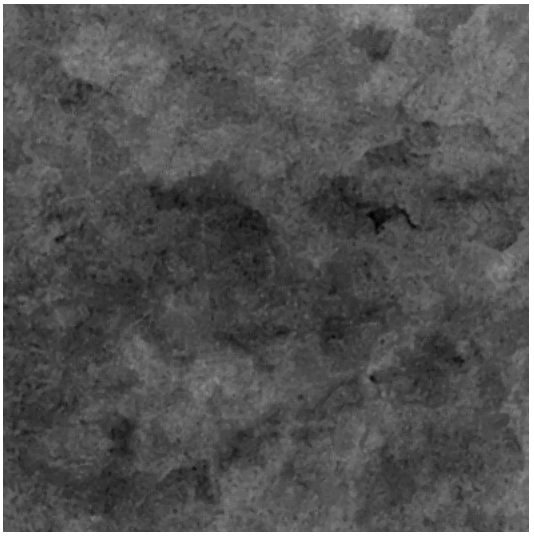


(e) Central London, from the London image (Original)

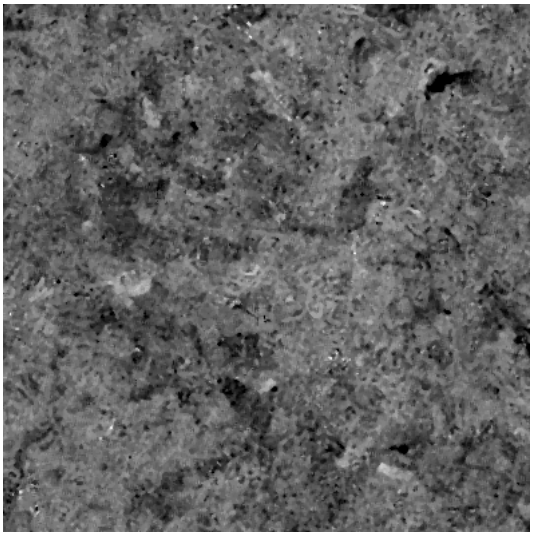


(f) Central London, from the London image (Corrected)

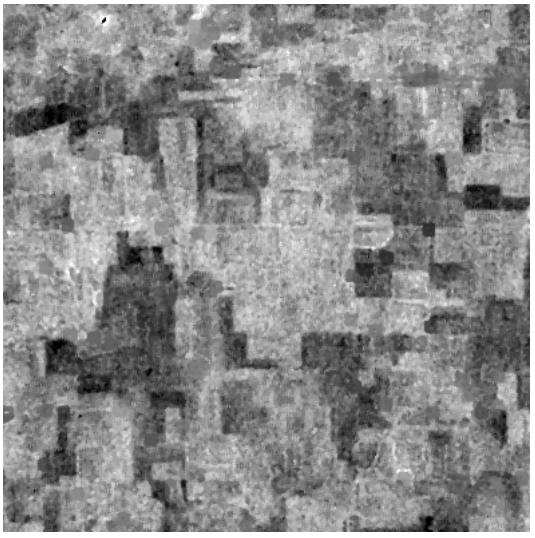
FIGURE 5.25: Examples of good HOT image corrections. In all images, high HOT values are shown in white, with low values in black.



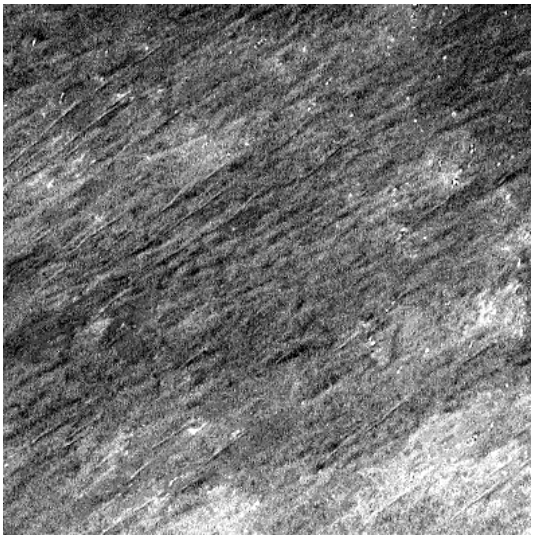
(a) Part of the New Forest in the SotonNewForest image



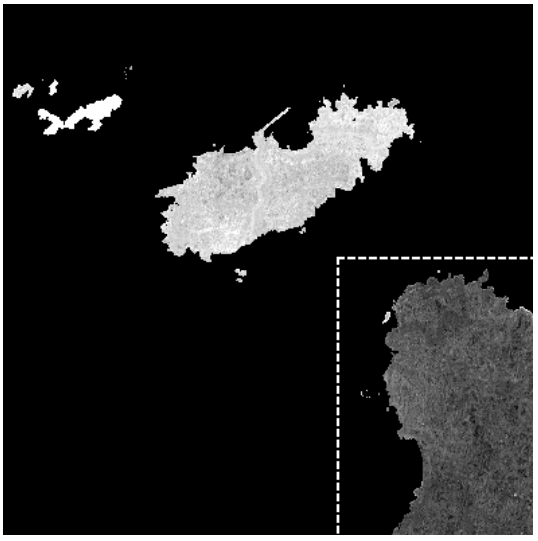
(b) Urban/Rural fringe in the London image



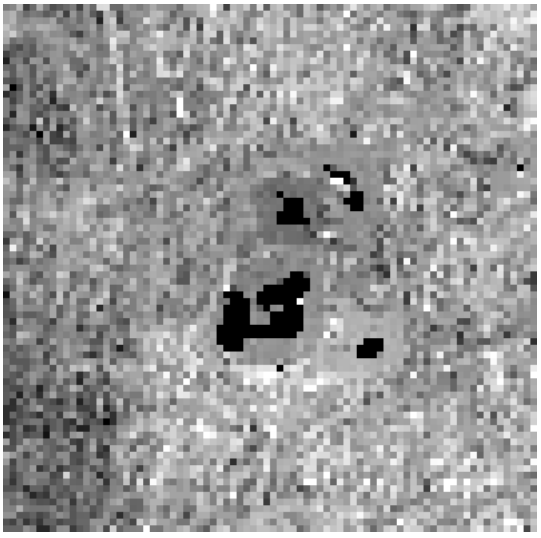
(c) Agricultural area in the Midwest image



(d) Linear dunes in the Sahara image



(e) The island of Alderney, with the French mainland (inset), from the SotonNewForest image



(f) A zoomed-in section of the Chilbolton image

FIGURE 5.26: Examples of issues with the HOT image correction In all images, high HOT values are shown in white, with low values in black.

This is particularly noticeable in Figure 5.26(b), where the edges of some of the objects used in the correction can be seen. In Figure 5.26(a) the effect is less noticeable, but relatively sharp borders between the darker and lighter areas of the corrected image can be seen, suggesting edge effects that have been removed by the median smoothing procedure. A further issue which can be seen in this example is the 'contamination' of parts of the corrected image by dark HOT anomalies. That is, most of the dark areas do not result from genuinely lower HOT values, but from the correction of a cluster of negative HOT anomalies. As the correction process uses border pixels to calculate the magnitude of correction to apply, a large group of anomalies bordering each other can start a feedback loop, where dark border values mean that the area is never fully corrected. During the iteration process it is possible - though rare - for this larger anomaly to spread to cover a broader area of the image.

Figures 5.26(c) and 5.26(d) show poor corrections caused by poor segmentation. In Figure 5.26(c) the segmentation did not detect the edges of all of the fields which caused negative HOT anomalies, and thus they weren't corrected. A similar issue occurred in Figure 5.26(d) which had no self-contained objects that could be extracted by the segmentation method. The linear dunes in the image had different HOT values due to the angular and surface differences between the stoss and lee side of the dunes, and as the segmentation algorithm does not extract these as self-contained objects, they can't be corrected, and so remain in the image. The broader-scale brighter and darker areas of the image are due to changes in the geology of the sand, but as these vary gradually - as one type of sand mixes gradually with another - they are not detected as objects by the algorithm, and so are not corrected.

The issues shown in Figures 5.26(e) and 5.26(f) have the same root cause: no data areas. Figure 5.26(e) shows that the island of Alderney has been corrected to have far higher values than the French mainland. The tiny island of Burhou nearby has been corrected to an even higher value. As water is masked at the start of the algorithm, both of these areas are entirely surrounded by 'no data' pixels. In Figure 5.26(f) the areas around the masked area - in this case, clouds - have become unnaturally uniform. These effects are both caused by the use of pixel information surrounding an object to determine how much to alter the pixels within the object. If there are a limited number of data pixels surrounding the object - or, in the case of an island, no data pixels - then the correction value will be skewed. This effect tends to get worse over multiple iterations of the method, and the values become more skewed.

Although there are many advantages associated with the OBIA approach to land cover correction, there are also disadvantages to this approach. The principal disadvantage is the relatively arbitrary nature of the parameterisation: T_{high} and T_{low} have to be estimated for each image location, and do not seem to be directly related to external factors such as the land-cover mosaic of the image. These parameters are important as the whole correction system is tightly-coupled - meaning that a small change in one of the segmentation parameters (or in any of the algorithms themselves) can have a large impact

on the resulting correction. The overall effect of this is that the algorithm and parameterisation may require 'tweaking' for use with new sensors, and potentially even for certain images from existing sensors. The approach taken in this chapter was designed to reduce the impact of these issues - which often cause problems for object-based approaches - by testing the algorithms on multiple LEDAPS images during the development phase, thus ensuring that the algorithm has as broad an applicability as possible.

In summary, the object-based HOT correction procedure works well where the anomalous HOT values form clearly-defined objects which can be segmented and then corrected. This is often the case in urban areas, where most anomalies will be due to buildings, and agricultural areas, where most anomalies will be due to crops in fields. The correction performs poorly in areas where landcovers mix and merge gradually - such as in deserts. There are also issues dealing with areas of no data, as there is less data surrounding each object for the correction algorithm to use to determine the correct offset to apply.

5.5 Conclusion

In this chapter, three major improvements have been made to the HOT:

1. The use of partially-corrected reflectance values (corrected for all atmospheric effects apart from aerosols), rather than raw DN values as used by Zhang *et al.* (2002a). This removes residual atmospheric effects and should make the HOT values of different images more comparable.
2. The estimation of the Clear Line using a Monte Carlo regression based upon the proportion of land covers in the image. This removes the need to estimate the Clear Line from a clear part of the image, which is impossible when trying to detect aerosols rather than haze, as no part of the image will have no aerosols.
3. The application of an HOT correction procedure based upon Object-based Image Analysis. This performs a similar correction to that used in the Advanced HOT (He *et al.*, 2010), but in a far more robust and effective manner.

With these improvements, the HOT is ready to be used to estimate AOT. The relationship between HOT and AOT, and the development of methods to model this relationship, will be examined in the next chapter.

Chapter 6

Investigating the HOT-AOT relationship

6.1 Introduction

The previous chapter developed an improved version of the HOT, designed to be applied to estimate AOT. In this chapter the relationship between these HOT values and AOT is investigated, and a method is developed to estimate AOT from HOT. First, the simulation of test images in which the AOT is known is described, followed by an investigation of the HOT-AOT relationship, and development of a method to model this for a given image. As this is the final stage in the development of the AOT retrieval algorithm, the chapter finishes with an overview of the whole algorithm, including its inputs, main processing stages and outputs.

6.2 Simulation of test images

Landsat images where the AOT is known for each pixel are needed to allow investigation of the HOT-AOT relationship. As discussed previously, there are currently no methods for deriving true per-pixel AOT from Landsat data, and thus there were two options: use lower resolution data, or produce simulated images. There are significant issues with the use of lower resolution data, as due to the scale differences the results are unlikely to be applicable to Landsat-resolution data. Therefore, simulated data were produced by combining the LEDAPS surface reflectance images with random AOT images to produce an image where the AOT for each pixel is known.

In all of the following, AOT refers to AOT_{550}: the Aerosol Optical Thickness at 550nm. The 6S model only allows parameterisation of the AOT at 550nm, therefore all simulations are based upon AOT_{550}, and the HOT-AOT relationship is actually the HOT-AOT_{550} relationship. Thus, the final output from the algorithm will be AOT_{550}.

6.2.1 Simulation of random AOTs

A key assumption of the HOT correction method is that the AOT varies smoothly across an image, whereas land cover effects produce sharp boundaries. Thus the random AOTs used in simulating the test images must be spatially-correlated. There are many methods to produce images containing spatially-correlated values, but many implementations fail to produce images which can cover a whole Landsat image (approximately 7,000 x 7,000 pixels, but varying between scenes) due to computational inefficiencies in the algorithms.

The fields package (Nychka *et al.*, 2014) for the R programming language (R Core Development Team, 2014) was used to generate a random Gaussian image using the Wood and Chan (1994) algorithm. The Matérn covariance function (Matérn, 1986) with $\theta = 0.7$ and $v = 0.9$ was used to define the spatial correlation and data were produced on an 8000 x 8000 pixel grid (this size was chosen so that the random image would be larger than any potential Landsat image). This was very computationally expensive, in terms of both CPU time and memory usage, and so only one random image was created, and this was used for all simulations.

The values in the resulting image are distributed according to the standard normal distribution ($\mu = 0$ and $\sigma = 1$), an unrealistic distribution for AOT. To find a realistic AOT distribution, the distribution of global AOT values was assessed by combining all AERONET Level 2.0 measurements from every site ever operated. This showed a global mean of 0.19, with 1st and 99th percentiles of 0.01 and 1.08 respectively. A key aspect of this AOT distribution is that all values are > 0, as a negative AOT is physically impossible. It is challenging to find a statistical distribution which can be parameterised to have a peak at around 0.2, but with no values < 0. A log-normal distribution was chosen as the most suitable, and the normal variates were convert to log-normal variates by the following formula:

$$X' = \exp(0.4X - 1.3) \tag{6.1}$$

A comparison of the distributions of the random AOTs and all AERONET data are shown in Figure 6.1: the peak value of the random AOTs is lower, but it was found to be impossible to raise this value without part of the distribution becoming negative, so this distribution was accepted. The resulting random image had a 'cloud-like' look, and varied smoothly at a realistic AOT magnitude (Figure 6.2).

6.2.2 Simulation of images with AOTs

Simulated radiance images were created by combining the random AOT image with each of the LEDAPS surface reflectance images. Py6S (Wilson, 2012) was used to simulate the top-of-atmopshere radiance for each pixel based upon the ground reflectance (from the LEDAPS image), the AOT (from the random AOT image), and a configurable aerosol type. The aerosol types used are those built-in to the 6S model (Continental, Maritime,

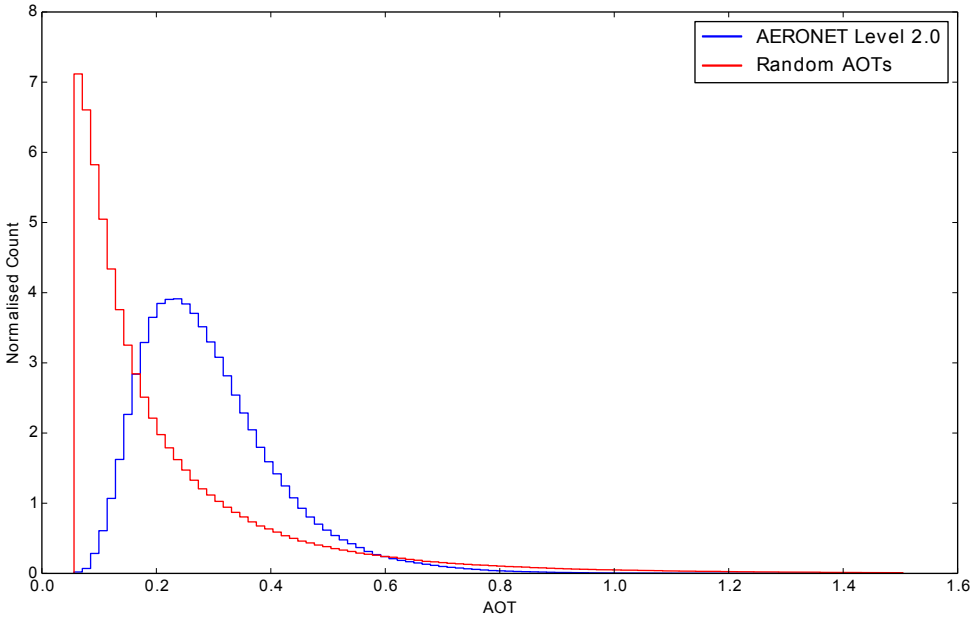


FIGURE 6.1: Comparison of the normalised distributions of all AERONET Level 2.0 data, and the generated Random AOT image.

Desert, Stratospheric and Urban), discussed in more detail in Vermote *et al.* (1997). All other Py6S parameters were set to standard values for mid-latitude areas.

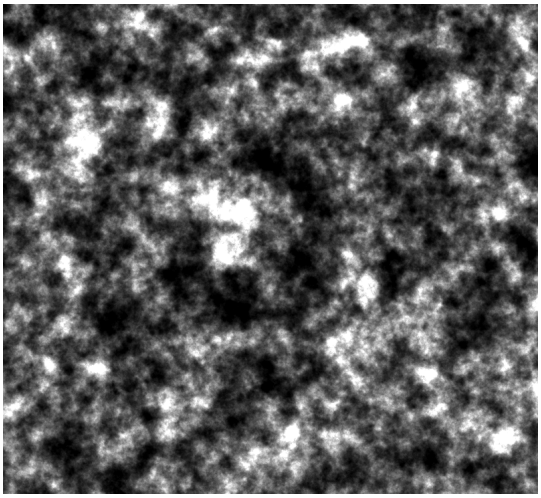
Running Py6S for a single parameterisation (one band, aerosol type, AOT and reflectance) takes approximately four seconds. Running this for each of the 38 million pixels in a Landsat scene would take almost five years, so it is essential to improve the computational efficiency of this simulation. Lookup Tables (LUTs) were used to reduce the time taken significantly, by pre-computing the results for a fixed set of parameters and then performing the simulation by interpolating between these parameters as needed.

A LUT was created for each band and each aerosol type, using two dimensional linear interpolation between values simulated at a range of AOTs (0.0–2.0 in steps of 0.1) and reflectances (0.0–1.0 in steps of 0.1). The LUT creation takes around an hour for each aerosol type, but this only needs to be done once, and then all 38 million pixels can be processed in under twenty seconds.

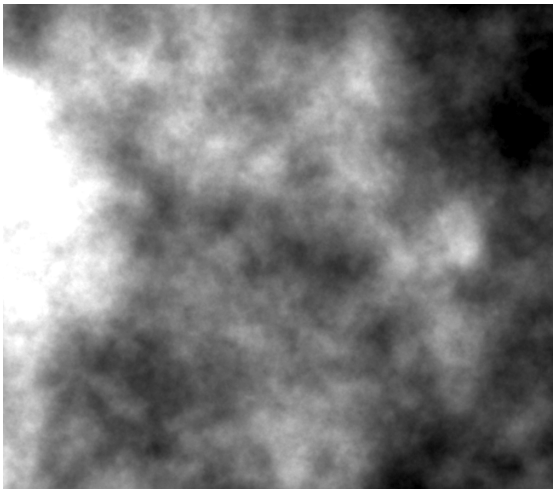
Using these LUTs, simulations were performed using each aerosol type for all of the LEDAPS test images. Example simulated images are shown in Figure 6.3. These simulated images were then processed through the algorithm described in the previous chapter, involving correction to pseudo-BOA reflectance, estimation of the Clear Line, and then creation and correction of the HOT image.

6.3 Assessing the HOT-AOT relationship

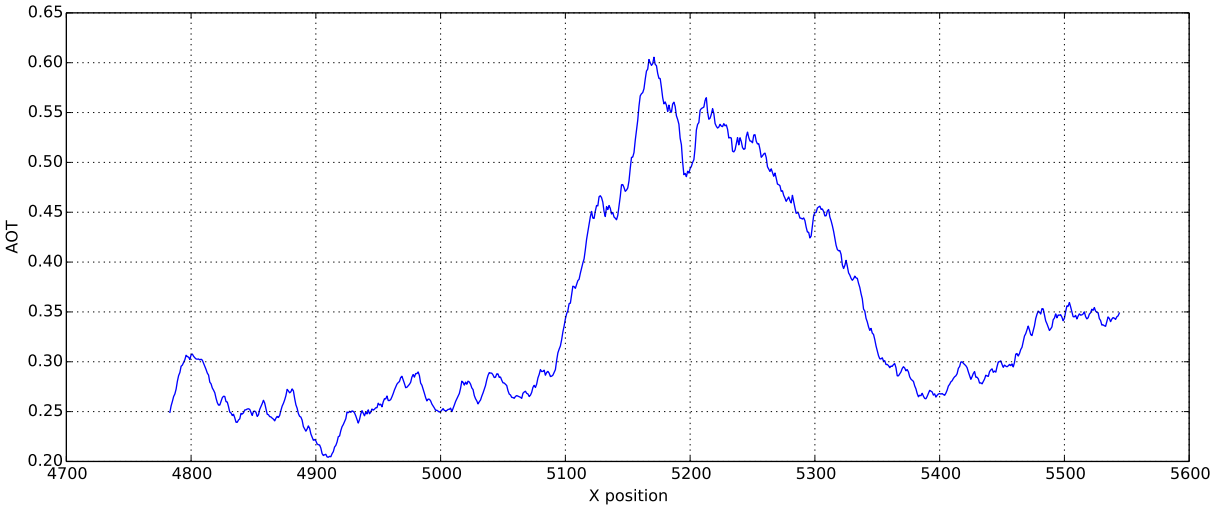
Plotting the HOT values against the AOT values for each of the simulated images (Figure 6.4), and examining correlations between these values (Table 6.1) shows that there



(a) The whole Random AOTs image



(b) A zoomed-in segment of the Random AOTs image

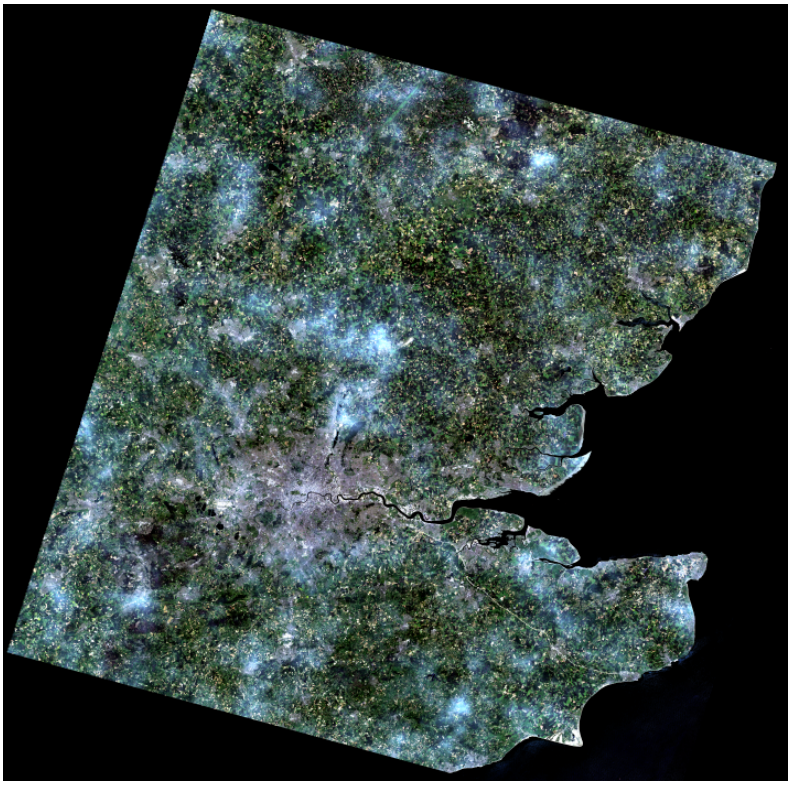


(c) A profile across the Random AOTs image, showing the smoothly, but realistically, varying values

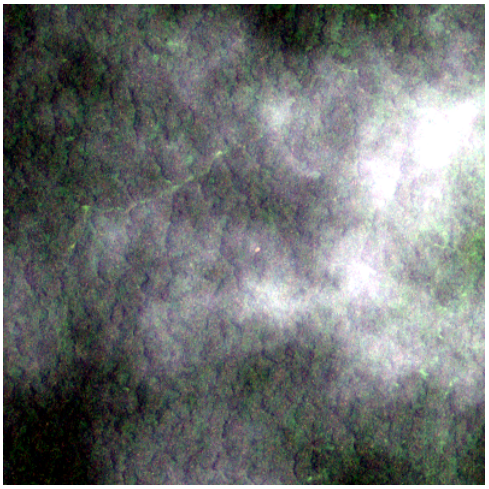
FIGURE 6.2: The Random AOT image

appears to be an approximately linear relationship for most images. Some, such as Amazon and London, have a strong correlation and a narrow cloud of points on the graph. Others, such as Botswana and Namibia, have a far weaker correlation with a very broad cloud of points. The accuracy of the estimation of AOT will be significantly better for those images with a narrow point cloud, and so it is important to understand what controls the breadth of the point cloud. This closely related to the question investigated in §4.4.2, where the breadth of the point cloud of clear pixels in blue-red feature-space was investigated, and the main control on the position of the points was found to be land cover. In the case of the HOT-AOT relationship the main control is also land cover: in this case, issues with land cover that were not removed by the HOT correction procedure.

The purpose of the HOT correction algorithm is to move anomalous pixels to the Clear Line, and therefore reduce the range of HOT values for a given AOT value. Thus, a high quality correction process will produce a better HOT-AOT relationship. This is



(a) London, Continental aerosol type



(b) Amazon, Continental aerosol type

(c) SotonNewForest, Continental aerosol type

(d) SotonNewForest, Maritime aerosol type

(e) SotonNewForest, Urban aerosol type

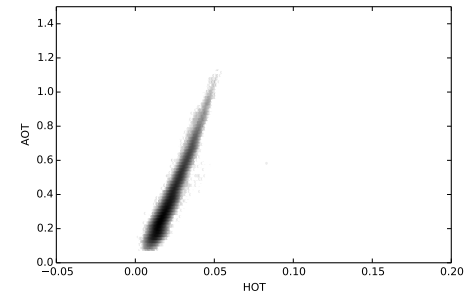
FIGURE 6.3: Examples of LEDAPS images simulated with AOTs

TABLE 6.1: Pearson correlation coefficients between HOT and AOT for each of the LEDAPS test images simulated with each aerosol type.

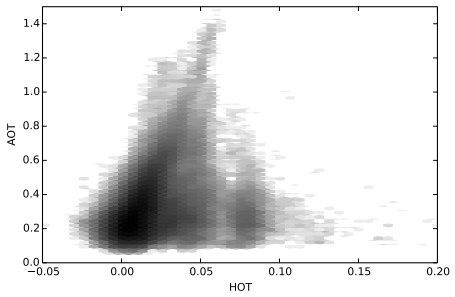
	Continental	**Desert**	**Maritime**	**Stratospheric**	**Urban**	**Mean**
Amazon	0.96	0.95	0.95	0.90	0.78	0.91
Botswana	0.31	0.30	0.32	0.22	0.04	0.24
Chilbolton	0.70	0.70	0.71	0.58	0.22	0.58
London	0.79	0.77	0.03	0.25	0.29	0.43
Midwest	0.82	0.82	0.82	0.74	0.28	0.70
Namibia	0.34	0.32	0.33	0.24	0.21	0.29
Sahara	0.65	0.52	0.58	0.38	0.48	0.52
SaudiArabia	0.54	0.43	0.58	0.41	-0.03	0.39
SotonNewForest	0.73	0.44	0.74	0.63	0.14	0.53
Mean	0.65	0.58	0.56	0.48	0.27	0.51

particularly obvious for the Amazon image, which consists almost entirely of forest and therefore requires very little correction, but also for other images on which the correction performs well, such as London, Midwest and SotonNewForest. Conversely, the images with the lowest mean correlation are Botswana, Namibia and SaudiArabia – all images for which the correction algorithm performs poorly (see §5.4.8). Botswana and Namibia have particularly broad point clouds, and this is likely to be due to the breadth of smoothly-varying land cover changes within these images: images such as Sahara and SaudiArabia consist almost entirely of sand, and thus do not require much correction, whereas Botswana and Namibia have a range of land covers which merge into one another without sharp boundaries, and thus the correction algorithm performs poorly.

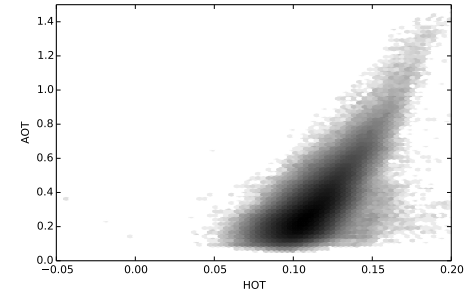
The HOT-AOT relationships for a given image with different aerosol types are very similar (Figure 6.5), although there are slight differences in the gradient of the relationship. The exception is the urban aerosol type, which generally has low correlations, and an almost vertical point cloud, showing no real relationship. In 6S, the urban aerosol type is defined as a mixture of dust-like (17%), water-soluble (61%) and soot (22%) aerosols. Soot, composed almost entirely of black carbon, has a low Single Scattering Albedo (SSA), meaning that it scatters very little light: instead, most of the light is absorbed. For comparison, the SSA of the 6S Urban aerosol type is 69% compared to 98% and 90% for the Maritime and Continental aerosol types respectively. The effect of this can be seen in Figure 6.3(e) which shows a darkening of a high AOT area when simulated with the urban aerosol type, as opposed to the brightening seen when the same area is simulated with the other aerosol types. The poor HOT-AOT relationship for the urban aerosol type raises concerns about the ability of this method to monitor urban air pollution: this will be investigated further when the monitoring of air pollution is investigated as a potential application in Chapter 8.2. However, recent research on aerosol properties in Beijing during severe urban pollution episodes shows an average SSA of 0.90 ± 0.03 - far higher than the 6S Urban aerosol types (Che *et al.*, 2014), suggesting that the 6S Urban aerosol type is not representative of actual urban aerosol conditions.



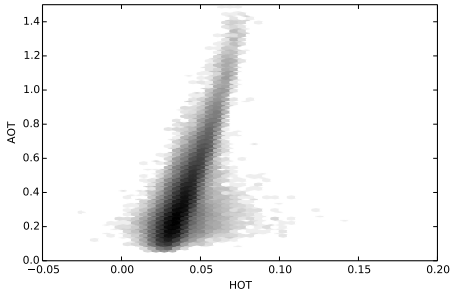
(a) Amazon



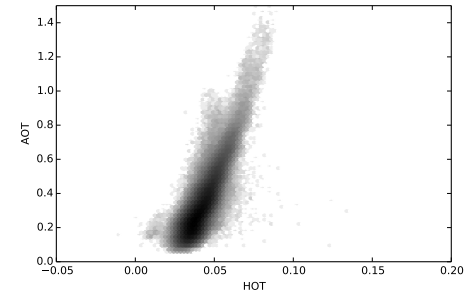
(b) Botswana



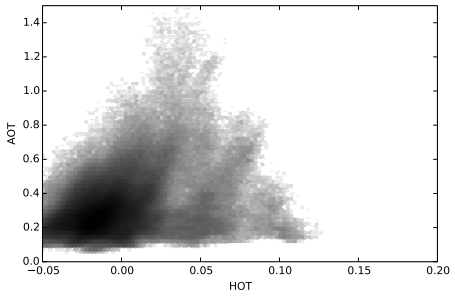
(c) Chilbolton



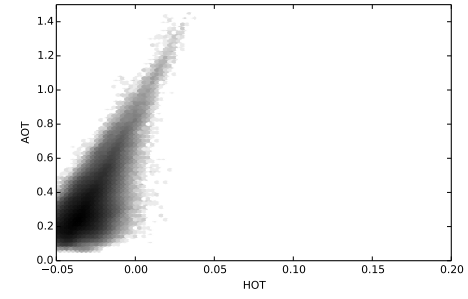
(d) London



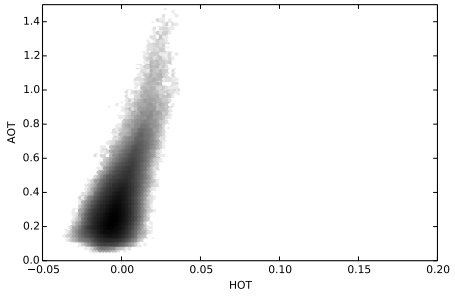
(e) Midwest



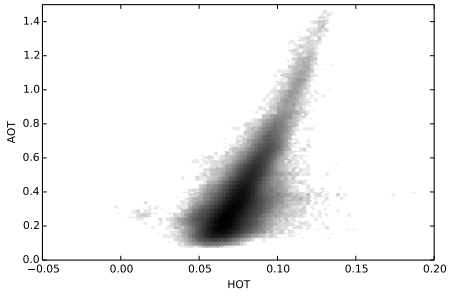
(f) Namibia



(g) Sahara

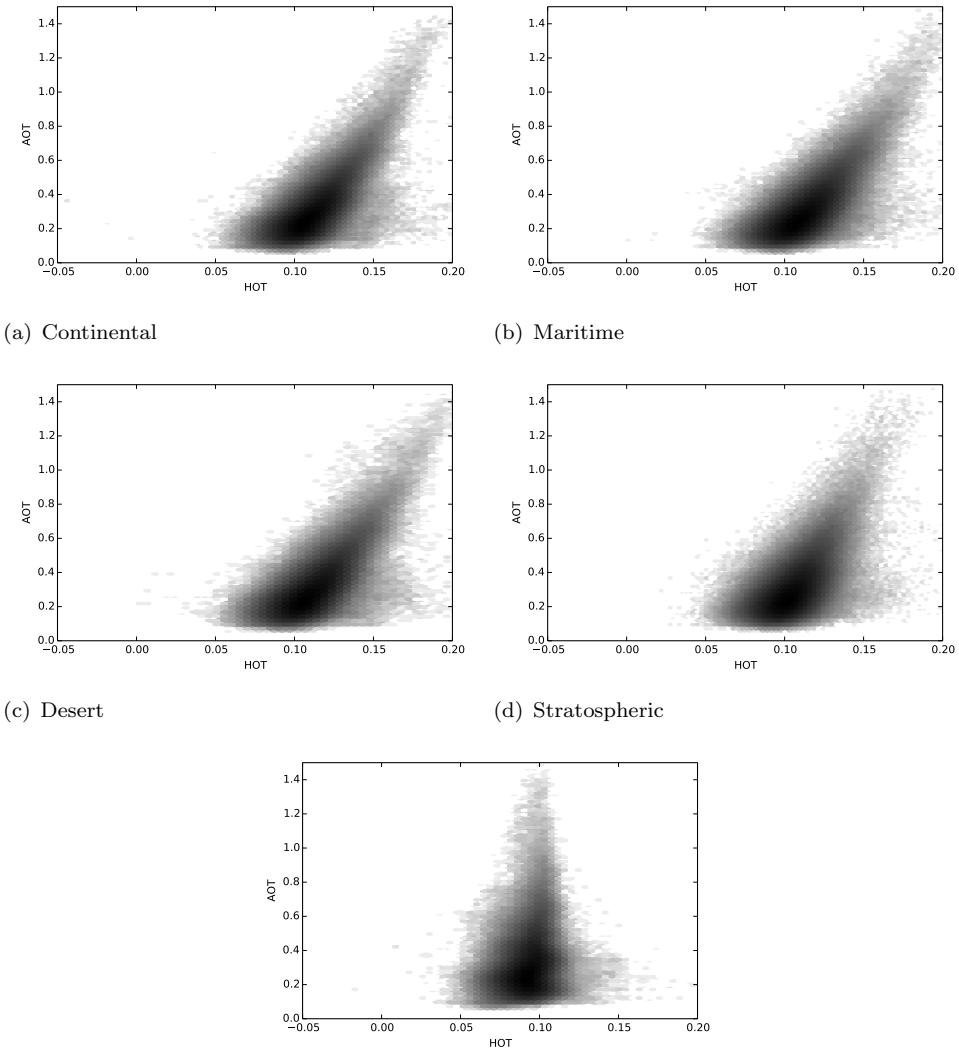


(h) SaudiArabia



(i) SotonNewForest

FIGURE 6.4: Density plots showing the relationship between HOT and AOT for each LEDAPS image, simulated with the Continental aerosol type



(a) Continental

(b) Maritime

(c) Desert

(d) Stratospheric

(e) Urban

FIGURE 6.5: Density plots showing the relationship between HOT and AOT for the Chilbolton LEDAPS image, for all simulated aerosol types.

6.3.1 HOT-AOT regression

Due to equal axes for all subplots within Figure 6.4, it is easy to see that both the gradient and intercept of the linear relationship varies between images. The variation in intercept (with a 5%–95% range of -6.2–0.4) is caused by the magnitude of the HOT values varying significantly between images (as can be seen by examining the x-axes of the graphs in Figure 6.4), with an AOT of 0.2 producing HOT values ranging from -0.04 (Sahara, Continental) to 0.1 (Chilbolton, Continental). The HOT magnitude variation (with a 5%–95% range of 8.0–56.4) seems to be related to land cover, with images with similar land covers (such as London, Midwest and SotonNewForest) having similar HOT magnitudes. Given these differences between images, the relationship between HOT and AOT will have to be estimated on a per-image basis. In this section regression methods will be used to quantify the actual relationship between HOT and AOT for the test images. However, it should be noted that another approach will be needed to estimate

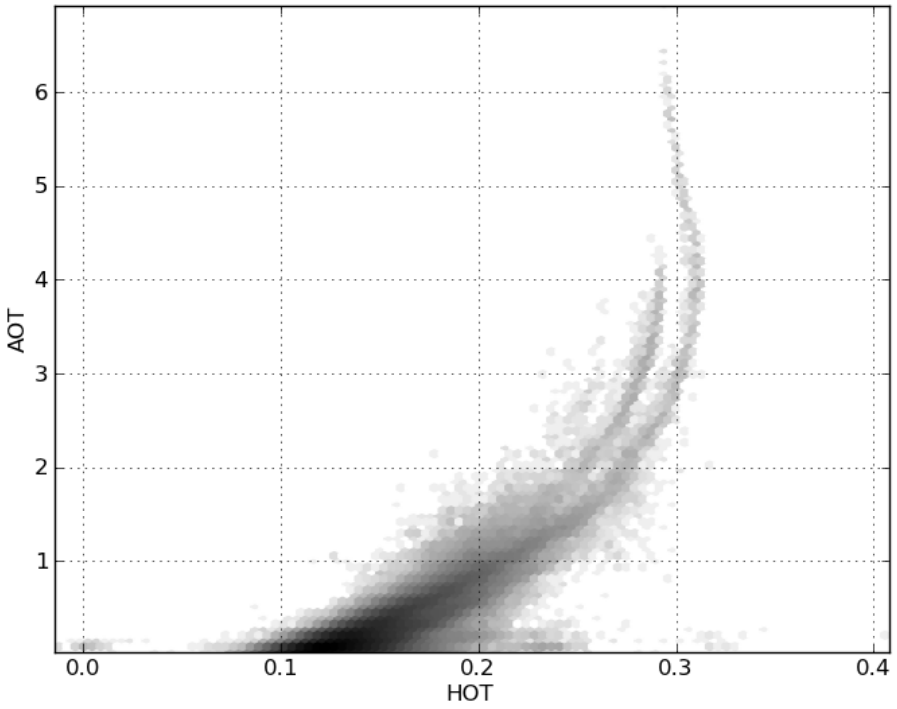


FIGURE 6.6: HOT-AOT relationship plot for SotonNewForest with Continental aerosols, and AOTs ranging from 0.0 to 6.0, showing non-linearity at high AOTs.

this relationship when the AOT is unknown, as will be the case in the final application of the algorithm.

Before using linear regression to model the relationship, the assumption of linearity must be examined. The plots in Figure 6.4 appear linear, but examining the relationship between HOT and AOT for a significantly larger range of AOTs (0–6) shows significant non-linearity at higher AOTs (Figure 6.6). The relationship becomes complex above an AOT of 2.0, with the cloud of points splitting into two sections which curve at different angles. However, focusing on the densest areas of the plot, the AOT appears to be linear up to an AOT of around 1.5. Examining the statistics of all AERONET Level 2.0 data again, only 1.29% of AOTs are greater than 1.0, and only 0.3% are greater than 1.5. This suggests that a linear HOT-AOT relationship will be valid for at least 99.7% of real-world AOT values, and so it was decided to focus on the simpler task of developing a linear relationship, rather than trying to parameterise the complex non-linear relationship found with higher AOTs.

Orthogonal Distance Regression (ODR) was used in §5.3 to estimate the Clear Line from the sample of LandsatAERONET points, as the line needed to fit the overall relationship best, minimising the error in both blue and red reflectances rather than just in one of these. The HOT-AOT situation is different: here we are trying to estimate AOT from HOT with the lowest error. Given this, standard linear regression would appear to be the

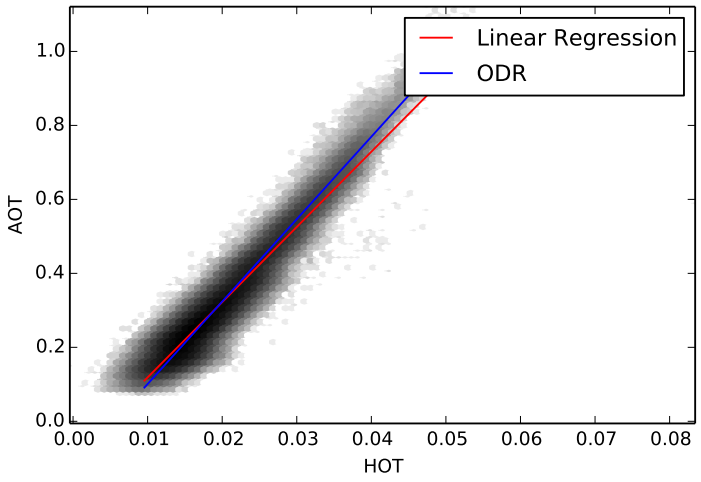
best approach, but Figure 6.7 shows that the lines produced by standard linear regression do not fit the whole point cloud well. This is partially caused by the distribution of AOTs used to create these test images: the majority of the AOTs are less than 0.6, and as the regression attempts to minimise the sum of squared error in AOT it focuses on these lower points, along with the outliers at the bottom of the cloud, and selects a shallower line.

Figure 6.8 shows the SotonNewForest image simulated with AOTs of different magnitudes, with standard linear regression and ODR lines. The standard linear regression line is different for each of these images - that is, it is dependent on AOT magnitude - whereas with ODR the regression line is almost constant across each image, and thus unaffected by AOT magnitude. Therefore, using ODR will significantly simplify the estimation of the HOT-AOT relationship in the final algorithm as the AOT magnitude - which will be unknown in real-world applications - will not influence the relationship.

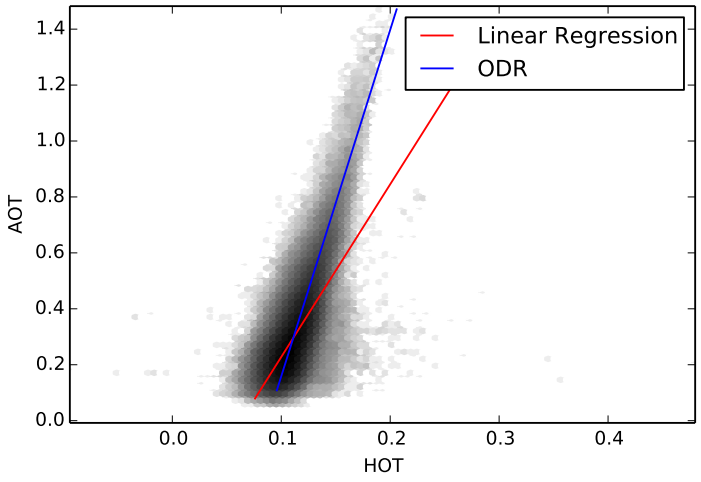
Once the relationship between HOT and AOT for each of the test images was estimated using ODR (Table 6.2), statistics were calculated to show how accurately HOT can be used to estimate AOT (Table 6.3). These show similar results to the correlations: high errors for images with poor corrections (often double that of well-corrected images), high errors for the urban aerosol type (often quadruple the error for the continental aerosol type), and medium/low errors for all other images. The error in the results of the whole algorithm will be examined in more detail in Chapter 7, where it will also be compared to the error from other satellite-derived measurements of AOT. However, in this context it is useful to note that the validation of the MODIS 10km AOT product states that 67% of values will be within $\pm 0.05 \pm 0.15\tau$, which for the global average AOT of 0.19, corresponds to an error of ± 0.08. The errors found for the well-corrected images, such as Chilbolton, London and Midwest, are comparable to this, with the percentage within ± 0.1 ranging from 59% to 79%. The majority of these images have over 90% of the values within ± 0.2. The error for the Amazon image is exceptionally low, with a median error of 0.03 and 98% of values within ± 0.1, due to the uniform land cover and good LEDAPS correction performance.

6.4 Modelling the HOT-AOT relationship

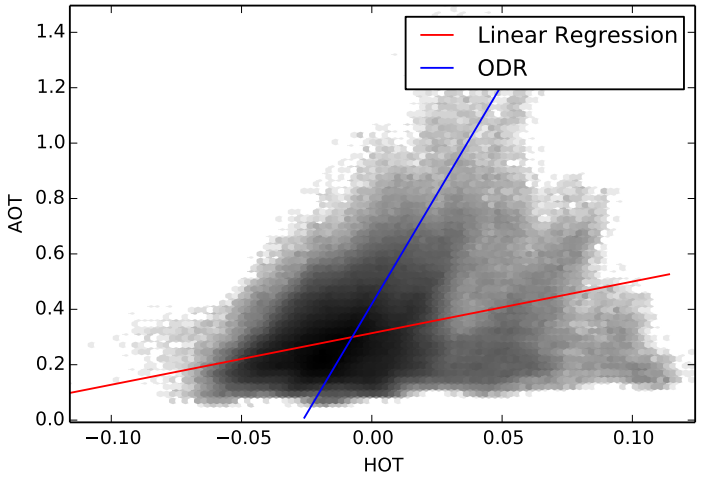
The previous section derived HOT-AOT relationships from test images where both the HOT and AOT values were known: these relationships will be referred to as the *actual* relationships. However, when applying this algorithm to real-world images, only the HOT will be known, and the goal is to estimate the AOT. Thus, a method must be developed to estimate the HOT-AOT relationship for a given image when the AOT is not known: the results of this estimation process will be referred to as the *modelled* relationships. From the investigations performed above, it seems that the relationship varies between images and aerosol types but, through the use of ODR, is constant across a range of AOT magnitudes. The primary difference between images - particularly after the pseudo-BOA correction - is land cover, and so it is assumed that this, combined with aerosol type, has the largest influence on the relationship.



(a) Amazon

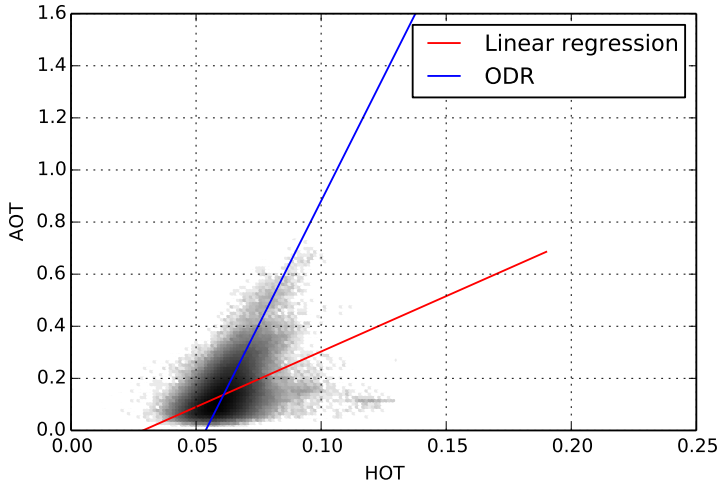


(b) Chilbolton

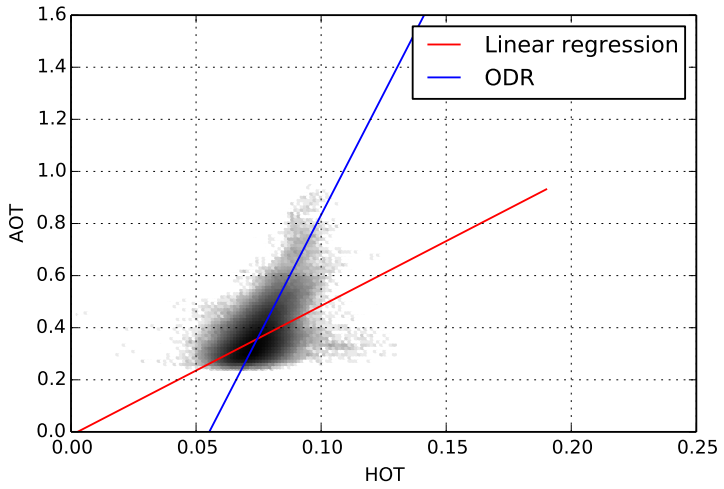


(c) Namibia

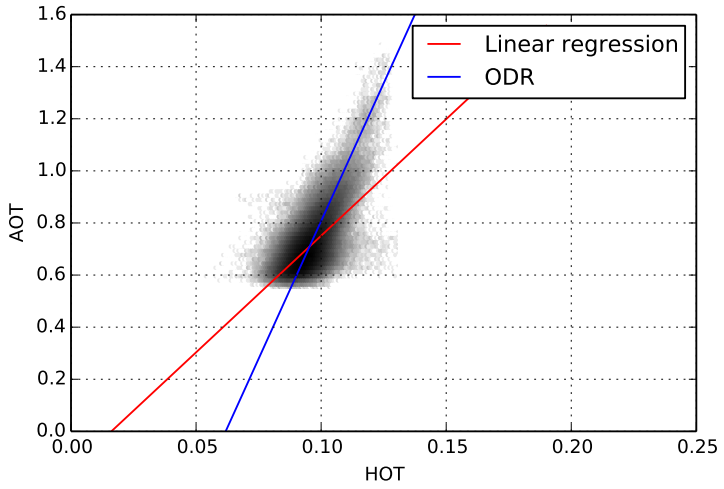
FIGURE 6.7: Examples of standard linear regression and Orthogonal Distance Regression (ODR) on the HOT-AOT relationships for LEDAPS images simulated with Continental aerosols



(a) Low (AOT range: 0.01–0.73)



(b) Medium (AOT range: 0.2–0.95)



(c) High (AOT range: 0.54–1.50)

FIGURE 6.8: HOT-AOT relationship density plots for the SotonNewForest image simulated with three different ranges of AOTs, with regression lines produced through standard linear regression and ODR.

TABLE 6.2: Parameters of the HOT-AOT relationship, defined as AOT = mHOT + c, as estimated by Orthogonal Distance Regression.

Image	**Aerosol**	**m**	**c**	R^2
Amazon	Continental	9.31	-0.75	0.92
	Desert	9.05	-0.72	0.92
	Maritime	9.27	-0.75	0.91
	Stratospheric	12.93	-1.03	0.83
	Urban	27.02	-2.27	0.63
Botswana	Continental	10.80	-0.93	0.09
	Desert	10.60	-0.90	0.09
	Maritime	9.20	-0.79	0.10
	Stratospheric	13.43	-1.16	0.04
	Urban	178.49	-15.34	0.01
Chilbolton	Continental	11.20	-1.13	0.49
	Desert	10.96	-1.09	0.48
	Maritime	10.72	-1.08	0.49
	Stratospheric	14.48	-1.44	0.32
	Urban	51.18	-5.18	0.04
London	Continental	11.62	-1.15	0.65
	Desert	11.15	-1.09	0.65
	Maritime	11.08	-1.10	0.66
	Stratospheric	15.25	-1.49	0.50
	Urban	52.60	-5.25	0.09
Midwest	Continental	10.54	-1.18	0.70
	Desert	10.04	-1.12	0.69
	Maritime	9.62	-1.08	0.70
	Stratospheric	13.49	-1.50	0.09
	Urban	57.30	-6.39	0.12
Namibia	Continental	7.80	-0.05	0.12
	Desert	9.03	-0.07	0.08
	Maritime	7.08	-0.02	0.11
	Stratospheric	10.87	-0.07	0.04
	Urban	18.59	-0.20	0.02
Sahara	Continental	9.82	0.25	0.36
	Desert	13.04	0.34	0.23
	Maritime	10.00	0.23	0.30
	Stratospheric	16.82	0.39	0.11
	Urban	25.35	0.73	0.10
SaudiArabia	Continental	15.37	-0.93	0.25
	Desert	18.60	-1.11	0.17
	Maritime	11.38	-0.71	0.34
	Stratospheric	17.77	-1.08	0.15
	Urban	-60.06	3.20	0.01
SotonNewForest	Continental	12.47	-1.42	0.56
	Desert	11.77	-1.34	0.51
	Maritime	11.42	-1.31	0.57
	Stratospheric	15.53	-1.73	0.40
	Urban	74.31	-8.26	0.03

TABLE 6.3: Errors in AOT when estimated from HOT using relationships from Orthogonal Distance Regression, for each simulated image

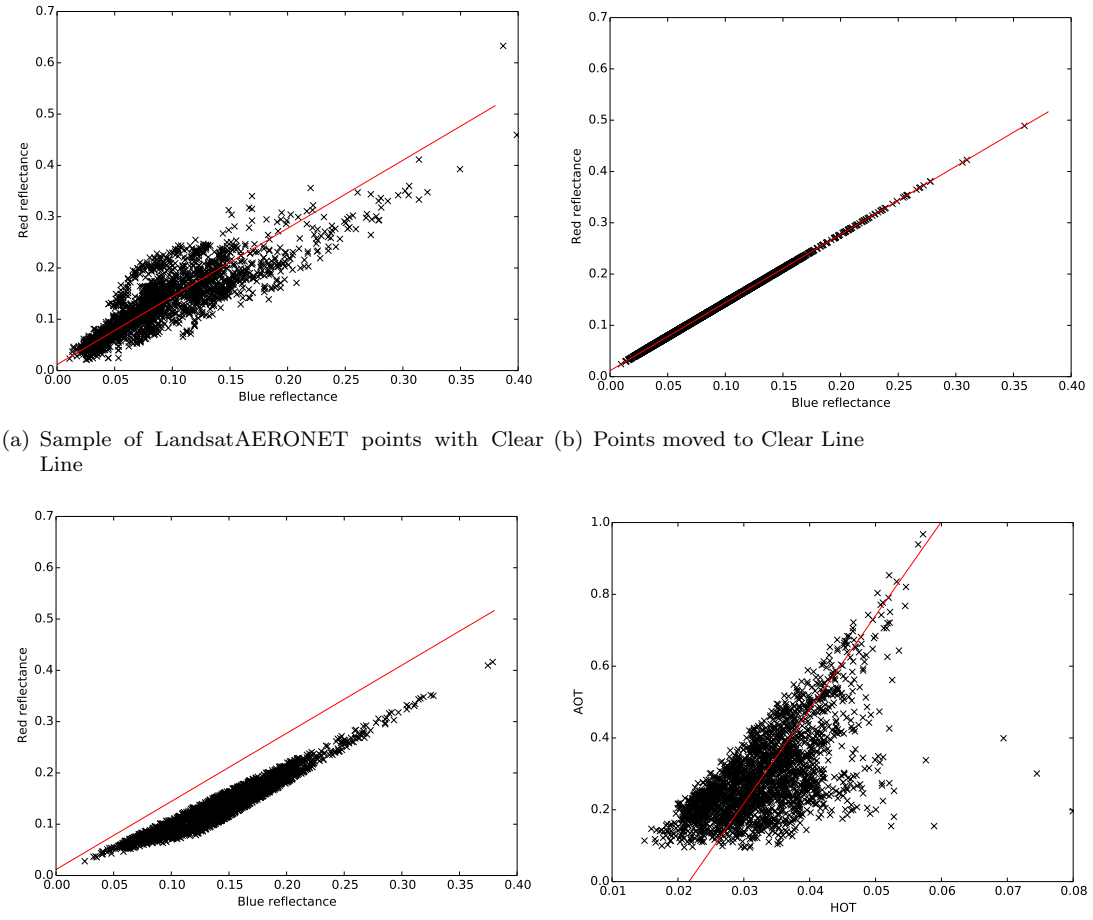
Image	Aerosol	Median	% within ±0.1	% within ±0.2	5th percentile	95th percentile
Amazon	Continental	0.03	98	100	0.00	0.08
	Desert	0.03	98	100	0.00	0.08
	Maritime	0.03	98	100	0.00	0.08
	Stratospheric	0.04	90	99	0.01	0.12
	Urban	0.06	69	95	0.01	0.20
Botswana	Continental	0.17	29	56	0.02	0.73
	Desert	0.18	28	56	0.02	0.73
	Maritime	0.16	30	60	0.02	0.68
	Stratospheric	0.25	20	40	0.03	1.04
	Urban	2.32	2	4	0.22	9.86
Chilbolton	Continental	0.08	59	88	0.01	0.26
	Desert	0.08	58	87	0.01	0.26
	Maritime	0.08	58	88	0.01	0.26
	Stratospheric	0.11	44	75	0.01	0.36
	Urban	0.33	16	31	0.03	1.05
London	Continental	0.05	79	96	0.00	0.17
	Desert	0.05	78	96	0.00	0.18
	Maritime	0.05	78	96	0.00	0.17
	Stratospheric	0.07	64	90	0.01	0.25
	Urban	0.21	24	47	0.02	0.74
Midwest	Continental	0.05	79	97	0.00	0.17
	Desert	0.05	79	97	0.00	0.17
	Maritime	0.05	79	97	0.00	0.17
	Stratospheric	0.07	65	92	0.01	0.22
	Urban	0.25	21	40	0.02	0.77
Namibia	Continental	0.20	26	51	0.02	0.64
	Desert	0.22	24	46	0.02	0.76
	Maritime	0.20	26	50	0.02	0.64
	Stratospheric	0.30	18	34	0.03	0.99
	Urban	0.37	14	27	0.04	1.23
Sahara	Continental	0.11	47	76	0.01	0.33
	Desert	0.15	36	62	0.01	0.46
	Maritime	0.13	40	69	0.01	0.38
	Stratospheric	0.23	24	45	0.02	0.66
	Urban	0.21	26	47	0.02	0.71
SaudiArabia	Continental	0.14	37	65	0.01	0.43
	Desert	0.18	29	53	0.02	0.54
	Maritime	0.12	43	73	0.01	0.35
	Stratospheric	0.19	27	52	0.02	0.56
	Urban	0.50	10	21	0.05	1.46
SotonNewForest	Continental	0.07	65	92	0.01	0.22
	Desert	0.07	65	93	0.01	0.22
	Maritime	0.07	65	92	0.01	0.22
	Stratospheric	0.10	51	82	0.01	0.30
	Urban	0.37	14	28	0.03	1.21

Given the success of the Monte Carlo regression approach used to estimate the Clear Line in §5.3, a similar approach has been applied to estimate the HOT-AOT relationship. The same method is used to select a representative sample of points from the LandsatAERONET dataset, but rather than fitting a line immediately, the points are simulated with randomly chosen AOT values, thus allowing the HOT to be calculated and a HOT-AOT relationship developed. In detail, the steps of algorithm, illustrated in Figure 6.9, are:

1. Select a sample of points from the LandsatAERONET dataset based upon land cover proportions within the image, calculated from the GlobCover dataset (as in §5.3)
2. Calculate the parameters of the Clear Line (as in §5.3)
3. Move all points to the closest location on the Clear Line, replicating the effect of a 'perfect' HOT correction algorithm
4. Generate a random AOT value for each point
5. Simulate the TOA radiance for each point given the reflectance from the LandsatAERONET sample, its associated random AOT, and a user-provided aerosol type
6. Correct each point to pseudo-BOA reflectance
7. Calculate the HOT for these simulated points (simply the distance from the Clear Line to each point)
8. Estimate the HOT-AOT relationship by performing Orthogonal Distance Regression on the set of HOT and AOT values

Several configurable extensions were made to the basic estimation process and tested to determine whether they improved the accuracy of the resulting HOT-AOT relationship. These were:

- **Sample size:** The original method sampled 5000 points from the LandsatAERONET dataset; increasing the sample size may improve the representativeness of the sample.
- **Monte Carlo repeats:** The original algorithm took a single sample of 5000 points from the LandsatAERONET dataset; again, taking multiple sets of 5000-point samples may improve the representative of the sample.
- **Partial movement to the Clear Line:** The original method moved all of the LandsatAERONET points so that they lay exactly upon the Clear Line. This was designed to replicate the effects of a perfect HOT correction algorithm, which would remove all land cover effects, and thus move all clear points to lie exactly upon the Clear Line. However, as discussed in §5.4.8, the HOT correction procedure is not perfect, and so not all clear points will be moved to lie exactly upon the Clear Line. Therefore, the algorithm was altered to allow the points to be moved a configurable fraction of their distance to the line: a value of 1.0 would move the points to sit exactly upon the line, a value of 0.5 would move the points half-way to the line, and a value of 0.0 would leave the points in their original positions.



(a) Sample of LandsatAERONET points with Clear Line

(b) Points moved to Clear Line

(c) pseudo-BOA reflectance for points simulated with random AOT values

(d) HOT-AOT relationship

FIGURE 6.9: Steps of the HOT-AOT relationship estimation algorithm

- **Distribution of random AOTs:** The random AOTs used in the original algorithm were taken from the same distribution used for the random AOT images (see §6.2.1). Due to the parameterisation of this distribution, the majority of the AOTs are less than 0.6 and there are very few high AOTs. This means that there will be few points at the top end of the regression line, potentially leading to an incorrect placement of the top of the line, and thus a poor estimation of the gradient of the relationship. To test this, an extension to use a uniform distribution over a reasonable range of AOTs (0.01–1.00) was implemented.

Error statistics were calculated for the modelled relationship created with various configurations of the estimation process (Figure 6.10). The percentage of pixels with an AOT value within ±0.1 of the true value is used here as this gives a good indication of the consistency of the performance of the relationship across the whole image, and is not skewed by very large or small errors. Results are shown for the Continental aerosol type - as all other aerosol types had similar patterns - and some extensions are excluded as they had very little effect. For example, altering the sample size did not change the results - presumably because the default sample size of 5,000 is already statistically representative -

and combining the various extensions produced only very minor differences.

Overall the accuracies found using the modelled relationship are lower than the accuracies found with the actual relationship. This is to be expected as the estimation procedure does not necessarily take into account all of the factors which affect the relationship. The reduction in accuracy is most noticeable with the desert images. These had a lower accuracy than the rest of the images when using the actual relationship, but when using the estimated relationship the accuracy becomes very poor, with images such as Namibia, Sahara and SaudiArabia having 0% of the points within ±0.2 and median absolute AOT errors of over 2.0.

This is due to a complete mis-estimation of the HOT-AOT relationship which places the modelled regression line well outside the cloud of points from the simulated image (Figure 6.11(a)). This poor relationship estimation is caused by issues with the original LEDAPS images from which the simulated images were produced. LEDAPS is known to perform poorly on images with limited DDV - such as desert areas - producing erroneous ground reflectance values (see §4.2.3 and Maiersperger *et al.*, 2013). This can be seen in Figure 6.11(b) where the blue and red reflectances in the LEDAPS image are very different from those in the LandsatAERONET sample. This means that the Clear Line estimated from the LandsatAERONET sample is not valid for the pixels in the LEDAPS image, and the offset causes negative HOT values, even for relatively high AOTs. This is not a problem in itself, but the HOTs calculated from the LandsatAERONET-based modelling are positive, and thus the estimated relationship is very different to the true relationship. These issues will only occur with the test images - as LEDAPS images will not be used in the application of the algorithm to real-world images - and thus these accuracies do not give a true impression of the algorithm performance over desert images. Given this, desert scenes (Botswana, Namibia, Sahara and SaudiArabia) are left out of the analyses that follow.

The differences between the methods are mostly small with the best method for each image increasing the percentage within ±0.1 by around 8–10 percentage points. However, the best method varies between images - each method is the best for at least one image - making it difficult to choose the best method overall. One key consideration is the reliability of the method across all of the images: the chosen method must perform relatively well for every image. Figure 6.12 shows boxplots of two statistics (percentage within ±0.1 and mean absolute error) across all images, for each method. Both of these statistics show that the Uniform AOTs extension has the narrowest range in values, with the most uniform performance across all images. Of course, this method would be a poor choice if it performed uniformly poorly but it has a relatively good accuracy, with the highest median percentage within ±0.1 of all methods, and the second highest median mean absolute error. Therefore, the estimation method with the uniform AOT extension was chosen to model the relationship between HOT and AOT.

The AOT errors for each simulated image resulting from the final modelled relationship

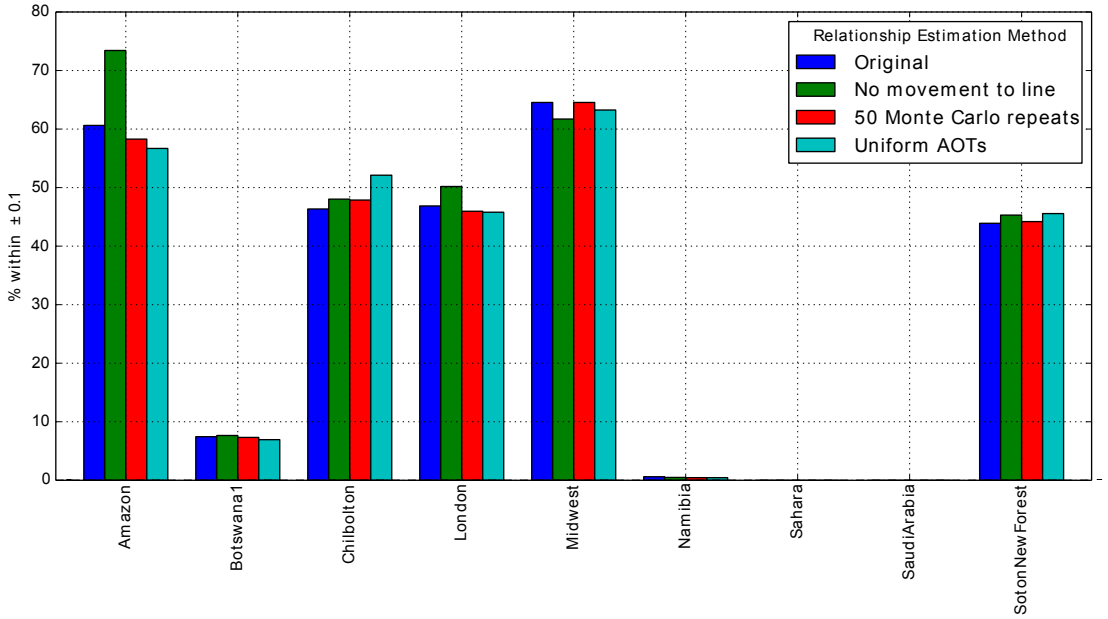


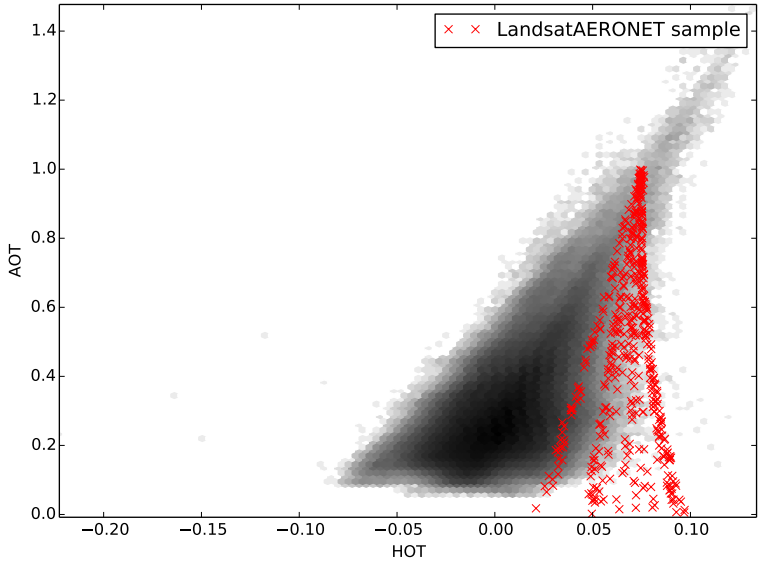
FIGURE 6.10: Percentage of AOT values within ±0.1 for each test image for four relationship estimation methods

are shown in Table 6.4. Ignoring desert images, as discussed above, the mean percentages within ±0.2, ±0.1 and ±0.05 for images simulated with the Continental aerosol type are 88%, 52% and 25%. The mean median error is 0.09, and the mean 95th percentile of the error is 0.25. Even though the errors from the modelled relationship are worse than the errors from the actual relationship, these still compare favourably with the MODIS 10km AOT product validation, which stated that 67% of the AOT values were within $0.05 \pm 0.15\tau$, equivalent to around ±0.08 for the global mean AOT (Levy *et al.*, 2010).

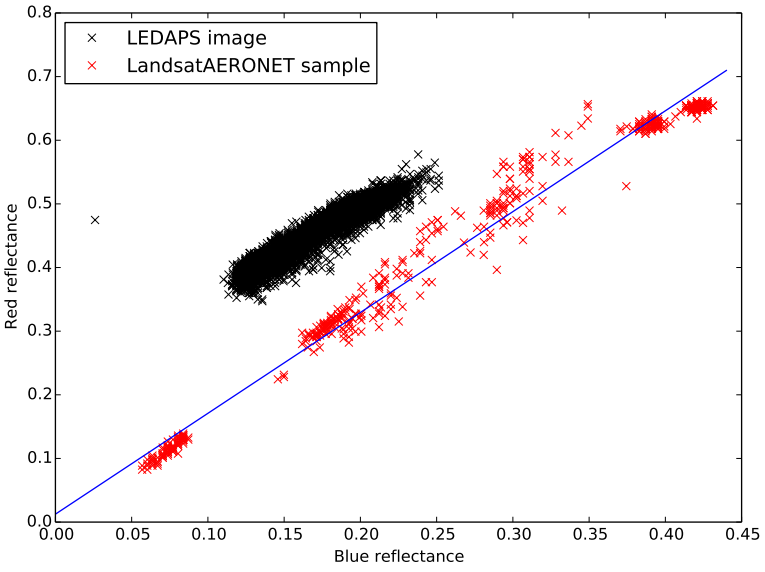
6.4.1 Aerosol type sensitivity analysis

An aerosol type is required by the method used to estimate the HOT-AOT relationship. This is used to parameterise the Py6S simulations which simulate TOA reflectance for each point in the LandsatAERONET sample under various AOTs. When the algorithm is applied to real-world images the user will set this parameter based on their estimation of the relevant aerosol type. This raises the question of the effect of the user choosing the incorrect aerosol type on the AOT accuracy. This can be investigated using the test images, by simulating an image with one aerosol type, and modelling the HOT-AOT relationship using a different aerosol type. The issues with the test images described above still apply, so the desert images are not discussed here.

Results for two images are shown in Figure 6.13, and the other non-desert images show very similar patterns. The results are surprising: using the correct aerosol type does not seem to give the highest accuracy. In fact, for images simulated with all aerosol types excluding Urban, estimating the relationship using the Maritime aerosol type seems to give the best result. The differences are striking: in the London image using the Maritime aerosol type to estimate the relationship for an image simulated using the Continental

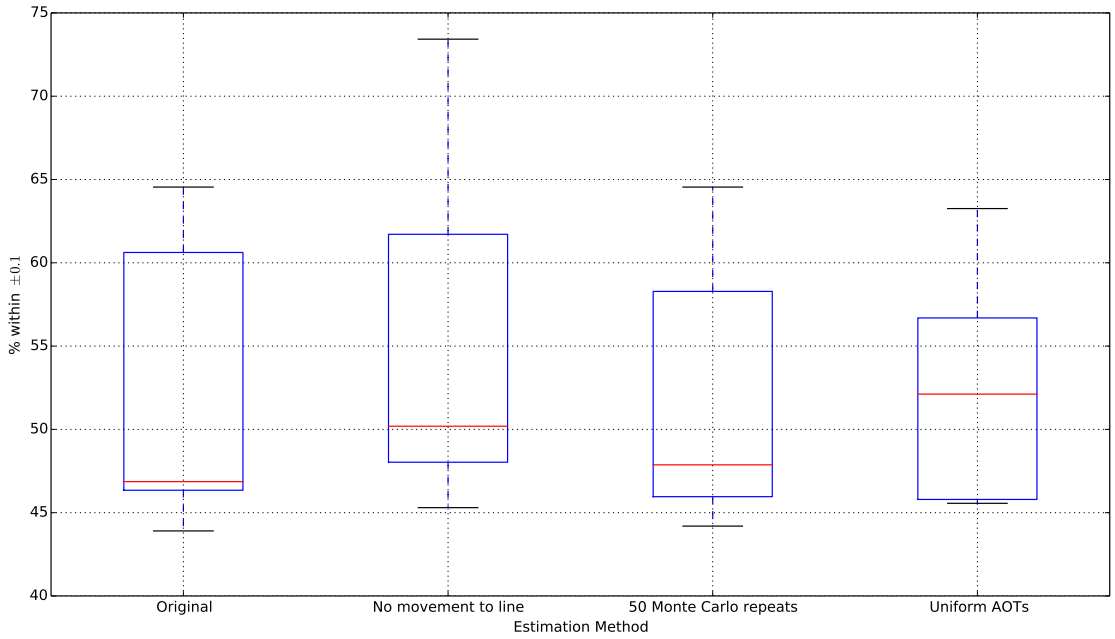


(a) HOT-AOT relationship for the simulated image based upon LEDAPS (black) and the simulated points based on LandsatAERONET (red)

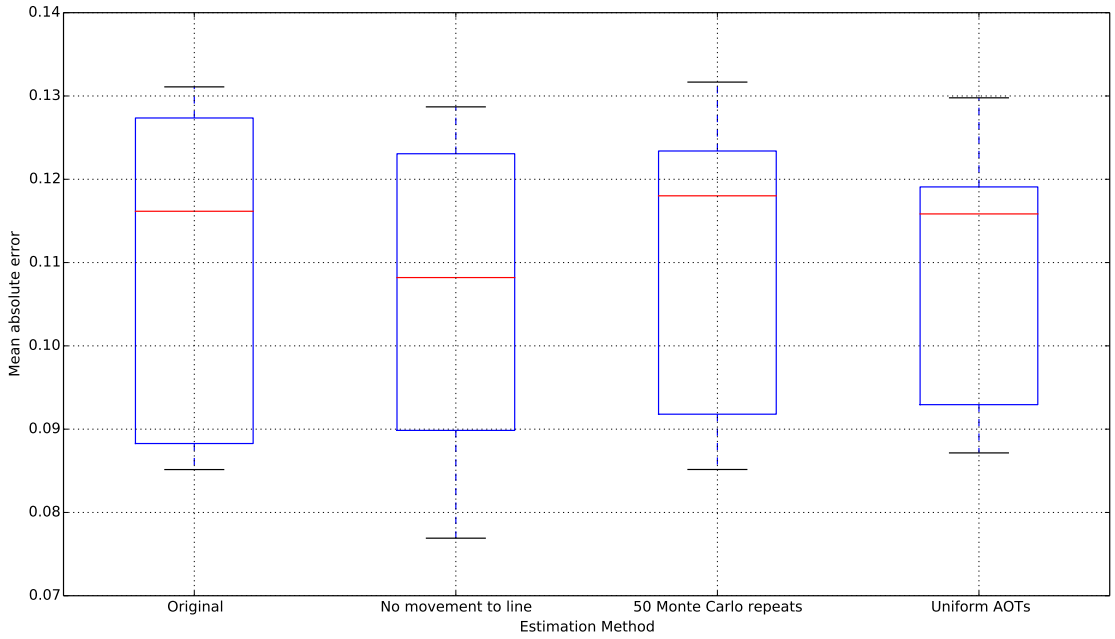


(b) Sahara LEDAPS image and LandsatAERONET sample shown in Blue-Red feature space, with the estimated Clear Line shown in blue

FIGURE 6.11: Examples of differences between LEDAPS images and LandsatAERONET data for the Sahara image



(a) Percentage within ±0.1



(b) Mean error

FIGURE 6.12: Boxplots showing error statistics calculated over all non-Desert test images for four relationship estimation methods

TABLE 6.4: Absolute errors in AOT when estimated from HOT using the modelled relationship derived using uniform AOTs

Image	**Aerosol**	**Median**	**% within ±0.1**	**% within ±0.2**	**5th percentile**	**95th percentile**
	Continental	0.09	57	98	0.01	0.18
	Maritime	0.11	40	96	0.03	0.19
Amazon	Desert	0.09	58	98	0.02	0.17
	Stratospheric	0.15	22	76	0.04	0.26
	Urban	0.36	15	29	0.03	0.94
	Continental	0.40	7	18	0.08	0.78
	Maritime	0.36	8	20	0.07	0.73
Botswana	Desert	0.40	7	17	0.08	0.77
	Stratospheric	0.51	6	13	0.09	1.01
	Urban	2.42	1	2	0.45	4.78
	Continental	0.10	52	84	0.01	0.29
	Maritime	0.08	58	88	0.01	0.26
Chilbolton	Desert	0.11	45	78	0.01	0.32
	Stratospheric	0.12	44	74	0.01	0.37
	Urban	2.00	2	5	0.21	4.63
	Continental	0.11	46	86	0.01	0.26
	Maritime	0.07	66	94	0.01	0.21
London	Desert	0.11	44	83	0.01	0.27
	Stratospheric	0.10	51	84	0.01	0.29
	Urban	1.23	2	4	0.24	2.11
	Continental	0.08	63	95	0.01	0.20
	Maritime	0.06	75	97	0.01	0.17
Midwest	Desert	0.08	63	94	0.01	0.21
	Stratospheric	0.08	58	90	0.01	0.24
	Urban	0.77	6	11	0.09	1.60
	Continental	2.75	0	1	1.13	3.76
	Maritime	1.72	1	1	0.69	2.37
Namibia	Desert	2.55	1	1	0.99	3.43
	Stratospheric	2.57	0	1	1.13	3.46
	Urban	2.15	0	1	0.91	3.05
	Continental	7.56	0	0	6.29	8.74
	Maritime	3.04	0	0	2.59	3.47
Sahara	Desert	11.40	0	0	9.67	13.02
	Stratospheric	4.58	0	0	3.91	5.24
	Urban	1.82	0	0	1.36	2.16
	Continental	5.58	0	0	4.41	6.81
	Maritime	1.89	0	0	1.52	2.30
SaudiArabia	Desert	7.17	0	0	5.71	8.74
	Stratospheric	2.96	0	0	2.39	3.60
	Urban	1.18	0	0	0.84	1.45
	Continental	0.11	46	80	0.01	0.32
	Maritime	0.09	56	87	0.01	0.26
SotonNewForest	Desert	0.13	40	68	0.01	1.02
	Stratospheric	0.13	41	71	0.01	0.37
	Urban	1.41	2	5	0.20	3.11

aerosol type leads increases the percentage within ±0.1 from 45% to 69%. There are two possible reasons for the Maritime aerosol type performing better for all images: either aerosol type contamination from the LEDAPS images, or a link between aerosol Single Scattering Albedo and the quality of the relationship.

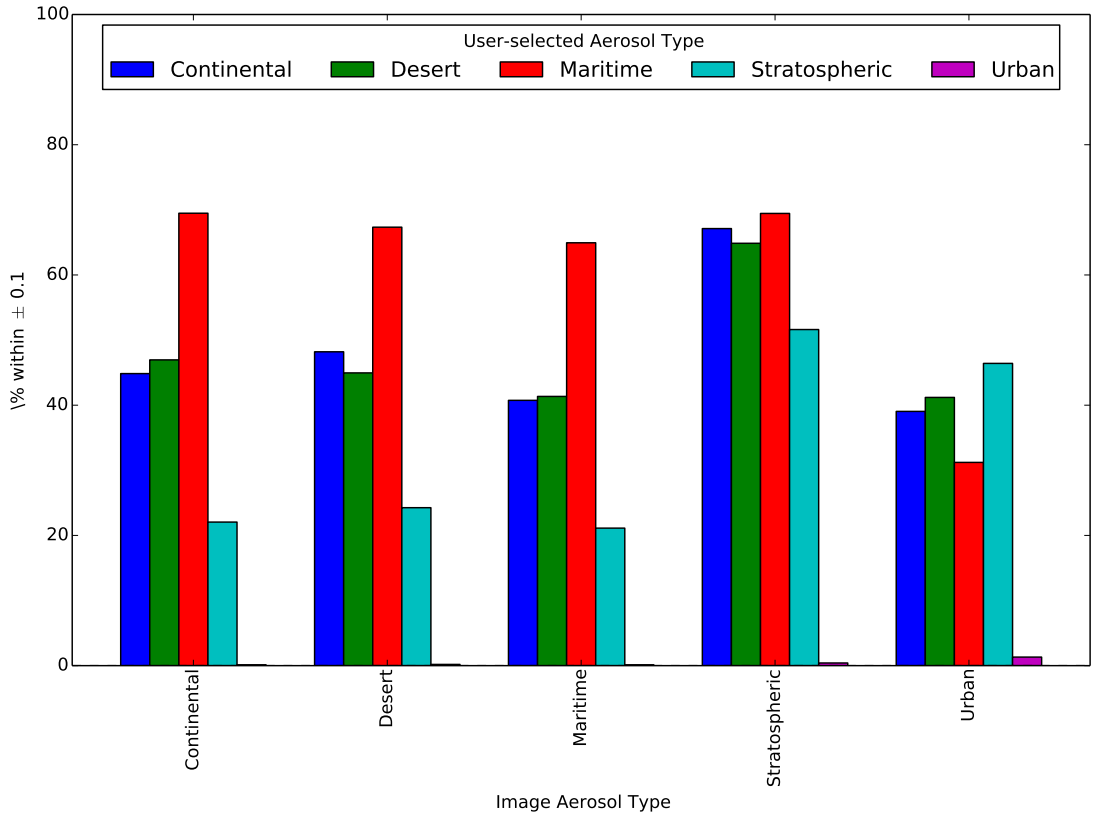
LEDAPS uses a fixed Continental aerosol type for all AOT retrievals (Maiersperger *et al.*, 2013), and this may cause errors when estimating AOT over images with a different aerosol type. Several of the test images are likely to have a Maritime aerosol type - particularly SotonNewForest and London - and so there may be the residual effects of Maritime aerosols left in the images, regardless of what aerosol type the test images were simulated with. These residual effects could then be picked up best when the relationship is estimated using the Maritime aerosol type, causing a low AOT error.

Alternatively the increased accuracy when using the Maritime aerosol type could be caused by the high SSA of this type of aerosol. Table 6.5 shows the SSA for each of the 6S aerosol types, excluding Stratospheric (a very unusual aerosol type as it does not reflect dominant boundary-layer aerosol, and is therefore rarely used), in order of decreasing SSA. Maritime has the highest SSA, and thus the effects of Maritime aerosols will be most easily noticeable in terms of HOT values. Calculating the mean percentage within ±0.1 for each aerosol type, and sorting by decreasing accuracy (Figure 6.5), shows the same ordering of aerosol types: there seems to be a relationship between SSA of an aerosol type and the accuracy of a HOT-AOT relationship developed using that aerosol type.

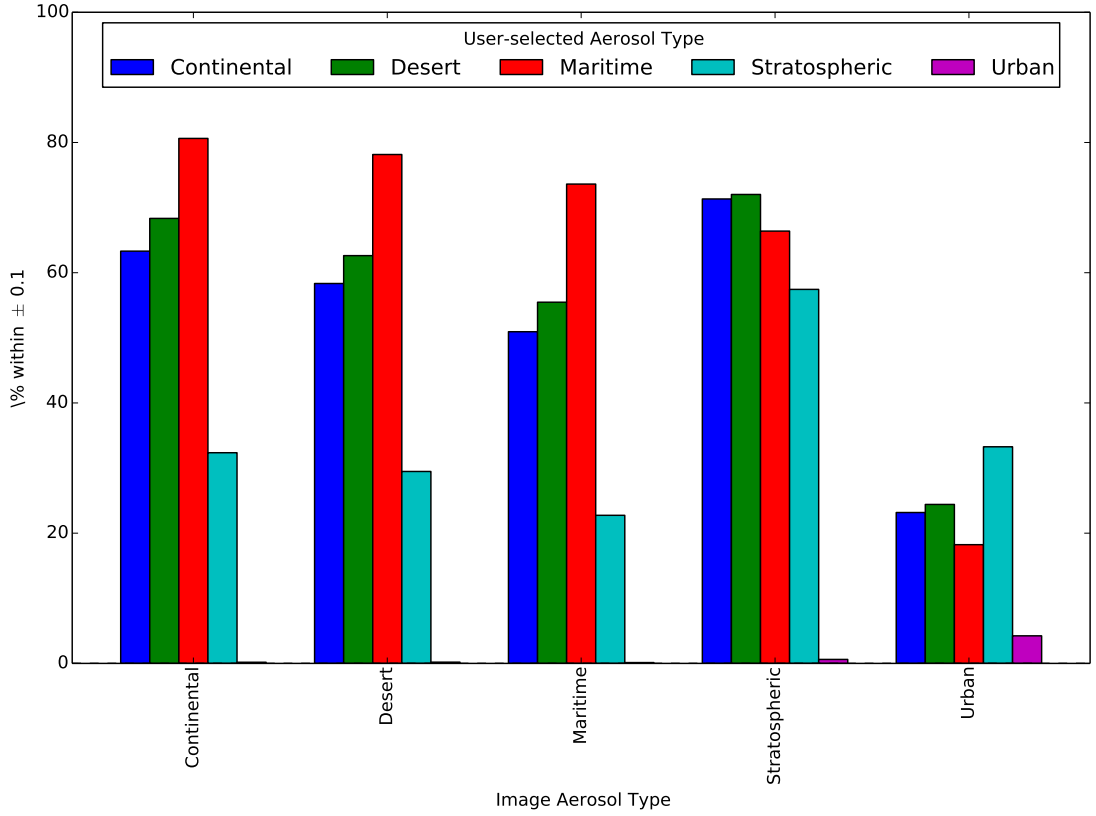
For a given AOT, simulations with an aerosol type with a high SSA (such as Maritime) will produce higher HOT values than those with a lower SSA (such as Continental). It is this increase in HOT magnitude which is key here: by increasing the magnitudes of the simulated HOT values we get a better relationship. This could be a genuine effect - due to HOT generally being over-estimated (perhaps due to a poor correction for Rayleigh scattering in the pseudo-BOA reflectance calculation process), or it could be another side-effect from the LEDAPS issues. For example, if LEDAPS under-estimates the AOT value for a pixel then the correction procedure will leave some residual atmospheric effects, which will still be present in the simulated test images. These residual effects will increase the HOT value for the pixel, and thus an artificially-increased HOT value produced using the Maritime aerosol type will produce a better-fitting relationship. The only way to test this would be to use a 'perfectly corrected' image which did not have the issues with the LEDAPS, or to use a very dense network of ground data to perform a very accurate correction over a small area of an image, over which the HOT aerosol type sensitivity could then be tested.

6.5 Algorithm summary

The development of a method to estimate the HOT-AOT relationship for a given image completes the development of an algorithm to estimate per-pixel AOT from Landsat images. From this point onwards, no changes have been made to the method: future



(a) London



(b) Midwest

FIGURE 6.13: Percentage of AOT values within ±0.1 for images simulated with one aerosol type (x-axis) and relationships modelled using a different aerosol type (colours)

TABLE 6.5: Single Scattering Albedo (SSA) and Mean % within ±0.1 for each aerosol type, excluding Stratospheric. Each table is sorted by the values within it, in descending order, and both tables have the same order.

(a) Single Scattering Albedo (SSA)

Aerosol	SSA
Maritime	0.990
Desert	0.952
Continental	0.899
Urban	0.693

(b) Mean % within ±0.1

Aerosol	Mean % within ±0.1
Maritime	28.7
Desert	25.2
Continental	24.7
Urban	01.9

chapters focus on validating and applying the results. The inputs, outputs and main steps in the method are summarised below:

Inputs:

- Raw Landsat ETM+ image, with pixel values as Digital Numbers (either stacked into a multi-band image, or each band as individual files)
- Corresponding Landsat ETM+ Metadata file
- T_{high} and T_{low} parameters for segmentation

Processing stages:

1. **Pre-process image:** Clouds, cloud shadows, water and snow are masked out of the input image, and the image is corrected for pseudo-BOA reflectance (that is, standard BOA reflectance but without any correction for aerosol effects). See §5.2.
2. **Estimate Clear Line:** The Clear Line is estimated by taking a sample of 5000 points from the LandsatAERONET dataset based on the land cover proportions found within the input image and performing an Orthogonal Distance Regression (ODR) on the blue and red reflectances of this sample. See §5.3.
3. **Create HOT image:** An image is created where each pixel contains the Haze Optimized Transform (HOT) value of the corresponding pixel in the input image, calculated as the distance from the pixel to the Clear Line in blue-red feature space. See §5.3.3.
4. **Correct HOT image:** The HOT image is corrected to the effects of land cover, using an Object-based Image Analysis approach which corrects sharp-bordered anomalies. See §5.4.
5. **Estimate HOT-AOT relationship:** The relationship between HOT values and AOT is estimated by taking the sample of LandsatAERONET points (as in step 2), and simulating the pseudo-BOA reflectance that would result if pixels with these reflectances were acquired with a range of random AOTs. The HOT value for each of these simulated pixels is then calculated, and Orthogonal Distance Regression is used to estimate the relationship between these HOT values and the AOTs used in the simulation. See §6.4.

6. **Convert image to AOT:** The HOT-AOT relationship developed in the previous step is applied to each pixel in the image to produce a final image containing AOT values.

Outputs:

- AOT image, with the same size and filetype as the input image
- Parameters file specifying all input and output files plus the parameters used for processing (including all internal parameters which are always kept at default values)
- Log file stating the results of specific processing stages (such as the parameters of the Clear Line) and the time taken to do the processing
- Lookup Table file, storing the lookup table used for simulating TOA radiance under differing AOTs as part of the estimation of the HOT-AOT relationship (this file can be automatically used in future processing of this image, removing the need to regenerate it and thus saving significant time)

The algorithm has been implemented in the Python programming language, using the following third-party libraries:

- `numpy`: provides efficient array processing for Python (Walt *et al.*, 2011)
- `scipy`: provides a range of scientific programming functionality built upon numpy, including interpolation and regression (Jones *et al.*, 2001)
- `pandas`: provides a data type that can store and process multiple columns of heterogeneously-typed data, as in a spreadsheet (McKinney, 2010)
- `GDAL`: provides functions to read and write almost all geospatial formats, including TIFF files, ENVI format images and ESRI Shapefiles (GDAL Development Team, 2014)
- `mahotas`: provides a range of image processing functions (Coelho, 2013). Here the functions relating to mathematical morphology are used.
- `scikit-image`: provides further image processing functions. Here the filtering, labelling and Otsu thresholding functions are used (van der Walt *et al.*, 2014).

The algorithm has been designed in a modular manner, with functions to carry out each of the individual stages in the overall process. To execute the algorithm a class is instantiated with the required parameters, and then `run` method called. A `seg_test` method also exists to run the segmentation process with a range of parameters, to allow the user to easily select the parameters that produce the best segmentation. After specification of the parameters, the processing is entirely automatic, and helper functions exist to allow the entire process to be run automatically from a compressed Landsat image downloaded directly from the USGS. The code runs effectively on high performance computing clusters, and this allows multiple images to be processed in parallel.

From here on, the new algorithm will be known as HOTBAR, the Haze Optimized Transform Based Aerosol Retrieval. As an example, HOTBAR was run using the Landsat ETM+ image shown in Figure 6.14, with $T_{high} = 0.95$ and $T_{low} = 0.20$, and the output

FIGURE 6.14: Example Landsat ETM+ image of New York (LE70130322002251EDC00), shown as a true colour composite

AOT image is shown in Figure 6.15. Comparison of the HOTBAR output (at 30m resolution) with AOT images from the MODIS MOD04 (10km resolution) and MAIAC (1km) algorithms show the level of detail which is provided by HOTBAR (Figures 6.16 and 6.17).

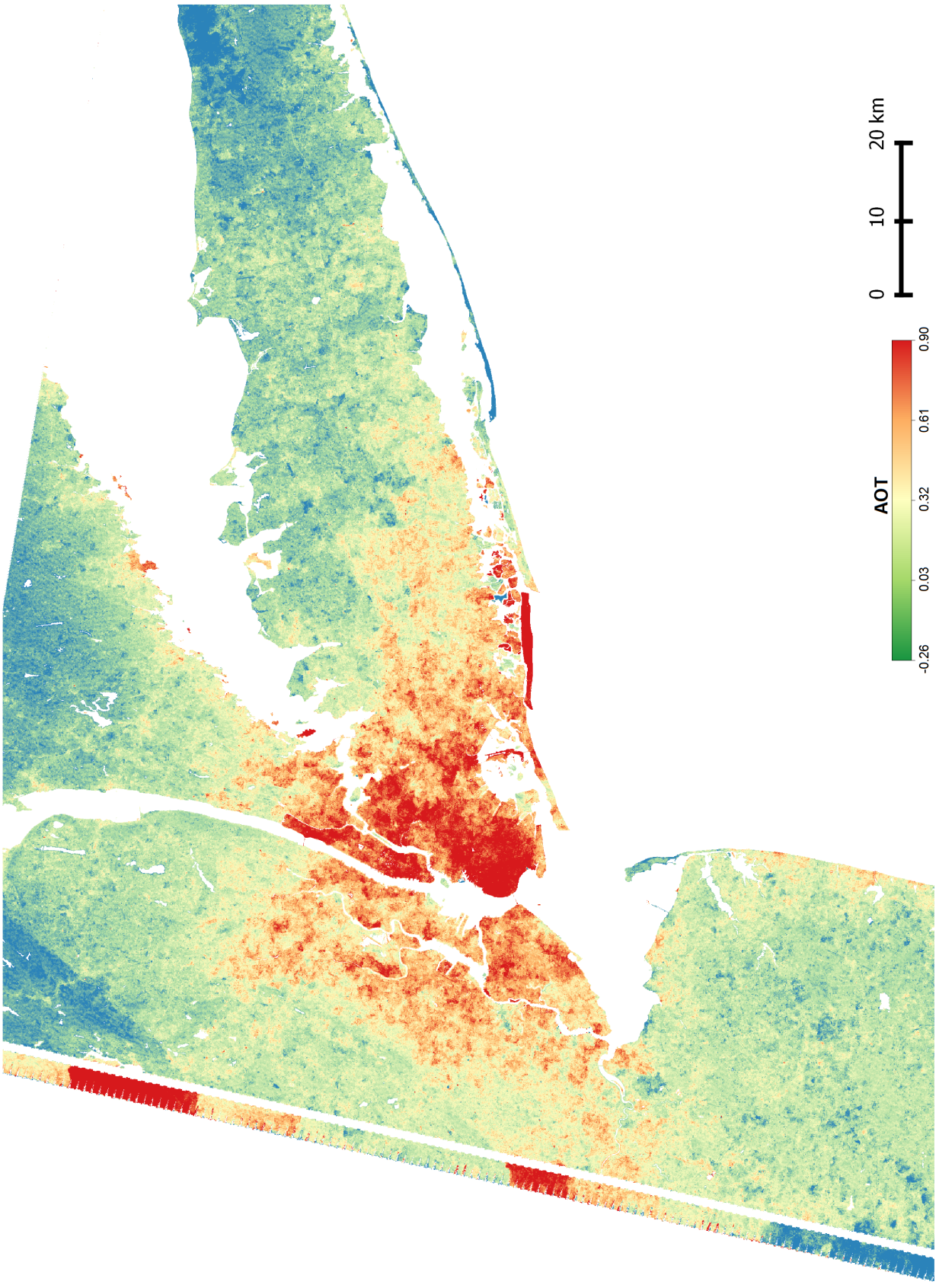


FIGURE 6.15: AOT image generated using the Landsat ETM+ image shown in Figure 6.14, using the parameters $T_{high} = 0.95$ and $T_{low} = 0.20$

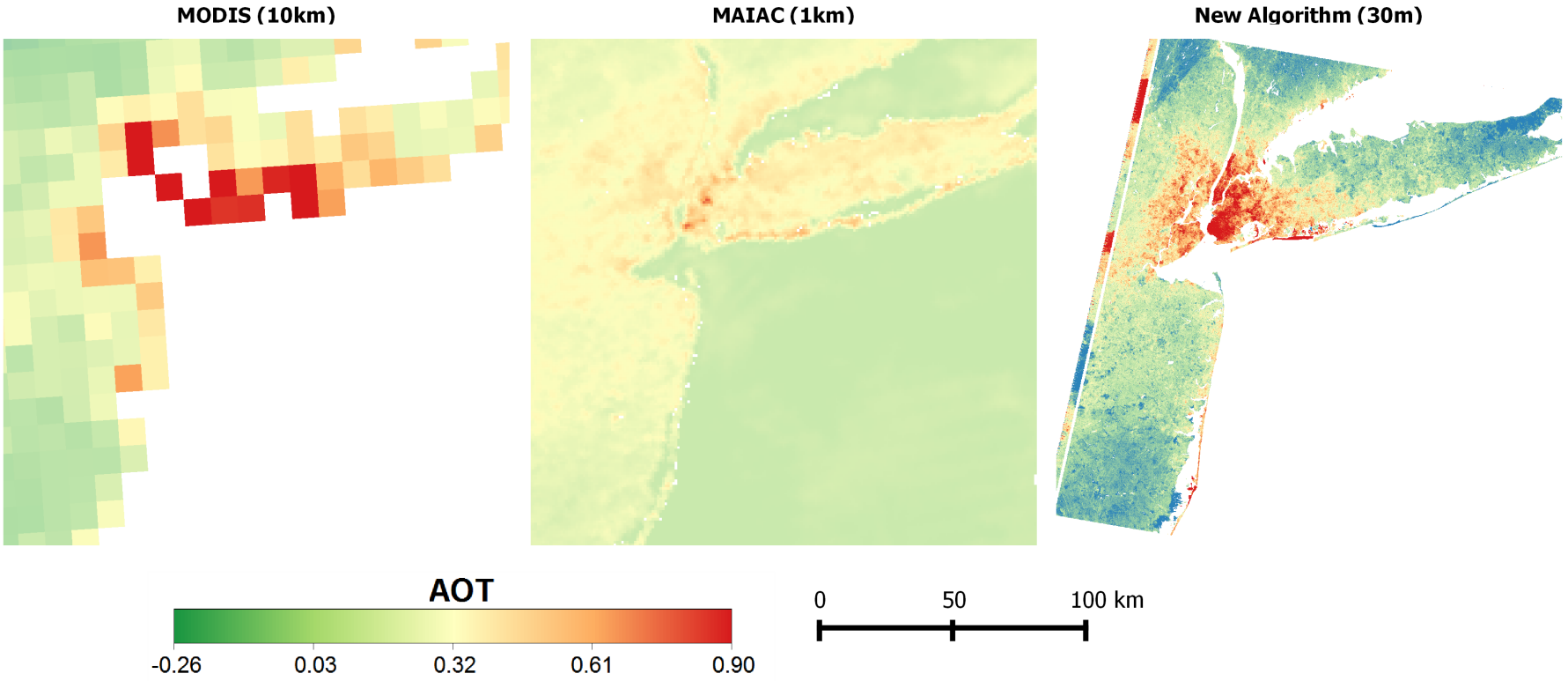


FIGURE 6.16: Comparison of AOT data over New York from three data sources: MODIS MOD04 (10km), MAIAC (1km) and the new algorithm (30m)

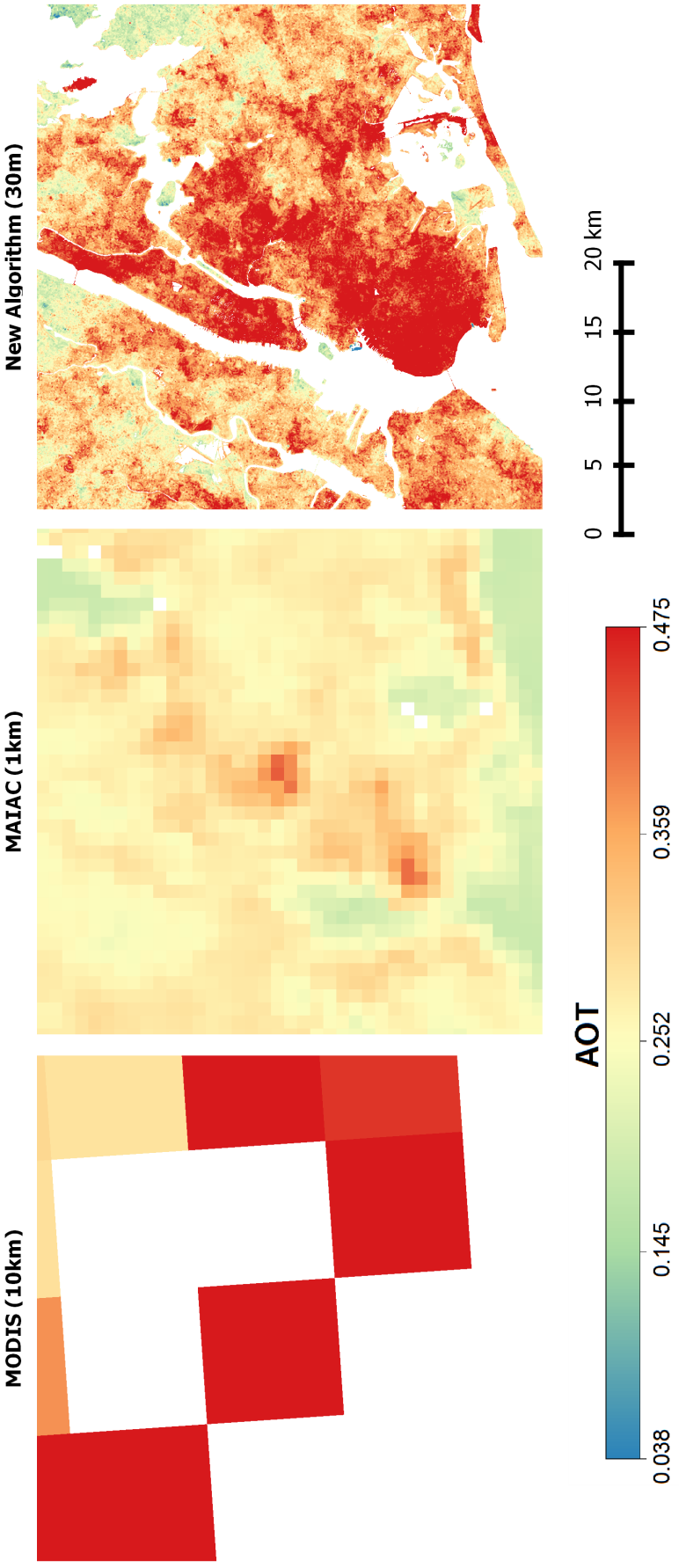


FIGURE 6.17: Comparison of AOT data over central New York from three data sources: MODIS MOD04 (10km), MAIAC (1km) and the new algorithm (30m)

Chapter 7

Validation

7.1 Introduction

An essential part of the development of a new method to retrieve physical measurements from satellite data is to assess the accuracy of the resulting retrievals through validation with other data sources. In the context of a high-resolution AOT retrieval method there are two principal considerations: the accuracy of the AOT values, and the 'spatial coherence' of these values. The latter refers to the relative accuracy of adjacent pixels: does the method show the correct spatial trend in AOT values, even if the values themselves are offset from the true value. These two aspects of the error will be assessed in this chapter, through comparison of HOTBAR results with AERONET ground-based AOT and MAIAC satellite-based AOT measurements, as well as further analysis of the AOT errors from the simulated images used in Chapter 6.

The purpose of the validation is to assess the accuracy of the AOT derived from HOTBAR, and thus it is useful to consider the potential sources of error in the AOT measurements. These are shown in Table 7.1, and can be split into two main categories: errors in the HOT values, and errors in the HOT-AOT relationship. Errors in the HOT values are likely to be the source of spatially-variable error, often caused by poor HOT correction, whereas errors in the HOT-AOT relationship are likely to be the source of magnitude-variable error, caused by an incorrect estimation of the slope of the HOT-AOT relationship, causing errors at high or low AOTs.

It is likely that some of the AOTs produced by the HOTBAR algorithm will be negative. A negative AOT is physically impossible, but this does not mean that these AOTs should be removed from the images. If the true AOT is 0.05 and HOTBAR has an error of ± 0.1 then HOTBAR could produce an AOT estimate anywhere between -0.05 and 0.15. Other AOT retrieval algorithms have also had to deal with this issue, for example, the MOD04 Algorithm Theoretical Basis Document (Remer *et al.*, 2006) states that negative AOT retrievals are not just allowed, but necessary: removing all negative retrievals will, by definition, bias the dataset. The validation of the MOD04 product states that the error is

TABLE 7.1: Categorisation of potential error sources within the HOTBAR algorithm

Error source	Examples
HOT value errors	
Pre-processing	Poor cloud-masking Inaccuracies in 6S correction to pseudo-BOA
HOT calculation	LandsatAERONET dataset not representative Poor regression for Clear Line estimation
HOT correction	Under-correction Over-correction Edge effects
HOT-AOT relationship errors	
Actual HOT-AOT relationship	Aerosol type (eg. urban) Non-linearity for high AOTs
Modelling of HOT-AOT relationship	LandsatAERONET dataset not representative Inaccuracies in 6S simulation or LUTs Poor regression

$0.05 \pm 0.15\tau$, and thus they treat any AOT between -0.05 and 0 as 'high quality', assuming that it represents an AOT of very close to zero. Retrievals between -0.10 and -0.05 are marked as 'low quality', and those below -0.10 are marked as 'out of range'. In this case, during the validation all negative values will be included, though extremely negative values will be noted, and equivalent thresholds for 'good' and 'bad' negative AOTs will be determined in the conclusion of this chapter.

Due to the Landsat 7 Scan Line Corrector (SLC) failure, images acquired since May 2003 have regularly spaced lines of missing pixels (Markham *et al.*, 2004). This missing data causes the segmentation stage of the object-based HOT correction to perform very poorly, and so data acquired in 'SLC-off' mode cannot be used with HOTBAR. Thus all validation must be carried out with data acquired before May 2003, which significantly limits the volume of validation data available.

7.2 AERONET validation

AERONET data are widely used as a 'ground truth' in validation of satellite AOT products due to its very low error (approximately $\pm$ 0.02 according to Eck *et al.*, 1999) and its high temporal resolution. All AERONET sites measure AOT at multiple wavelengths, allowing the AOT data to be interpolated to estimate AOT at 550nm, the wavelength at which HOTBAR estimates AOT.

Most previous validations of satellite AOT against AERONET AOT have used the Ichoku *et al.* (2002) method to take into account the resolution difference between a coarse-resolution satellite measurement and a single-point ground measurement. Ichoku *et al.* (2002) assumed that aerosols moved at around 50km/hr and suggested comparing a

one hour temporal average of AERONET data with a 50x50km spatial average of satellite data, and these parameters have been used for almost all satellite AOT validations (including Liu *et al.*, 2014; Benas and Chrysoulakis, 2013; Bréon *et al.*, 2011; Wong *et al.*, 2013). However, when validating high-resolution data the main reason for using the Ichoku *et al.* (2002) method is invalidated: instead of comparing a point measurement with a large (many kilometer) pixel, we are comparing it with a 30m pixel. To apply the Ichoku method with the standard parameters to this high-resolution data would require spatially-averaging over two million pixels: an absurd idea. Given this, the Ichoku *et al.* (2002) method is not used in this validation, and ground measurements are compared to the directly-overlaying 30m pixel in the satellite AOT.

In 2003 there were 207 AERONET sites, but only 119 of these collected data over at least six months of the year (many were temporary sites used for testing or specific campaigns) - this compares unfavourably with 2013, in which there were 443 and 292 sites respectively. HOTBAR requires parameterisation with the high and low thresholds (T_{high} and T_{low}) for segmentation. These can be estimated by testing a range of different parameterisations for each image location: a time-consuming activity when processing many images. Thus, validation was only performed over selected AERONET sites: these were chosen to cover a range of different areas of the world and types of land cover, as well as attempting to choose places where multiple AERONET sites would be located within one Landsat image. The Landsat path/row locations listed in Table 7.2, and shown in Figure 7.1, were chosen, covering 12 AERONET sites - around 10% of the sites regularly collecting data during 2003.

Table 7.2: Landsat path/row locations used for AERONET validation, along with chosen segmentation parameters

Path	**Row**	$T_{\textbf{high}}$	$T_{\textbf{low}}$
015	033	0.95	0.20
041	036	0.85	0.75
199	024	0.95	0.20
199	025	0.85	0.75
199	030	0.95	0.20

All images with less than 10% cloud cover acquired during the whole period of Landsat 7 SLC-on operation (May 1999 until May 2003) were downloaded from the USGS and processed through HOTBAR. Segmentation parameters were determined by testing a range of parameters on one image for each path/row location used (the chosen parameters are shown in Table 7.2), and then applied to all other images at that location. AOT values from the satellite image were extracted at all AERONET sites within the image - regardless of whether the AERONET site was operating at the time of the image or not - and then AERONET measurements within 30 minutes of the image acquisition time for these sites were extracted. Error, calculated as the AERONET AOT minus the satellite AOT, was then calculated for each site for which both measurements were found.

FIGURE 7.1: Locations of the Landsat images used for AERONET validation

AERONET Level 2.0 data were used, as this is the highest quality data available. Tests were made with Level 1.5 data but no significant increases in the data volume were found, and the validation statistics remained similar. AOT data from HOTBAR were produced with and without the 3x3 median smoothing described in §5.4.8, to determine whether the smoothing improved the real-world performance of the AOT product.

7.2.1 Results

Matched AERONET and satellite observations were found for 12 sites, with a total of 27 observations. Summary statistics (Table 7.3) and a histogram (Figure 7.2) of the error show a higher error than that found from the simulations in the previous chapter (33% within ±0.1 compared to 40–50% for the majority of non-desert simulated images). This could be due to the greater complexity of real-world images, poor representativeness of the simulated images, or simply due to the limited sample of AERONET sites which did not cover areas with particularly low error. HOTBAR generally over-estimates AOT values, although negative AOTs contribute to the 33% of values which are underestimated. Smoothing of the image data improved almost all of the statistics, and significantly reduced the maximum error (from 0.75 to 0.59) as anomalously high pixels were reduced by the smoothing process. The median error increased, but this is almost certainly an anomaly caused by the internal distribution of the data.

Examining the results on a per-site basis (Figure 7.3) shows a very large range of errors for Rogers_Dry_Lake, which includes both the largest positive and negative errors in the dataset. This is the only desert site, and desert sites performed badly with the simulated images (§6.4). The large range is likely due to the poor performance of the HOT correction on desert areas: if no correction (or a very poor correction) is performed on the images then the AOT values estimated at a single point over time will be almost entirely due to changes in surface reflectance. For the rest of the sites, median errors range from +0.06 to -0.33. Only two sites under-estimate the majority of their data: The_Hague (for

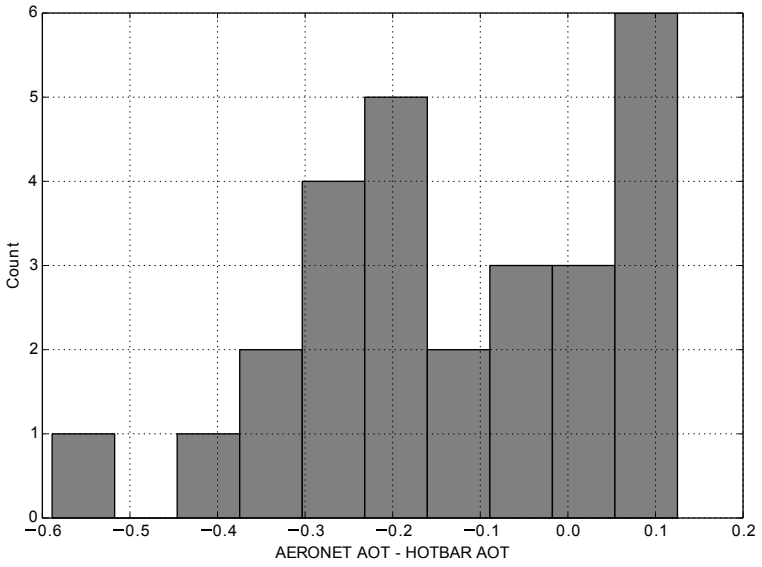


FIGURE 7.2: Histogram of AOT error from AERONET validation, based upon the smoothed AOT image

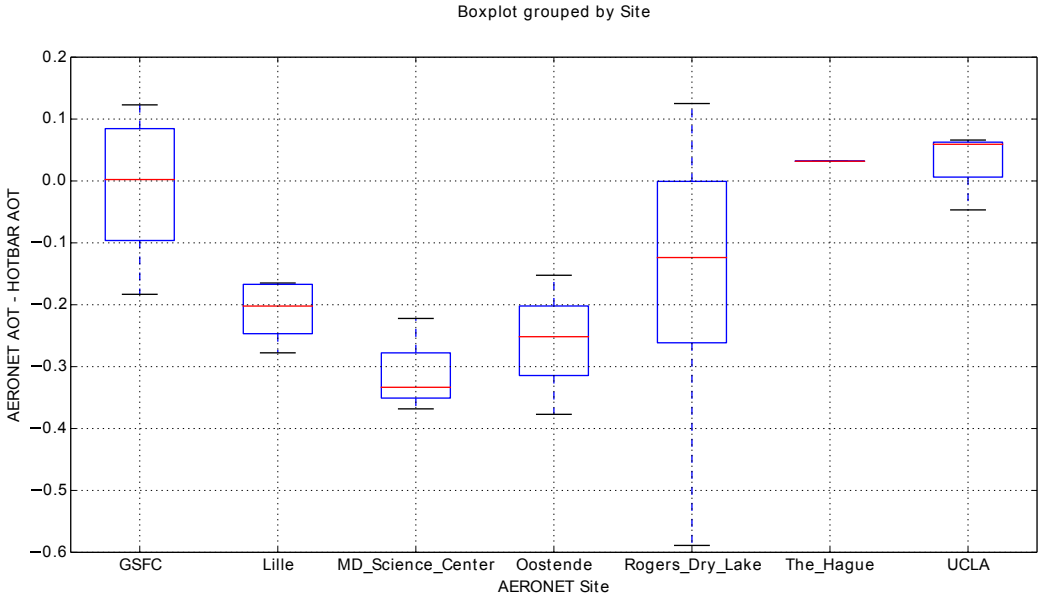


FIGURE 7.3: Boxplots of AOT error from AERONET validation, based upon the smoothed AOT image, grouped by AERONET site

TABLE 7.3: Summary statistics of absolute AOT error according to the AERONET validation for the raw image and the image with a 3x3 Median filter applied

	Original	**Smoothed**
Min	0.047	0.002
Mean	0.179	0.173
Median	0.138	0.152
Max	0.745	0.589
% within ± 0.1	30%	33%
% overestimated	74%	67%
RMSE	0.232	0.219

which there is only one data point, hence the narrow boxplot) and UCLA; all of the rest over-estimate.

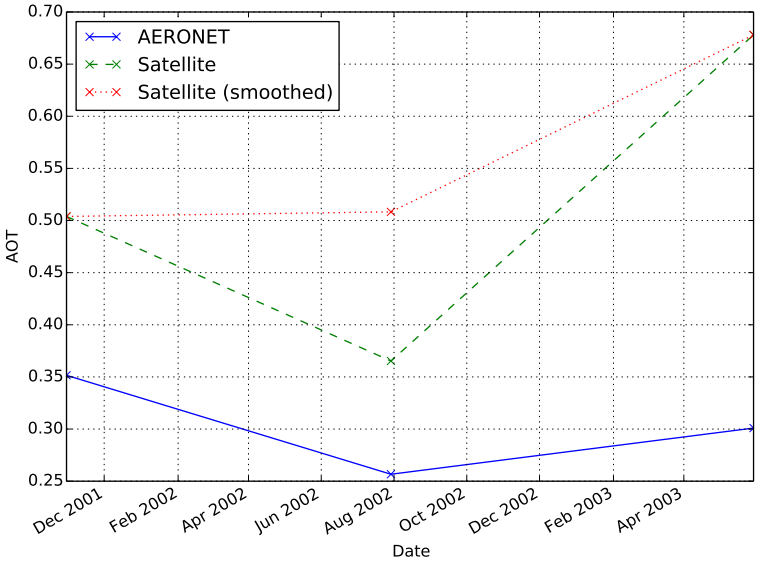
Examining the AOT derived from AERONET and satellite measurements over time shows very similar trends for some sites (Figure 7.4), but differing trends for others (Figure 7.5). Again, the desert site (Figure 7.5(b)) performs badly, with very high AOT errors around spring and early summer. These are likely caused by seasonal changes in the surface reflectance which have not been removed by the HOT correction procedure. It is harder to explain the poor performance of some other sites (Figures 7.5(a) and 7.5(c)), although the significant differences between the raw and smoothed satellite data in Figure 7.5(c) suggest issues with correction in the pixel directly over the AERONET site (as the smoothed data has a far closer trend to the AERONET data).

Comparing the error at different AOT magnitudes (Figure 7.6) shows no trend, suggesting approximately constant error across the whole range of AOTs used for the validation. It should be noted, however, that this range is limited (approximately 0.02–0.35), and there may be significant trends in error for AOTs outside of this range.

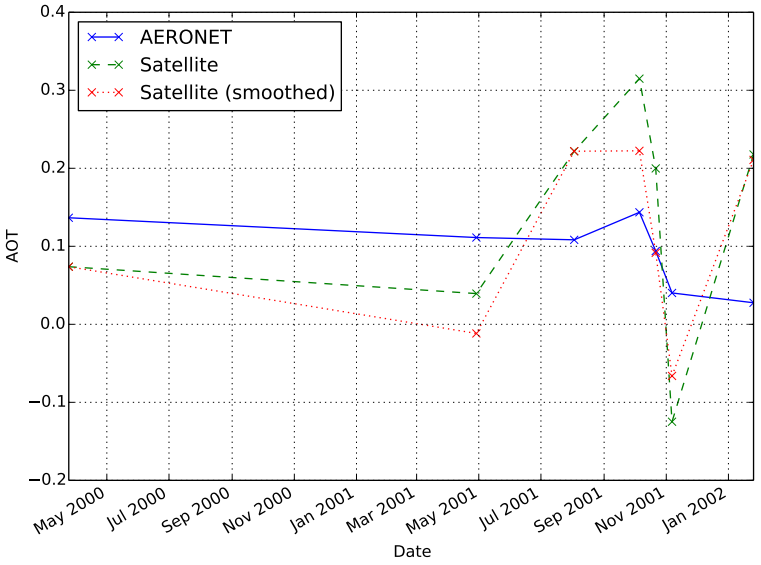
7.3 MAIAC validation

The fundamental problem with the AERONET validation is that the spatial coverage is poor, which makes it very difficult to assess the 'spatial coherence' of the HOTBAR results: the only way to assess this using real-world data is by comparison with AOT images derived using a different method. There are no other 30m AOT products currently available, so the MAIAC 1km AOT product (Lyapustin *et al.*, 2011) was used, as it is the highest resolution AOT data currently available to researchers (results from the SYNTAM method (Tang *et al.*, 2005), which also produces 1km AOT data, are not currently available for academic use).

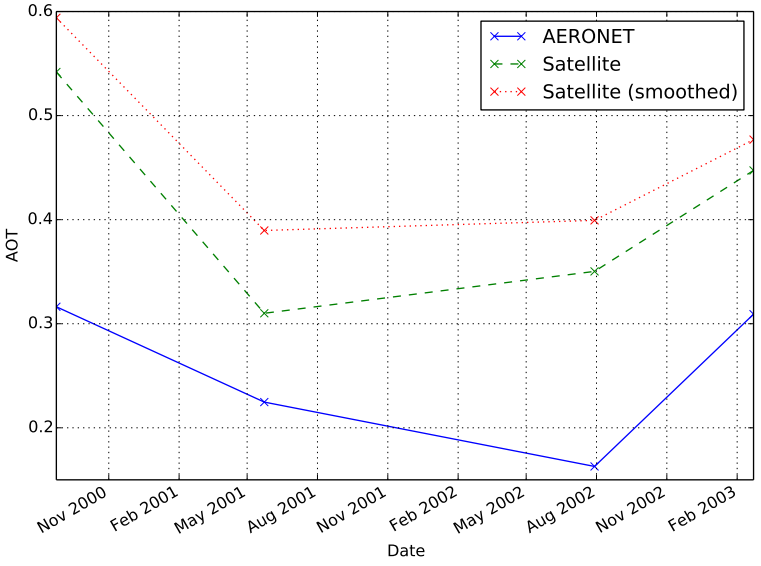
MAIAC data were acquired for September 8th 2002, on the same date and covering the same area as the Landsat image shown in Figure 6.14 (MAIAC tile h09v08, Landsat image LE70130322002251EDC00). This was chosen as the image has a mix of urban and rural land covers, and was one of the few areas where MAIAC data was available on the same day as a Landsat 7 SLC-off acquisition. The HOTBAR algorithm was run on the



(a) Oostende

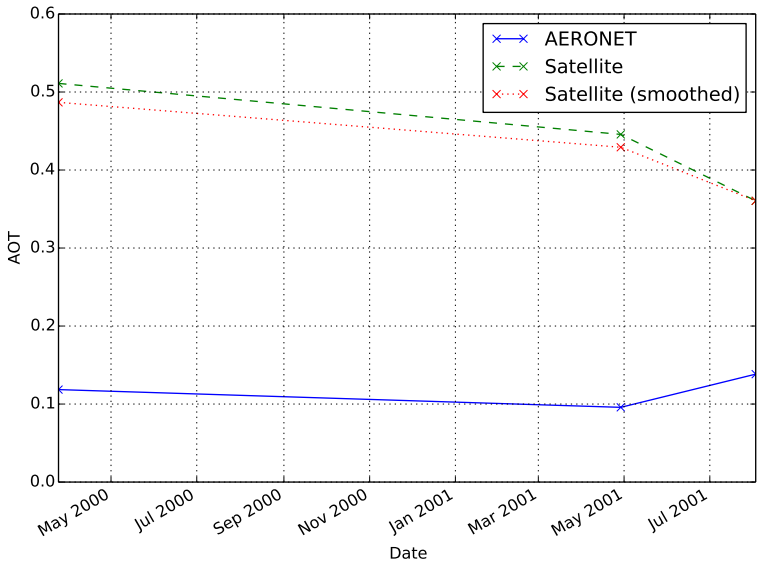


(b) GSFC

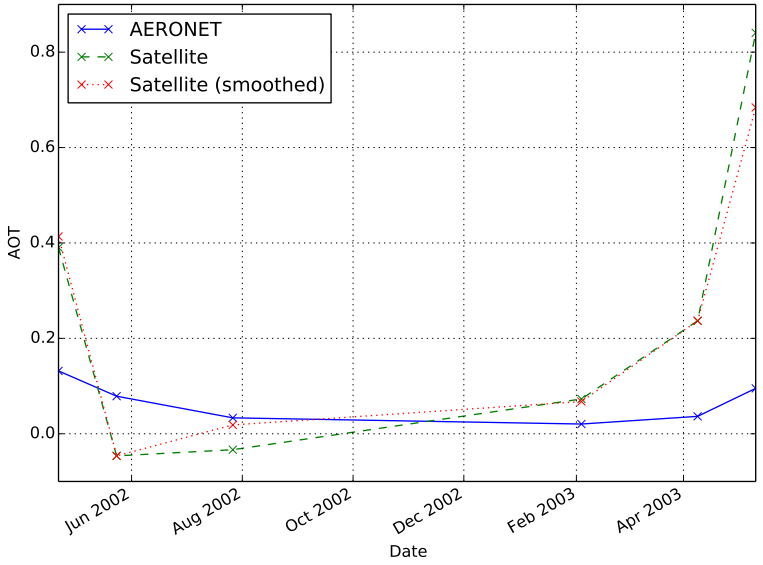


(c) Lille

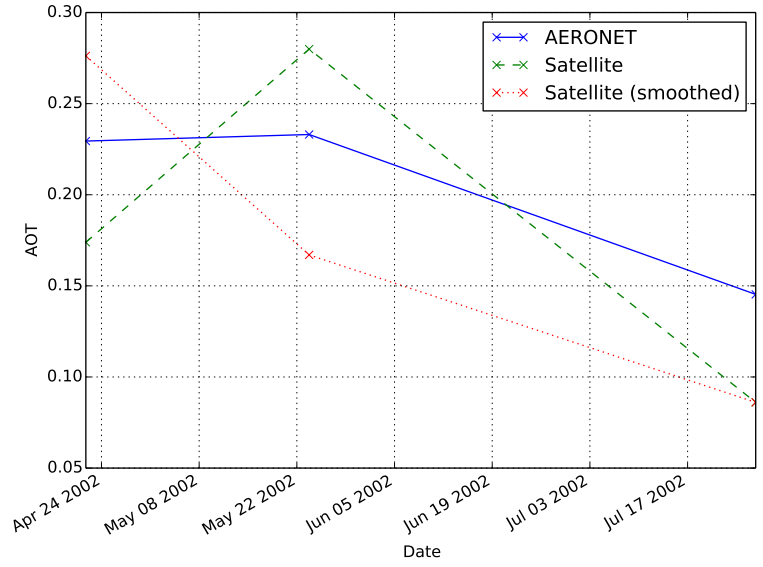
FIGURE 7.4: Time series plots for AERONET validation sites with good matches in trends



(a) MD_Science_Center



(b) Rogers_Dry_Lake



(c) UCLA

FIGURE 7.5: Time series plots for AERONET validation sites with poor matches in trend

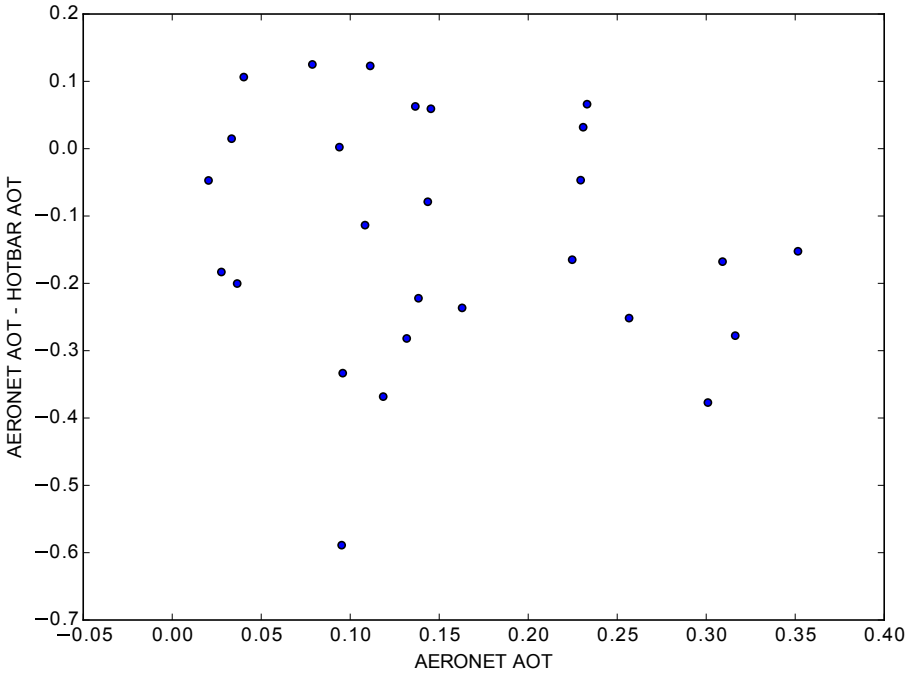


FIGURE 7.6: Relationship between error and the AERONET AOT value, showing no significant trend between AOT magnitude and the error

Landsat image, and the AOT result was resampled to a 1km resolution using a median aggregation (Figure 7.7) to allow comparison with the MAIAC data. The AOT error was then calculated as the MAIAC AOT minus the HOTBAR AOT.

There are several issues with comparing the HOTBAR AOT with the MAIAC AOT:

1. The AOTs are measured at different wavelengths: MAIAC measures AOT at 470nm, whereas HOTBAR measures AOT at 550nm. For the AERONET validation the AOT was interpolated to produce an estimated value at 550nm, however this is not possible here as MAIAC only measures AOT at one wavelength and thus there is not enough data for the interpolation.
2. There is a time offset between the acquisition of the Landsat image at 15:21 and the MODIS image at 17:30 (all times in UTC). During this time the AOT, and its spatial pattern, is likely to have changed.
3. The resampling of the HOTBAR AOT to match the MAIAC resolution may introduce artefacts into the resulting 1km images (such as anomalous pixels in areas where HOTBAR had limited data availability due to the presence of cloud, water or snow), and thus affect the error.

However, this is currently the only feasible method to assess spatial error across the area of a Landsat image, so it used with these issues taken into account.

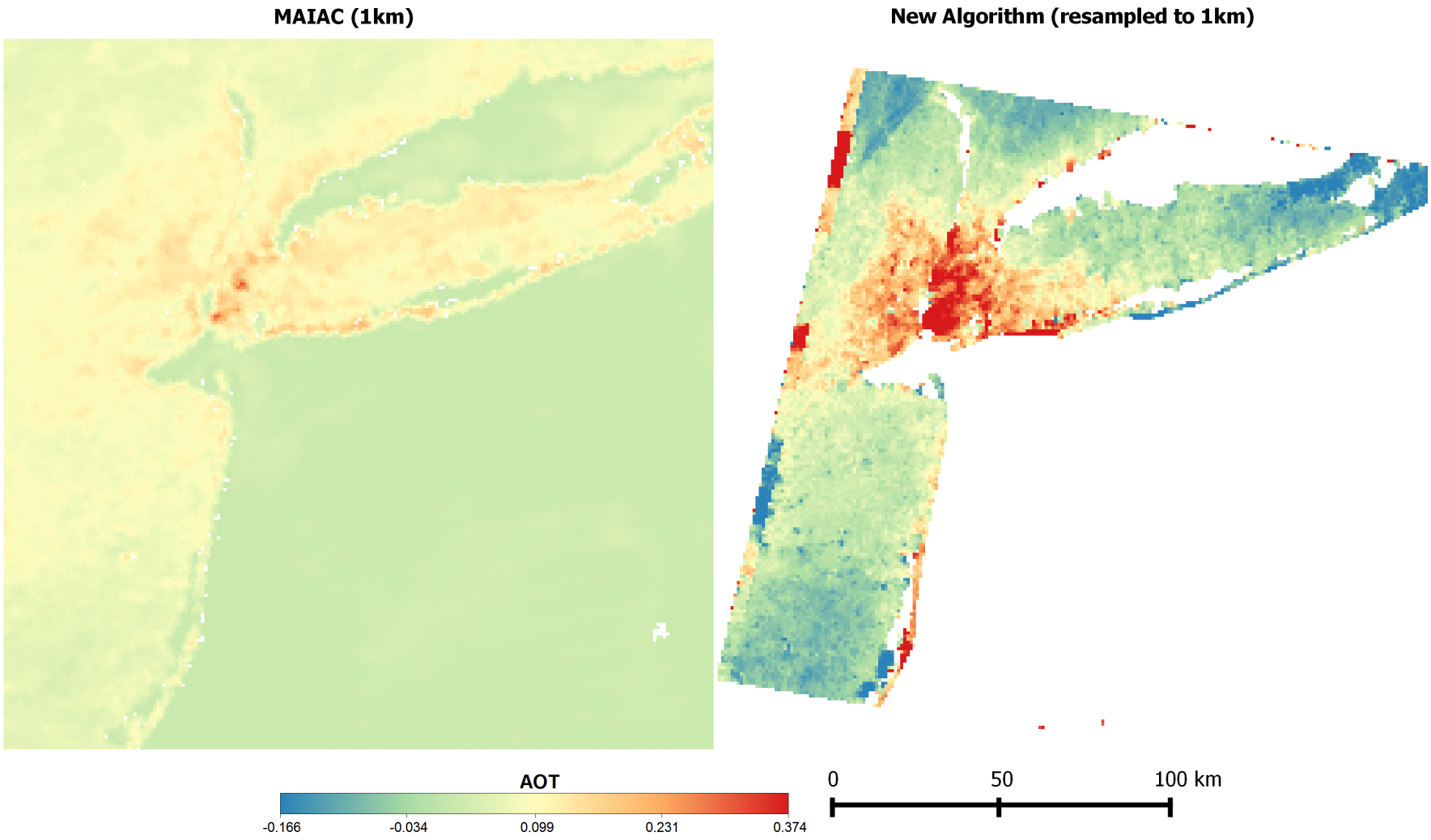


FIGURE 7.7: MAIAC AOT (at 1km resolution) and HOTBAR AOT (originally at 30m, resampled to 1km) over New York, using the same images as in Figure 6.16. The original 30m AOT image is shown in Figure 6.15.

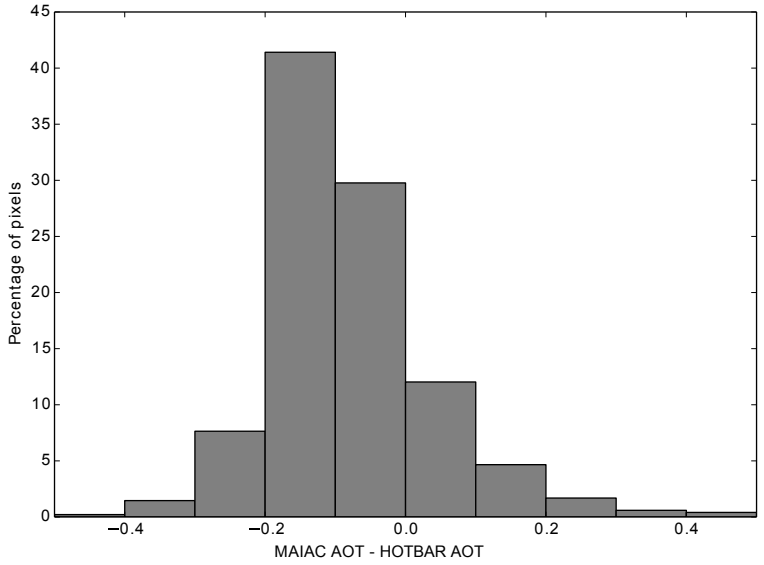


FIGURE 7.8: Histogram of MAIAC validation error, calculated as MAIAC AOT - HOTBAR AOT

7.3.1 Results

The error histogram (Figure 7.8), and summary statistics (Table 7.4) show that 42% and 87% of the HOTBAR AOT values are within ±0.1 and ±0.2, respectively, of the corresponding MAIAC values. This is a slightly lower accuracy than that found with the simulated images in §6.4, but better than the AERONET validation, which showed only 33% within ±0.1. As with the AERONET validation, the majority of pixels (80%) are over-estimated, but the mean and median absolute errors are slightly better than AERONET (Mean: 0.124 compared to 0.173, Median: 0.116 compared to 0.152). The maximum error is significantly worse, at 1.82 compared to 0.589 for AERONET, but this is likely to be due to resampling artefacts (visible as isolated dark purple pixels in the error map, shown in Figure 7.9).

TABLE 7.4: Summary statistics of absolute AOT error according to the MAIAC validation

Min	0.000
Mean	0.124
Median	0.116
Max	1.82
% within ± 0.1	42%
% overestimated	80%
RMSE	0.147

Examining the error for different AOT magnitudes (Figure 7.10) shows a possible trend with positive errors (under-estimation) at low AOTs, and negative errors (over-estimation) at higher AOTs. However, this may not be a genuine trend due to the small range of MAIAC AOTs (approximately 0.05–0.30), and the uneven distribution of AOT magnitudes throughout the image (with far fewer high AOT pixels, and thus a narrower distribution for higher AOTs). The magnitude dependence is confirmed by the error map,

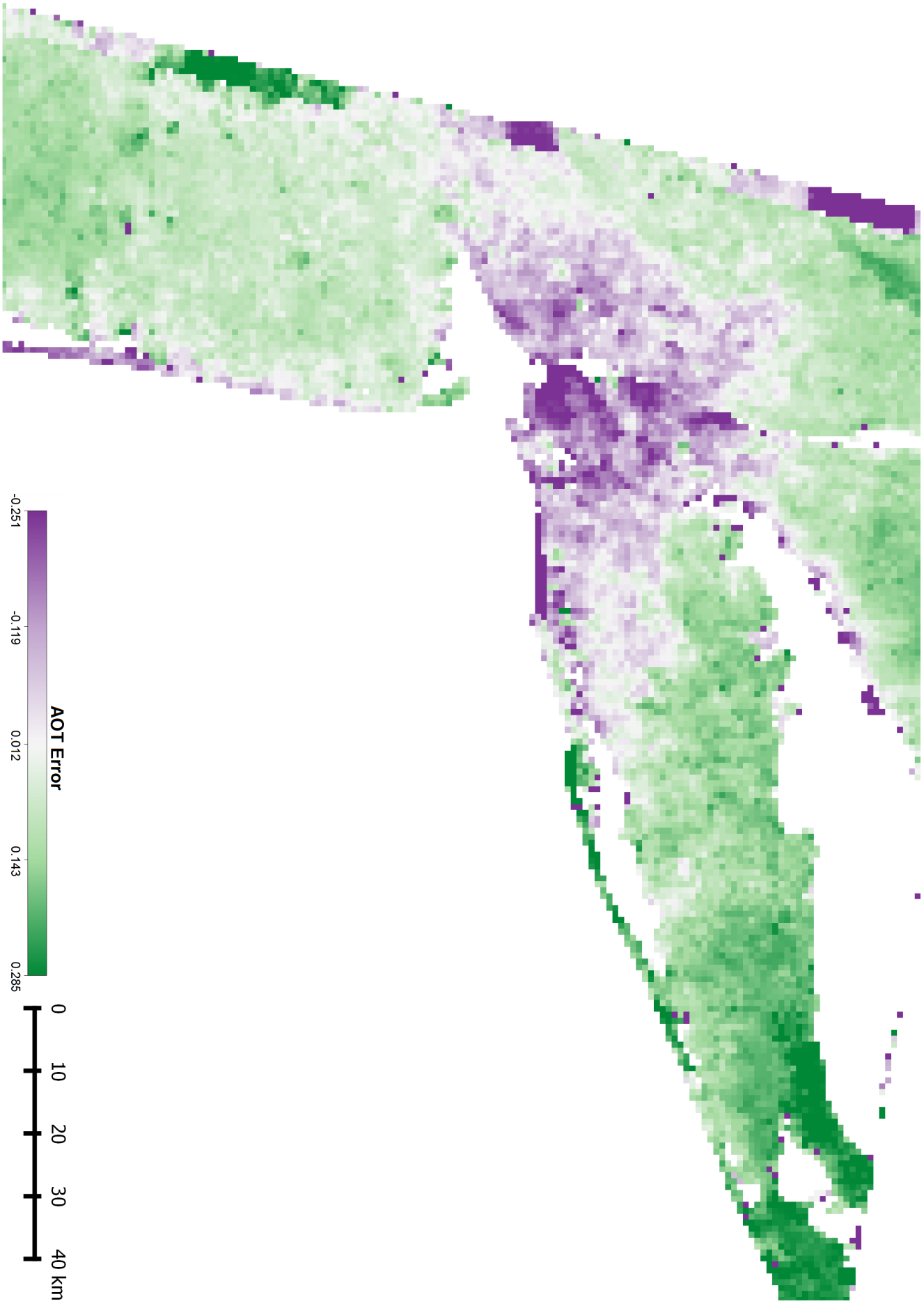


FIGURE 7.9: MAIAC validation error image, showing MAIAC AOT - HOTBAR AOT

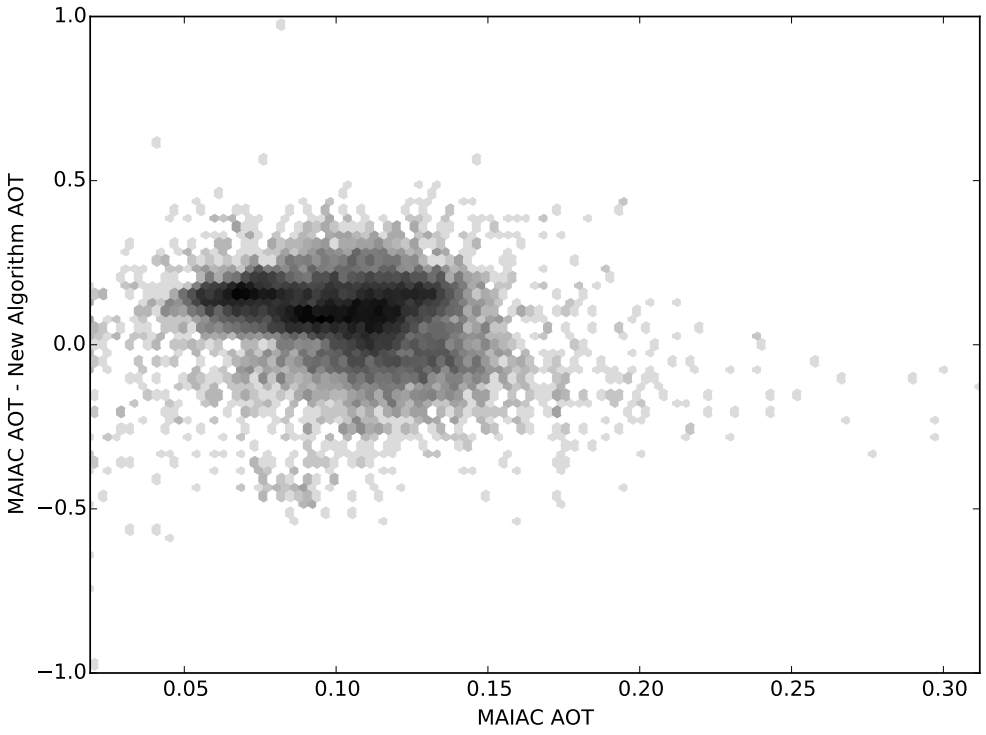


FIGURE 7.10: Relationship between error and the MAIAC AOT value

with over-estimation within the centre of New York City (where AOTs are generally high), and under-estimation over the rest of the image (where AOTs are lower). This also affects the histogram of errors (Figure 7.8), as there tends to be a negative error for low AOTs, and the majority of the image consists of relatively low AOTs.

The errors from comparison with MAIAC data must be considered in light of the MAIAC data itself. According to Emili *et al.* (2011), MAIAC has a RMSE of 0.05 when compared with AERONET sites, and thus the true RMSE of the HOTBAR results could be anywhere between 0.097 and 0.197. Similar adjustments will be needed for the other error statistics, but, as these are not reported for the MAIAC validation, exact values cannot be provided.

7.4 Simulated images validation

The only way to truly assess the spatial coherence of the AOT values is through the use of the simulated images which were created and used in Chapter 6, as this is the only situation in which we know the true AOT value for each pixel in the image. Due to the difficulties in creating the random AOT images (see §6.2.1), the same images as used in Chapter 6 are used here for validation. All of the validation in this section has been performed on the Chilbolton image simulated with Continental aerosols, except where otherwise specified, as this image covers a wide range of landcovers and all other images produce very similar results.

Comparing the error at different AOT magnitudes (Figure 7.11(a)) appears to show a

TABLE 7.5: Summary statistics of absolute AOT error according to validation against the Chilbolton image simulated with Continental aerosols

Min	0.000
Mean	0.123
Median	0.104
Max	5.333
% within ± 0.1	48%
% overestimated	67%
RMSE	0.157

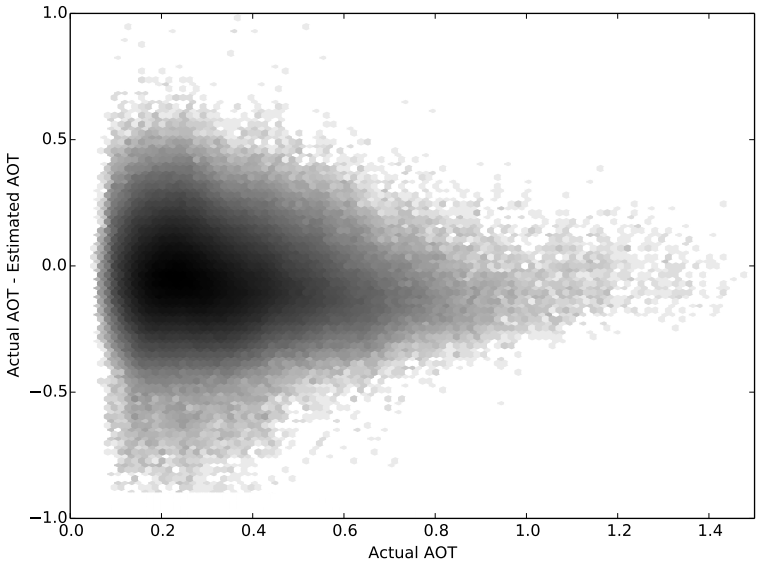
strong inverse relationship, with error reducing as AOT values increase. However, this apparent relationship is mostly due to the distribution of AOTs within the simulated image: there are very few pixels with high AOT values, and thus the range of error appears to be lower at this magnitude. Taking an equally-weighted sample of 100,000 points from each AOT interval (0.0–0.1, 0.1–0.2, and so on) shows a far weaker relationship (Figure 7.11(b)), with only a slight decrease in error as AOT increases. The directionality of the error at varying AOT magnitudes can be seen more clearly in Figure 7.12, which shows under- and over-estimation at lower AOTs, with some extremely over-estimated values (caused by poor HOT correction), and general over-estimation at higher AOT values.

7.4.1 Profile comparisons

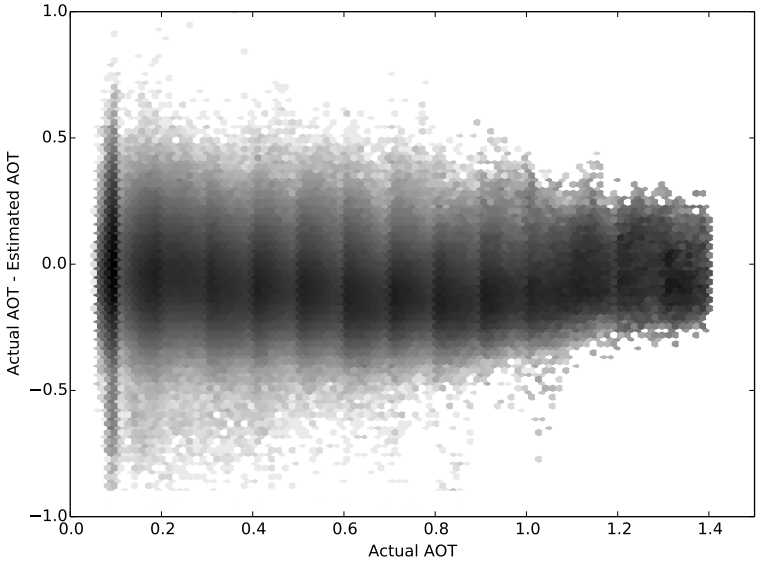
Profiles of actual and estimated AOT (raw and smoothed) from two images are shown in Figure 7.13. These show that the smoothed AOT replicates the trend in actual AOT relatively well: the overall shape of the curves is very similar. On a pixel level the match is not as good, but it is remarkable how many small-scale variations in AOT are captured by the HOTBAR AOT for the Amazon image (Figure 7.13(a)). The results for the SotonNewForest image (Figure 7.13(b)) are worse, and the effects of several areas of poor correction can be easily seen: however, even with these anomalies, the overall trend matches well. Both graphs show the importance of adding smoothing as a final step within the HOT correction procedure, as it is very effective in removing many sharp variations in AOT caused by individual poorly-corrected pixels.

7.4.2 Spatial variation of the error

A key aspect of the spatial variation of the error is the effect of the HOT correction procedure: if pixels which have been modified during the correction process have a higher error than other pixels then that would suggest that the correction process was increasing, rather than reducing, the error. Figure 7.14 shows that pixels modified by the correction process have a wider error range, with more outliers (2% of data outside of the range of Figure 7.14(b), compared to 0.8% for unmodified pixels) than unmodified pixels. However, when these outliers are removed from the dataset, the changed pixels have a very similar error range to the unchanged pixels, with a slightly lower median. This suggests that the correction process performs well on the majority of pixels, but around 1% of pixels (the



(a) Random sample from all points



(b) Sampled with 100,000 points in each AOT interval (the apparent banding is due to the discrete intervals used for sampling)

FIGURE 7.11: Relationship between error and the actual AOT value for the Chilbolton image simulated with Continental aerosols, shown with (a) a random sample from all points, and with (b) an equally-weighted sample (to remove the effect of the distribution of AOT values in the image)

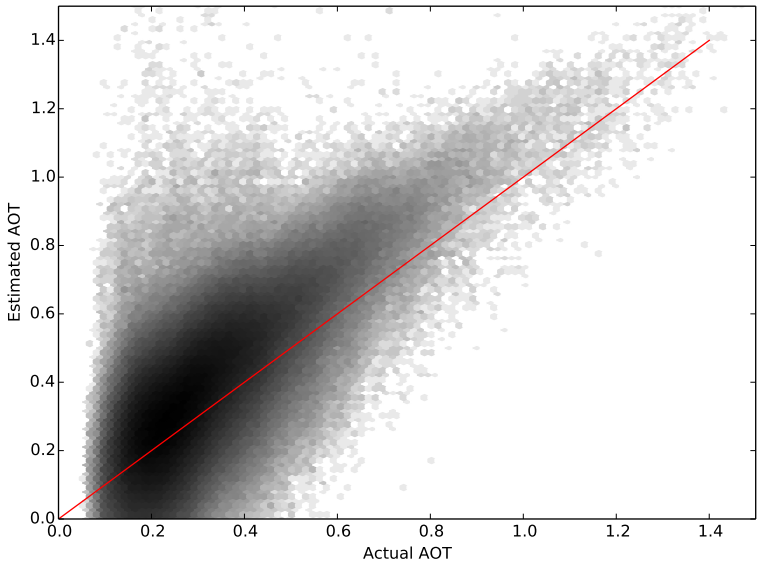


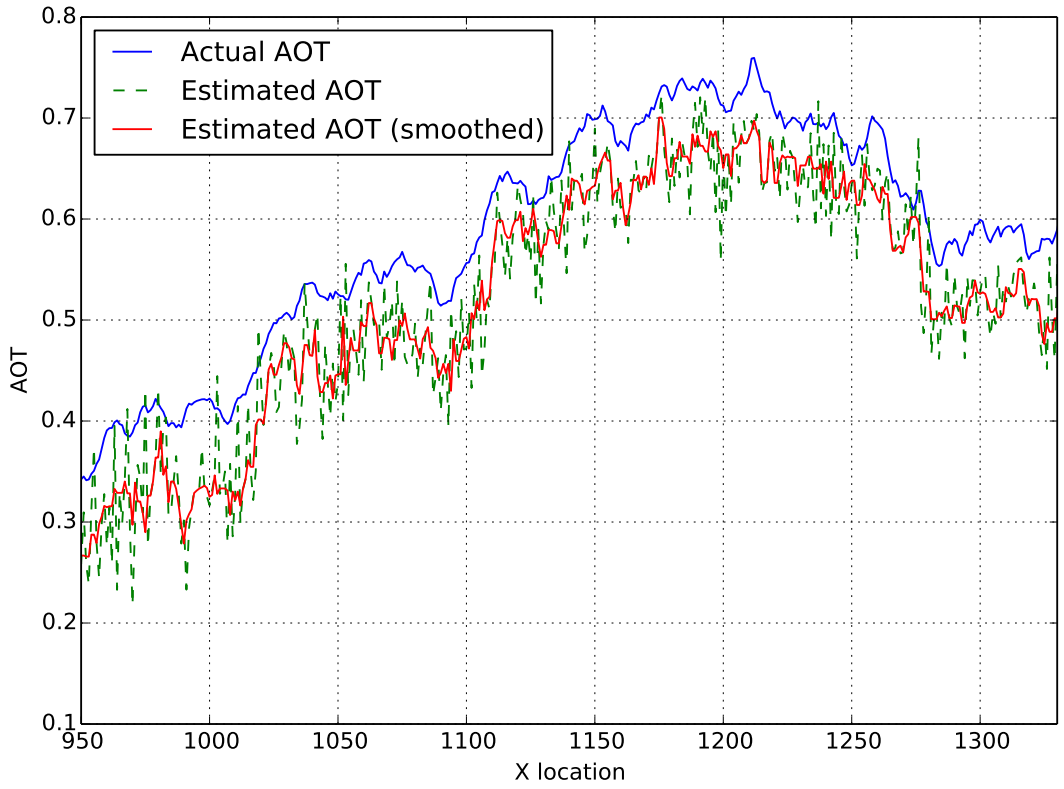
FIGURE 7.12: Actual AOT compared to Estimated AOT for the Chilbolton image, simulated with Continental aerosols, with the 1:1 line shown

difference in high error rates between the two categories of pixels) are modified too much or not enough, producing high errors.

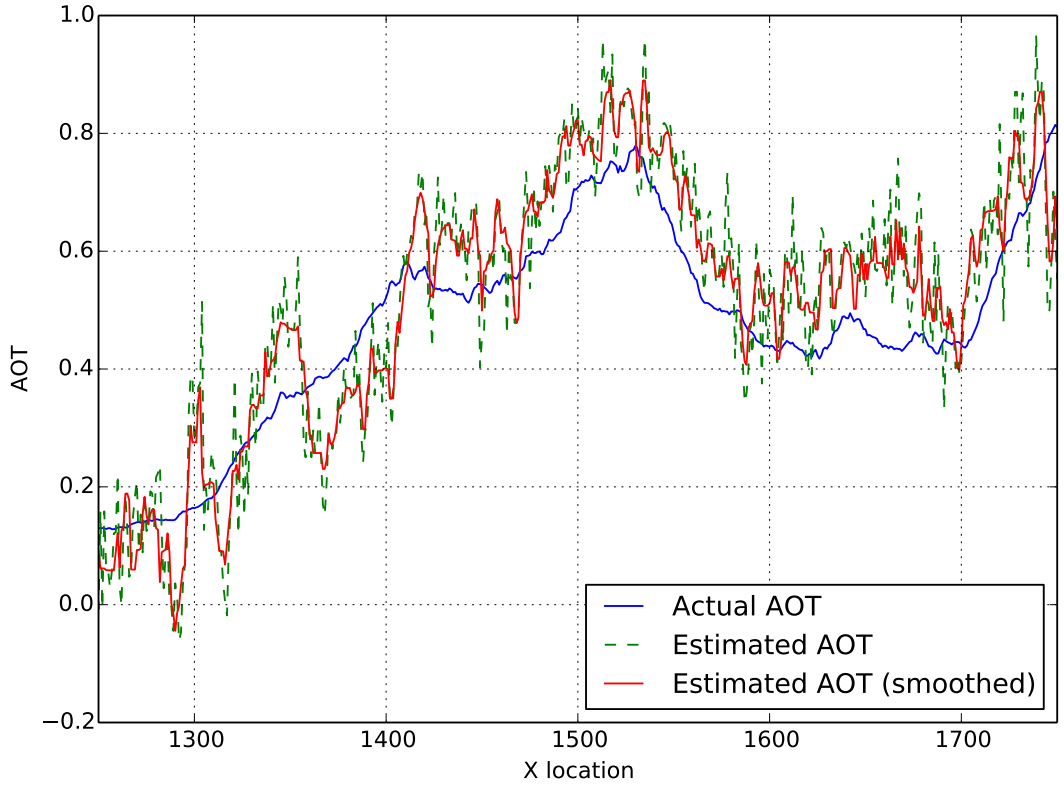
In this context it would be appropriate to assess the error over a range of land covers to determine if the algorithm performs particularly well or poorly over certain types of surface (for example, very bright surfaces, which are often troublesome for aerosol retrieval algorithms; Hsu *et al.*, 2004). However, a land cover analysis would not provide accurate results due to the issues with the LEDAPS images that were used to create the simulated images. As discussed in §4.2.3, the accuracy of the LEDAPS product varies significantly with land cover, and these errors are carried through to the simulated images, and the HOTBAR AOTs, thus, it would be impossible to tell whether differing errors over different land covers were due to issues in the HOTBAR algorithm or LEDAPS errors.

7.5 Conclusions

Examining the statistics from each of the three validation approaches (Table 7.6) shows that the AERONET validation produces the lowest accuracy, and the simulated image validation the highest accuracy. An overall range for average error (both median and mean) is approximately ± 0.10–0.15. Ideally there needs to be a specified uncertainty that can be taken into account for each pixel in the AOT image produced by the HOTBAR algorithm. It is difficult to estimate one single number, but taking into account the issues with each of the validations, an overall uncertainty of approximately ±0.12 seems appropriate. The validation of the spatial coherence of the image showed that the method captures the spatial variation in AOT relatively well: and thus the relative error between adjacent pixels will normally be significantly lower than this. Given this error, negative AOTs greater than -0.15 could be just erroneous estimations of very low (near-zero)

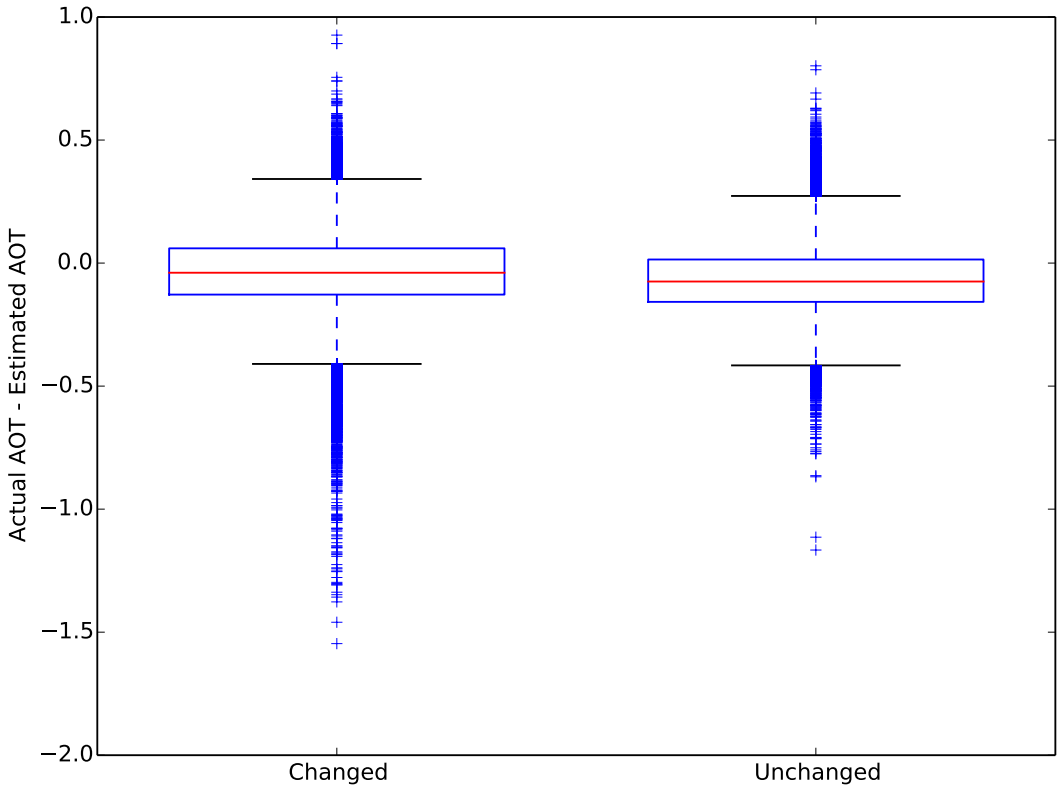


(a) Amazon

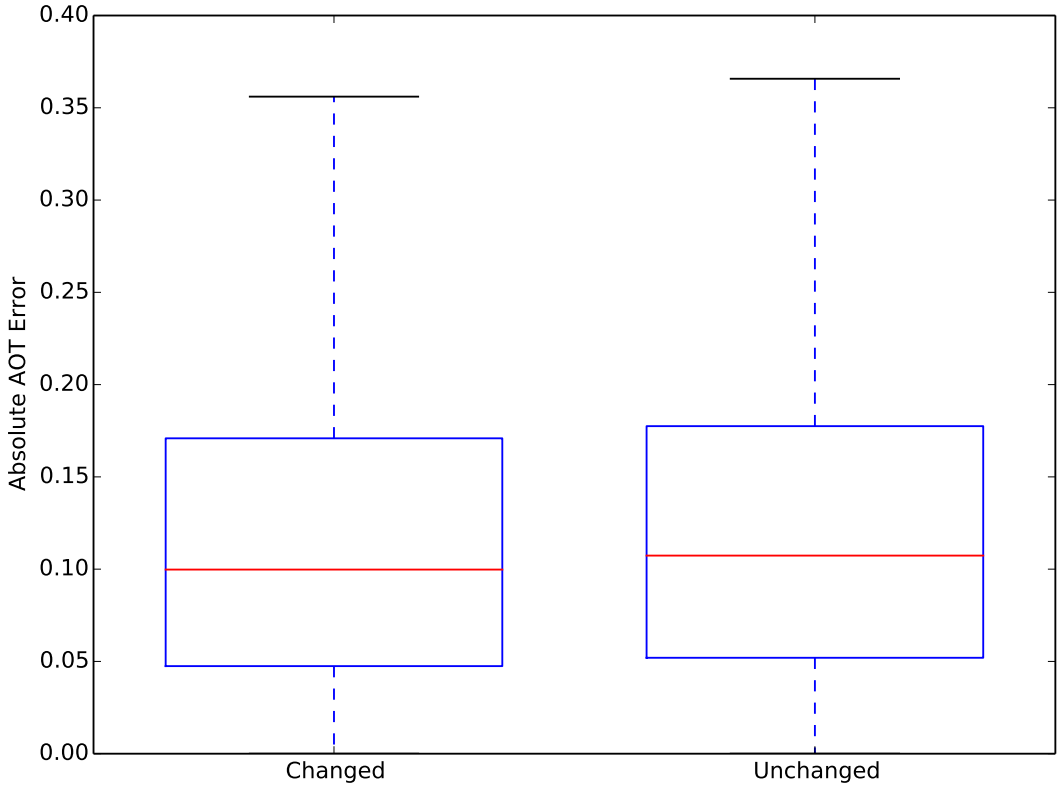


(b) SotonNewForest

FIGURE 7.13: Comparison of actual AOT, raw estimated AOT and smoothed estimated AOT for profiles across two images simulated with the Continental aerosol type



(a) Raw error, shown across full range



(b) Absolute error, with no outliers shown

FIGURE 7.14: Boxplots showing the differences in error between pixels where the HOT value was changed during the HOT correction procedure, and pixels where the HOT value was unaltered. (a) shows the whole boxplot, and (b) shows the boxplot for absolute error in more detail, without outliers

AOTs, and thus AOTs greater than -0.15 will be accepted as valid, but with those between -0.15 and -0.1 marked as potentially low quality.

TABLE 7.6: Summary statistics from each of the validations

	Mean	**Median**	**% within ± 0.1**	**RMSE**
AERONET	0.173	0.152	33	0.219
MAIAC	0.124	0.116	42	0.147
Simulated	0.123	0.104	48	0.157

It should also be noted that some pixels will have significantly higher error than others. The reasons for this are not fully understood, and so it is impossible to produce a per-pixel error estimate at present. Pixels whose values have changed as part of the HOT correction procedure do not have a significantly higher error than those which have remained unchanged, but it is likely that poor correction still has an impact on the error. The effect of the AOT magnitude on the error is difficult to discern: the error seems to be higher the lower the true AOT value is, but this is most noticeable at very low AOTs. The magnitude of the path radiance due to aerosol scattering is very low when the AOT is very low, and thus the signal to noise ratio (SNR) is also low, as the noise (in this case the ground reflectance) stays relatively constant. This low SNR makes retrieval difficult, and leads to high errors.

It is difficult to compare the accuracy of AOT retrieval methods as the statistics are often reported differently for each method. All of the methods listed in Chapter 3 Table 3.3 provide data at a significantly lower resolution, and thus their accuracy would be expected to be higher. However, the RMSE of OMI AOT according to Ahn *et al.* (2008) is 0.15, very similar to the RMSEs calculated from comparison of HOTBAR with MAIAC and simulated data (0.147 and 0.157 respectively). The MODIS 10km AOT product validation states that 67% of the retrievals are within $\pm 0.05 \pm 0.15\tau$, which corresponds to an error of ±0.08 for the global average AOT (0.19) and ±0.13 for a higher AOT (0.5). The recently-released high resolution MODIS product produces 3km data, albeit with higher error: with 67% of the retrievals within an error of ±0.1 for the global average AOT and ±0.18 for a higher AOT. In comparison, depending on the data source used for validation, between 33% and 48% of the retrievals from the new method are within ±0.1. This is lower than the percentages for the MODIS products, but impressive given that HOTBAR produces data at 10,000 times the resolution of even the high-resolution MODIS product.

As none of the high-resolution AOT retrieval methods discussed in §3.4.5 have been are operationally deployed, the validation campaigns are often limited, and so the few statistics that are available may not be representative of the global algorithm performance. Liu *et al.* (2002) found a RMSE of 0.122 from 20m AOT data retrieved from SPOT images, which compares well with the RMSE from the HOTBAR algorithm. However, HOTBAR has a significantly lower accuracy than the reported accuracy for the Lyapustin *et al.* (2004) method (±0.03), although the authors admit that this may not be a realistic

assessment of the accuracy as it was performed over a very small area with stable atmospheric conditions.

Overall, the HOTBAR algorithm produces results with a reasonable accuracy, comparable to that of the MODIS products (particularly the new 3km 'high-resolution' product). The AERONET validation produced the lowest accuracy, with the MAIAC and simulated images validations producing better results. As discussed in the introduction to this chapter, the current validation is significantly limited by the requirement to use Landsat 7 SLC-on data, which is only available between 1999 and 2003. Plans for a more in-depth validation are discussed in detail in §9.4.

Chapter 8

Example applications

This chapter shows two example applications of the HOTBAR algorithm for estimating AOT which was developed and validated in the previous chapters. These applications show the fulfilment of the objectives set out earlier in the thesis: the development of a per-pixel atmospheric correction method using the HOTBAR AOT product satisfies the challenge at the end of Chapter 2, and the application of the HOTBAR AOT product to monitor the distribution of $PM_{2.5}$ air pollution at high resolution fulfils the challenge laid down by Hoff and Christopher (2009).

8.1 Per-pixel atmospheric correction

8.1.1 Introduction

Chapter 2 investigated the effect of spatial variability in AOT values on the accuracy of atmospheric correction, and found that performing a full spatially-variable atmospheric correction, with a per-pixel AOT product, was essential to produce accurate results. However, this was impossible for high-resolution images (such as those from Landsat) as no methods existed to produce an AOT product at these resolutions. In this dissertation, the development of a method for creating a per-pixel AOT product from a Landsat image (HOTBAR) has been described, and this section applies this AOT data to atmospheric correction of the Landsat image, performing a true per-pixel atmospheric correction of a Landsat image (that is, a correction where each pixel uses an AOT value estimated from that individual pixel, rather than an interpolated value as used in LEDAPS) for the first time.

8.1.2 Methods

A per-pixel atmospheric correction method based upon Py6S was implemented in Python. The 6S model has a built-in atmospheric correction mode, which estimates surface reflectance from a TOA radiance value and an atmospheric parameterisation (including AOT, PWC, aerosol type, atmospheric model) at a given wavelength. As well as estimating the surface reflectance for the specific TOA radiance given, 6S also produces

three coefficients, x_a, x_b and x_c, which allow the surface reflectance to be estimated for other TOA radiance values using the following formula:

$$\rho = \frac{x_a L - x_b}{1 + x_c(x_a L - x_b)}$$

where L is the TOA radiance and ρ is the surface reflectance. These coefficients can be used to perform a uniform atmospheric correction (with a constant AOT value across the image) based upon a single Py6S simulation per band. However, to perform a spatially-variable atmospheric correction (with a per-pixel AOT value) these coefficients must be estimated for every AOT in the image. As discussed in §6.2.2, it would be very time-consuming to run the model for each pixel in the image, so a Look-up Table (LUT) is used, in which the coefficients are estimated for a number of AOTs, and the results interpolated to the AOT values required. For the implementation of this atmospheric correction tool, LUTs were produced for each band based on AOTs from 0.1 to 1.5 in steps of 0.1, with an additional extreme AOT added at each end (0.001 and 3.0) to allow interpolation for any extremely low or high AOT values.

As this is a relatively short example application of the method, a simple parameterisation for the other atmospheric constituents was used, with ozone amount estimated from the van Heuklon (1979) model, a fixed sea-level altitude and geometry based upon the date, time and location of the image. The aerosol model, Precipitable Water Content, and AOT were user-configurable. Code was implemented to perform corrections based upon a constant AOT (in which case a single AOT is used to correct each pixel of the image) or a spatially-variable AOT (in which case each pixel of the image is corrected using the AOT value from the corresponding pixel of a specified AOT image).

The only way to truly validate an atmospheric correction method is to compare the atmospherically-corrected surface reflectances to the true surface reflectance, as measured using a field spectrometer at the same time as the satellite overpass. As found when validating the AOT product (Chapter 7), the requirement to use data collected before May 2003 significantly limits the ground reflectance data which can be used. After extensive searching, the best dataset to use for validation was found to be ground reflectances collected around the Shunyi Institute of Agricultural Sciences, Beijing in Spring 2001 (Yan *et al.*, 2003). Many measurements were taken over twelve sampling sites, but after matching with Landsat acquisitions it was found that the only matching data covered five sites and one Landsat image, and there was a five hour offset between the Landsat overpass and the ground data collection.

Three ground measurements were taken for each site, covering the canopy, soil and leaves of the plants within the sample site. The measurements were acquired with an Analytical Spectral Devices field spectrometer (model unknown), with a spectral resolution of approximately 1.8nm, covering the wavelength range 420–2500nm. For comparison with the satellite data, the spectra were convolved with the Landsat ETM+ Spectral Response

Functions (taken from the built-in 6S data, at 2.5nm resolution) to produce reflectance estimates for each of the Landsat ETM+ bands. Limited metadata for the measurements was available, so it was unknown exactly what land covers were present in each Landsat pixel. Given this, spectra for each site were produced by averaging the canopy reflectance and the soil reflectance, based on an assumption that each pixel would consist of a mixture of vegetation and bare soil, and these spectra were used for comparison with the satellite data.

The matching image (LE71230322001091EDC00) was downloaded from the USGS and converted to radiance using the parameters specified in the image metadata. The AOT estimation algorithm was then run with $T_{high} = 0.85$ and $T_{low} = 0.70$ to produce the spatially-variable AOT used for the correction. The atmospheric correction procedure was run using a Continental aerosol model and a PWC of 0.9cm (taken from the Beijing AERONET site measurement closest to the time of satellite image acquisition).

Two other atmospheric correction approaches were used for comparison with the spatially-variable correction: a uniform atmospheric correction (with an AOT value measured at the Beijing AERONET site and interpolated to $\mathrm{AOT}_{550} = 0.85$) implemented without the use of a LUT, and a LEDAPS-corrected version of the image (acquired directly from the USGS).

Comparisons were made between the spectrometer data and the satellite data pixel directly over the location of ground data collection. The location of the ground data may not be accurate enough to locate the exact Landsat pixel in which the data were acquired, but averaging pixels around the ground data location was ruled out due to significant spatial variability in the data. This potential positional inaccuracy, combined with the time difference between image and ground data acquisition and possible ground data sampling issues means that the calculated absolute difference between the satellite and ground data is likely to have significant error. However, this is the best pre-2003 data available, and so it is used with these issues taken into account, and with a focus on relative comparisons of atmospheric correction methods (as the above issues should affect all methods equally), rather than absolute error.

8.1.3 Results & Discussion

The AOT image (Figure 8.1) shows very high AOTs (up to 0.9) in the urban area of Beijing, with significantly lower AOTs - including some negative values - in the mountainous areas to the north and west. Figure 8.2 shows the visual improvement in the atmospherically-corrected image, with the blue hazy look to the radiance image having been removed. The mountainous areas look a lot darker than expected: this is due to the combination of areas masked due to genuine cloud, commission errors in the cloud-masking algorithm (errors caused by shadows are a particular problem in this mountainous region) and areas where atmospheric correction failed due to negative AOTs. At present, the algorithm produces no result (and sets the pixel value to NaN, which is

displayed as black) where a negative AOT is used, as negative AOTs are physically impossible. This is acceptable for an example application like this, but would need further work in future.

Comparing spectra from the radiance and atmospherically-corrected images (Figure 8.3) shows that the scattering effects in the lower bands that are present in the radiance image have been removed. The vegetation spectra, particularly, exhibits the characteristic features of live vegetation (low blue, higher green, and very high near infra-red reflectances) in Figure 8.3(b), features which are not present in Figure 8.3(a).

The absolute percentage error between the ground reflectance and atmospherically-corrected reflectance for the three atmospheric correction methods (using constant AOT, spatially-variable HOTBAR 30m AOT and LEDAPS) is shown separately for each band in Figure 8.4. As discussed above, the absolute error is unlikely to be accurate, but the relative differences in error between the three methods can be compared. This relative comparison shows that the atmospheric correction using the spatially-variable 30m AOT produces the lowest overall errors (around 5–10%) for bands 1 and 2 (blue and green), however this is not the case for the other bands. LEDAPS performs best for band 3 (red), constant AOT performs best for band 4 (NIR), although the errors for this band are particularly high (40–60%). Constant AOT and the spatially-variable 30m AOT perform very similarly for band 5, but the maximum error is higher than for bands 1 and 2.

As aerosol scattering is wavelength-dependent, the largest aerosol effects will be in the blue and green bands, and these are the bands in which the 30m AOT product performs best (with errors as low as 2%). As expected, this shows that the spatially-variable 30m AOT correction produces better results than assuming a uniform AOT (as shown through simulation in Chapter 2), but it also shows that the HOTBAR AOT produces better atmospheric correction results than LEDAPS - at least for this image. This is due to the ability of HOTBAR to retrieve AOT over all surfaces, as opposed to LEDAPS which can only retrieve AOT over DDV, and thus must interpolate between - often sparse - areas of DDV.

The high absolute error in band 4 is likely due to positional uncertainties in the measurements, as the NIR reflectance of vegetation varies significantly over space due to changes in plant health. The relative differences are also likely to be affected by this, as use of a constant AOT is likely to lead to a larger correction over this area (as the AOT was derived from an AERONET station in a more central area of Beijing), and this significant reduction in NIR value brings the value closer to the erroneously-positioned ground value. Another contributor to the errors in bands 3–5 are the effects of other atmospheric constituents (such as water vapour and ozone) which have absorption features primarily within these bands. These constituents were parameterised very simply in the atmospheric correction algorithm, and also vary spatially, likely reducing the accuracy of the atmospheric correction.

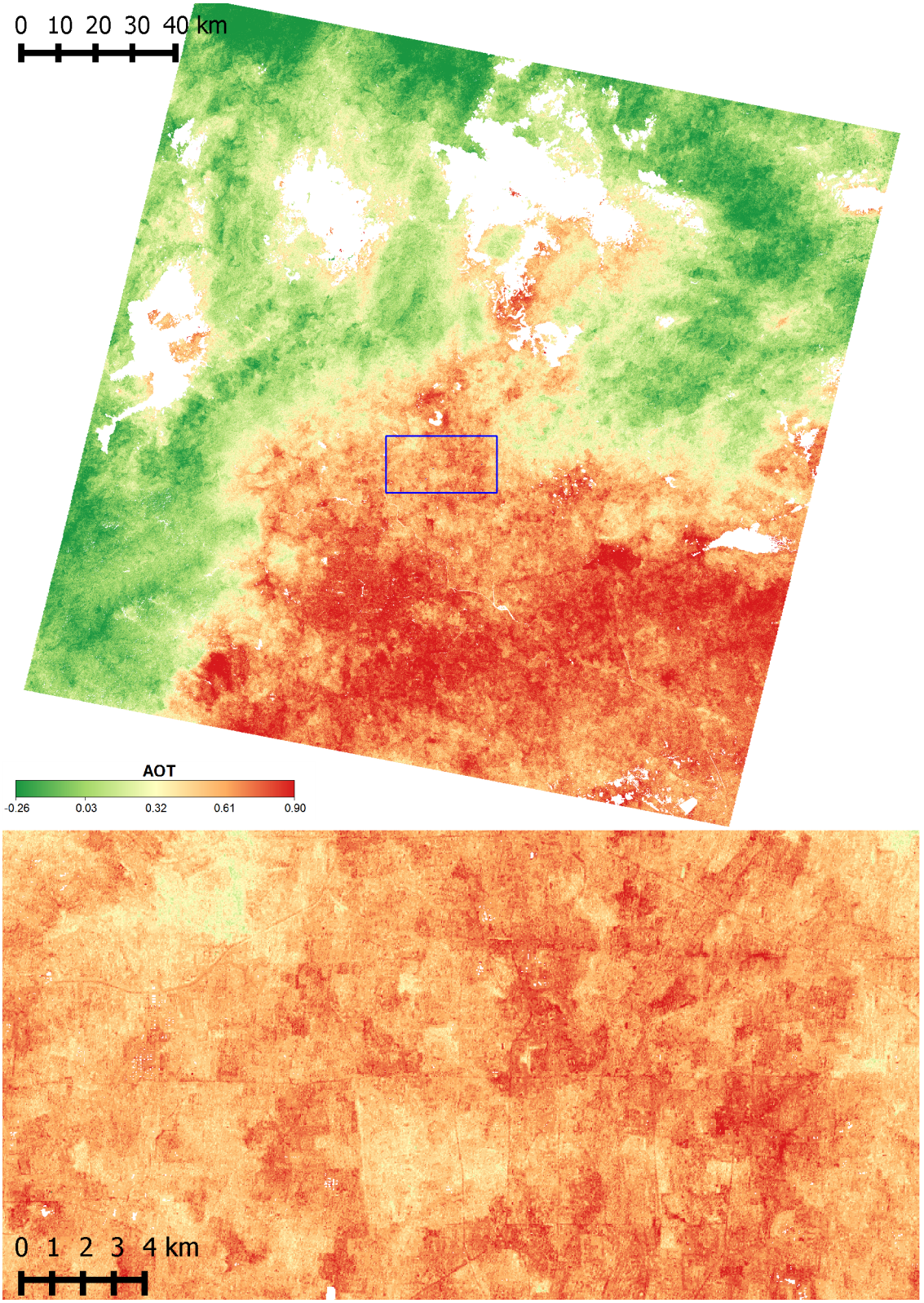


FIGURE 8.1: AOT image derived from the Beijing Landsat 7 image (LE71230322001091EDC00), and used for the spatially-variable atmospheric correction shown in Figure 8.2.

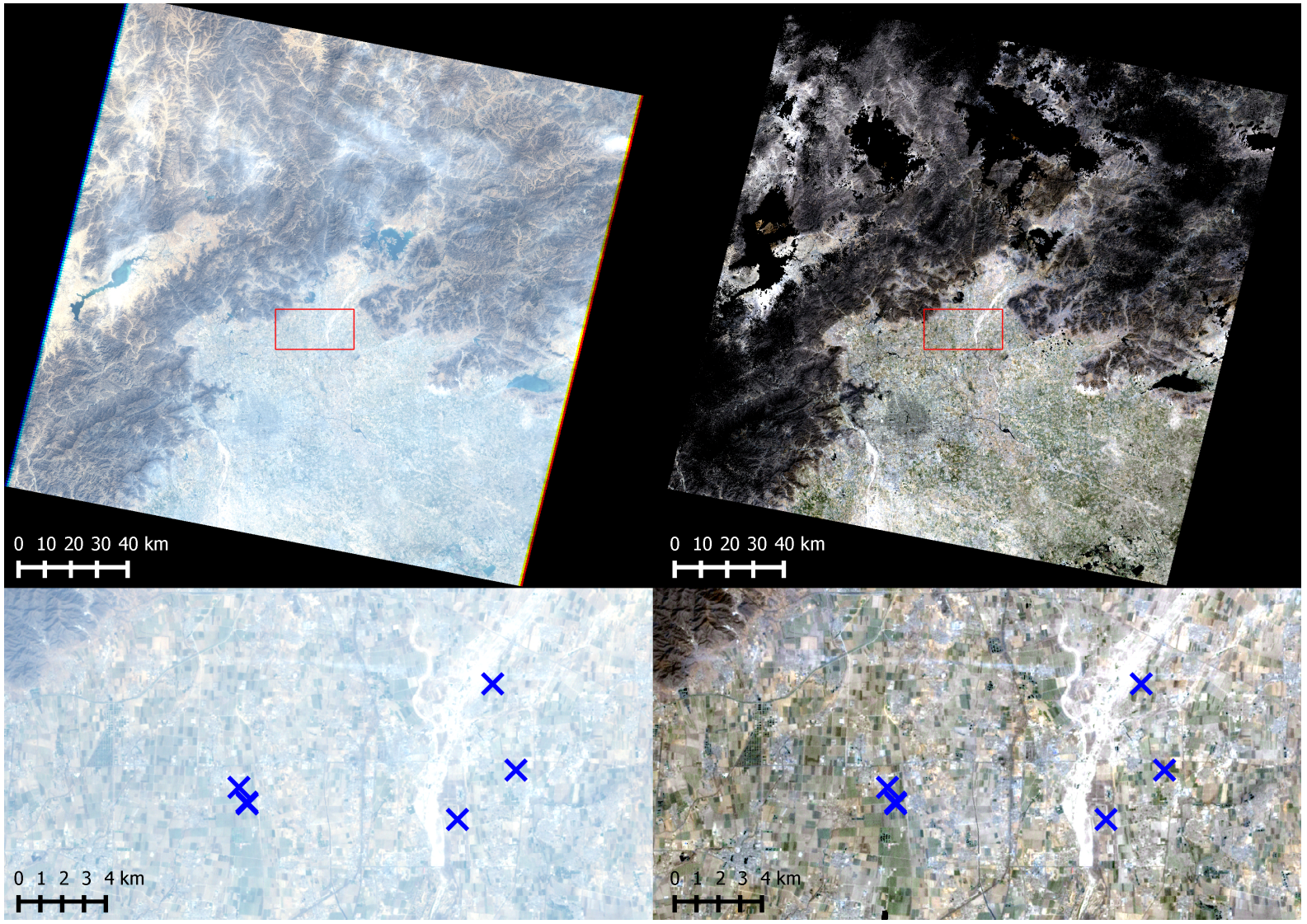
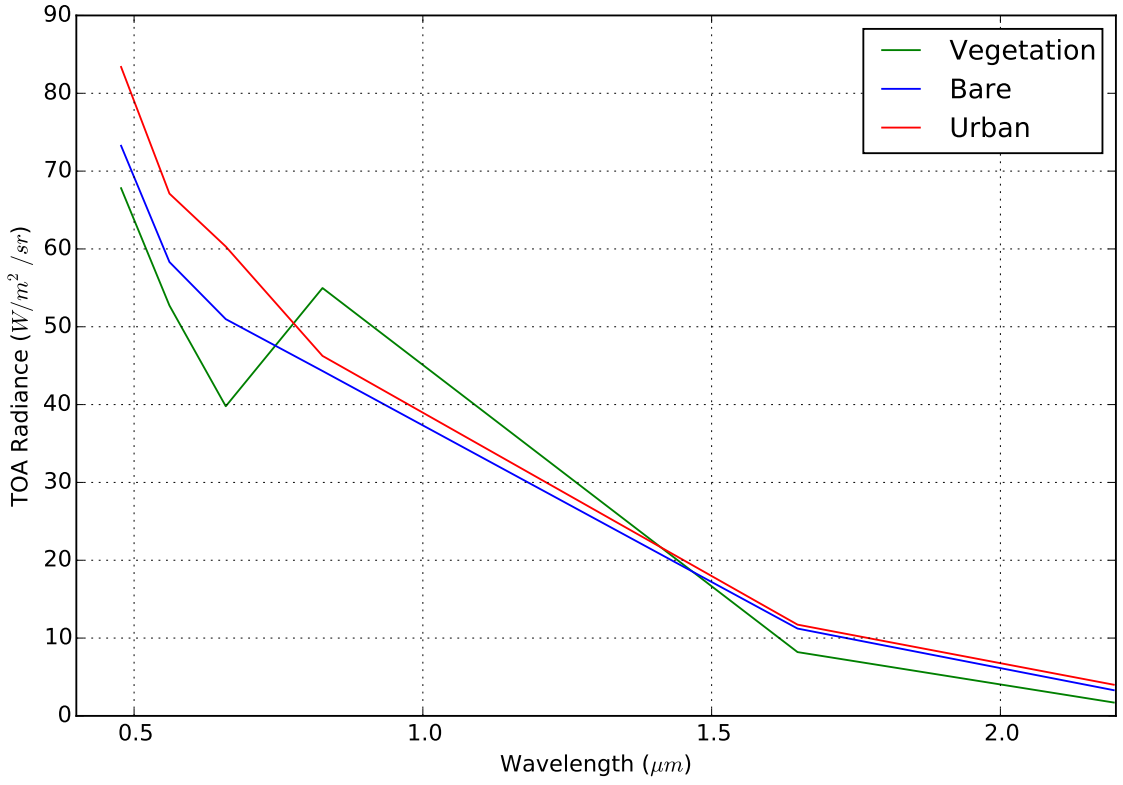
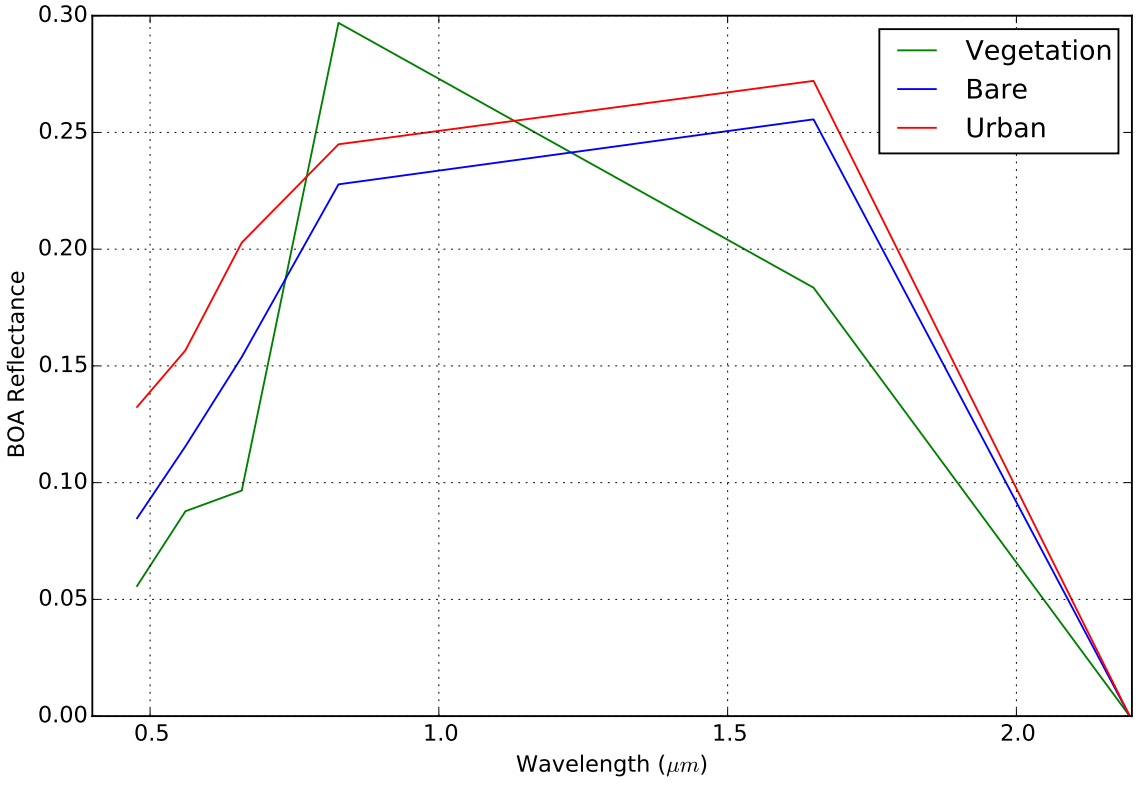


FIGURE 8.2: The Landsat 7 Beijing image shown as radiance (left) and atmospherically-corrected reflectance (right), with a zoomed view shown underneath, and validation sites marked.

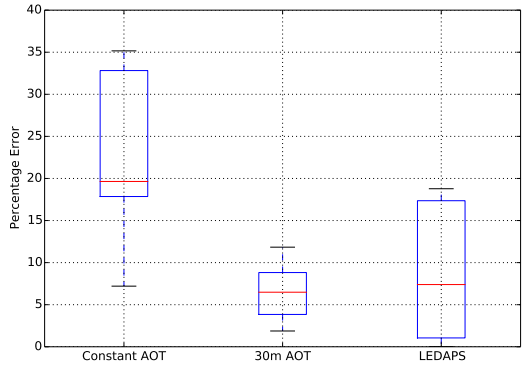


(a) Radiance

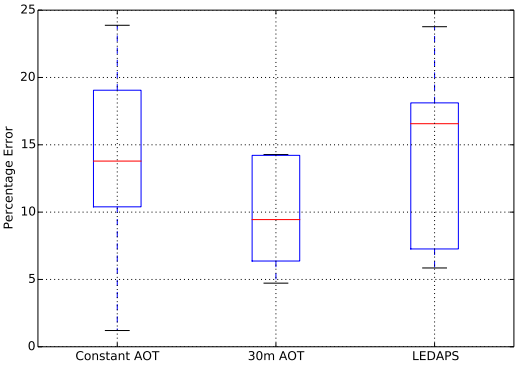


(b) Reflectance

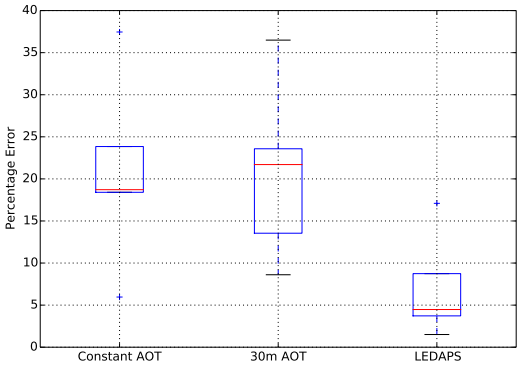
FIGURE 8.3: Comparison of spectra for three land covers (Bare, Vegetation and Urban) from (a) the original radiance image and (b) the spatially-variable atmospherically corrected image



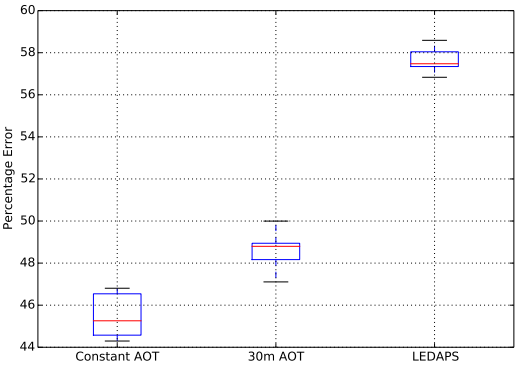
(a) Band 1



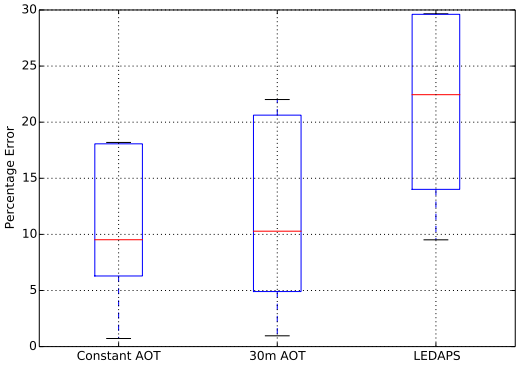
(b) Band 2



(c) Band 3



(d) Band 4



(e) Band 5

FIGURE 8.4: Boxplots showing the absolute percentage error between ground reflectance and atmospherically-corrected reflectance for bands 1–5. Error is shown for three different atmospheric correction methods: using a constant AOT across the image, the 30m HOTBAR AOT product and the LEDAPS method.

8.1.4 Conclusions & Potential Extensions

This example application shows the potential for performing accurate atmospheric correction using the per-pixel HOTBAR AOT. The validation of the atmospheric correction is limited, due to the requirement to use pre-2003 data, but suggests that atmospheric correction of the blue and green bands using the new AOT is more accurate than using either a constant AOT or the LEDAPS algorithm. It is likely that the method also performs better for other bands, but positional issues with the validation dataset hide this, making it hard to draw firm conclusions. This provides a preliminary fulfilment of the challenge given at the end of Chapter 2, and objective 5 of the PhD (as listed in §1.4) although further work is required to produce an operational method.

The following potential further work is currently under consideration:

- **Improve the atmospheric correction parameterisation:** In this simple example, the benefits of using a true per-pixel AOT product have been shown, but little attention has been given to other atmospheric constituents. Improving the parameterisation of the water vapour, ozone and aerosol model (the latter using a modelling approach such as that developed by Segura *et al.*, 2013) should reduce the errors found in the red and NIR bands. A wide range of datasets (including other satellite data and meteorological re-analysis products) can be used to perform these parameterisations, and an semi-automated method for parameterisation should be straightforward to implement.
- **Improve the usability of the atmospheric correction:** Currently the atmospheric correction algorithm is run through various Python scripts. These are not particularly complex to use, but they limit the usability of the algorithm to those who have experience writing Python code. A simple command-line or Graphical User Interface (GUI) could be created relatively easily, and this would enable use of per-pixel atmospheric correction by end-users.
- **Perform a better validation:** The validation of the atmospherically-corrected ground reflectances so far has been very limited. Once the AOT retrieval algorithm has been extended to work with more modern sensors (such as Landsat 8) validation can be performed using a far broader set of data. Data that have already been collected by various groups, such as the SPectra bARrax Campaign (SPARC) data (Moreno *et al.*, 2004) could be used, or a specialised field campaign could be run to collect exactly the data necessary for validating the atmospheric correction.

8.2 Estimating $PM_{2.5}$ air pollution

8.2.1 Introduction

As discussed in §3.7.1, particulate matter pollution (principally $PM_{2.5}$, but also PM_{10})is a significant health issue worldwide. Ground-based monitoring networks exist in many countries, but these are expensive to maintain and provide relatively poor spatial coverage. Furthermore, many countries with severe pollution problems - such as China and India - have very limited monitoring networks. Satellite data could provide an effective way to perform this monitoring with a far lower cost, but the low resolution of satellite data is a significant problem. Hoff and Christopher (2009) stated that health monitoring requires data on a scale of 1–100m, which neither surface monitoring or satellite measurements could provide. However, a 30m AOT product has been developed in this thesis, and this has the potential to be used to fulfil the requirements of Hoff and Christopher (2009) and assess particulate pollution at its natural scale.

From a detailed search of the literature, it seems that this will be the first time in which $PM_{2.5}$ pollution will have been monitored at a very high resolution, and this enables a broad range of potential analyses including detailed investigation of spatial patterns, mapping of sources and high-resolution population-weighted exposures. This section will estimate $PM_{2.5}$ from the 30m HOTBAR AOT product developed earlier in this thesis, validate the resulting estimates against ground measurements, and provide simple examples of the analyses that can be performed on the resulting data.

8.2.2 Methods

Ground-based air quality measurements are acquired in the UK by the Automatic Urban-Rural Network (AURN), which currently consists of over 109 sites. Very few sites measured $PM_{2.5}$ during the period of Landsat 7 SLC-on operation (1999–2003), but three such sites were found within one Landsat image area (path 202, row 024): two located within the London urban area (Bloomsbury and Marylebone) and one in a rural area near Oxford (Harwell). These three sites (shown in Figure 8.6) measure $PM_{2.5}$ with Tapered Element Oscillating Microbalance (TEOM) instruments (Air Quality Expert Group, 2005). These draw air through a filter on the end of a glass tube, then the tube is made to oscillate, and the resonant frequency of oscillation is proportional to the square root of the mass of particles on the filter. For comparison, traditional methods for $PM_{2.5}$ measurement involve drawing a known volume of air through a filter that only traps particles less than 2.5 µm in diameter, and then manually weighing the filter at regular intervals to determine the mass of particles per unit volume (referred to as the gravimetric method). This approach is very labour intensive, and in the past decade automated TEOM instruments have come into widespread use. These automated instruments allow near-real-time monitoring, and measurements from the AURN are provided as hourly or daily averages; the hourly averages are used here to allow the closest comparison with satellite data.

However, TEOM instruments tend to produce different measurements than the reference

gravimetric method, and so correction factors are normally applied to TEOM results (Page, 2013; Fu *et al.*, 2014; Allen *et al.*, 1997; Ayers *et al.*, 1999). European Union Standard EN14907:2005 defines requirements for $PM_{2.5}$ measurement instruments, but focuses entirely on gravimetric methods. However, the UK Government Department for Environment Food and Rural Affairs have demonstrated the equivalence of data from a range of automated measurement devices, including the TEOM instruments used in the AURN, to this standard (Harrison *et al.*, 2006; Harrison, 2010).

Four cloud-free Landsat ETM+ images were available for this location during the period of SLC-on operation, and these were acquired and processed through the HOTBAR algorithm using the parameters $T_{high} = 0.8$ and $T_{low} = 0.2$. The resulting AOT estimates were then converted to estimates of $PM_{2.5}$ concentrations by multiplying the AOT values by the van Donkelaar *et al.* (2010) conversion factors (described further in §3.7.1). Rather than using a single conversion factor for an image, these conversion factors are provided as a series of twelve global images (one for each month) at 0.1° resolution, and exhibit significant spatial variation over the UK (Figure 8.5). The conversion factor images were resampled to 30m resolution and multiplied with the HOTBAR AOT data to produce the $PM_{2.5}$ estimates.

As in previous sections, validation was performed between ground measurements and directly coincident pixels, with the ground measurement closest to the time of Landsat overpass used. Inaccuracies in the georeferencing of either the ground-based pixels or the Landsat image itself could cause issues with this approach, but using pixel averages over a window centred on the measurement site was not appropriate due to the significant spatial variability in $PM_{2.5}$, particularly in urban areas (the so-called 'urban-canyon' effect; Vardoulakis *et al.*, 2003).

8.2.3 Results & Discussion

Example maps showing $PM_{2.5}$ distributions over the whole image (Figure 8.6) and London (Figure 8.7) show the level of detail which this new method provides. For comparison, the 10km pixel grid used for the MOD04 AOT product is overlain in blue, showing that there would only be one MOD04 measurement over the whole of central London. The spatial pattern of $PM_{2.5}$ across the whole image shows the expected pattern of higher values over built-up areas, and lower values in rural areas. The so-called 'M4 corridor' is visible as a line of moderately-high pollution stretching East-West across the lower part of the image - and within this the town of Reading has particularly concentrated areas of high pollution. Interestingly, the western and north-western parts of the image - around Oxford and surrounding settlements such as Bicester - appear to have the highest levels of pollution. This is likely due to transport of particulate pollution rather than local sources - in this case the prevailing south-westerly wind bringing pollution from the industrial areas surrounding Bristol towards Oxford.

The spatial distribution of $PM_{2.5}$ across London (Figure 8.7) shows a significant difference

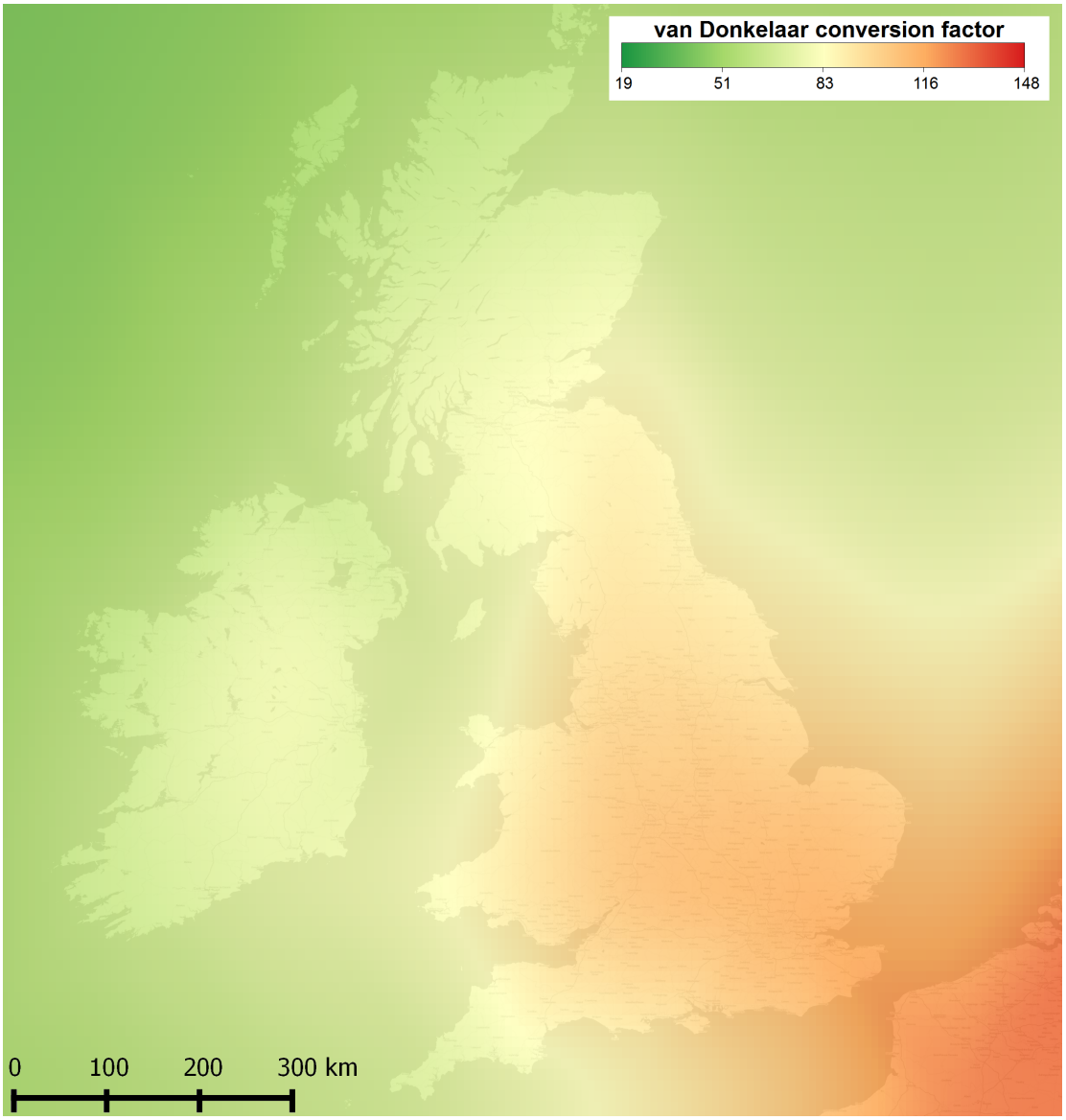


FIGURE 8.5: Map showing the spatial variation in the van Donkelaar *et al.* (2010) conversion factor used to estimate $PM_{2.5}$ from AOT for dates in March

in pollution north and south of the River Thames, with significantly higher pollution in the north. Particularly high concentrations are found in Kensington, Euston Rd, Westminster and the industrial area near Bromley-by-Bow. All except the last of these are likely to be caused by traffic: for example, the A4 passing through Kensington and Euston Road are two of the major routes into London. Green space in Central London (such as Hyde Park and Regent's Park) still have significant amounts of $PM_{2.5}$ pollution, but generally less than the surrounding areas - this is particularly noticeable for the western part of Hyde Park.

Time series of ground measurements and coincident (with a time difference of less than thirty minutes) satellite estimates of $PM_{2.5}$ are shown in Figure 8.8, and absolute error statistics in Table 8.1. The rural site (Harwell) has the lowest average error, potentially reflecting the effect of positional accuracy in these error calculations: the concentrations

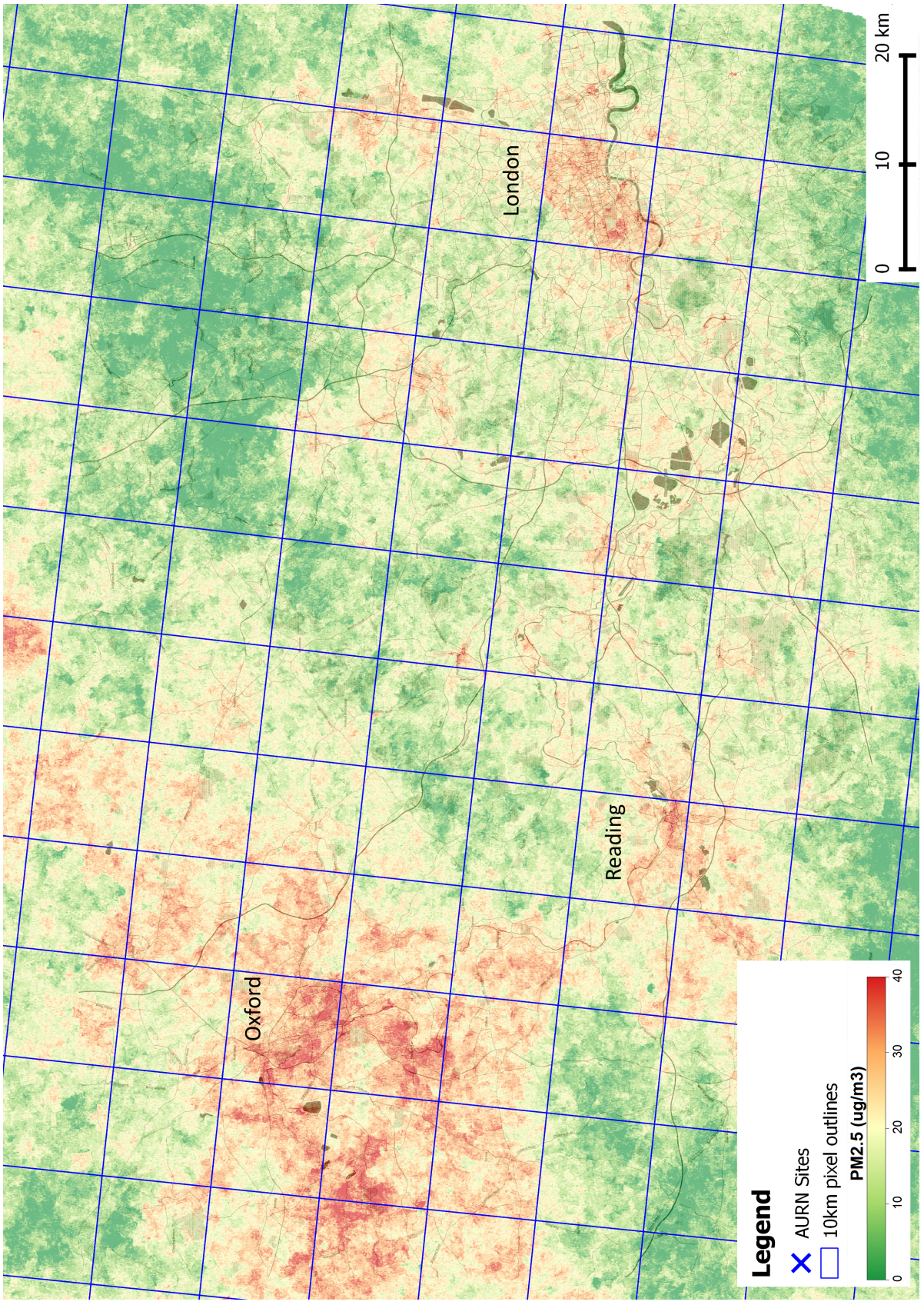


FIGURE 8.6: $PM_{2.5}$ estimated using the van Donkelaar *et al.* (2010) conversion factors on AOT derived from a Landsat image of southern England (LE72020242002087EDC00), acquired on 28th March 2002. The AOT is shown overlain on a black and white background map (the Open Street Map Stamen Toner map), with the Automatic Urban-Rural Network sites used for validation, MODIS 10km pixels also shown, and place labels also shown.

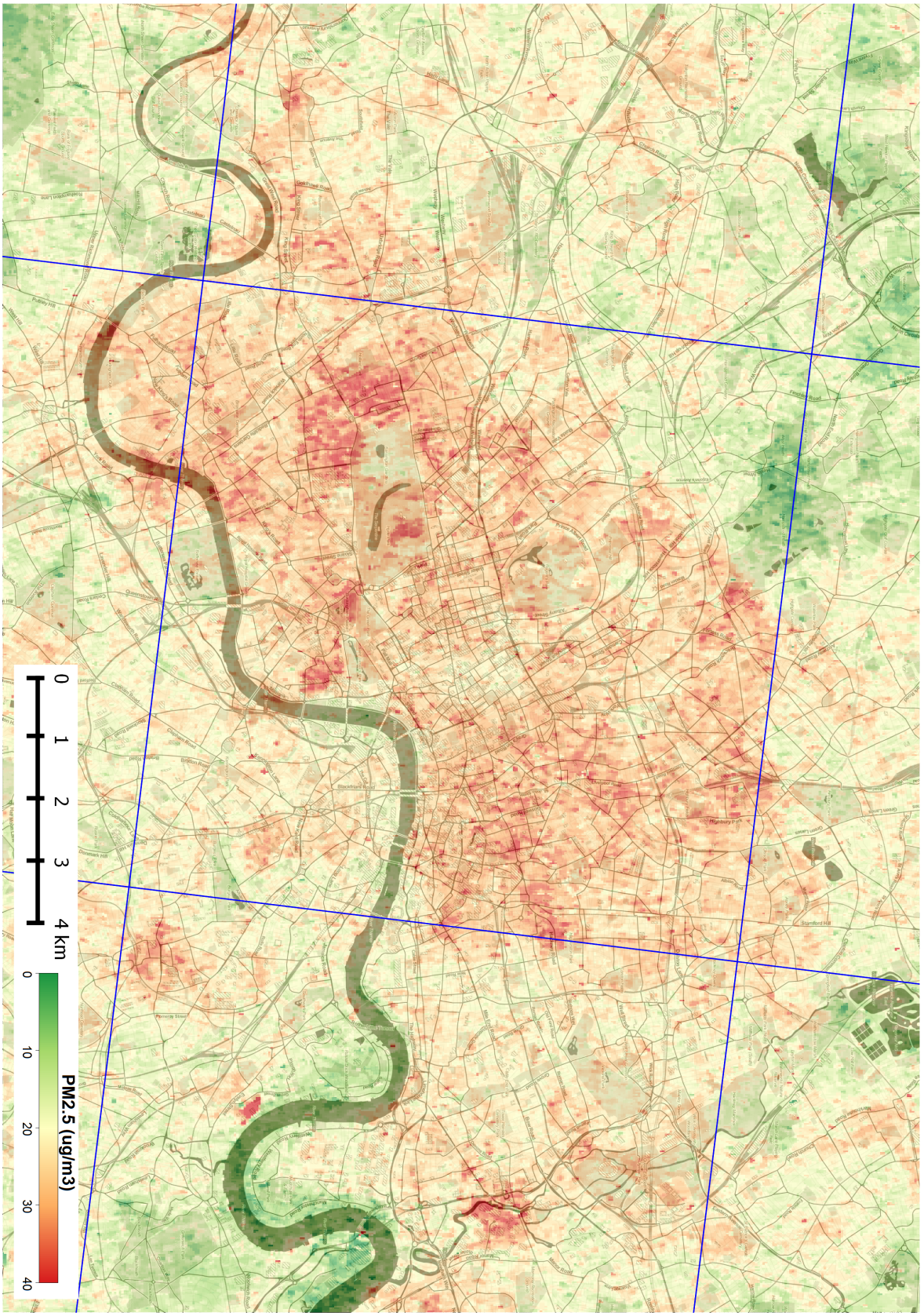


FIGURE 8.7: A detailed view of $PM_{2.5}$ over London, from the same image as Figure 8.6.

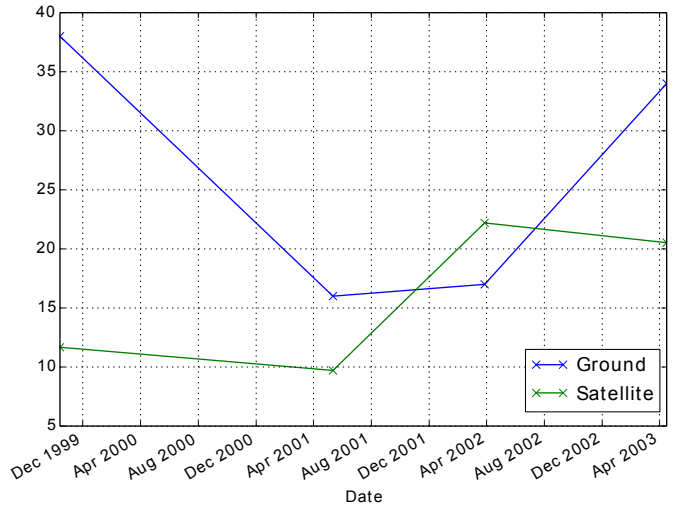
TABLE 8.1: Absolute error statistics, in $\mu g\,m^{-3}$, for each of the three AURN sites used for validation, plus the overall error

	Bloomsbury	**Harwell**	**Marylebone**	**All**
Min	5	2	4	2
Mean	13	11	14	12
Median	10	5	13	6
Max	26	31	26	31

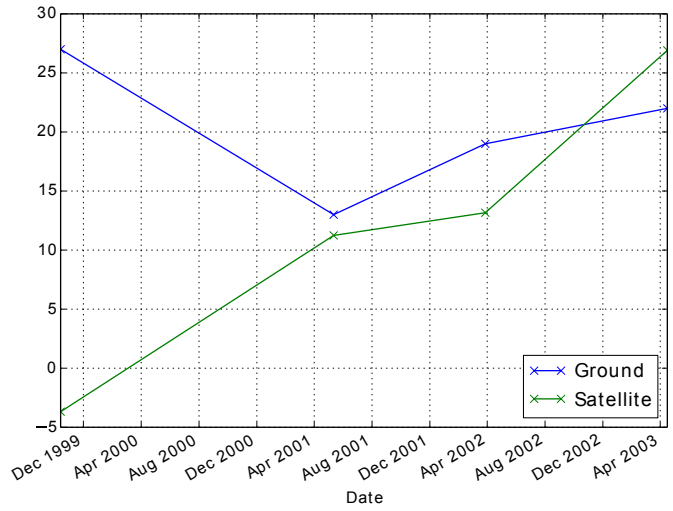
around Harwell are very uniform in comparison to those around the London sites, and the land cover is far less variable. The overall mean and median errors of $12\mu g\,m^{-3}$ and $6\mu g\,m^{-3}$ respectively compare favourably with the results from a global analysis performed by van Donkelaar *et al.* (2010) at a far lower resolution (10km), which produced a global population-weighted mean uncertainty of $6.7\mu g\,m^{-3}$. Examining the time-series plots shows a significant variation in error over time, with particularly high error in all images for the image collected in December 1999 (ranging from 20–$30\mu g\,m^{-3}$). The reason for this high error is not fully understood at this time, but as it was acquired very early in the satellite's period of operation, it could be due to instrument calibration issues. Removing the data from the December 1999 image from the analysis reduces the mean absolute error to $8\mu g\,m^{-3}$.

The overall estimate of HOTBAR AOT error found in Chapter 7 was ± 0.12, and this is consistent with these $PM_{2.5}$ errors, as over the validations sites the van Donkelaar *et al.* (2010) conversion factors have an approximate range of 65–90, and thus a $PM_{2.5}$ error of between $7.8\mu g\,m^{-3}$ and $10.8\mu g\,m^{-3}$ would be expected. In this context the errors inherent in the van Donkelaar *et al.* (2010) AOT-$PM_{2.5}$ conversion are ignored, and these would be expected to increase the error, so the mean $PM_{2.5}$ error found in this validation is very good. Little information is available on the accuracy of ground-based measurements of $PM_{2.5}$ as, unlike with trace gas measuring instruments, it is very difficult to perform reference measurements in a laboratory (Air Quality Expert Group, 2005). However, the manufacturer of one of the instruments used in the AURN states an estimated accuracy of $\pm 2.5\mu g\,m^{-3}$ (Thermo Scientific, 2009), around a third of the average error found from the satellite estimates.

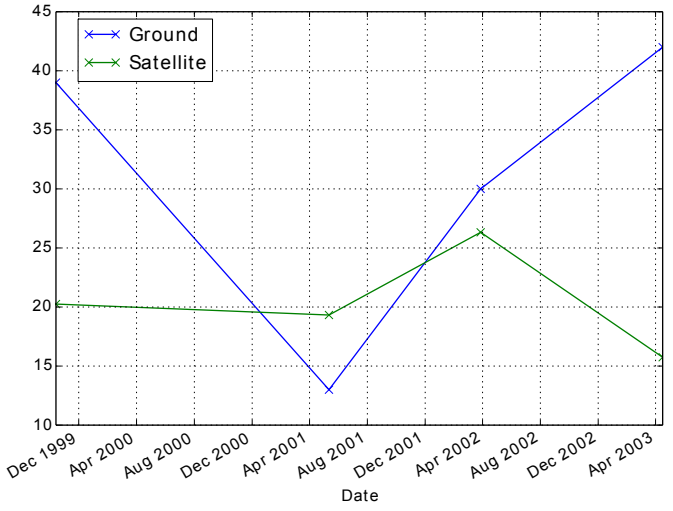
It is difficult to compare the $PM_{2.5}$ values from these images to the National Air Quality objectives (Defra, 2010), as the objectives are stated in terms of the annual mean. However, analysis has been carried out based upon an average of the images used in this study (excluding the December 1999 image with high error) to allow a indicative comparison to be carried out, albeit with high uncertainty. Figure 8.9 highlights the pixels in London in which the mean estimated $PM_{2.5}$ concentration exceeded the National Air Quality objective of $25\mu g\,m^{-3}$, and Figure 8.10 shows the percentage of the area of each London borough which exceeded the objective. As expected, the boroughs in the middle of central London (Westminster, City of London and Kensington and Chelsea) perform worst, with each exceeding the objective over more than 20% of their area.



(a) Bloomsbury



(b) Harwell



(c) Marylebone

FIGURE 8.8: Time series comparisons of $PM_{2.5}$ measured by the Automatic Urban-Rural Network monitoring stations and estimated from AOT

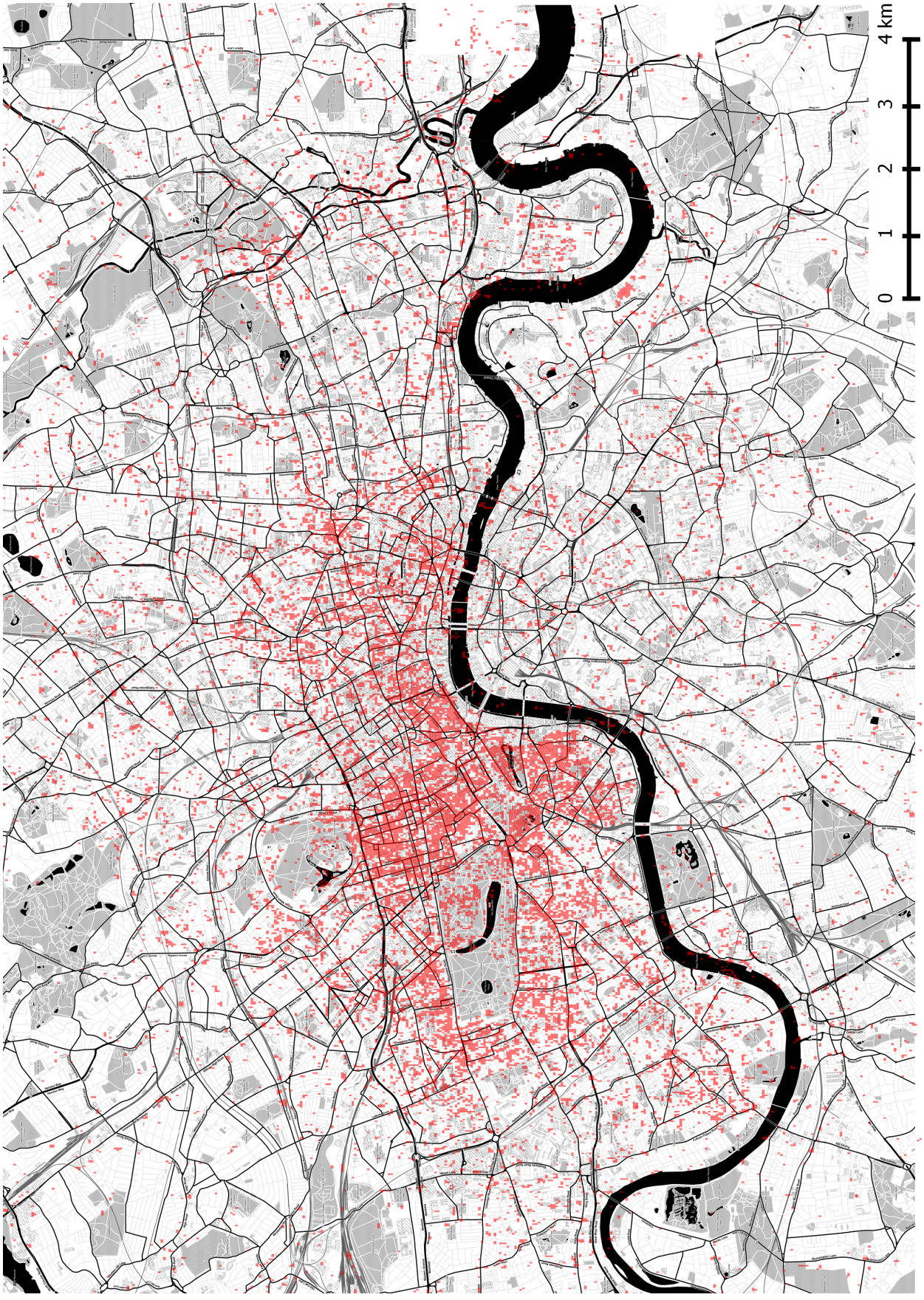


FIGURE 8.9: Map of London showing areas where the average $PM_{2.5}$ value over the three images used in this study exceeded the UK National Air Quality objective of $25 \mu g\,m^{-3}$

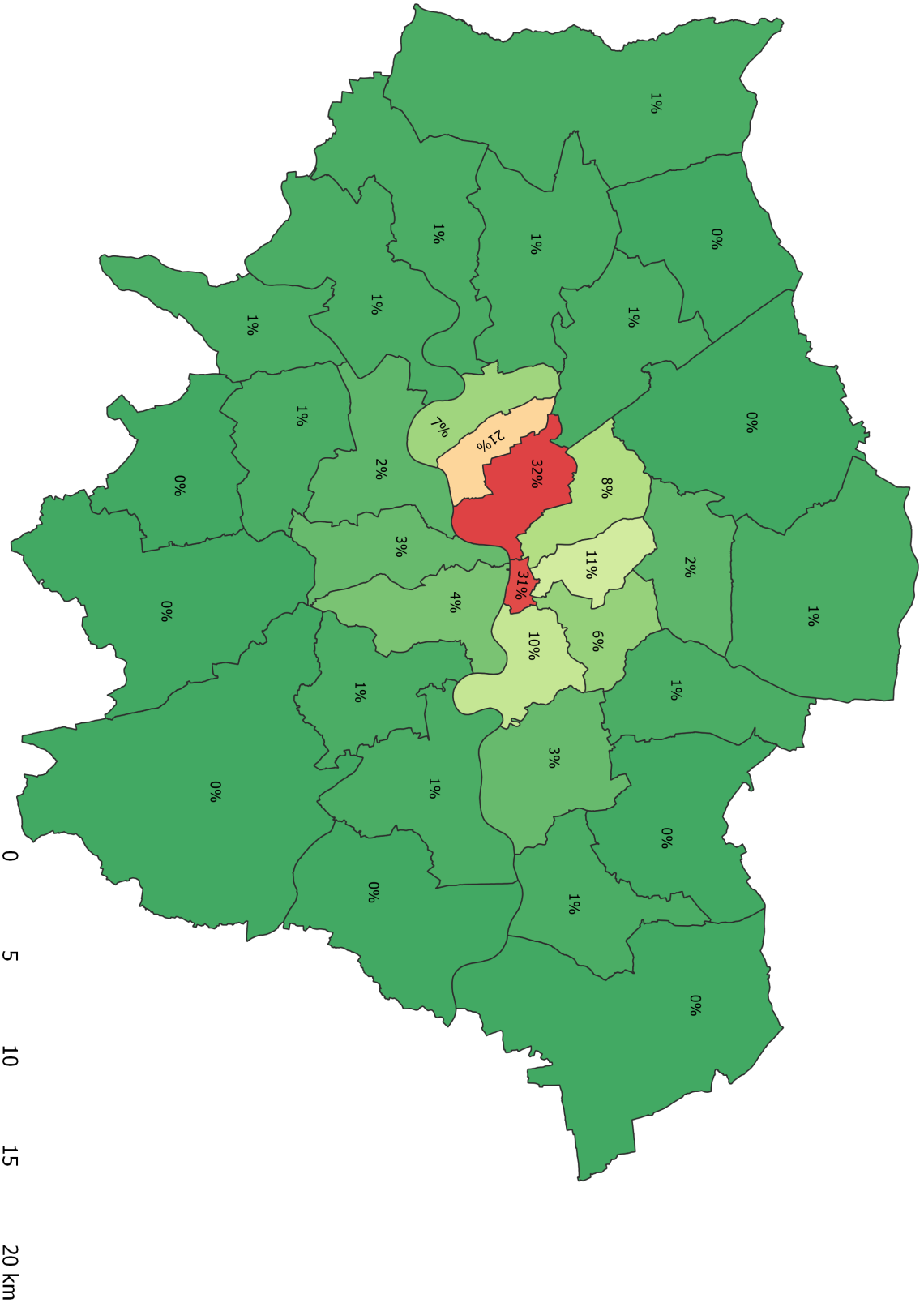


FIGURE 8.10: Map of London showing the percentage of the area of each borough in which the average $PM_{2.5}$ value over the three images used in this study exceeded the UK National Air Quality objective of 25µg m^{-3}

8.2.4 Conclusions & Potential Extensions

This example application shows the potential for monitoring particulate matter pollution (principally $PM_{2.5}$, but potentially also PM_{10}) using the new high-resolution AOT product (HOTBAR) developed in this thesis. The estimates have an uncertainty of around $\pm 8\ \mu g\, m^{-3}$, which compares favourably with lower-resolution methods of estimating $PM_{2.5}$, and is around three times the estimated error of ground instruments.

To the author's knowledge, this method provides the highest resolution data on particulate matter ever produced (fulfilling the requirement of Hoff and Christopher, 2009), and this high resolution allows the effective use of satellite data for local-scale particulate matter pollution monitoring for the first time. From a regulatory perspective, Local Authorities within the UK have a statutory duty to monitor air pollution, but have been discouraged from using satellite data due to the low resolution: previously available data provided no more detail than the limited network of ground measurement stations already in operation. However, the availability of high-resolution data changes this, and preliminary discussions with Local Authorities about the application of this method in their monitoring work have led to significant interest.

The following potential further work is currently under consideration:

- **Integrate ground and satellite measurements:** Integrating ground-based measurements (with high accuracy) and satellite-based measurements (with high spatial resolution) through a data assimilation approach would allow the production of detailed measurements with high accuracy.
- **Improve temporal resolution:** Landsat images are acquired over the same location every 16 days, but cloud cover means that the practical temporal resolution is far lower. Extending the method to work with a broader range of sensors (see §9.4.1) would allow a far higher temporal resolution either by use of sensors with shorter revisit periods (such as SPOT-5 or the soon-to-be-launched Sentinel-2) or by combining data from a range of sensors into one time-series.
- **Investigate a wider range of applications:** There are many previously impossible applications which are made possible by the use of high-resolution $PM_{2.5}$ data. These should be investigated in collaboration with air quality experts and spatial epidemiologists, and may include detailed population-weighted exposures, links between air quality and health, and validation of high-resolution air quality models.
- **Extend to estimate PM_{10} :** There is also a significant relationship between AOT and PM_{10} (Hoff and Christopher, 2009; Péré *et al.*, 2009), and similar methods to van Donkelaar *et al.* (2010) could be developed and used to allow estimation of both PM_{10} and $PM_{2.5}$ at high resolution.
- **Perform a better validation:** The present validation has been limited by the need to use data collected before 2003 - a period in which relatively few $PM_{2.5}$ measurement stations were operating. Once the AOT retrieval algorithm has been

extended to work with more modern sensors (such as Landsat 8, see §9.4.1), a validation can be performed with a far broader set of data.

Chapter 9

Conclusions & Further Work

9.1 Summary & Novel Contributions

This dissertation has described the development of a novel method to retrieve high spatial resolution AOT from satellite data. The overall aim of developing a method to produce a per-pixel AOT product from Landsat images has been achieved through the development of the HOTBAR algorithm, which is summarised in §6.5. The PhD objectives (originally listed in §1.4) have been achieved:

1. **To assess the validity of the assumptions behind the Haze Optimized Transform:** Achieved in Chapter 4, in which the assumptions were found to be valid.
2. **To develop an improved version of the Haze Optimized Transform, focusing particularly on the correction procedure:** Achieved in Chapter 5, in which the HOT was improved by pre-processing the data, developing a new method to model the Clear Line and implementing a significantly improved correction procedure based on Object-based Image Analysis.
3. **To develop a method to estimate AOT from the Haze Optimized Transform:** Achieved in Chapter 6, where the inherent relationship between HOT and AOT was assessed, and then a method to model this relationship for an arbitrary image was developed.
4. **To assess the accuracy of the resulting AOT values both in terms of absolute accuracy and 'spatial coherence' (the relative accuracy between adjacent pixels):** Achieved in Chapter 7, where the algorithm results were compared to AERONET ground measurements, the MAIAC 1km satellite product and simulated images.
5. **To apply the new method to some potential application areas where high-resolution AOT data are needed:** Achieved in Chapter 8 through the application of the new data to $PM_{2.5}$ air quality monitoring and per-pixel atmospheric correction.

The significant novel contribution of this research is the development of a new algorithm to estimate high-resolution AOT from a single optical satellite image (in this case, a Landsat image, producing AOT data at 30m resolution). The method requires only two visible bands (blue and red), allowing it to be easily extended to work with a wide range of other sensors, and is known as the Haze Optimized Transform Based Aerosol Retrieval (HOTBAR) algorithm.

As part of the development of this algorithm a number of separate novel contributions have been made:

- An assessment of the error involved in performing a uniform atmospheric correction across a satellite image acquired over southern England. Even when images were acquired on clear days, the variation in atmospheric conditions was found to be significant, and the resulting error high. Full details are in Chapter 2.
- The development of a Python interface to the 6S atmospheric radiative transfer model, allowing multiple simulations to be performed easily and computationally-efficiently. This has been released as open-source software, and is now in use by a broad range of research groups at universities across the world. Full details are in Appendix A.
- The development of a method to create an 'as close to perfect as possible' dataset of atmospherically-corrected satellite image pixel reflectances. This dataset is referred to in this thesis as LandsatAERONET but can be easily created from other data, and will be useful for a broad range of remote sensing research. Full details are in §4.2.4.
- The development of a method to define the Clear Line without requiring a clear part of the image, using a land-cover-based Monte Carlo approach to orthogonal distance regression. The original Clear Line estimation technique is impossible when using the HOT to estimate AOT, but the same approach may be useful within the standard HOT procedure as it removes the need to manually find a clear part of the image from which the Clear Line can be derived. Full details are in §5.3.
- The development of a significantly improved HOT correction procedure to remove land cover effects from HOT images. This method uses Object-based Image Analysis to take into account local context when removing anomalies, and could be applied separately to the rest of the algorithm in place of the correction procedure used in the Advanced HOT (He *et al.*, 2010). Full details are in §5.4.
- The application of the new high-resolution AOT data to perform a true per-pixel atmospheric correction of a Landsat image for the first time. Full details are §8.1.
- The application of the new high-resolution AOT data to produce high-resolution maps of $PM_{2.5}$ concentrations for the first time. Full details are in §8.2.

The literature review (Chapter 3) categorised AOT estimation methods by the manner in which they split the at-sensor radiance in to the ground radiance and path radiance. HOTBAR doesn't explicitly estimate the path radiance, so the method doesn't easily fit into one of the categories. However, the HOT is a proxy for the path radiance, and thus

the new algorithm could be categorised as using a combination of *Estimate the ground reflectance from another band* (as the HOT uses deviation from the correlation between blue and red reflectance to estimate a proxy for path radiance; see §3.4.2.2), and a new category *Use spatial context* (as the removal of land cover effects is performed using spatial context in the object-based image analysis HOT correction procedure).

9.2 Strengths & weaknesses

HOTBAR uses a novel approach to retrieving AOT from satellite images, and thus the strengths and weaknesses of the overall approach should be examined. It should be noted that these are not the strengths and weaknesses of this particular implementation, many of which are dealt with in the list of potential future work in §9.4.

The strengths of this approach are that:

- It produces high-resolution AOT data with reasonable accuracy
- It has very lenient requirements for input data, and requires only a blue and red band, making it applicable to a range of sensors
- It only requires one image of the area to be processed (as opposed to a time-series of images, or a reference image with very low AOT)
- It uses context to take into account the perturbing effects of land cover, thus reducing the signal-to-noise ratio issues experienced by other high-resolution AOT retrieval methods

The weaknesses of this approach are that:

- It performs very poorly over some areas where land covers with significantly different reflectances merge without sharp boundaries (such as deserts), as the segmentation procedure cannot extract objects to be corrected from these areas
- It requires a reasonable historical archive of images (lasting around six months to a year, depending on the sensor temporal resolution) to produce a LandsatAERONET-equivalent dataset which is representative of a range of land covers
- It performs poorly over areas where the LandsatAERONET database is not representative (for example, due to limited sampling from desert areas)

9.3 Broader implications

The development of HOTBAR has implications for two groups: those involved in the development of AOT retrieval methods, and those who are end-users of AOT retrievals.

The broader implications for the development of AOT retrieval methods generally are based upon the effectiveness of certain parts of the methodology which will be relevant in the development of future algorithms. Specifically:

1. The OBIA-based HOT correction (§5.4), which applies spatial context in the

process of removing land cover effects for the first time, providing an effective way to deal with the signal-to-noise ratio issues caused by land cover anomalies.
2. The development and use of the LandsatAERONET dataset (§4.2.4) as a representative sample of 'perfectly' atmospherically-corrected Landsat surface reflectance, which has broad applications across a range of remote sensing research.
3. The success of using the HOT to estimate AOT shows the benefits of 're-imagining' work performed in the early days of remote sensing (the HOT is closely linked to work from the 1970s and 1980s) in a modern context.

There are also broader implications for end-users. The results of the self-contained study in Chapter 2 have important practical implications for remote sensing researchers: the conclusions show the importance of performing a spatially-variable atmospheric correction, as opposed to a uniform correction. Similarly, the development of HOTBAR has significant implications for end-users of AOT data: the method could be operationally applied to produce high-resolution AOT data for a range of applications including the assessment of $PM_{2.5}$ air pollution, atmospheric correction of satellite imagery, monitoring of smoke from fires, and monitoring of dust transport. The current legislative framework in the UK and EU requires $PM_{2.5}$ measurements used for monitoring adherence to air quality regulations to be acquired from ground monitoring sites (Defra, 2010), thus limiting the official application of satellite-based measurements. However, it is likely that, once high-resolution satellite measurements have been applied more broadly in health research projects, the legislation will be changed to allow a combination of satellite and ground measurements to be used for official monitoring.

9.4 Further Work

The work described in this thesis is the early stage of the development of HOTBAR: further work could be carried out to both improve both the algorithm itself, and the validation and application work described in the later chapters of this thesis. Most other aerosol retrieval algorithms took many years to develop fully: for example, MAIAC took over five years (Lyapustin *et al.*, 2011) and work started on the MOD04 product in 1992, seven years before the MODIS sensor was first launched (King *et al.*, 1992), and so the need for further work is to be expected.

9.4.1 Extending to other sensors

One of the major issues with the algorithm implemented in this research is that it requires Landsat ETM+ SLC-on data, which was only acquired between May 1999 and May 2003. As well as removing the ability of the algorithm to measure AOT from current imagery, this also significantly reduces the volume of data available for validation. The algorithm was originally developed using Landsat TM and ETM+ images, but it was found that high noise levels in the Landsat TM data caused the HOT correction to perform poorly (see §5.4.7), and SLC-off data caused major problems with the segmentation algorithm, leaving only Landsat ETM+ SLC-on data for use by the algorithm. Thus the highest priority

further work is to extend the method to work with Landsat 8 data, as detailed in §5.4.7.

9.4.2 Improved validation

The validation so far has been limited by the requirement to use Landsat ETM+ SLC-on data, which significantly reduced the potential data (both from AERONET and other satellite products) which could be used for comparison. After the algorithm has been extended to work with Landsat 8 data it will be possible to perform a far more comprehensive validation, consisting of:

- **Broad AERONET validation:** A key aspect of any validation is to ensure that it is as representative as possible of the real-world conditions in which the method will be used. With the far broader availability of AERONET data during the period of Landsat 8 data acquisition, it should be possible to perform validation over most - if not all - of the AERONET sites worldwide.
- **Validation against dense ground measurement networks:** As discussed in the validation (Chapter 7) the 'spatial coherence' of the resulting AOT data is important, and this is very difficult to assess as ground networks are not dense enough and other satellite data are of low resolution. There are two possibilities to extend this part of the validation: the use of DRAGON data (§3.3.1.2) and the development of a new high-resolution ground-based dataset through the SkySci project (§3.3.1.3)).

9.4.3 Algorithm improvements

Although the algorithm produces relatively good results, there are a number of potential areas for improvement. These are detailed below:

- **HOT Correction:** The HOT correction algorithm is a significant improvement on previous methods, but does not perform well in all situations. Improvements could be made by:
  - Developing a method to automatically choose segmentation parameters, thus allowing the entire algorithm to be automated, either through standard OBIA parameter selection approaches (for example Dragut *et al.*, 2010), or by redeveloping the edge detection and linking stages so that a single parameterisation provides good results for all images
  - Improving the removal of edge effects by developing an adaptive thresholding method
  - Solving issues with background values and No Data areas, to remove the issues with poor correction over islands
  - Improving the iteration termination conditions so that over- and under-correction issues do not occur
- **HOT-AOT Relationship:** Similarly, the algorithm used to estimate the HOT-AOT relationship performs well most of the time, but poorly in certain situations. Improvements could be made by:

  - Using different parameterisations for different images, which should lead to final accuracies closer to those found from the actual relationships (§6.3) rather than the modelled relationship (§6.4)
  - Implementing more complex aerosol type parameterisation, to allow different areas of the image to have different aerosol types, for example: urban aerosols over a city and maritime aerosols near the coast
- **Computational efficiency:** Currently the algorithm takes approximately 2 hours to process a Landsat image (on a 3.2Ghz PC with 16Gb of RAM), depending on the complexity of the image (as this influences the time taken to perform the HOT correction procedure). This delay increases the time taken to iteratively develop and improve the algorithm, as well as making large-scale production of AOT data challenging. Speed improvements could be made by storing HOT values as integers rather than floating point values, following the guidance in Gorelick and Ozsvald (2014) regarding efficient use of `numpy` and implementing parallelisation where appropriate.
- **Outputs:** Currently the algorithm produces the resulting AOT image, along with a log file and a list of parameters used. Several other outputs which would be very useful to end-users could be produced including a Quality Assurance image (providing quality and error information for each pixel) and an 'atmospheric correction ready' AOT image with negative AOTs either masked out or set to a near-zero value (depending on the magnitude, as discussed in §7.5) and an option to interpolate small no data areas.

9.5 Conclusion

This dissertation has described the development of Haze Optimized Transform Based Aerosol Retrieval (HOTBAR), a new algorithm to estimate high-resolution AOT. The development of this method, including the development of several novel stages in the processing methodology, is the major novel contribution of this PhD. The method is currently implemented on Landsat images, providing AOT with a spatial resolution of 30m at a comparable accuracy to the operational MODIS 10km and 3km AOT products. Example applications have shown how this high-resolution data can be applied for monitoring of $PM_{2.5}$ air pollution at an unprecedented resolution, and for per-pixel atmospheric correction of satellite imagery.

Appendix A

Py6S

The 6S Radiative Transfer Model (Vermote *et al.*, 1997) is used extensively in this thesis to simulate atmospheric radiative transfer. The standard 6S interface is very difficult to use, and manual modifications must be made to the input file to run the same simulation for multiple wavelengths, angles or other parameters. To deal with this issue, and to provide a good foundation for future work, a modern interface to 6S was developed during the early stages of the PhD research. This interface was developed through the Python programming language, and is called Py6S.

A paper describing Py6S was published in Computers and Geosciences (Wilson, 2012), and the post-peer-review text of this paper is presented in this appendix.

A.1 Introduction

Radiative Transfer Models (RTMs) are widely used to simulate the passage of solar radiation through atmospheres on Earth and other planets. They have a range of uses including atmospheric research and solar energy system design, and are widely used within remote sensing and Earth observation, but they are often seen as difficult to use with respect to the numerous input and outputs parameters. This paper outlines Py6S, a Python interface to the 6S RTM (Vermote *et al.*, 1997) designed to address these issues. Py6S allows the user to write simple Python scripts which can set 6S input variables, run simulations and access the outputs. Methods are also provided to perform common tasks, such as running a simulation for a range of wavelengths. It is envisaged that Py6S will provide a useful framework in which research on atmospheric radiative transfer can be conducted using 6S, as well as opening the use of 6S to a wider audience including students.

A.2 6S

Second Simulation of the Satellite Signal in the Solar Spectrum (6S; Vermote *et al.*, 1997) is a radiative transfer model which has established itself as one of the standard RTMs used for both remote sensing research and the creation of operational products. The model is intermediate in complexity, between simple RTMs such as SPCTRAL2 (Bird and Riordan, 1993) and FAR (Seidel *et al.*, 2010), which do not produce results of the required accuracy for many applications, and very complex and computationally-intensive models such as SCIATRAN (Rozanov *et al.*, 2005), libRADTRAN (Mayer and Kylling, 2005) and LIDORT/VLIDORT (Spurr, 2008). It has been used in the development of new algorithms and spectral indices (for example Ceccato *et al.*, 2002) and is often combined with other models to produce fully-integrated models - for example it was used in the Kuusk and Nilson (2000) integrated forest reflectance model.

The current version is 6SV1.1, a vector version of the original 6S code which can simulate the atmospheric radiative transfer of polarised and non-polarised visible and infra-red radiation under different atmospheric conditions. Parameters include the atmospheric conditions, altitude of the sensor and target, wavelength and ground reflectance (with the ability to use a number of built-in BRDF models). An atmospheric correction mode allows the calculation of a ground reflectance, given an at-sensor radiance or reflectance value and a set of atmospheric parameters.

The primary simulation outputs are at-sensor reflectance and radiance, broken down into their individual components, as well as a number of other calculated atmospheric parameters. Validation has shown differences of less than 0.1% when compared with MODTRAN4 (Kotchenova *et al.*, 2006; Kotchenova and Vermote, 2007), and Kotchenova *et al.* (2008) studied a number of standard RTMs and found that 6S demonstrated the best agreement with a Monte Carlo benchmark (within 1%). The ability of the latest version of 6S to take into account polarisation of light is thought to be behind its

```
0 (User defined)
40.0 100.0 45.0 50.0 7 23 (geometrical conditions)
8 option for Water Vapor and Ozone
3.0 3.5 Water Vapor and Ozone
4 User's Components
0.25 0.25 0.25 0.25
0
0.5 value
-0.2  (target level, negative value)
-3.3 (sensor level)
-1.5 -3.5 (water vapor and ozone)
0.25 (aot)
11 (chosen band)
1 (Non homogeneous surface)
2 1 0.5 (ro1 ro2 radius)
1 BRDF
-0.1 radiance (positive value)
```

Listing A.1: Sample 6S input file

increased accuracy in many situations (Kotchenova *et al.*, 2008).

6S is used operationally as part of the atmospheric correction procedure for Landsat TM (Ouaidrari and Vermote, 1999), and for generating lookup tables in the MODIS atmospheric correction procedure (Vermote and Vermeulen, 1999). 6S is also frequently used for atmospheric correction of images from a number of sensors by end-users (for example Alencar *et al.*, 2011; Steven *et al.*, 2003).

A.2.1 Limitations in the interface

The user interface to 6S is provided through text input and output files (see Listings A.1 and A.2) which have a number of issues:

- Every parameter in the input file is specified using a number, even categorical parameters, making the file difficult to read and edit.
- The input file must have exactly the correct format (including whitespace) if it is to be read correctly by 6S, thus any small errors can lead to problems ranging from software crashes to subtly incorrect outputs.
- It is only possible to specify one parameter set in the 6S input file, thus running 6S for a range of parameter values (for example multiple wavelengths, or multiple atmospheric conditions) requires manually editing the input file between each simulation.
- The format of the 6S output file is easy for humans to read, but hard to extract values from automatically.

A.3 Py6S

Py6S is a Python interface to the 6S model which has been developed to address the limitations described above. By not re-implementing the model itself, we can ensure that results produced using Py6S will be exactly the same as results produced using 6S by itself, thus significantly reducing the amount of testing and validation required of the Py6S code.

```
******************************************************************************
*                                                                            *
*                        integrated values of  :                             *
*                        --------------------                                *
*                                                                            *
*       apparent reflectance  0.0330894  appar. rad.(w/m2/sr/mic)   12.749    *
*                  total gaseous transmittance  0.675                        *
*                                                                            *
******************************************************************************
*                                                                            *
*                        coupling aerosol -wv  :                             *
*                        --------------------                                *
*           wv above aerosol :   0.033     wv mixed with aerosol :   0.033    *
*                       wv under aerosol :   0.033                           *
******************************************************************************
*                                                                            *
*                        integrated values of  :                             *
*                        --------------------                                *
*                                                                            *
*       app. polarized refl.  0.0014    app. pol. rad. (w/m2/sr/mic)    0.065 *
*             direction of the plane of polarization -27.40                  *
*                  total polarization ratio     0.043                        *
*                                                                            *
******************************************************************************
*                                                                            *
*                        int. normalized  values  of  :                      *
*                        ----------------------------                        *
*                     % of irradiance at ground level                        *
*     % of direct  irr.    % of diffuse irr.    % of enviro. irr              *
*              0.773               0.221               0.005                 *
*                     reflectance at satellite level                         *
*     atm. intrin. ref.   environment ref.  target reflectance               *
*              0.015               0.004               0.014                 *
*                                                                            *
*                        int. absolute values of                             *
*                        -----------------------                             *
*                     irr. at ground level (w/m2/mic)                        *
*     direct solar irr.    atm. diffuse irr.    environment  irr              *
*             453.572             127.136               3.157                *
*                    rad at satel. level (w/m2/sr/mic)                       *
*     atm. intrin. rad.    environment rad.    target radiance                *
*              5.649               1.633               5.468                 *
*                                                                            *
*                                                                            *
*          int. funct filter (in mic)              int. sol. spect (in w/m2)  *
*             0.1174545                                 185.589               *
*                                                                            *
******************************************************************************

******************************************************************************
*                                                                            *
*                         integrated values of  :                            *
*                         --------------------                               *
*                                                                            *
*                             downward        upward          total          *
*      global gas. trans. :     0.68965         0.97248         0.67513       *
*      water    "     "    :     0.98573         0.98623         0.97592       *
*      ozone    "     "    :     0.70609         0.99079         0.70008       *
*      co2      "     "    :     1.00000         1.00000         1.00000       *
*      oxyg     "     "    :     0.99344         0.99533         0.99179       *
*      no2      "     "    :     1.00000         1.00000         1.00000       *
*      ch4      "     "    :     1.00000         1.00000         1.00000       *
*      co       "     "    :     1.00000         1.00000         1.00000       *
*                                                                            *
*                                                                            *
*      rayl.  sca. trans. :     0.96494         0.93809         0.90520       *
*      aeros. sca.   "    :     0.72090         0.82111         0.59194       *
*      total  sca.   "    :     0.69208         0.81074         0.56110       *
*                                                                            *
*                                                                            *
*                                                                            *
*                             rayleigh       aerosols         total          *
*                                                                            *
*      spherical albedo   :     0.04939         0.04918         0.06820       *
*      optical depth total:     0.05550         0.42021         0.47570       *
*      optical depth plane:     0.01848         0.23606         0.25454       *
*      reflectance I      :     0.01098         0.01327         0.02175       *
*      reflectance Q      :     0.00118         0.00037         0.00122       *
*      reflectance U      :    -0.00156         0.00000        -0.00173       *
*      polarized reflect. :     0.00195         0.00037         0.00212       *
*      degree of polar.   :       17.77            2.76            9.75       *
*      dir. plane polar.  :      -26.48            0.00          -27.43       *
*      phase function I   :     1.26026         0.27565         0.39051       *
*      phase function Q   :    -0.21911        -0.00611        -0.03096       *
*      phase function U   :    -1.19913        -0.15957        -0.28084       *
*      primary deg. of pol:    -0.17386        -0.02215        -0.07927       *
*      sing. scat. albedo :     1.00000         0.52284         0.57850       *
*                                                                            *
*                                                                            *
******************************************************************************
```

Listing A.2: Extract from a sample 6S output file

A.3.1 Features

Py6S provides a superset of the 6S features: any parameter that can be set manually in a standard 6S input file can also be set through Py6S. Thus, the description of features here will not focus on the scientific features of the standard 6S model, but will focus on the improvements that Py6S provides.

- **User-friendly parameter setting, with easily-accessible documentation:** 6S parameters can be set using a simple Python interface rather than a cryptic input file. For example, to set the wavelength for the simulation to that of Landsat TM band 1, you simply run `s.wavelength = Wavelength(PredefinedWavelengths.Landsat_TM_B1)`. Extensive documentation is provided regarding the parameters which can be set, and this documentation can be accessed interactively through the Python interpreter.
- **Helper functions making common operations simple:** Manually running 6S simulations for many wavelengths across a certain wavelength range is a common need, but it is normally very time-consuming as it requires much manual editing of the 6S input files. In Py6S this can be accomplished with a single call to `run_vnir`, or `run_landsat_tm`. Similarly, running a simulation with many solar or view angles to produce a polar plot showing directional reflectance effects can be accomplished with a single call to `plot_and_run_360`.
- **Plotting capabilities:** Py6S links with the Matplotlib plotting library (Hunter, 2007), allowing the results from 6S simulations to be easily plotted using functions such as `plot_wavelengths` and `plot_360`.
- **Access to all other Python functionality:** Py6S does not provide a Graphical User Interface (GUI) to 6S, as MODO (Schläpfer, 2001) does for MODTRAN4, but instead provides an API for the Python language, which allows a lot more flexibility. GUIs are easy to use for simple tasks, but can make it very difficult to perform more complicated tasks which the author may not have anticipated. By providing a Python API, code using Py6S can do anything that is possible within the Python language. For example, it can use all of the built-in functionality of the Python Standard Library, as well as access other Python modules commonly used in scientific computing (for example, `numpy`, `scipy`, `matplotlib` and `python-statslib`), allowing analysis of 6S outputs to be performed within the Python environment.
- **Ability to import parameters from external data sources:** Py6S allows detailed 6S parameterisation from real-world measurements. Currently supported sources are radiosonde data from the University of Wyoming Atmospheric Sciences department (available at `http://weather.uwyo.edu/upperair/sounding.html`) and sun photometry data from the AERONET network (Holben *et al.*, 1998). Importing this data manually would require interpolation, unit conversion and date/time subsetting, all of which is done automatically by the Py6S functions.
- **Reproducibility:** There has been a increased emphasis recently on improving the reproducibility of research conducted using computational approaches in many fields

(for example Baiocchi, 2007; Vandewalle *et al.*, 2009). By allowing a whole series of 6S simulations to be run from a single script, and linking with other Python modules for further analysis, Py6S allows entire research projects using 6S to be reproducible from a single Python script.

A.3.2 Usage

A.3.2.1 Installation

As Py6S is purely an interface to 6S, the original 6S executable is required to run Py6S. Full instructions for 6S compilation on Windows, Mac OS X and Linux are included in the Py6S documentation. Py6S and its dependencies can be installed from the Python Package Index with a single call to the `pip` utility.

A.3.2.2 Example Py6S scripts

This section provides a number of examples to introduce Py6S functionality; the first example shows how to set a number of parameters through Py6S, run the simulation, and extract some outputs:

```
from Py6S import *
# Create an object to hold the 6S parameters
s = SixS()
# Set the atmospheric profile to Tropical
s.atmos_profile = AtmosProfile.PredefinedType(AtmosProfile.Tropical)
# Set the wavelength to 0.357um
s.wavelength = Wavelength(0.357)
# Run the model and print some outputs
s.run()
print s.outputs.pixel_radiance
print s.outputs.background_radiance
print s.outputs.single_scattering_albedo
print s.outputs.transmittance_water.downward
```

This may not seem significantly easier than writing 6S input files manually – although it is less error-prone – but another example shows the power of Py6S when running a number of simulations for many wavelengths and plotting the results, shown in Figure 1.

```
from Py6S import *
s = SixS()
# Run the 6S simulation defined by this SixS object across the
# whole VNIR wavelength range, extracting the pixel reflectance
# from the model output
wavelengths, results = SixSHelpers.Wavelengths.run_vnir(s,
        output_name="pixel_radiance")
# Plot these results, with the y axis label set to "Pixel Radiance"
SixSHelpers.Wavelengths.plot_wavelengths(wavelengths, results,
        r"At-sensor Spectral Radiance ($W/m^2\!/\mu m$)")
```

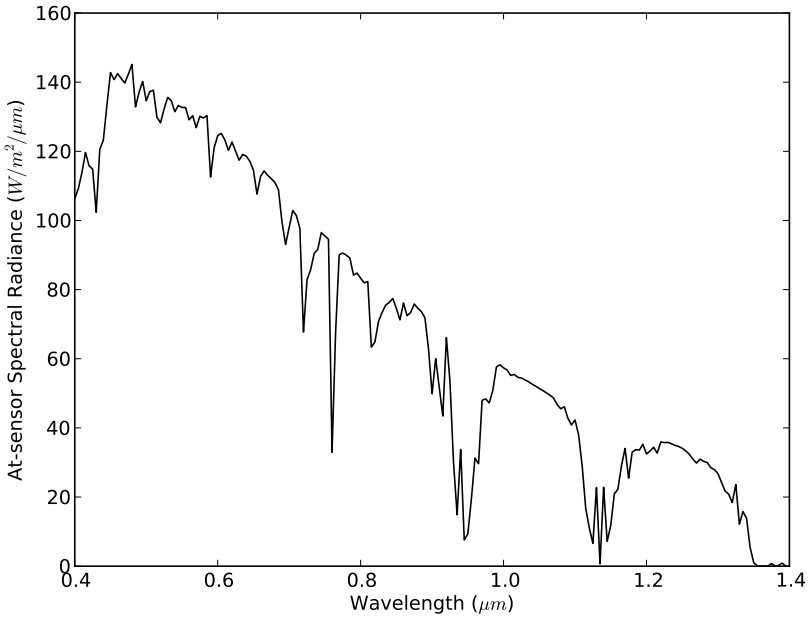


FIGURE A.1: An example output from the Py6S commands `run_vnir` and `plot_wavelengths`.

Similarly, simulations can be run for many angles, with the results plotted as a polar contour plot (shown in Figure 2), a time-consuming task which would require much editing of input files using the standard 6S interface, but which is far simpler in Py6S.

```
from Py6S import *
s = SixS()
# Set solar azimuth and zenith angles and wavelength
s.geometry.solar_a = 0
s.geometry.solar_z = 30
s.wavelength = Wavelength(0.550)
# Set the directional ground reflectance to be modeled
# by the Roujean BRDF model, using parameters for a pine forest
# (parameters taken from Roujean et al., 1992)
s.ground_reflectance = GroundReflectance.HomogeneousRoujean(0.037, 0.0, 0.133)
# Run the model and plot the results, varying the view angle (the other
# option is to vary the solar angle) and plotting the pixel radiance.
SixSHelpers.Angles.run_and_plot_360(s, 'view', 'pixel_radiance',
        colorbarlabel=r"At-sensor Spectral Radiance ($W/m^2\!/\mu m$)")
```

Further examples including importing of real-world data and more detailed parameterisations are provided in the documentation.

A.3.3 Design and Implementation

Py6S is a set of Python classes combined into the module `Py6S`. The main class is `SixS` which has attributes for setting parameters, and a `run` method to run the model and parse the outputs. When this method is called, the parameters are written to a temporary

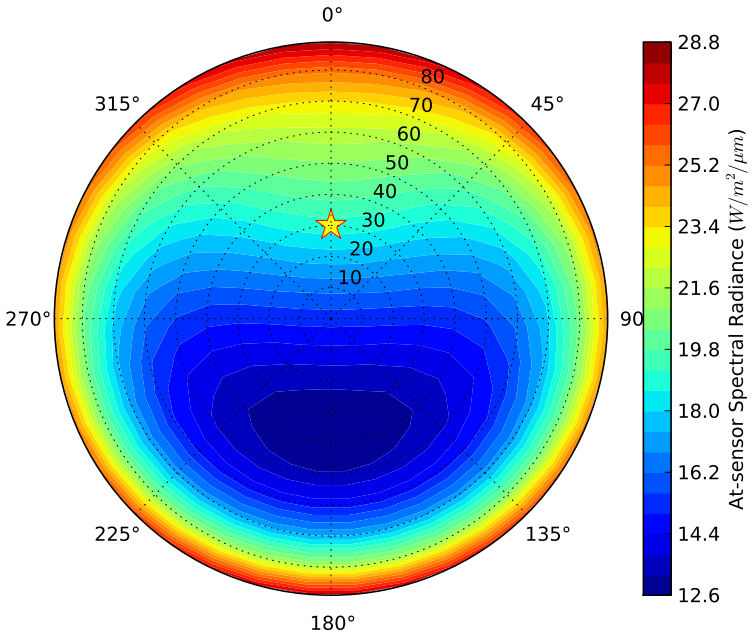


FIGURE A.2: An example output from the Py6S command `run_and_plot_360`. This shows the radiance of a surface parameterised with the Roujean Bidirectional Reflectance Distribution Function model for a pine forest ($k_0 = 0.037$, $k_1 = 0$, $k_2 = 0.133$), at 500nm. Simulations were performed every 10 deg for both azimuth and zenith, and the yellow star denotes the location of the sun.

input file and the standard 6S executable is then called on the input file. The full text output from 6S is captured and parsed to extract the numerical values of each output parameter which are then stored in the `values` dictionary.

Python's flexibility as a dynamically-typed language, allows functions to respond to a number of different types of input parameters. This has been used to provide a simple interface to the user, for example the `Wavelength` function can be called in any of the following manners:

```
# Wavelength of 0.43um
Wavelength(0.43)
# Band from 0.43-0.50um, with a flat response function of 1.0
Wavelength(0.43, 0.50)
# Band from 0.4-0.41um, with a custom response function
Wavelength(0.400, 0.410, [0.7, 0.9, 1.0, 0.3])
# A pre-defined sensor band wavelength
Wavelength(PredefinedWavelengths.LANDSAT_TM_B1)
```

This is far simpler to understand than a number of separate functions for setting different types of wavelength ranges. Similarly, although the output values are stored in a Python dictionary, they can be accessed as if they were attributes of the output class. For example:

```
# The standard way to access an item from a dictionary
```

```
s.outputs.values['pixel_reflectance']
# A cleaner and simpler way to access the output
s.outputs.pixel_reflectance
```

A.4 Conclusions

In conclusion, Py6S provides a modern environment for the scripting of 6S, a respected radiative transfer model in the remote-sensing community. Its features allow easy setting and modification of input variables and parameters and running of 6S simulations, it provide methods to import real-world measurements to 6S parameters, and it makes common operations such as running a simulation for all bands of a particular sensor easy. Py6S is released under the Lesser GNU Public License, and is available from `www.rtwilson.com/academic/py6s`.

Appendix B

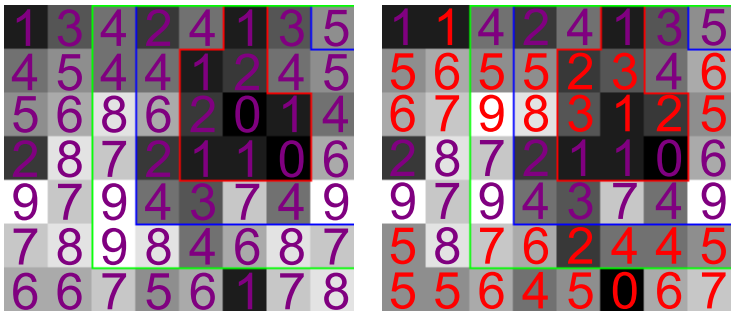
Example of noise effects on correction algorithm

The effect of the noise in Landsat TM data on the quality of the correction is discussed in §5.4.7. In this Appendix, an example on a very small image will be shown, so the effect of the noise on the calculations can be seen.

A small original HOT image was created by hand, with a single negative HOT anomaly present within the image. Noise was then added to this image, in a similar way to the Landsat TM noise: some random noise, plus two bands of increased and decreased values. All image values were integers in the range 0–10. The original image, with and without noise, are shown in Figure B.1. The noise is almost impossible to see in Figure B.1(d), but still has an effect on the correction.

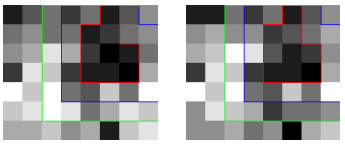
The segmentation was manually performed and the image object is outlined in red, with the border and offset border shown in blue and green respectively.

The algorithm was applied manually, and the results are shown in Figure B.2. The edge effect is clearly visible in Figure B.2(d), particularly when compared to Figure B.2(c). This edge effect is caused by the noise affecting both the average pixel value of the object (hence why the noise effects were reduced significantly when using the median as opposed to the mean - as the median is less sensitive to noise) and the Otsu threshold choice, meaning that more edge pixels are included in the group of pixels which are modified. The net result of this is an increase in edge effects - in this case positive edge effects (that is, increases in pixel value), although the same would occur as negative edge effects for positive anomalous areas.



(a) No noise

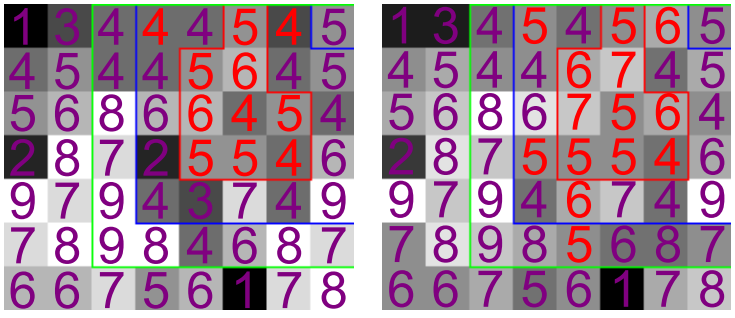
(b) With noise, altered pixels highlighted in red



(c) No noise (without annotations)

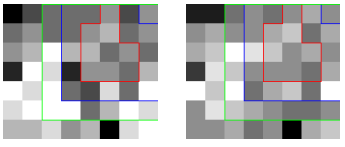
(d) With noise (without annotations)

FIGURE B.1: Original images, with the segmentation shown as a red line around the main object, a blue line around the immediate border, and a green line around the offset border.



(a) No noise

(b) With noise



(c) No noise (without annotations)

(d) With noise (without annotations)

FIGURE B.2: Corrected images, with altered pixels shown in red. The edge effect is particularly visible within the blue border area.

Bibliography

Acharya, Y. (2005). Spectral and emission characteristics of LED and its application to LED-based sun-photometry. *Optics & Laser Technology 37*(7), 547–550.

Ahn, C., O. Torres, and P. K. Bhartia (2008). Comparison of Ozone Monitoring Instrument UV Aerosol Products with Aqua/Moderate Resolution Imaging Spectroradiometer and Multiangle Imaging Spectroradiometer observations in 2006. *Journal of Geophysical Research 113*(D16), D16S27.

Air Quality Expert Group (2005). Particulate Matter in the United Kingdom: Summary. Technical Report PB10580a, Department for the Environment, Food and Rural Affairs, London.

Albrecht, B. A. (1989). Aerosols, cloud microphysics, and fractional cloudiness. *Science 245*(4923), 1227.

Alencar, A., G. Asner, D. Knapp, and D. Zarin (2011). Temporal variability of forest fires in eastern Amazonia. *Ecological Applications 21*(7), 2397–2412.

Allen, G., C. Sioutas, P. Koutrakis, R. Reiss, F. W. Lurmann, and P. T. Roberts (1997). Evaluation of the TEOM® Method for Measurement of Ambient Particulate Mass in Urban Areas. *Journal of the Air & Waste Management Association 47*(6), 682–689.

American Meteorological Society (2013). Glossary of Meteorology.

Anderson, K. and E. J. Milton (2006). On the temporal stability of ground calibration targets: implications for the reproducibility of remote sensing methodologies. *International Journal of Remote Sensing 27*(16), 3365–3374.

Arino, O., D. Gross, F. Ranera, L. Bourg, M. Leroy, P. Bicheron, J. Latham, A. Di Gregorio, C. Brockman, and R. Witt (2007). GlobCover: ESA service for global land cover from MERIS. In *IEEE International Geoscience and Remote Sensing Symposium, 2007*, pp. 2412–2415.

Ashpole, I. and R. Washington (2012). An automated dust detection using SEVIRI: A multiyear climatology of summertime dustiness in the central and western Sahara. *Journal of Geophysical Research: Atmospheres 117*(D8), n/a–n/a.

Ayers, G. P., M. D. Keywood, and J. L. Gras (1999). TEOM vs. manual gravimetric methods for determination of PM2.5 aerosol mass concentrations. *Atmospheric Environment 33*(22), 3717–3721.

Baiocchi, G. (2007). Reproducible research in computational economics: guidelines, integrated approaches, and open source software. *Computational Economics 30*(1), 19–40.

Bäumer, D. and B. Vogel (2007). An unexpected pattern of distinct weekly periodicities in climatological variables in Germany. *Geophysical Research Letters 34*, L03819.

Becker, M., M. Kirchner, P. Häfele, W. Söhne, and G. Weber (2003). Near real-time tropospheric signal delay from EPN and German permanent GPS sites. In *Proceedings of the EUREF Symposium, Toledo, Ohio, 4-7 June*.

Benas, N. and N. Chrysoulakis (2013). Comparative Analysis of MERIS/AATSR Synergy Algorithm Aerosol Retrievals Versus MODIS Aerosol Product and Validation Against AERONET Observations. In C. G. Helmis and P. T. Nastos (Eds.), *Advances in Meteorology, Climatology and Atmospheric Physics*, Springer Atmospheric Sciences, pp. 911–916. Springer Berlin Heidelberg.

Berk, A., G. P. Anderson, L. S. Bernstein, P. K. Acharya, H. Dothe, M. W. Matthew, S. M. Adler-Golden, J. H. Chetwynd Jr, S. C. Richtsmeier, B. Pukall, *et al.* (1999). MODTRAN 4 radiative transfer modeling for atmospheric correction. In *Proceedings of SPIE - The International Society for Optical Engineering*, Volume 3756, pp. 348–353.

Berk, A., L. S. Bernstein, and D. C. Robertson (1987). MODTRAN: A moderate resolution model for LOWTRAN. Technical Report AFGL-TR-87-0220, Air Force Geophysics Laboratory, Massachusetts.

Bevis, M., S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware (1992). GPS meteorology - Remote sensing of atmospheric water vapor using the Global Positioning System. *Journal of Geophysical Research 97*(D14), 15787–15801.

Bey, I., D. J. Jacob, R. M. Yantosca, J. A. Logan, B. D. Field, A. M. Fiore, Q. Li, H. Y. Liu, L. J. Mickley, and M. G. Schultz (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research: Atmospheres 106*(D19), 23073–23095.

Bird, R. and C. Riordan (1993). Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the Earth's surface for cloudless atmospheres. *SPIE Milestone Series 54*, 171–181.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing 65*(1), 2–16.

Blaschke, T., C. Burnett, and A. Pekkarinen (2004). Image segmentation methods for object-based analysis and classification. In *Remote sensing image analysis: including the spatial domain*, pp. 211–236. Springer.

Boggs, P. T., J. R. Donaldson, and R. B. Schnabel (1989). Algorithm 676: ODRPACK: software for weighted orthogonal distance regression. *ACM Transactions on Mathematical Software 15*(4), 348–364.

Boggs, P. T. and J. E. Rogers (1990). Orthogonal distance regression. *Contemporary Mathematics 112*, 183–194.

Bréon, F., A. Vermeulen, and J. Descloitres (2011). An evaluation of satellite aerosol products against sunphotometer measurements. *Remote Sensing of Environment 115*(12), 3102–3111.

Brooks, D. R. and F. M. Mims (2001). Development of an inexpensive handheld LED-based Sun photometer for the GLOBE program. *Journal of Geophysical Research 106*(D5), 4733–4740.

Brown, O. B., P. J. Minnett, R. Evans, E. Kearns, K. Kilpatrick, A. Kumar, R. Sikorski, and A. Závody (1999). MODIS Infrared Sea Surface Temperature Algorithm Algorithm Theoretical Basis Document Version 2.0. Technical report, University of Miami.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (6), 679–698.

Canty, M. J., A. A. Nielsen, and M. Schmidt (2004). Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sensing of Environment 91*(3–4), 441–451.

Ceccato, P., N. Gobron, S. Flasse, B. Pinty, and S. Tarantola (2002). Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach. *Remote Sensing of Environment 82*(2), 188–197.

Chan, P. W. (2009). Comparison of aerosol optical depth (AOD) derived from ground-based LIDAR and MODIS. *The Open Atmospheric Science Journal 3*, 131–137.

Chand, D., D. Jaffe, E. Prestbo, P. C. Swartzendruber, W. Hafner, P. Weiss-Penzias, S. Kato, A. Takami, S. Hatakeyama, and Y. Kajii (2008). Reactive and particulate mercury in the Asian marine boundary layer. *Atmospheric Environment 42*(34), 7988–7996.

Chatterjee, A., A. M. Michalak, R. A. Kahn, S. R. Paradise, A. J. Braverman, and C. E. Miller (2010). A geostatistical data fusion technique for merging remote sensing and ground-based observations of aerosol optical thickness. *Journal of Geophysical Research 115*(D20), D20207.

Chavez, P. S. (1975). Atmospheric, solar, and MTF corrections for ERTS digital imagery. In *American Society of Photogrammetry Fall Convention Proceedings*, Phoenex, pp. 69.

Chavez, P. S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment 24*(3), 459–479.

Che, H., X. Xia, J. Zhu, H. Wang, Y. Wang, J. Sun, X. Zhang, and G. Shi (2014). Aerosol optical properties under the condition of heavy haze over an urban site of Beijing, China. *Environmental Science and Pollution Research*, 1–11.

Chen, J., J. Fu, and M. Zhang (2011). An Atmospheric Correction Algorithm for Landsat/TM Imagery Basing on Inverse Distance Spatial Interpolation Algorithm: A Case Study in Taihu Lake. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4*(4), 882–889.

Chen, L. X., B. Zhang, W. Q. Zhu, X. J. Zhou, Y. F. Luo, Z. J. Zhou, and J. H. He (2009). Variation of atmospheric aerosol optical depth and its relationship with climate change in China east of 100 degrees East over the last 50 years. *Theoretical and Applied Climatology 96*(1), 191–199.

Chin, M., P. Ginoux, S. Kinne, O. Torres, B. N. Holben, B. N. Duncan, R. V. Martin, J. A. Logan, A. Higurashi, and T. Nakajima (2002). Tropospheric Aerosol Optical Thickness from the GOCART Model and Comparisons with Satellite and Sun Photometer Measurements. *Journal of the Atmospheric Sciences 59*(3), 461–483.

Chin, M., R. B. Rood, S. Lin, J. Müller, and A. M. Thompson (2000). Atmospheric sulfur cycle simulated in the global model GOCART: Model description and global properties. *Journal of Geophysical Research 105*(D20), 24,671–24,687.

Christopher, S. A., N. Feng, A. Naeger, B. Johnson, and F. Marenco (2012). Satellite remote sensing analysis of the 2010 Eyjafjallajökull volcanic ash cloud over the North Sea during 4–18 May 2010. *Journal of Geophysical Research 117*(D00U20).

Chudnovsky, A., C. Tang, A. Lyapustin, Y. Wang, J. Schwartz, and P. Koutrakis (2013). A critical assessment of high resolution aerosol optical depth (AOD) retrievals for fine particulate matter (PM) predictions. *Atmospheric Chemistry and Physics 13*, 10907–10917.

Coelho, L. P. (2013). Mahotas: Open source software for scriptable computer vision. *Journal of Open Research Software 1*(1).

Cooley, T., G. Anderson, G. Felde, M. Hoke, A. Ratkowski, J. Chetwynd, J. Gardner, S. Adler-Golden, M. Matthew, A. Berk, L. Bernstein, P. Acharya, D. Miller, and P. Lewis (2002). FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation. In *IEEE International Geoscience and Remote Sensing Symposium, 2002*, Volume 3, pp. 1414 – 1418.

Crist, E. P. (1985). A spectral haze diagnostic feature for normalizing Landsat Thematic Mapper data. In *18th International Symposium on Remote Sensing of Environment, Paris, France*, pp. 735–744.

Crist, E. P. and R. C. Cicone (1984). A physically-based transformation of Thematic Mapper Data - the TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing 22*(3), 256 –263.

d'Almeida, G. A., P. Koepke, and E. P. Shettle (1991). *Atmospheric aerosols: global climatology and radiative characteristics*. Deepak Publications, Hampton, VA, USA, 561pp.

Damoah, R., N. Spichtinger, C. Forster, P. James, I. Mattis, U. Wandinger, S. Beirle, T. Wagner, and A. Stohl (2004). Around the world in 17 days - hemispheric-scale transport of forest fire smoke from Russia in May 2003. *Atmospheric Chemistry and Physics 4*(5), 1311–1321.

Dave, J. V. (1978). Effect of aerosols on the estimation of total Ozone in an atmospheric column from the measurements of Its ultraviolet radiance. *Journal of the Atmospheric Sciences 35*(5), 899–911.

Davidson, C. I., R. F. Phalen, and P. A. Solomon (2005). Airborne particulate matter and human health: A review. *Aerosol Science and Technology 39*(8), 737–749.

de Vries, C., T. Danaher, R. Denham, P. Scarth, and S. Phinn (2007). An operational radiometric calibration procedure for the Landsat sensors based on pseudo-invariant target sites. *Remote Sensing of Environment 107*(3), 414–429.

Defra (2010). UK National Air Quality Directives. Technical report, Department for the Environment, Food and Rural Affairs, London.

Dey, S., L. Di Girolamo, A. van Donkelaar, S. Tripathi, T. Gupta, and M. Mohan (2012). Variability of outdoor fine particulate (PM2.5) concentration in the Indian Subcontinent: A remote sensing approach. *Remote Sensing of Environment 127*, 153–161.

Diner, D. J., W. A. Abdou, T. P. Ackerman, K. Crean, H. R. Gordon, R. A. Kahn, J. V. Martonchik, S. McMuldroch, S. R. Paradise, and B. Pinty (2001). MISR Level 2 aerosol retrieval Algorithm Theoretical Basis. Technical Report JPL-D114000, Revision G, Jet Propulsion Laboratory.

Dragut, L., D. Tiede, and S. R. Levick (2010). ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science 24*(6), 859–871.

Draxler, R. R. and G. D. Rolph (2003). HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) model. http://ready.arl.noaa.gov/HYSPLIT.php.

Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, A. Meygret, F. Spoto, O. Sy, F. Marchese, and P. Bargellini (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment 120*, 25–36.

Du, Y., B. Guindon, and J. Cihlar (2002). Haze detection and removal in high resolution satellite image with wavelet analysis. *IEEE Transactions on Geoscience and Remote Sensing 40*(1), 210 –217.

Dubovik, O. and M. D. King (2000). A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. *Journal of Geophysical Research 105*(D16), PP. 20,673–20,696.

Eck, T. F., B. N. Holben, J. S. Reid, O. Dubovik, A. Smirnov, N. T. O'Neill, I. Slutsker, and S. Kinne (1999). Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *Journal of Geophysical Research 104*(D24), 31333–31349.

El-Metwally, M., S. C. Alfaro, M. M. Abdel Wahab, A. S. Zakey, and B. Chatenet (2010). Seasonal and inter-annual variability of the aerosol content in Cairo (Egypt) as deduced from the comparison of MODIS aerosol retrievals with direct AERONET measurements. *Atmospheric Research 97*(1–2), 14–25.

Emili, E., A. Lyapustin, Y. Wang, C. Popp, S. Korkin, M. Zebisch, S. Wunderle, and M. Petitta (2011). High spatial resolution aerosol retrieval with MAIAC: Application to mountain regions. *Journal of Geophysical Research 116*(D23), D23211.

Engel-Cox, J. A., R. M. Hoff, and A. Haymet (2004). Recommendations on the Use of Satellite Remote-Sensing Data for Urban Air Quality. *Journal of the Air & Waste Management Association 54*(11), 1360–1371.

Evans, J., A. van Donkelaar, R. V. Martin, R. Burnett, D. G. Rainham, N. J. Birkett, and D. Krewski (2013). Estimates of global mortality attributable to particulate air pollution using satellite imagery. *Environmental Research 120*, 33–42.

Evans, K. F. and G. L. Stephens (1991). A new polarized atmospheric radiative transfer model. *Journal of Quantitative Spectroscopy and Radiative Transfer 46*(5), 413–423.

Ezzati, M., A. D. Lopez, A. Rodgers, S. Vander Hoorn, and C. J. Murray (2002). Selected major risk factors and global and regional burden of disease. *The Lancet 360*(9343), 1347–1360.

Finarelli, M. G. (1998). GLOBE: A Worldwide Environmental Science and Education Partnership. *Journal of Science Education and Technology 7*(1), 77–84.

Flowerdew, R. J. and J. D. Haigh (1995). An approximation to improve accuracy in the derivation of surface reflectances from multi-look satellite radiometers. *Geophysical Research Letters 22*(13), 1693–1696.

Flowerdew, R. J. and J. D. Haigh (1996). Retrieval of aerosol optical thickness over land using the ATSR-2 dual-look satellite radiometer. *Geophysical Research Letters 23*(4), 351–354.

Fry, J. A., G. Xian, S. Jin, J. A. Dewitz, C. G. Homer, L. Yang, C. A. Barnes, N. D. Herold, and J. D. Wickham (2011). National Land Cover Dataset of the Conterminous United States. *Photogrammetric Engineering and Remote Sensing 77*(9), 859–864.

Fu, L., T. Nunifu, and B. Leung (2014). A Two-step Approach for Relating TEOM and Dichotomous Air Sampler PM2.5 Measurements. *Journal of the Air & Waste Management Association 64*(10), 1195–1203.

Fuller, G. W. and D. Green (2006). Evidence for increasing concentrations of primary PM10 in London. *Atmospheric Environment 40*(32), 6134–6145.

Gao, B. and A. Goetz (1990). Column atmospheric water vapor and vegetation liquid water retrievals from airborne imaging spectrometer data. *Journal of Geophysical Research: Atmospheres 95*(D4), 3549–3564.

Gao, B. C. and Y. J. Kaufman (2003). Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels. *Journal of Geophysical Research 108*(D13), 4389.

Gautam, R., Z. Liu, R. P. Singh, and N. C. Hsu (2009). Two contrasting dust-dominant periods over India observed from MODIS and CALIPSO data. *Geophysical Research Letters 36*(6), L06813.

GCOS (2004). Implementation Plan for the Global Observing Systems for Climate in support of the UNFCCC. Technical Report GCOS-92, San Diego.

GDAL Development Team (2014). *GDAL - Geospatial Data Abstraction Library, Version 1.11*. Open Source Geospatial Foundation.

Ginoux, P., M. Chin, I. Tegen, J. M. Prospero, B. Holben, O. Dubovik, and S. Lin (2001). Sources and distributions of dust aerosols simulated with the GOCART model. *Journal of Geophysical Research 106*(D17), 20255–20,273.

Goloub, P., D. Tanré, J. Deuze, M. Herman, A. Marchand, and F. Breon (1999). Validation of the first algorithm applied for deriving the aerosol properties over the ocean using the POLDER/ADEOS measurements. *IEEE Transactions on Geoscience and Remote Sensing 37*(3), 1586 –1596.

González, C. R., M. Schaap, G. D. Leeuw, P. J. H. Builtjes, and M. van Loon (2003). Spatial variation of aerosol properties over Europe derived from satellite observations and comparison with model calculations. *Atmospheric Chemistry and Physics 3*, 521–533.

Gorelick, M. and I. Ozsvald (2014). *High Performance Python*. Sebastopol, California: O'Reilly Media. 370pp.

Green, A., M. Berman, P. Switzer, and M. Craig (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing 26*(1), 65–74.

Grell, G. A., S. E. Peckham, R. Schmitz, S. A. McKeen, G. Frost, W. C. Skamarock, and B. Eder (2005). Fully coupled "online" chemistry within the WRF model. *Atmospheric Environment 39*(37), 6957–6975.

Griggs, M. (1975). Measurements of atmospheric aerosol optical thickness over water using ERTS-1 data. *Journal of the Air Pollution Control Association 25*(6), 622–626.

Gupta, P., S. A. Christopher, J. Wang, R. Gehrig, Y. Lee, and N. Kumar (2006). Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment 40*(30), 5880–5892.

Gupta, P., F. Patadia, and S. Christopher (2008). Multisensor Data Product Fusion for Aerosol Research. *IEEE Transactions on Geoscience and Remote Sensing 46*(5), 1407 –1415.

Hadjimitsis, D. G. and C. Clayton (2009). Determination of aerosol optical thickness through the derivation of an atmospheric correction for short-wavelength Landsat TM and ASTER image data: an application to areas located in the vicinity of airports at UK and Cyprus. *Applied Geomatics 1*(1), 31–40.

Hall, D., G. Riggs, and V. Salomonson (1995). Development of methods for mapping global snow cover using Moderate Resolution Imaging Spectroradiometer data. *Remote Sensing of Environment 54*(2), 127–140.

Han, Y., X. Dai, X. Fang, Y. Chen, and F. Kang (2008). Dust aerosols: A possible accelerant for an increasingly arid climate in North China. *Journal of Arid Environments 72*(8), 1476–1489.

Harrison, D. (2010). Assessment of the UK AURN Particulate Matter monitoring equipment against the January 2010 guide to demonstration of equivalence. Technical Report AGG04003328/BV/AQ/DH/2658/V3, Bureau Veritas, London.

Harrison, D., R. Maggs, and J. Booker (2006). UK Equivalence Programme for Monitoring of Particulate Matter. Technical Report BV/AQ/AD202209/DH/2396, Bureau Veritas, London.

Hay, G. J. and G. Castilla (2008). Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In *Object-based image analysis*, pp. 75–89. Berlin: Springer.

Haywood, J. M., S. R. Osborne, P. N. Francis, A. Keil, P. Formenti, M. O. Andreae, and P. H. Kaye (2003). The mean physical and optical properties of regional haze dominated by biomass burning aerosol measured from the C-130 aircraft during SAFARI 2000. *Journal of Geophysical Research 108*(D13), 8473.

He, J., Y. Zha, J. Zhang, J. Gao, and Q. Wang (2014). Synergetic retrieval of terrestrial AOD from MODIS images of twin satellites Terra and Aqua. *Advances in Space Research 53*(9), 1337–1346.

He, X. Y., J. B. Hu, W. Chen, and X. Y. Li (2010). Haze removal based on advanced haze-optimized transformation (AHOT) for multispectral imagery. *International Journal of Remote Sensing 31*(20), 5331–5348.

Henderson, B. and P. Chylek (2005). The Effect of Spatial Resolution on Satellite Aerosol Optical Depth Retrieval. *IEEE Transactions on Geoscience and Remote Sensing 43*(9), 1984 – 1990.

Hoff, R. M. and S. A. Christopher (2009). Remote sensing of particulate pollution from space: Have we reached the promised land? *Journal of the Air & Waste Management Association 59*(6), 645–675.

Holben, B., T. Eck, I. Slutsker, D. Tanré, J. Buis, A. Setzer, E. Vermote, J. Reagan, Y. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov (1998). AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. *Remote Sensing of Environment 66*(1), 1–16.

Holben, B., E. Vermote, Y. Kaufman, D. Tanré, and V. Kalb (1992). Aerosol retrieval over land from AVHRR data-application for atmospheric correction. *IEEE Transactions on Geoscience and Remote Sensing 30*(2), 212 –222.

Holben, B. N., A. Smirnov, T. F. Eck, I. Slutsker, N. Abuhassan, W. W. Newcomb, J. S. Schafer, D. Tanré, B. Chatenet, and F. Lavenu (2001). An emerging ground-based aerosol climatology - Aerosol optical depth from AERONET. *Journal of Geophysical Research 106*(D11), 12067–12097.

Horvath, H. (1971). On the applicability of the Koschmieder visibility formula. *Atmospheric Environment 5*(3), 177–184.

Horvath, H. (1981). Atmospheric visibility. *Atmospheric Environment 15*(10), 1785–1796.

Hsu, N., S. Tsay, M. King, and J. Herman (2004). Aerosol properties over bright-reflecting source regions. *IEEE Transactions on Geoscience and Remote Sensing 42*(3), 557–569.

Hsu, N. C., J. R. Herman, P. K. Bhartia, C. J. Seftor, O. Torres, A. M. Thompson, J. F. Gleason, T. F. Eck, and B. N. Holben (1996). Detection of biomass burning smoke from TOMS measurements. *Geophysical Research Letters 23*(7), 745–748.

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering 9*(3), 90–95.

Ichoku, C., D. A. Chu, S. Mattoo, Y. J. Kaufman, L. A. Remer, D. Tanré, I. Slutsker, and B. N. Holben (2002). A spatio-temporal approach for global validation and analysis of MODIS aerosol products. *Geophysical Research Letters 29*(12), 8006.

IPCC (2007). Summary for Policymakers. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate change*. Cambridge University Press.

Iqbal, M. (1983). *An introduction to solar radiation*. Academic Press, Toronto. 408pp.

Irish, R. R. (2000). Landsat 7 automatic cloud cover assessment. In *AeroSense 2000, International Society for Optics and Photonics*, pp. 348–355.

Irons, J. R., J. L. Dwyer, and J. A. Barsi (2012). The next Landsat satellite: The Landsat Data Continuity Mission. *Remote Sensing of Environment 122*, 11–21.

ITC (2014). Database of Satellites and Sensors.

Jackson, J. M., H. Liu, I. Laszlo, S. Kondragunta, L. A. Remer, J. Huang, and H. Huang (2013). Suomi-NPP VIIRS aerosol algorithms and data products. *Journal of Geophysical Research: Atmospheres 118*(22), 12,673–12,689.

Jethva, H., S. Satheesh, J. Srinivasan, and K. Moorthy (2009). How Good is the Assumption About Visible Surface Reflectance in MODIS Aerosol Retrieval Over Land? A Comparison With Aircraft Measurements Over an Urban Site in India. *IEEE Transactions on Geoscience and Remote Sensing 47*(7), 1990–1998.

Jones, E., T. Oliphant, and P. Peterson (2001). *SciPy: Open source scientific tools for Python*.

Ju, J., D. P. Roy, E. Vermote, J. Masek, and V. Kovalskyy (2012). Continental-scale validation of MODIS-based and LEDAPS Landsat ETM+ atmospheric correction methods. *Remote Sensing of Environment 122*, 175–184.

Kahn, R. A. (2005). Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations. *Journal of Geophysical Research 110*(D10).

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Sana, G. White, J. Woollen, *et al.* (1996). The NCEP/NCAR 40-Year Re-analysis Project. *Bulletin of the American Meteorological Society 77*(3), 437–471.

Kanaroglou, P., N. Soulakellis, and N. Sifakis (2002). Improvement of satellite-derived pollution maps with the use of a geostatistical interpolation method. *Journal of Geographical Systems 4*(2), 193–208.

Kaufman, Y., L. Remer, D. Tanré, R. Li, R. Kleidman, S. Mattoo, R. Levy, T. Eck, B. Holben, C. Ichoku, J. Martins, and I. Koren (2005). A critical examination of the residual cloud contamination and diurnal sampling effects on MODIS estimates of aerosol over ocean. *IEEE Transactions on Geoscience and Remote Sensing 43*(12), 2886–2897.

Kaufman, Y., A. Wald, L. Remer, B. Gao, R. Li, and L. Flynn (1997). The MODIS 2.1um channel-correlation with visible reflectance for use in remote sensing of aerosol. *IEEE Transactions on Geoscience and Remote Sensing 35*(5), 1286 –1298.

Kaufman, Y. J. (1993). Aerosol optical thickness and atmospheric path radiance. *Journal of Geophysical Research 98*(D2), 2677–2692.

Kaufman, Y. J. and C. Sendra (1988). Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. *International Journal of Remote Sensing 9*(8), 1357–1381.

Kaufman, Y. J. and D. Tanré (1996). Strategy for direct and indirect methods for correcting the aerosol effect on remote sensing: From AVHRR to EOS-MODIS. *Remote Sensing of Environment 55*(1), 65–79.

Kaufman, Y. J., D. Tanré, A. Chu, S. Mattoo, and L. A. Remer (1998). Algorithm for remote sensing of tropospheric aerosol from MODIS: Collection 004. Technical Report MOD04-1998, National Aeronautics and Space Administration, Goddard Space Flight Center.

Kauth, R. J. and G. S. Thomas (1976). The tasselled cap - A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In *Symposium on Machine Processing of Remotely Sensed Data, 29th June - 1st July*, Purdue University.

King, M., Y. Kaufman, W. Menzel, and D. Tanré (1992). Remote sensing of cloud, aerosol, and water vapor properties from the moderate resolution imaging spectrometer (MODIS). *IEEE Transactions on Geoscience and Remote Sensing 30*(1), 2–27.

King, M. D., Y. J. Kaufman, D. Tanré, and T. Nakajima (1999). Remote sensing of tropospheric aerosols from space: Past, present, and future. *Bulletin of the American Meteorological Society 80*(11), 2229–2259.

Kneizys, F. X., E. P. Shettle, and W. O. Gallery (1981). Atmospheric transmittance and radiance: The LOWTRAN 5 Code. *SPIE Atmospheric Transmission 277*, 1–19.

Koelemeijer, R., C. Homan, and J. Matthijsen (2006). Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. *Atmospheric Environment 40*(27), 5304–5315.

Kondratyev, K. Y., L. S. Ivlev, V. F. Krapivin, and C. A. Varostos (2005). *Atmospheric aerosol properties: Formation, processes and impacts*. Berlin: Springer. 572pp.

Koschmieder, H. (1924). Theorie der horizontalen Sichtweite. *Beitrage zur Physik der Freien Atmosphare 12*, 33–55.

Kotchenova, S. and E. Vermote (2007). Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part II. Homogeneous Lambertian and anisotropic surfaces. *Applied Optics 46*(20), 4455–4464.

Kotchenova, S., E. Vermote, R. Levy, and A. Lyapustin (2008). Radiative transfer codes for atmospheric correction and aerosol retrieval: intercomparison study. *Applied Optics 47*(13), 2215–2226.

Kotchenova, S., E. Vermote, R. Matarrese, F. Klemm Jr, *et al.* (2006). Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path radiance. *Applied Optics 45*(26), 6762–6774.

Krijger, J. M., M. Van Weele, I. Aben, and R. Frey (2007). Technical Note: The effect of sensor resolution on the number of cloud-free observations from space. *Atmospheric Chemistry and Physics 7*(11), 2881–2891.

Kuusk, A. and T. Nilson (2000). A Directional Multispectral Forest Reflectance Model. *Remote Sensing of Environment 72*(2), 244 – 252.

Lambeck, P. F. (1977). Signature Extension Preprocessing for Landsat MSS Data. Technical Report ERIM l22700—32—F, Environmental Research Institute of Michigan, Ann Arbor, Michigan.

Langley, S. P. (1884). Researches on solar heat and its absorption by the earth's atmosphere: A report of the Mount Whitney expedition. Technical report, Professional papers of the Signal Service, USA War Department.

Langmann, B., A. Folch, M. Hensch, and V. Matthias (2012). Volcanic ash over Europe during the eruption of Eyjafjallajökull on Iceland, April–May 2010. *Atmospheric Environment 48*, 1–8.

Lavreau, J. (1991). De-hazing Landsat Thematic Mapper images. *Photogrammetric Engineering and Remote Sensing 57*(10), 1297–1302.

Le Treut, H. (2012). Greenhouse Gases, Aerosols and Reducing Future Climate Uncertainties. *Surveys in Geophysics 33*(3), 723–731.

Leckner, B. (1978). The spectral distribution of solar radiation at the Earth's surface—elements of a model. *Solar Energy 20*(2), 143–150.

Lentile, L. B., Z. A. Holden, A. M. S. Smith, M. J. Falkowski, A. T. Hudak, P. Morgan, S. A. Lewis, P. E. Gessler, and N. C. Benson (2006). Remote sensing techniques to assess active fire characteristics and post-fire effects. *International Journal of Wildland Fire 15*(3), 319–345.

Levy, R. C., L. A. Remer, R. G. Kleidman, S. Mattoo, C. Ichoku, R. Kahn, and T. F. Eck (2010). Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. *Atmospheric Chemistry and Physics 10*(21), 10399–10420.

Li, Y., Y. Xue, X. He, and J. Guang (2012). High-resolution aerosol remote sensing retrieval over urban areas by synergetic use of HJ-1 CCD and MODIS data. *Atmospheric Environment 46*, 173–180.

Li, Z., J. P. Muller, and P. Cross (2003). Comparison of precipitable water vapor derived from radiosonde, GPS, and Moderate-Resolution Imaging Spectroradiometer measurements. *Journal of Geophysical Research 108*(10.1029).

Liang, S., H. Fang, and M. Chen (2001). Atmospheric correction of Landsat ETM+ land surface imagery. I. Methods. *IEEE Transactions on Geoscience and Remote Sensing 39*(11), 2490 –2498.

Liang, S., H. Fang, J. T. Morisette, M. Chen, C. J. Shuey, C. L. Walthall, and C. S. T. Daughtry (2002). Atmospheric correction of Landsat ETM+ land surface imagery. II. Validation and applications. *IEEE Transactions on Geoscience and Remote Sensing 40*(12), 2736–2746.

Lim, S. S., T. Vos, A. D. Flaxman, G. Danaei, K. Shibuya, H. Adair-Rohani, *et al.* (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet 380*(9859), 2224–2260.

Liu, C., J. Hu, Y. Lin, S. Wu, and W. Huang (2011). Haze detection, perfection and removal for high spatial resolution satellite imagery. *International Journal of Remote Sensing 32*(23), 8685–8697.

Liu, C. H. and G. R. Liu (2009). Aerosol Optical Depth Retrieval for SPOT HRV Images. *Journal of Marine Science and Technology 17*(4), 300–305.

Liu, G. R., A. J. Chen, T. H. Lin, and T. H. Kuo (2002). Applying SPOT data to estimate the aerosol optical depth and air quality. *Environmental Modelling & Software 17*(1), 3–9.

Liu, H., L. A. Remer, J. Huang, H. Huang, S. Kondragunta, I. Laszlo, M. Oo, and J. M. Jackson (2014). Preliminary evaluation of S-NPP VIIRS aerosol optical thickness. *Journal of Geophysical Research: Atmospheres 119*(7), 3942–3962.

Liu, Y., J. A. Sarnat, B. A. Coull, P. Koutrakis, and D. J. Jacob (2004). Validation of Multiangle Imaging Spectroradiometer (MISR) aerosol optical thickness measurements using Aerosol Robotic Network (AERONET) observations over the contiguous United States. *Journal of Geophysical Research 109*(D6), D06205.

Liu, Z., M. S. Wong, J. Nichol, and P. W. Chan (2011). A multi-sensor study of water vapour from radiosonde, MODIS and AERONET: a case study of Hong Kong. *International Journal of Climatology*.

Livingston, J. M., J. Redemann, P. B. Russell, O. Torres, B. Veihelmann, P. Veefkind, R. Braak, A. Smirnov, L. Remer, R. W. Bergstrom, O. Coddington, K. S. Schmidt, P. Pilewskie, R. Johnson, and Q. Zhang (2009). Comparison of aerosol optical depths from the Ozone Monitoring Instrument (OMI) on Aura with results from airborne sunphotometry, other space and ground measurements during MILAGRO/INTEX-B. *Atmospheric Chemistry and Physics 9*(18), 6743–6765.

Lobser, S. E. and W. B. Cohen (2007). MODIS tasselled cap: land cover characteristics expressed through transformed MODIS data. *International Journal of Remote Sensing 28*(22), 5079–5101.

Lohmann, U. and J. Feichter (2005). Global indirect aerosol effects: a review. *Atmospheric Chemistry and Physics 5*(3), 715–737.

Loughner, C., D. Lary, L. Sparling, R. Cohen, P. DeCola, and W. Stockwell (2007). A Method to Determine the Spatial Resolution Required to Observe Air Quality From Space. *IEEE Transactions on Geoscience and Remote Sensing 45*(5), 1308 –1314.

Lyapustin, A., Y. Wang, I. Laszlo, R. Kahn, S. Korkin, L. Remer, R. Levy, and J. S. Reid (2011). Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *Journal of Geophysical Research 116*(D3), D03211.

Lyapustin, A., D. L. Williams, B. Markham, J. Irons, B. Holben, and Y. Wang (2004). A method for unbiased high-resolution aerosol retrieval from Landsat. *Journal of the Atmospheric Sciences 61*(11), 1233–1244.

Lyapustin, A. I., Y. Wang, I. Laszlo, T. Hilker, F. G.Hall, P. J. Sellers, C. J. Tucker, and S. V. Korkin (2011). Multi-angle implementation of atmospheric correction for MODIS (MAIAC): 3. Atmospheric correction. *Remote Sensing of Environment*.

Maiersperger, T., P. Scaramuzza, L. Leigh, S. Shrestha, K. Gallo, C. Jenkerson, and J. Dwyer (2013). Characterizing LEDAPS surface reflectance products by comparisons with AERONET, field spectrometer, and MODIS data. *Remote Sensing of Environment 136*, 1–13.

Markham, B. (1985). The Landsat sensors' spatial responses. *IEEE Transactions on Geoscience and Remote Sensing 6*, 864–875.

Markham, B., J. Storey, D. L. Williams, and J. Irons (2004). Landsat sensor performance: history and current status. *IEEE Transactions on Geoscience and Remote Sensing 42*(12), 2691–2694.

Martin, R. V. (2008). Satellite remote sensing of surface air quality. *Atmospheric Environment 42*(34), 7823–7843.

Martins, J. V., D. Tanré, L. Remer, Y. Kaufman, S. Mattoo, and R. Levy (2002). MODIS Cloud screening for remote sensing of aerosols over oceans using spatial variability. *Geophysical Research Letters 29*(12), 8009.

Masek, J. G., M. Honzak, S. N. Goward, P. Liu, and E. Pak (2001). Landsat-7 ETM+ as an observatory for land cover: Initial radiometric and geometric comparisons with Landsat-5 Thematic Mapper. *Remote Sensing of Environment 78*(1–2), 118–130.

Masek, J. G., E. F. Vermote, N. E. Saleous, R. Wolfe, F. G. Hall, K. F. Huemmrich, F. Gao, J. Kutler, and T. Lim (2006). A Landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters 3*(1), 68–72.

Matérn, B. (1986). *Spatial variation*. Berlin: Springer. 151pp.

Matthias, V., D. Balis, J. Bösenberg, R. Eixmann, M. Iarlori, L. Komguem, I. Mattis, A. Papayannis, G. Pappalardo, and M. R. Perrone (2004). Vertical aerosol distribution over Europe: Statistical analysis of Raman lidar data from 10 European Aerosol Research Lidar Network (EARLINET) stations. *Journal of Geophysical Research 109*(18), D18201.

Mayer, B. and A. Kylling (2005). Technical note: The libRadtran software package for radiative transfer calculations: description and examples of use. *Atmospheric Chemistry and Physics 5*(2), 1319–1381.

McGowan, H. and A. Clark (2008). Identification of dust transport pathways from Lake Eyre, Australia using HYSPLIT. *Atmospheric Environment 42*(29), 6915–6925.

McKinney, W. (2010). Data Structures for Statistical Computing in Python. pp. 51–56.

McMurry, P. H. (2000). A review of atmospheric aerosol measurements. *Atmospheric Environment 34*(12–14), 1959–1999.

Mélin, F., G. Zibordi, and S. Djavidnia (2007). Development and validation of a technique for merging satellite derived aerosol optical depth from SeaWiFS and MODIS. *Remote Sensing of Environment 108*(4), 436–450.

Middleton, W. (1952). *Vision through the atmosphere*. University of Toronto Press. 250pp.

Milton, E. J., F. Baret, P. Rossello, E. Anderson, and E. Rockall (2011). A multistage database of field measurements and synoptic remotely sensed data to support model validation and testing in Earth observation. *Computers and Geosciences 37*(9), 1511–1514.

Mims, F. M. (2003). A 5-year study of a new kind of Photosynthetically Active Radiation sensor. *Photochemistry and Photobiology 77*(1), 30–33.

Mishchenko, M. I., B. Cairns, J. E. Hansen, L. D. Travis, R. Burg, Y. J. Kaufman, J. Vanderlei Martins, and E. P. Shettle (2004). Monitoring of aerosol forcing of climate from space: analysis of measurement requirements. *Journal of Quantitative Spectroscopy and Radiative Transfer 88*(1–3), 149–161.

Mishchenko, M. I. and L. D. Travis (1997). Satellite retrieval of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight. *Journal of Geophysical Research 102*(D14), 16–989.

Mook, R. (1978). Minimum precipitable water in the atmosphere above Skibotn, Northern Norway. *Pure and Applied Geophysics 117*(4), 573–582.

Moran, M. S., R. D. Jackson, P. N. Slater, and P. M. Teillet (1992). Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output. *Remote Sensing of Environment 41*(2), 169–184.

Moreno, J., L. Alonso, G. Fernández, J. C. Fortea, S. Gandía, L. Guanter, J. C. García, J. M. Martí, and J. Melia (2004). The SPECTRA Barrax Campaign (SPARC): Overview and first results from CHRIS data. In *Proceedings of 2nd CHRIS/PROBA Workshop, 28-30 April*, Frascati, Italy.

Moro, G. D. and L. Halounova (2007). Haze removal for high-resolution satellite data: a case study. *International Journal of Remote Sensing 28*(10), 2187–2205.

Moroz, B. E., H. L. Beck, A. Bouville, and S. L. Simon (2010). Predictions of dispersion and deposition of fallout from nuclear testing using the NOAA HYSPLIT meteorological model. *Health Physics 99*(2), 252–269.

Morys, M., F. M. Mims, S. Hagerup, S. E. Anderson, A. Baker, J. Kia, and T. Walkup (2001). Design, calibration, and performance of MICROTOPS II handheld ozone monitor and Sun photometer. *Journal of Geophysical Research 106*(D13), 14573–14,582.

Mukai, S. and I. Sano (1999). Retrieval algorithm for atmospheric aerosols based on multi-angle viewing of ADEOS/POLDER. *Earth Planets and Space 51*(11), 1247–1254.

Munchak, L. A., R. C. Levy, S. Mattoo, L. A. Remer, B. N. Holben, J. S. Schafer, C. A. Hostetler, and R. A. Ferrare (2013). MODIS 3 km aerosol product: applications over land in an urban/suburban region. *Atmospheric Measurement Techniques 6*, 1747–1759.

Nagaraja Rao, C. R., L. L. Stowe, and E. P. McClain (1989). Remote sensing of aerosols over the oceans using AVHRR data Theory, practice and applications. *International Journal of Remote Sensing 10*(4-5), 743–749.

Natural Environment Research Council (2012). NERC British Isles Continuous GNSS Facility. http://bigf.ac.uk/.

Nichol, J. E. and V. Vohora (2004). Noise over water surfaces in Landsat TM images. *International Journal of Remote Sensing 25*(11), 2087–2093.

Nirala, M. (2008). Technical Note: Multi-sensor data fusion of aerosol optical thickness. *International Journal of Remote Sensing 29*(7), 2127–2136.

Nychka, D., R. Furrer, and S. Sain (2014). *fields: Tools for spatial data*. R package version 7.1.

Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica 11*(285-296), 23–27.

Ouaidrari, H. and E. F. Vermote (1999). Operational Atmospheric Correction of Landsat {TM} Data. *Remote Sensing of Environment 70*(1), 4 – 15.

Page, S. J. (2013). Bias Sources and Corrections in TEOM Technology. *Journal of Occupational and Environmental Hygiene 10*(5), 233–242.

Penner, J. E., R. J. Charlson, S. E. Schwartz, J. M. Hales, N. S. Laulainen, L. Travis, R. Leifer, T. Novakov, J. Ogren, and L. F. Radke (1994). Quantifying and Minimizing Uncertainty of Climate Forcing by Anthropogenic Aerosols. *Bulletin of the American Meteorological Society 75*(3), 375–400.

Péré, J. C., V. Pont, M. Mallet, and B. Bessagnet (2009). Mapping of PM10 surface concentrations derived from satellite observations of aerosol optical thickness over South-Eastern France. *Atmospheric Research 91*(1), 1–8.

Petty, G. W. (2006). *A first course in atmospheric radiation.* Madison, WI: Sundog Pub. 459pp.

Planchon, O. and F. Darboux (2002). A fast, simple and versatile algorithm to fill the depressions of digital elevation models. *CATENA 46*(2–3), 159–176.

Poulsen, C., R. Siddans, G. Thomas, A. Sayer, R. G. Grainger, O. Altaratz, O. Portela-Arjona, and P. Deschamps (2009). ESA GlobAerosol: Final Validation and Intercomparison Report. Technical Report GlobAerosol FVIR v3.2, European Space Agency, Frascati, Italy.

Powell, S. L., D. Pflugmacher, A. A. Kirschbaum, Y. Kim, and W. B. Cohen (2007). Moderate resolution remote sensing alternatives: a review of Landsat-like sensors and their applications. *Journal of Applied Remote Sensing 1*(1), 012506–012506–16.

Prasad, A. K. and R. P. Singh (2007). Changes in aerosol parameters during major dust storm events (2001–2005) over the Indo-Gangetic Plains using AERONET and MODIS data. *Journal of Geophysical Research 112*(D9), D09208.

R Core Development Team (2014). R: A Language and Environment for Statistical Computing. Technical report, R Foundation for Statistical Computing, Vienna, Austria.

Randel, D. L., T. J. Greenwald, T. H. Vonder Haar, G. L. Stephens, M. A. Ringerud, and C. L. Combs (1996). A New Global Water Vapor Dataset. *Bulletin of the American Meteorological Society 77*(6), 1233–1246.

Remer, L. A., R. G. Kleidman, R. C. Levy, Y. J. Kaufman, D. Tanré, S. Mattoo, J. V. Martins, C. Ichoku, I. Koren, and H. Yu (2008). Global aerosol climatology from the MODIS satellite sensors. *Journal of Geophysical Research 113*(D14), D14S07.

Remer, L. A., S. Mattoo, R. C. Levy, and L. Munchak (2013). MODIS 3 km aerosol product: algorithm and global perspective. *Atmospheric Measurement Techniques 6*, 1829–1844.

Remer, L. A., D. Tanré, Y. J. Kaufman, R. Levy, and S. Mattoo (2006). Algorithm for remote sensing of tropospheric aerosol from MODIS: Collection 005. Technical report, National Aeronautics and Space Administration, Goddard Space Flight Center.

Retalis, A., D. G. Hadjimitsis, S. Michaelides, F. Tymvios, N. Chrysoulakis, C. R. I. Clayton, and K. Themistocleous (2010). Comparison of aerosol optical thickness with in situ visibility data over Cyprus. *Natural Hazards and Earth Systems Science 10*, 421–428.

Richter, R. (1996a). A spatially adaptive fast atmospheric correction algorithm. *International Journal of Remote Sensing 17*(6), 1201–1214.

Richter, R. (1996b). Atmospheric correction of satellite data with haze removal including a haze/clear transition region. *Computers and Geosciences 22*(6), 675–681.

Richter, R. (2004). ATCOR: Atmospheric and Topographic Correction software. Technical report, DLR-German Aerospace Center. Remote Sensing Data Center, Munich.

Richter, R. (2012). The Haze-Algorithm in ATCOR for ERDAS Imagine. Technical report, Geosystems White Paper, Munich.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C. J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, and M. Bosilovich (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society 85*(3), 381–394.

Roy, D. P., J. Ju, I. Kommareddy, M. Hansen, E. Vermote, C. Zhang, and A. Kommareddy (2011). Web Enabled Landsat Data (WELD) Products - Algorithm Theoretical Basis Document. Technical Report WELD-ATBD-v1.0, National Aeronautics and Space Administration, Goddard Space Flight Center.

Rozanov, A., V. Rozanov, M. Buchwitz, A. Kokhanovsky, and J. Burrows (2005). SCIATRAN 2.0 – A new radiative transfer model for geophysical applications in the 175–2400nm spectral region. *Advances in Space Research 36*(5), 1015–1019.

Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. Reeves, and H. Hashimoto (2004). A continuous satellite-derived measure of global terrestrial primary production. *Bioscience 54*(6), 547–560.

Ryder, C. L., E. J. Highwood, P. D. Rosenberg, J. Trembath, J. K. Brooke, M. Bart, A. Dean, J. Crosier, J. Dorsey, H. Brindley, J. Banks, J. H. Marsham, J. B. McQuaid, H. Sodemann, and R. Washington (2013). Optical properties of Saharan dust aerosol and contribution from the coarse mode as measured during the Fennec 2011 aircraft campaign. *Atmospheric Chemistry and Physics 13*(1), 303–325.

Saleska, S. R., K. Didan, A. R. Huete, and H. R. Da Rocha (2007). Amazon forests green-up during 2005 drought. *Science 318*(5850), 612–612.

Samanta, A., S. Ganguly, H. Hashimoto, S. Devadiga, E. Vermote, Y. Knyazikhin, R. R. Nemani, and R. B. Myneni (2010). Amazon forests did not green-up during the 2005 drought. *Geophysical Research Letters 37*(5), L05401.

Santer, R., V. Carrere, P. Dubuisson, and J. C. Roger (1999). Atmospheric correction over land for MERIS. *International Journal of Remote Sensing 20*(9), 1819–1840.

Santer, R., D. Ramon, J. Vidot, and E. Dilligeard (2007). A surface reflectance model for aerosol remote sensing over land. *International Journal of Remote Sensing 28*(3-4), 737–760.

Scaramuzza, P. L., B. L. Markham, J. A. Barsi, and E. Kaita (2004). Landsat-7 ETM+ on-orbit reflective-band radiometric characterization. *IEEE Transactions on Geoscience and Remote Sensing 42*(12), 2796–2809.

Schaap, M., A. Apituley, R. M. A. Timmermans, R. B. A. Koelemeijer, and G. de Leeuw (2009). Exploring the relation between aerosol optical depth and PM2.5 at Cabauw, the Netherlands. *Atmospheric Chemistry and Physics 9*(3), 909–925.

Schaepman-Strub, G., M. Schaepman, T. Painter, S. Dangel, and J. Martonchik (2006). Reflectance quantities in optical remote sensingDefinitions and case studies. *Remote Sensing of Environment 103*(1), 27–42.

Schepanski, K., I. Tegen, B. Laurent, B. Heinold, and A. Macke (2007). A new Saharan dust source activation frequency map derived from MSG-SEVIRI IR-channels. *Geophysical Research Letters 34*(18), L18803.

Schläpfer, D. (2001). MODO: an interface to MODTRAN for the simulation of imaging spectrometry at-sensor signals. In *10th JPL Airborne Earth Science Workshop. JPL, Pasadena (CA), Vol. Publication*, pp. 02–1.

Schmid, B., C. Matzler, A. Heimo, and N. Kampfer (1997). Retrieval of optical depth and particle size distribution of tropospheric and stratospheric aerosols by means of Sun photometry. *IEEE Transactions on Geoscience and Remote Sensing 35*(1), 172 –182.

Schmid, B. and C. Wehrli (1995). Comparison of Sun photometer calibration by use of the Langley technique and the standard lamp. *Applied Optics 34*(21), 4500–4512.

Schott, J. R., C. Salvaggio, and W. J. Volchok (1988). Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment 26*(1), 1–16.

Seaton, A., D. Godden, W. MacNee, and K. Donaldson (1995). Particulate air pollution and acute health effects. *The Lancet 345*(8943), 176–178.

Segura, S., V. Estellés, M. Pilar Utrillas, A. Raquel Esteve, and J. Antonio Martínez-Lozano (2013). Proposal of a simple model for the characterization of aerosols in relation to the dominant air masses. *International Journal of Remote Sensing 34*(9), 3625–3635.

Seidel, F. C., A. A. Kokhanovsky, and M. E. Schaepman (2010). Fast and simple model for atmospheric radiative transfer. *Atmospheric Measurement Techniques 3*(4), 1129–1141.

Shaw, G. E. (1983). Sun Photometry. *Bulletin of the American Meteorological Society 64*(1), 4–10.

Shih, F. Y. and S. Cheng (2004). Adaptive mathematical morphology for edge linking. *Information Sciences 167*(1–4), 9–21.

Siddans, R., C. Poulsen, B. J. Kerridge, G. Thomas, S. M. Dean, E. Carboni, and R. G. Grainger (2007). Algorithm Theoretical Basis Document for the creation of a Merged Global Aerosol Product. Technical Report GlobAerosol-ATBD-v2.1, Rutherford Appleton Laboratory, Harwell, UK.

Sifakis, N. and P. Deschamps (1992). Mapping of air pollution using SPOT satellite data. *Photogrammetric Engineering and Remote Sensing 58*, 1433–1433.

Sifakis, N. I. and C. Iossifidis (2014). CHRISTINE Code for High Resolution Satellite Mapping of Optical Thickness and Ångstrom Exponent PART I: Algorithm and Code. *Computers and Geosciences*.

Silva, D. (1992). Determination of'tasseled cap' transformation parameters for images obtained by the SPOT satellite. In *24th International Symposium on Remote Sensing of Environment, Rio de Janeiro, Brazil*, pp. 291–300.

Skamarock, W. C., M. L. Weisman, J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers (2005). A description of the Advanced Research WRF version 2. *NCAR Tech. Note, NCAR/TN-468+ STR*.

Slater, P. N. (1980). *Remote sensing: optics and optical systems*. Addison-Wesley Publishing Co., Inc., Advanced Book Program, Reading, MA. 575pp.

Smirnov, A., B. Holben, T. Eck, O. Dubovik, and I. Slutsker (2000). Cloud-Screening and Quality Control Algorithms for the AERONET Database. *Remote Sensing of Environment 73*(3), 337–349.

Smirnov, A., B. N. Holben, O. Dubovik, N. T. O'Neill, L. A. Remer, T. F. Eck, I. Slutsker, and D. Savoie (2000). Measurement of atmospheric optical parameters on U.S. Atlantic coast sites, ships, and Bermuda during TARFOX. *Journal of Geophysical Research 105*(D8), 9887–9901.

Smirnov, A., B. N. Holben, T. F. Eck, I. Slutsker, B. Chatenet, and R. T. Pinker (2002). Diurnal variability of aerosol optical depth observed at AERONET (Aerosol Robotic Network) sites. *Geophysical Research Letters 29*(23), 2115.

Smith, G. M. and E. J. Milton (1999). The use of the empirical line method to calibrate remotely sensed data to reflectance. *International Journal of Remote Sensing 20*(13), 2653–2662.

So, C. K., C. M. Cheng, and K. C. Tsui (2005). Weather and environmental monitoring using MODIS AOD data in Hong Kong, China. In *Proceedings of the First International Symposium on Cloud-prone and Rainy Areas Remote Sensing, The Chinese University of Hong Kong*, Volume 6, pp. 50–57.

Song, C., C. E. Woodcock, K. C. Seto, M. P. Lenney, and S. A. Macomber (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment 75*(2), 230–244.

Spurr, R. (2008). LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. *Light Scattering Reviews 3*, 229–275.

Steven, M., T. Malthus, F. Baret, H. Xu, and M. Chopping (2003). Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment 88*(4), 412–422.

Stevens, B. and O. Boucher (2012). Climate science: The aerosol effect. *Nature 490*(7418), 40–41.

Storey, J. C. (2001). Landsat 7 on-orbit modulation transfer function estimation. In *Proceedings of SPIE Sensors, Systems and Next Generation Satellites, Toulouse, France, 17th September*, Volume 4540, pp. 50–61.

Stowe, L. L., A. M. Ignatov, and R. R. Singh (1997). Development, validation, and potential enhancements to the second-generation operational aerosol product at the National Environmental Satellite, Data, and Information Service of the National Oceanic and Atmospheric Administration. *Journal of Geophysical Research 102*(D14), 16923–16,934.

Strahler, A. H., C. E. Woodcock, and J. A. Smith (1986). On the nature of models in remote sensing. *Remote Sensing of Environment 20*(2), 121–139.

Swap, R. J., H. J. Annegarn, J. T. Suttles, M. D. King, S. Platnick, J. L. Privette, and R. J. Scholes (2003). Africa burning: A thematic analysis of the Southern African Regional Science Initiative (SAFARI 2000). *Journal of Geophysical Research 108*(D13), 8465.

Tang, J., Y. Xue, T. Yu, and Y. Guan (2005). Aerosol optical thickness determination by exploiting the synergy of Terra and Aqua MODIS. *Remote Sensing of Environment 94*(3), 327–334.

Tanré, D., P. Y. Deschamps, C. Devaux, and M. Herman (1988). Estimation of Saharan aerosol optical thickness from blurring effects in thematic mapper data. *Journal of Geophysical Research 93*(D12), 15955–15,964.

Tao, J., M. Zhang, L. Chen, Z. Wang, L. Su, C. Ge, X. Han, and M. Zou (2012). A method to estimate concentrations of surface-level particulate matter using satellite-based aerosol optical thickness. *SCIENCE CHINA Earth Sciences*, 1–12.

Tao, M., L. Chen, X. Xiong, M. Zhang, P. Ma, J. Tao, and Z. Wang (2014). Formation process of the widespread extreme haze pollution over northern China in January 2013: Implications for regional air quality and climate. *Atmospheric Environment 98*, 417–425.

Tegen, I. and I. Fung (1994). Modeling of mineral dust in the atmosphere: Sources, transport, and optical thickness. *Journal of Geophysical Research 99*(D11), 22897–22,914.

Themistocleous, K. and D. G. Hadjimitsis (2013). Development of an image based integrated method for determining and mapping aerosol optical thickness (AOT) over urban areas using the darkest pixel atmospheric correction method, RT equation and GIS: A case study of the Limassol area in Cyprus. *ISPRS Journal of Photogrammetry and Remote Sensing 86*, 1–10.

Themistocleous, K., D. G. Hadjimitsis, A. Retalis, and N. Chrysoulakis (2012). The development of air quality indices through image-retrieved AOT and PM10 measurements in Limassol Cyprus. Volume 8534, pp. 85340B–85340B–13.

Thermo Scientific (2009). Filter Dynamics Measurement System, FDMS 8500 Product Specifications. Technical report, Thermo Scientific, Franklin, MA, USA.

Thomas, C., X. Briottet, and R. Santer (2011). OSIS: remote sensing code for estimating aerosol optical properties in urban areas from very high spatial resolution images. *Applied Optics 50*(28), 5408–5421.

Thomas, G., C. Poulsen, R. Siddans, E. Carboni, A. Sayer, and D. Grainger (2010). The GlobAEROSOL dataset: Using a multi-instrument satellite aerosol dataset. In *European Geophysical Union General Assembly*, Volume 12, pp. 11081.

Toledano, C., Y. Bennouna, V. Cachorro, J. Ortiz de Galisteo, A. Stohl, K. Stebel, N. Kristiansen, F. Olmo, H. Lyamani, M. Obregón, V. Estellés, F. Wagner, J. Baldasano, Y. González-Castanedo, L. Clarisse, and A. de Frutos (2012). Aerosol properties of the Eyjafjallajökull ash derived from sun photometer and satellite observations over the Iberian Peninsula. *Atmospheric Environment 48*, 22–32.

Torres, O., P. K. Bhartia, J. R. Herman, Z. Ahmad, and J. Gleason (1998). Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis. *Journal of Geophysical Research 103*(D14), 17099–17,110.

Torres, O., A. Tanskanen, B. Veihelmann, C. Ahn, R. Braak, P. K. Bhartia, P. Veefkind, and P. Levelt (2007). Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview. *Journal of Geophysical Research 112*(D24), D24S47.

Tregoning, P., R. Boers, D. O'Brien, and M. Hendy (1998). Accuracy of absolute precipitable water vapor estimates from GPS observations. *Journal of Geophysical Research 103*(D22), 28701–28710.

Twomey, S. (1974). Pollution and the planetary albedo. *Atmospheric Environment 8*(12), 1251–1256.

UK Met Office (2006). MIDAS Land Surface Stations data (1853-current) at NCAS British Atmospheric Data Centre. http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_ukmo-midas.

Uno, I., N. Sugimoto, A. Shimizu, K. Yumimoto, Y. Hara, and Z. Wang (2014). Record heavy PM2.5 air pollution over China in January 2013: vertical and horizontal dimensions. *Sola 10*, 136–140.

van der Walt, S., J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu (2014). scikit-image: image processing in Python. *PeerJ 2*(e453).

van Donkelaar, A., R. V. Martin, M. Brauer, R. Kahn, R. Levy, C. Verduzco, and P. J. Villeneuve (2010). Global estimates of ambient fine particulate matter concentrations from satellite-based Aerosol Optical Depth: development and application. *Environmental Health Perspectives 118*(6), 847–855.

van Heuklon, T. K. (1979). Estimating atmospheric ozone for solar radiation models. *Solar Energy 22*(1), 63–68.

Vandewalle, P., J. Kovacevic, and M. Vetterli (2009). Reproducible research in signal processing. *IEEE Signal Processing Magazine 26*(3), 37–47.

Vardoulakis, S., B. E. A. Fisher, K. Pericleous, and N. Gonzalez-Flesca (2003). Modelling air quality in street canyons: a review. *Atmospheric Environment 37*(2), 155–182.

Veefkind, J. P., G. d. Leeuw, P. A. Durkee, P. B. Russell, P. V. Hobbs, and J. M. Livingston (1999). Aerosol optical depth retrieval using ATSR-2 and AVHRR data during TARFOX. *Journal of Geophysical Research 104*(D2), 2253–2260.

Vermote, E., N. El Saleous, C. Justice, Y. Kaufman, J. Privette, L. Remer, J. Roger, and D. Tanré (1997). Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation. *NASA Publications*, 31.

Vermote, E. and A. Vermeulen (1999). Atmospheric correction algorithm: spectral reflectances (MOD09), Algorithm Theoretical Basis Document. Technical report, University of Maryland.

Vermote, E. F., D. Tanré, J. L. Davis, M. Herman, and J. J. Morcette (1997). Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: an overview. *IEEE Transactions on Geoscience and Remote Sensing 35*(3), 675 –686.

Wainwright, M. (2010). Volcanic ash cloud cancels flights across Britain. *The Guardian, 16th May 2010*.

Walt, S., S. C. Colbert, and G. Varoquaux (2011). The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering 13*(2), 22–30.

Wang, J., L. Zhang, A. Dai, T. Van Hove, J. Van Baelen, *et al.* (2007). A near-global, 2-hourly data set of atmospheric precipitable water from ground-based GPS measurements. *Journal of Geophysical Research 112*(D11), D11–107.

Wang, K., S. E. Franklin, X. Guo, Y. He, and G. J. McDermid (2009). Problems in Remote Sensing of Landscapes and Habitats. *Progress in Physical Geography 33*(6), 747–768.

Wang, Q., Y. Zha, J. Gao, and D. Shen (2013). Estimation of atmospheric particulate matter based on MODIS haze optimized transformation. *International Journal of Remote Sensing 34*(5), 1855–1865.

Wang, Y., X. Zhang, and R. R. Draxler (2009). TrajStat: GIS-based software that uses various trajectory statistical analysis methods to identify potential sources from long-term air pollution measurement data. *Environmental Modelling & Software 24*(8), 938–939.

Wanner, W., X. Li, and A. H. Strahler (1995). On the derivation of kernels for kernel-driven models of bidirectional reflectance. *Journal of Geophysical Research 100*(D10), 21077–21089.

Wehrli, C. (1985). WRC Reference Spectrum. World Radiation Center Publication 615, World Radiation Center, Davos, Switzerland.

Wielicki, B. A., T. Wong, N. Loeb, P. Minnis, K. Priestley, and R. Kandel (2005). Changes in Earth's albedo measured by satellite. *Science 308*(5723), 825–825.

Wilson, R. T. (2012). Py6S: A Python interface to the 6S Radiative Transfer Model. *Computers and Geosciences 51*, 166–171.

Wilson, R. T. and E. J. Milton (2010). Automated Selection of Suitable Atmospheric Correction Sites. In *Annual Conference of the Remote Sensing and Photogrammetry Society. (RSPSoc 2010)*, Cork, Ireland.

Wilson, R. T., E. J. Milton, and J. M. Nield (2014). Spatial variability of the atmosphere over southern England, and its effect on scene-based atmospheric corrections. *International Journal of Remote Sensing 35*(13), 5198–5218.

Wiscombe, W. J. (1980). Improved Mie scattering algorithms. *Applied Optics 19*(9), 1505–1509.

Wolfe, D. E. and S. I. Gutman (2000). Developing an operational, surface-based, GPS, water vapor observing system for NOAA: Network design and results. *Journal of Atmospheric and Oceanic Technology 17*(4), 426–440.

Wolfe, R., J. Masek, N. Saleous, and F. Hall (2004). LEDAPS: mapping North American disturbance from the Landsat record. In *IEEE International Geoscience and Remote Sensing Symposium*, Volume 1.

Wong, M. S., M. I. Shahzad, J. E. Nichol, K. H. Lee, and P. Chan (2013). Validation of MODIS, MISR, OMI, and CALIPSO aerosol optical thickness using ground-based sunphotometers in Hong Kong. *International Journal of Remote Sensing 34*(3), 897–918.

Wood, A. T. A. and G. Chan (1994). Simulation of Stationary Gaussian Processes in [0, 1] d. *Journal of Computational and Graphical Statistics 3*(4), 409–432.

Xu, Q., Z. Obradovic, B. Han, Y. Li, A. Braverman, and S. Vucetic (2005). Improving aerosol retrieval accuracy by integrating AERONET, MISR and MODIS data. In *2005 8th International Conference on Information Fusion*, Volume 1, pp. 7.

Yan, G., L. Jiang, J. Wang, L. Chen, and X. Li (2003). Thermal bidirectional gap probability model for row crop canopies and validation. *Science in China Series D: Earth Sciences 46*(12), 1241–1249.

Zha, Y., J. Gao, J. Jiang, H. Lu, and J. Huang (2012). Normalized difference haze index: a new spectral index for monitoring urban air pollution. *International Journal of Remote Sensing 33*(1), 309–321.

Zhang, J. and S. A. Christopher (2003). Longwave radiative forcing of Saharan dust aerosols estimated from MODIS, MISR, and CERES observations on Terra. *Geophysical Research Letters 30*(23), 2188.

Zhang, Y. and B. Guindon (2003). Quantitative assessment of a haze suppression methodology for satellite imagery: effect on land cover classification performance. *IEEE Transactions on Geoscience and Remote Sensing 41*(5), 1082 – 1089.

Zhang, Y., B. Guindon, and J. Cihlar (2002a). An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. *Remote Sensing of Environment 82*(2–3), 173–187.

Zhang, Y., B. Guindon, and J. Cihlar (2002b). Development of a robust haze removal algorithm: assessment using temporally invariant targets. In *IEEE International Geoscience and Remote Sensing Symposium*, Volume 1, pp. 101 – 103.

Zhong, B. (2011). Improved estimation of aerosol optical depth from Landsat TM/ETM+ imagery over land. In *IEEE International Geoscience and Remote Sensing Symposium*, pp. 3304–3307.

Zhu, Z. and C. E. Woodcock (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment 118*, 83–94.

Zubko, V., G. Leptoukh, and A. Gopalan (2010). Study of Data-Merging and Interpolation Methods for Use in an Interactive Online Analysis System: MODIS Terra and Aqua Daily Aerosol Case. *IEEE Transactions on Geoscience and Remote Sensing 48*(12), 4219 –4235.