

**Application of Stereo-Photogrammetric
Methods to the Advanced Along Track
Scanning Radiometer for the Atmospheric
Sciences**

Presented by:

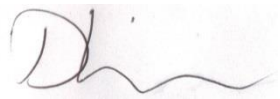
DANIEL FISHER

For the award of:

DOCTOR OF PHILOSOPHY

UCL

I, Daniel Fisher, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

A handwritten signature in black ink, appearing to read 'D. Fisher', with a wavy line underneath.

Daniel Fisher 21st October 2013

Abstract

This thesis studies photogrammetric techniques applied to the ATSR instruments for the extraction of atmospheric parameters with the objective of generating new scientific datasets. The atmospheric parameters under observation are cloud top height, smoke plume injection height, and tropospheric wind components. All have important applications in various tasks, including the initialisation and validation of climate models.

To generate accurate stereo measurements from the ATSR imagery the forward and nadir views need to be accurately co-registered. Currently this is not the case, with differences of up to 2 pixels in both axes recorded. In this thesis an automated image tie-pointing and image warping algorithm that improves ATSR co-registration to ≤ 1 pixel is presented. This thesis also identifies the census stereo matching algorithm for application to the ATSR instruments. When compared against a collocated DEM, census outperforms the previous stereo matching algorithm applied to the ATSR instrument, known as M4, significantly: RMSE $\sim 700\text{m}$ vs. $\sim 1200\text{m}$; bias $\sim 60\text{m}$ vs $\sim 600\text{m}$; $R^2 \sim 0.9$ vs ~ 0.7 . Furthermore, this thesis reviews the M6 algorithm developed for application within the ESA ALANIS Smoke Plume project.

Using census a climatological cloud fraction by altitude dataset over Greenland is generated and demonstrated to agree well with current observational datasets from MISR, MODIS and AATSR. The $11\mu\text{m}$ channel stereo output provides insights into high cloud characteristics over Greenland and appears to be, in comparison with CALIOP, practically unbiased. The ALANIS Smoke plume project is introduced and the inter-comparison of the M6 algorithm against MISR and CALIOP is presented. M6 demonstrates some ability for determining smoke plumes injection heights above 1km in elevation. However, the smoke plume masking approach currently employed is demonstrated to be lacking in quality. Finally, this thesis presents the determination of cloud tracked tropospheric winds from the ATSR2-AATSR tandem operation using the Farneback optical flow algorithm. This algorithm offers accuracy on the order of 0.5 ms^{-1} at full image resolution, which is unprecedented in comparison to similarly derived datasets.

Contents

1 Introduction.....	14
1.1 Motivation.....	14
1.2 Research Objectives.....	14
1.3 Outline.....	17
2 ATSR	18
2.1 Introduction.....	18
2.2 The Along Track Scanning Radiometer.....	19
2.3 Forward to Nadir Co-registration.....	21
2.3.1 ATSR Co-registration Problem.....	21
2.3.2 Co-registration Correction Methodology.....	22
2.3.3 Transformation Coefficients	29
2.3.4 Global Evaluation	33
2.3.5 Results and Analysis	36
2.3.6 Discussion.....	40
2.4 Chapter Summary	43
3 Stereo-Photogrammetry	44
3.1 Introduction.....	44
3.2 Stereo Photogrammetric Concepts.....	45
3.2.1 The Imaging System	45
3.2.2 The Stereo Matching Algorithm	47
3.2.3 The Camera Model.....	51
3.3 Stereo Matching Challenges	53
3.3.1 Stereo Image Dissimilarity.....	53
3.3.2 Real-world Algorithm Assessments.....	56
3.4 ATSR Stereo Matching Algorithms.....	57
3.4.1 M4 Algorithm	57
3.4.2 Census Algorithm	60
3.4.3 M6 Algorithm	61
3.5 Algorithm Assessment	64
3.5.1 Window Size Assessment	64
3.5.2 DEM Inter-comparison	66
3.6 Chapter Summary	70
4 Greenland Cloud Study.....	71
4.1 Introduction.....	71

4.1.1	Clouds and the Earth’s Climate System.....	71
4.1.2	Macrophysical Cloud Properties from EO.....	82
4.1.3	EO Macrophysical Cloud Climatologies	85
4.2	Scientific Aims and Objectives.....	92
4.3	Cloud Climatology Development	93
4.3.1	L2 Processing.....	93
4.3.2	Level 2 Validation.....	97
4.3.3	L2 Summary.....	104
4.3.4	L3 Processing.....	104
4.4	L3 Inter-Comparison Study	105
4.4.1	GEWEX Inter-comparison Datasets	105
4.4.2	CF Inter-comparison	107
4.4.3	CTH Inter-comparison.....	108
4.5	Summary.....	110
4.5.1	Objective Review.....	110
4.5.2	Future Work	114
4.5.3	Conclusions.....	114
5	ALANIS: Atmosphere Land Interaction Study	116
5.1	Introduction.....	116
5.2	Scientific Aims and Objectives.....	118
5.3	SPIH Determination.....	118
5.3.1	SPIH Product Generation.....	118
5.3.2	Processing Chain.....	119
5.4	SPIH Validation.....	120
5.4.1	CALIOP Validation Approach.....	120
5.4.2	CALIOP Inter-comparison Results	124
5.4.3	MISR Inter-comparison Approach.....	124
5.4.4	MISR Inter-comparison Results.....	127
5.4.5	Inter-comparison Discussion.....	128
5.5	SPM Determination	129
5.5.1	SPM Technique.....	129
5.5.2	SPM Evaluation Approach and Results	130
5.5.3	SPM Discussion	134
5.6	Summary and Future Work.....	134
5.6.1	Objective Review	134
5.6.2	Future Work	135

5.6.3 Chapter Summary	136
6 ATSR Tandem Atmospheric Motion Vectors.....	137
6.1 Introduction.....	137
6.2 Scientific Objectives	140
6.3 ATSR-2 AATSR AMV Extraction.....	140
6.3.1 Benefits of a Tandem System	140
6.3.2 Tandem AMV Retrieval Method	143
6.3.3 ATSR Tandem Wind Dataset	146
6.4 AMV Accuracy Assessment	146
6.5 Reanalysis Inter-comparison.....	147
6.5.1 MERRA Dataset	147
6.5.2 Inter-comparison Method.....	147
6.5.3 Results.....	148
6.5.4 Discussion	150
6.6 Summary and Future Work.....	150
6.6.1 Objective Review	150
6.6.2 Future Work	151
6.6.3 Chapter Summary	152
7 Concluding Remarks.....	153
7.1 Thesis Review	153
7.2 Contributions and Originality	155
7.2.1 Thesis Contributions	155
7.2.2 Peer Review Journal Papers	156
7.2.3 Conference Proceedings.....	156
7.3 Future Work.....	157

List of Figures

Figure 2-1. An image and cutaway of the ATSR instrument	18
Figure 2-2. ATSR viewing geometry	19
Figure 2-3. ATSR forward and nadir imagery	20
Figure 2-4. ATSR scene for tie-point derivation.....	21
Figure 2-5. SIFT feature detection process	23
Figure 2-6. Example SURF feature detection on an ATSR image.....	24
Figure 2-7. SURF matching threshold effect	25
Figure 2-8. Matched SURF features.....	26
Figure 2-9. Example tie-points from the tie-point selection process.....	28
Figure 2-10. 2008 Global co-registration evaluation process.....	34
Figure 2-11. Elevation screening process.....	35
Figure 2-12. ATSR co-registration errors by channel.....	37
Figure 2-13. ATSR co-registration errors by latitude.....	39
Figure 2-14. ATSR co-registration errors by longitude	41
Figure 3-1. The parallax effect	45
Figure 3-2. ATSR forward and nadir image differences	46
Figure 3-3. Azimuthal angle effects on parallax	48
Figure 3-4. Epipolar geometry	51
Figure 3-5. ATSR viewing geometry	52
Figure 3-6. Non-lambertian visible image effects	54
Figure 3-7. Thermal path length effects	54
Figure 3-8. M4 normalisation output	58
Figure 3-9. M6 normalisation output	62
Figure 3-10. DEM evaluation subset.....	64
Figure 3-11. Matching window size effects	65
Figure 3-12. 11 μm stereo assessment.....	67
Figure 3-13. 1.6 μm stereo assessment.....	68
Figure 3-14. 0.55 μm stereo assessment.....	68
Figure 4-1. Radiative interactions between the atmosphere, the ocean and the land surface	72

Figure 4-2. Climate radiative forcing agents	73
Figure 4-3. Radiation and its transmission through the atmosphere	74
Figure 4-4. Cloud radiative effects.....	78
Figure 4-5. L2 stereo outputs by channel	95
Figure 4-6. 10-year latitudinally averaged wind speed from MISR	96
Figure 4-7. AATSR CTH vs MISR.....	98
Figure 4-8. AATSR CTH vs CALIOP	101
Figure 4-9. 2008 latitudinally averaged CFbA for each AATSR channel.....	106
Figure 4-10. 2008 CF climatologies.....	107
Figure 4-11. 2008 CTH climatologies.....	108
Figure 4-12. Monthly mean 2008 CTH.....	109
Figure 4-13. May-August 2008 polar jet positions	113
Figure 5-1. Map of the ALANIS study region	117
Figure 5-2. AATSR observations and CALIOP paths	121
Figure 5-3. AATSR and CALIOP transects	122
Figure 5-4. AATSR SPIH vs. CALIOP	124
Figure 5-5. ALANIS mean reanalysis wind speeds	125
Figure 5-6. MISR plume height projects.....	126
Figure 5-7. AATSR SPIH vs. MISR	127
Figure 5-8. SPIH outputs.....	130
Figure 5-9. SPIH analysis 23 rd April 2010	131
Figure 5-10. SPIH analysis 23 rd April 2010.....	131
Figure 5-11. SPIH analysis 17 th May 2010	132
Figure 5-12. SPIH analysis 17 th May 2010	132
Figure 5-13. SPIH analysis 17 th July 2010	133
Figure 5-14. SPIH analysis 17 th July 2010.....	133
Figure 6-1. MERRA zonal and meridional latitudinal wind speed averages	141
Figure 6-2. ATSR2 AATSR tandem images.....	142
Figure 6-3. Farneback, SURF and census feature tracking outputs	143
Figure 6-4. Zero wind outputs.....	146

Figure 6-5. Zero wind histograms.	147
Figure 6-6. ATSR tandem vs MERRA AMVs.....	149
Figure 6-7. ATSR tandem vs MERRA zonal wind speed histograms.....	149
Figure 6-8 ATSR tandem vs MERRA meridional wind speed histograms..	150

List of Tables

Table 2-1. Uncorrected ATSR image co-registration statistics.	31
Table 2-2. Corrected ATSR image co-registration statistics	32
Table 3-1. Stereo matching algorithm quality statistics.	67
Table 4-1. AATSR vs CALIOP layer detection efficiencies and optical depths.....	102
Table A1. Co-registration warping coefficients.	175

Acknowledgements

I would like to thank my supervisor Prof. Jan-Peter Muller for his generosity, many useful suggestions, and excellent guidance provided during the course of this thesis. I would also like to thank my second supervisor Dr. Caroline Poulsen for providing useful advice and IDL code which got me up and running at the start of this process. Also, thanks to NERC, NCEO and ESA for providing the financial support that allowed this work to be undertaken.

Without the staff at the NEODC at RAL, much of the work carried out in this thesis would have been impossible to complete. Particular thanks go to Steve Donegan who provided me with much of the ATSR data employed in this thesis.

Furthermore I would like to thank Dr. Valdimir Yershov and Catherine Moroney for the useful discussions, comments and advice given over the last few years. Thanks also go to the ALANIS consortium members, working with whom was both a pleasure and very insightful.

Finally I would like to thank my family. My Dad, for always having an open ear, for the untold times he has given me support and advice, and for the many enjoyable real ale sessions when I have been home over the last four years. My Mum for always putting a smile on my face, filling my belly, and taking an interest in what I have been doing. My brother, James, for his selflessness in shuffling me around Surrey, despite it being a 250 mile round trip and for letting me stay over and make a mess in his kitchen without too much grumbling. My sisters Jess and Luce for keeping my feet on the ground and making me laugh, and Astrid, for her patience, understanding and support at the times when I have most needed it!

Acronyms

3D Three Dimensional

AATSR Advanced Along Track Scanning Radiometer

ALANIS Atmosphere Land Interaction Study

ALSC Adaptive Least Squares Correlation

API Application Programming Interface

AMV Atmospheric Motion Vector

ATSR Along Track Scanning Radiometer

ATSR-1 Along Track Scanning Radiometer 1

ATSR-2 Along Track Scanning Radiometer 2

BEAM Basic ERS and & Envisat AATSR and MERIS toolbox

BT Brightness Temperature

CALIOP Cloud-Aerosol Lidar with Orthogonal Polarization

CALIPSO Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations

CBH Cloud Base Height

CERES Clouds and the Earth's Radiant Energy System

CF Cloud Fraction

CFbA Cloud Fraction by Altitude

COVE CEOS Visualisation Environment

CPU Central Processing Unit

CRF Cloud Radiative Forcing

CTH Cloud Top Height

CTM Chemical Transport Model

CTP Cloud Top Pressure

CTT Cloud Top Temperature

DEM Digital Elevation Model

ECMWF European Centre for Medium Range Weather Forecasting

ECS Equilibrium Climate Sensitivity

EMS Electro-Magnetic Spectrum

EO Earth Observation

ERBE Earth Radiation Budget Experiment

ERS European Remote-Sensing Satellite

ERS-1 European Remote-Sensing Satellite 1

ERS-2 European Remote-Sensing Satellite 2

ESA European Space Agency

GCM General Circulation Model

GEWEX Global Energy and Water Exchanges Project

GMAO Global Modelling and Assimilation Office

GOES Geostationary Observational Environmental Satellite

GPU Graphics Processing Unit

IDL Interactive Data Language

IGBP International Geosphere-Biosphere Programme

IPCC International Panel on Climate Change

ISCCP International Cloud Climatology Project

IR Infrared

LW Long-wave

MERRA Modern-era Retrospective Analysis for Research and Applications

MINX MISR Interactive Explorer

MISR Multi-angle Imaging Spectro-Radiometer

MODIS Moderate Resolution Imaging Spectro-radiometer

MSSL Mullard Space Science Laboratory

NASA National Aeronautics and Space Administration

NCC Normalised Cross Correlation

NEODC National Earth Observation Data Centre

NIR Near-Infrared

NWP Numerical Weather Prediction

OD Optical Depth

ORAC Oxford RAL Aerosol Cloud

RMSD Root Mean Square Difference

RMSE Root Mean Square Error

RT Radiative Transfer

RTM Radiative Transfer Model

SAD Sum of the Absolute Differences

SADIST Synthesis of ATSR Data Into Sea Surface Temperature

SIFT Scale Invariant Feature Transform

SLSTR Sea and Land Surface Temperature Radiometer

SPIH Smoke Plume Injection Height

SPM Smoke Plume Mask

SSD Sum of the Square Differences

SST Sea Surface Temperature

SURF Speeded Up Robust Features

SW Short-Wave

SZN Solar Zenith Angle

TOA Top of Atmosphere

USGS United States Geological Service

UTC Coordinated Universal Time

WTA Winner Takes All

1 INTRODUCTION

1.1 Motivation

The ATSR instruments have been exploited variously throughout their period of operation from the launch of ATSR-1 on ERS-1 in 1991 through to the demise of AATSR on Envisat in 2012. The science applications have ranged from the determination of sea surface temperature at an unprecedented degree of accuracy (Handoll et al., 1991) to the mapping of wildfire occurrences (Arino et al., 2001). One area where potential has been demonstrated (Lorenz, 1985; Prata and Turner, 1997; Cawkwell et al., 2001; Muller et al., 2007), but left somewhat unfulfilled, is in the application of stereo-photogrammetric techniques to the ATSR imaging systems to determine cloud top height. Stereo-photogrammetry is the art of determining the 3D location of a feature within an image. The specific requirements are few: two or more observations of the feature (e.g. a cloud) from differing positions, an approach to locate and measure the distance between the feature across the images, and a model of the observing system which enables conversion of the observed distance into some real-world measurement (in the case of the previous ATSR studies, the height of the cloud above an Earth ellipsoid). Despite the relative simplicity of the approach and demonstrations of its potential for application to ATSR and the generation of important long-term climatological datasets in other instruments (e.g. Moroney et al., 2002 and Di Girolamo et al., 2010, generating stereo derived cloud top height climatologies from MISR), as of yet, no long-term stereo derived cloud top height datasets obtained from ATSR exist in the scientific domain. The motivation behind this thesis is to rectify this current situation, generating the first long-term stereo cloud parameter climatology from the ATSR instruments, and also to demonstrate other potential applications of the stereo-photogrammetric approach in combination with the ATSR instruments.

1.2 Research Objectives

The overall aim of the thesis is to evaluate the use of stereo image matching algorithms in combination with the ATSR instruments for various science dataset generation tasks for application within the atmospheric sciences. The work undertaken can be divided into two main areas: one focusing on the algorithms and the other on the applications. The aim of the first area is to identify or develop stereo image matching algorithms which perform effectively on the ATSR instruments for the retrieval of atmospheric features. As mentioned, there has been some previous work on the development of stereo

matching algorithms for application to ATSR-1 and ATSR-2 (Lorenz, 1985; Prata and Turner, 1997; Cawkwell et al., 2001; Muller et al., 2007), however the justifications behind the algorithm selected in these studies are mostly limited. This may in part have been due to a lack of available research on assessing the application of stereo image matching algorithms to real-world image matching tasks, perhaps caused by the inherent computational cost of stereo matching and the limited computational facilities available in the 1990s and early 2000s. These limitations were (and continue to be) further compounded by the predisposition of the imaging science community to focus on controlled, laboratory evaluations, of stereo matching algorithms (Scharstein and Szeliski, 2002). Such assessments do not give the necessary insights into real-world algorithm performance, and as such do not allow for effective and justifiable algorithm selection for application to real-world stereo image matching tasks. Fortunately, recent research undertaken by Hirschmuller and Scharstein (2009) has filled in this missing link between the laboratory and the real-world, providing extensive and in-depth evaluations of number of widely employed stereo image matching algorithms on imagery with the radiometric distortions commonly introduced during imaging in uncontrolled real-world environments. This important research enables informed and justifiable decisions to be made on algorithm selection. One of the research aims of this thesis therefore, is to apply the findings of Hirschmuller and Scharstein (2008) to the ATSR instruments in order to identify an improved image stereo matching algorithm or to justify the use of the techniques applied previously.

Following the identification of an improved stereo-matching algorithm or justification of the selection of an algorithm previously employed on the ATSR instruments the second aim of the thesis is both to build on existing, and to generate new atmospheric datasets from the ATSR instruments through application of stereo-photogrammetric techniques. Previously identified avenues for atmospheric datasets derived using stereo-photogrammetric approaches applied to the ATSR instruments include: the determination of cloud top height and wind from ATSR-1 (Lorenz, 1985; Prata and Turner, 1997); the determination of cloud top height from ATSR-2 (Muller et al., 2007); multi-layer cloud detection with ATSR-2 (Naud et al., 2007); and the potential for effective cloud top height and amount determination in polar regions (Cawkwell et al., 2001). The evident focus on clouds is in part due to the specific ATSR instrument characteristics, which, in combination with stereo-photogrammetric approaches, give excellent cloud elevation assignment ability, and also in part due to the significant role of clouds in influencing the Earth's radiation budget and the difficulty in their quantification and characterisation, leading to significant uncertainties in their overall impacts on the Earth's climate. The most recent report from the international panel on climate change (IPCC), assessment report five (AR5), predicts with a high level of confidence that a warming of between 1.5K and 4.5K would be required to restore the Earth to a state of climatic equilibrium (so that the amount of energy entering the Earth's climate system is the same as that leaving) in response to doubling of atmospheric CO₂ (Stocker et al., 2013). The prediction uncertainty primarily arises in the feedback processes, the internal processes of the climate system that either amplify or dampen the effect of a climate forcing (Webb et al., 2006; Randall et al., 2007). Of the climate feedback processes, the greatest uncertainty is associated with those of clouds. In the 4th IPCC report (Solomon et al., 2007) it was demonstrated that in the presence of three of the most critical feedback processes, that of water vapour, lapse rate and surface albedo feedbacks, the general circulation models analysed predicted

warming in the range of $1.9 \text{ K} \pm 0.15 \text{ K}$ to restore climate equilibrium following a doubling of atmospheric CO_2 . When cloud feedbacks were incorporated into the same models the mean warming increased significantly as did the standard deviation to 3.2 K and 0.7 K respectively. This indicates a strong disagreement on the radiative impact of clouds on the Earth's climate system. Reducing this uncertainty has been identified by the IPCC as the key requirement for improving consensus between climate projections and therefore, gaining a better understanding of the future state of the climate (Randall et al., 2007). This topic is explored in far greater detail in chapter 4.

Despite the cloud detection and height assignment capabilities of the ATSR instruments and the critical importance of improving understanding about the role of clouds in the climate system, none of the stereo-photogrammetric studies undertaken with the ATSR instruments have been translated into the development of long-term cloud climatological datasets. This lack of progression may be due to other, typically radiometric (whereas stereo-image matching is geometric), approaches for the retrieval of cloud properties generally experiencing greater favour within the scientific community (e.g. Fischer and Grassl, 1991; Rossow and Gardner, 1993; Watts 1998; Wylie and Menzel, 1999; Menzel et al., 2008; Minnis et al., 2011; Poulsen et al., 2012). However, stereo image matching derived cloud datasets from the ATSR instruments have much potential to contribute to the current understanding, by providing new and valuable insights into cloud characteristics. This then leads to another research aim of this thesis, to generate the first long-term macrophysical (cloud and height and amount) cloud dataset from the ATSR instruments and to define the potential contributions which such a dataset may provide to the current suite of cloud climatologies.

This final research objective of this thesis is to exploit ATSR data for the generation of new scientific datasets. One such dataset (Fisher et al., 2013), conceived externally to this thesis, is the generation of stereo-derived smoke plume injection height observations from AATSR. This new dataset was created as part of the ESA Atmosphere Land Interaction Study (ALANIS) Smoke plumes project. This project aimed to improve the prediction accuracy of the TM5 CTM (Krol et al., 2005), through assimilation of a number of novel observational datasets, which included stereo derived smoke plume injection height (SPIH) estimates, over the Eurasian boreal regions from the Advanced-ATSR instrument. The author's main undertaking within the ALANIS project was the validation of the stereo-derived injection heights, and as such, that work forms the major part of the report given in this thesis. Another avenue for the exploitation of ATSR data is offered by the tandem operation of ATSR-2 and AATSR, where the two satellites flew in a shared orbit separated by approximately 30 minutes. This tandem operation offers the potential to extract wind motion and elevation to a high degree of accuracy through the tracking of cloud features across the ATSR imagery. Such observations of tropospheric winds are of significant importance in defining the current and predicting the future state of the atmosphere in modelling processes (e.g. Bormann and Thépaut, 2004).

1.3 Outline

There are a total of seven chapters in this thesis, which are divided into two sections. The first section, comprising Chapter 2 and 3, describes the ATSR instruments and the evaluation of stereo-photogrammetric algorithms for application to them. Chapter 2 in addition to describing each ATSR instrument also introduces a technique to correct for a co-registration problem between the nadir and forward views which is common to all three sensors. Correction for co-registration errors is vital for effective stereo image matching, the basic principles of which are described in Chapter 3. Furthermore, Chapter 3 also describes the particular challenges present in real world stereo matching scenarios, and the particular techniques and approaches which are employed to overcome these challenges. This discussion, along with the finding from Hirschmuller and Scharstein (2009), is used to identify a robust stereo matching algorithm suitable for application to the ATSR instruments.

Chapter 4, 5 and 6, which comprise the second section, document the exploitation of the work undertaken in Chapters 2 and 3 for the generation of new atmospheric datasets. Chapter 4 focuses on the development and validation of a stereo derived cloud climatology over Greenland from the entire AATSR time-series. This chapter describes the importance of observational cloud climatologies and includes extensive validation of the derived L2 AATSR stereo products against collocated, higher accuracy, CTH observations obtained from the stereo capable MISR instrument and the CALIOP lidar. The L2 products are formed into stratified cloud fraction by altitude products, which are inter-compared against other observational cloud climatologies derived from various passive imaging systems. Chapter 5 documents the validation effort of the smoke plume injection heights derived from AATSR for the ALANIS Smoke plume project. As with Chapter 4, the validation is undertaken against higher resolution injection height observations from MISR and CALIOP. In Chapter 6 the ATSR-2 AATSR tandem winds assessment is presented. Various algorithms are evaluated for cloud feature tracking, and an optical flow algorithm which is capable of providing pixel resolution wind estimates is identified. Assessment of the identified algorithm in terms of zero wind retrieval ability (i.e. ground shift analysis), and inter-comparisons of the outputs against reanalysis wind profiles are undertaken. Finally, Chapter 7 reviews the major findings of the thesis and presents some recommendations for future work.

2 ATSR

2.1 Introduction

The ATSR instruments, with the original mission goal of obtaining measurements of sea surface temperature to an unprecedented degree of accuracy (Handoll et al., 1991), have been applied to numerous additional scientific tasks throughout their period of operation. Examples of such tasks include, but are not limited to: fire event mapping (Arino et al., 2001), aerosol retrieval (Thomas et al., 2010), land surface temperature measurement (Ghent and Remedios, 2013), and, as is the focus of this thesis, applications of a stereo-photogrammetric nature (Lorenz 1985; Prata and Turner 1997; Denis et al. 2007; Muller et al. 2007; Fisher et al., 2013). The ATSR instruments have a time-series which extends from 1991 through to 2012, giving more than 20 years of observational data, which makes them an excellent platform for climatological observations. Furthermore, a future mission based on ATSR, the Sea and Land Surface Temperature Radiometer (SLSTR), due for launch in 2014, will build upon the already extensive dataset. Given their varied application and soon to be continued period of operation, the scientific importance of these instruments is evident.

FIGURE REMOVED DUE TO THIRD PARTY CONFLICTS

Figure 2-1. ATSR instrument image and cutaway. Adapted from Llewellyn-Jones et al. (2001)

In this chapter, the general and specific characteristics of the ATSR instruments will be introduced. Following this introduction, the other main focus of this chapter, improving the co-registration between the forward and nadir views of the ATSR instruments will be addressed. Poor co-registration between the ATSR views is one of the main limitations of the instrument series for stereo-photogrammetric and other full resolution dual view scientific applications. Therefore, whilst this work is carried out under the premise of providing more accurate stereo observations, it ultimately has consequences for any study which necessitates full resolution dual view imaging capabilities. The ATSR co-registration methods and results presented in this chapter are also presented in a more concise fashion in Fisher and Muller (2013).

2.2 The Along Track Scanning Radiometer

The ATSR instruments (see Figure 2-1) acquired, near continuously, remotely sensed data of the Earth from the launch of ATSR-1 (Edwards et al., 1990; Mutlow et al., 1999) in 1991 on-board the ESA Earth Remote Sensing Satellite (ERS-1), through to 2012 when the Advanced-ATSR (Llewellyn-Jones et al., 2001, AATSR) stopped sending data due to the failure of the Envisat satellite. Amid ATSR-1 (1991-2000) and AATSR (2002-2012), ATSR-2 (Mutlow et al., 1999) provided data from its launch in 1995 through to 2008, and was carried on-board the ERS-2 satellite. The ATSR time-series therefore covers more than 20 years of observations, and represents one of the longest-term observational satellite datasets.

FIGURE REMOVED DUE TO THIRD PARTY CONFLICTS

Figure 2-2. The ATSR viewing geometry. Note that due to the conical scanning mechanism the resolution varies across the image cone angle. Adapted from Llewellyn-Jones et al. (2001).

The three instruments share a common imaging geometry, comprised of a dual-view conical scanning set up (see Figure 2-2), with an initial observation in the forward direction along the satellite track at a 55° viewing zenith angle (decreasing to 47° viewing zenith angle at the edges of the forward scan). Then,

approximately 120 seconds later (for the sub-satellite pixel), a second observation is at nadir at a viewing zenith angle of 2° (increasing to 22° viewing zenith angle at the edges of the nadir scan). The raw scanned data comprise a 500km swath sampled across 371 overlapping pixels in the forward direction and 555 pixels in the nadir direction. Both scans are resampled into the regularly gridded Envisat N1 geo-referenced swath format, which has a final product nominal pixel resolution of 1km across a 512-pixel swath for both views. It must be noted however, that whilst the resultant swaths cover the same geographical extent, the observations from each view comprise different image resolutions across the field of view (FoV) as can be seen in Figure 2-3.

The radiometric configuration of ATSR-1 comprises four channels at $1.6\mu\text{m}$, $3.7\mu\text{m}$, $11\mu\text{m}$ and $12\mu\text{m}$. For ATSR-2 and AATSR, three additional visible/NIR channels at $0.55\mu\text{m}$, $0.67\mu\text{m}$ and $0.87\mu\text{m}$ are available. On-board calibration of the thermal channels is achieved using twin blackbodies at known temperatures, which are observed during each scan cycle giving brightness temperatures correct to 0.05°K (Smith et al. 2001). The instruments also provide on-board visible channel calibration (Smith et al. 2002) and the visible channels are reported to be calibrated to an accuracy of better than 4% (Smith and Poulsen 2008). This gives the ATSR instruments excellent long-term stability in terms of radiometric calibration.

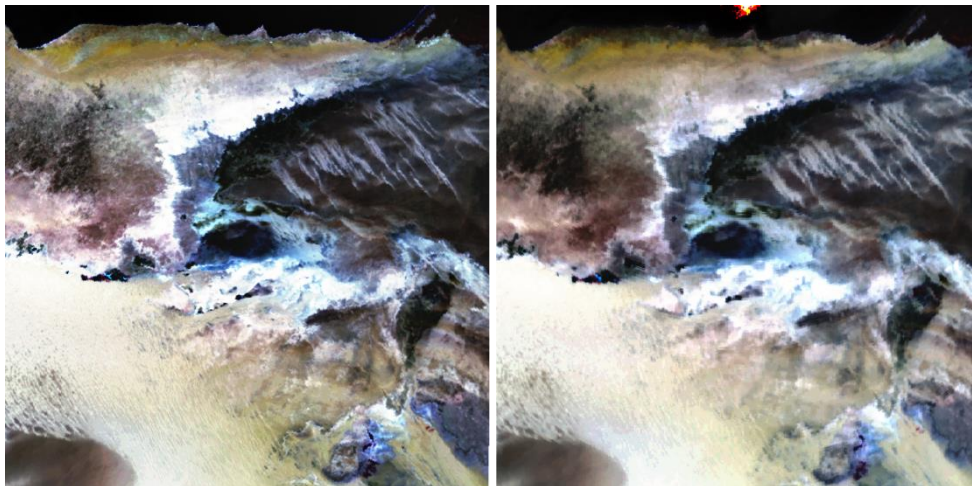


Figure 2-3. The left image is a histogram equalised false colour composite derived from an AATSR nadir view of the North African coast with the following channel combination, R: $0.87\mu\text{m}$, G: $0.67\mu\text{m}$, B: $0.55\mu\text{m}$. The right image is a histogram equalised false colour composite derived from the forward view of the same location using the same channel combination. Note the differing resolutions of the two images, evidenced by an apparent blurring in the forward imagery. (Taken from AATSR orbit: 32768, 6th June 2008)

The ERS and Envisat satellites which housed the ATSR instruments were maintained in sun synchronous polar orbits, with a period of 100 minutes, a mean orbital altitude of $\sim 800\text{km}$, and a descending node with an equatorial overpass of 10:30 a.m. local solar time. The repeat cycle for ERS-1 was dependent on the operational mode of the satellite and comprised of either 3, 35, or 168 day repeat cycles, and thereafter the general operational cycle became 35 days.

Absolute and relative geo-referencing accuracies are not well documented for all three instruments. For ATSR-2, an absolute geo-location accuracy (i.e. relative to the Earth) of 2km is reported (Merchant, 2003; Seiz, 2003), and similar accuracy is assumed for ATSR-1 and AATSR. Relative geo-location accuracy, the accuracy of the registration between the forward and nadir views, is a known weakness of the ATSR instruments, with differences of greater than 2 pixels in some instances as shown in Figure 2-4 (Seiz, 2003). As will become clear in Chapter 3, accurate co-registration between the views is of critical importance for accurate stereo-photogrammetric measurements; the remainder of this chapter is therefore devoted to describing a technique which has been developed (also documented in Fisher and Muller, 2013) to correct this short-coming of the ATSR instrument series.

2.3 Forward to Nadir Co-registration

2.3.1 ATSR Co-registration Problem

The ATSR co-registration problem was not identified for the first 13 years of the ATSR mission. This lack of detection was probably due to the fact that most applications did not make use of the forward channel and those that did were mostly focused on low resolution, high accuracy SST applications (25km and lower) which did not necessitate accurate co-registration of the views (e.g. Handoll et al., 1991).



Figure 2-4. A colour composite from the North African coast comprised of R: AATSR nadir 0.55 μ m, G: Null values, B: AATSR forward 0.55 μ m. The apparent pink edges along the coast line are due to poor co-registration between the nadir and forward views. (Taken from AATSR orbit: 32768, 6th June 2008)

The problem was first uncovered during analysis of oceanic eddy structures, where spurious edge effects were observed when differencing AATSR full resolution forward and nadir 3.7 μ m observations (Embury, 2004 in Corlett et al., 2009). Correction for these errors was achieved by shifting the forward view in both the across and along track directions by two pixels. Whilst this led to a significant reduction in edge effects no analysis was performed to determine either the quality of the correction applied, or the suitability of the correction for application across the entire AATSR time-series. Further histogram analysis on similar AATSR scenes attributed the co-registration error to an incorrect definition of the measured scan angle (Corlett et al., 2009). The conclusion drawn was that this would lead to a roughly consistent offset of two pixels along track and two pixels across track between the forward and nadir

views. Again, no analysis of the accuracy of the proposed corrections or assessment of general applicability was provided.

More recently a concerted effort to better understand the characteristics of the ATSR co-registration errors and provide a more rigorous assessment of their behaviour was undertaken by Casadio (2010). This study attempted to characterise the errors for all three ATSR instruments by cross correlation matching of six globally distributed islets with asymmetry relative to the instrument viewing geometry (i.e. circular in form). A number of important findings were presented. Firstly the along and across-track co-registration errors were found to have a high dependence on across-track position for all instruments and, in the case of ATSR-2 and AATSR, also on time. Secondly, the co-registration error dependency was shown to be linear with time for all three instruments. For ATSR-1, the change in error through time was shown to be less than a pixel. For ATSR-2 and AATSR the error was found to vary by approximately a pixel through both mission time-series. Thirdly, a parabolic across track error, dependent upon across-track position, was demonstrated for the AATSR instrument. The magnitude of this error was shown to be around a pixel. Lastly, the findings also presented slight channel dependent error variations, though the magnitudes never exceeded a fraction of a pixel.

This most recent study has demonstrated that co-registration errors affect all three ATSR instruments, albeit presenting with different characteristics. The following sections provide a description of a suitable method for obtaining globally (i.e. irrespective of location) applicable corrections for these errors and also a detailed statistical assessment of the improvements gained in co-registration following application of corrections.

2.3.2 Co-registration Correction Methodology

As ATSR data are resampled onto a regular rectangular grid the most appropriate method for correcting the errors in co-registration is through application of a raster based warp. The standard approach for this technique is to locate corresponding image locations between the images to be co-registered. The correspondences are then used to derive a transfer function, which relates the raster coordinates in one image to those in another. The approach can be split into a three-stage process: the first step is to generate correspondence between the images to be warped - this is hereafter referred to as tie-pointing; following tie-pointing, the matched points can be used in a least squares minimisation to generate a set of polynomial warping coefficients - the transfer function; finally, the transfer function used with an appropriate interpolation method maps one image onto the raster grid of the other.

2.3.2.1 Automated tie-pointing

Previous work, which provided improved ATSR-1 forward to nadir co-registration through such a three stage process as outlined above, employed dynamic programming to detect coastlines for use as a surrogate for tie-points (Shin et al. 1997). Once detected, the coastlines were used to derive polynomial transfer functions for raster based warping of pre-corrected SADIST (Synthesis of ATSR Data Into Sea Surface Temperature; Bailey, 1994) imagery. Whilst this method was shown to achieve pixel level accuracy, the coastline method employed does not provide a well-distributed set of tie-points throughout

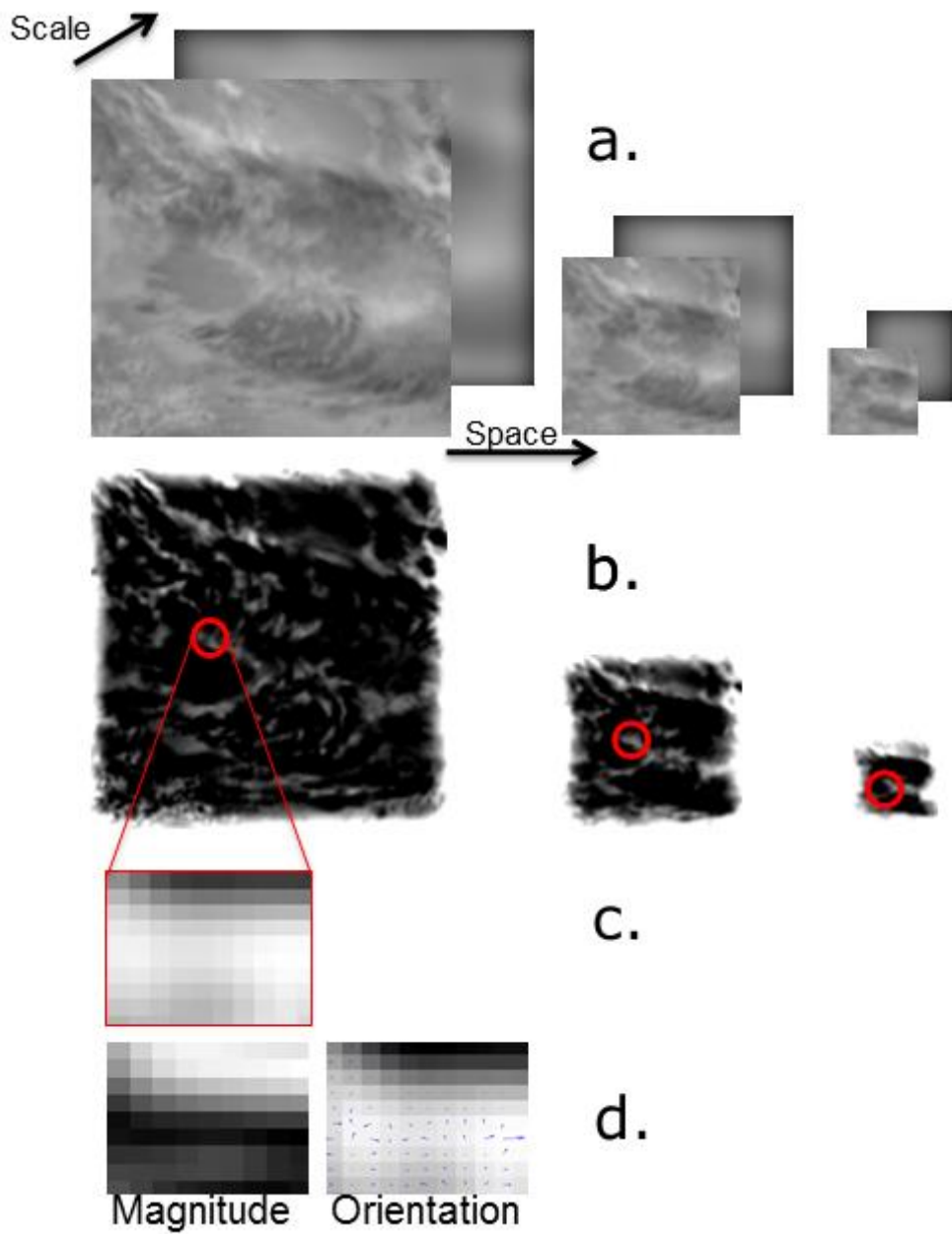


Figure 2-5. This figure shows a diagrammatical representation of the feature detection and description process employed by the scale invariant feature transform (SIFT). Similar processes are employed by the speeded up robust features algorithm (SURF), but they are highly optimised so are less intuitive when visualized. The first stage in the process (a), involves detection of maxima through scale and space, this is represented in (b) through a feature centred inside the red circle. The local neighbourhood around the feature is depicted in (c), and (d) demonstrates the type of information that is used to generate the feature descriptors used in the matching process.

the scenes to be co-registered. This is essential for effective polynomial-based warping as it ensures accurate co-registration across the raster.

In order to provide a better tie-point distribution, a state-of-the-art automated feature detection algorithm, Speeded Up Robust Features (SURF; Bay et al., 2008), is used in this thesis. Image features are defined as highly characteristic pixels within the image. In the case of SURF a feature pixel is defined as the maximum valued pixel within a local neighbourhood, and persists as a maximum within this neighbourhood through varying image scales¹, where image scale is varied through application of box filters of differing dimensions. A diagrammatical representation of this process for the scale invariant feature transform (SIFT; Lowe, 2004) which influenced SURF and many other similar algorithms is shown in Figure 2-5. The first 1000 maxima detected by SURF when applied to AATSR imagery are shown in Figure 2-6.

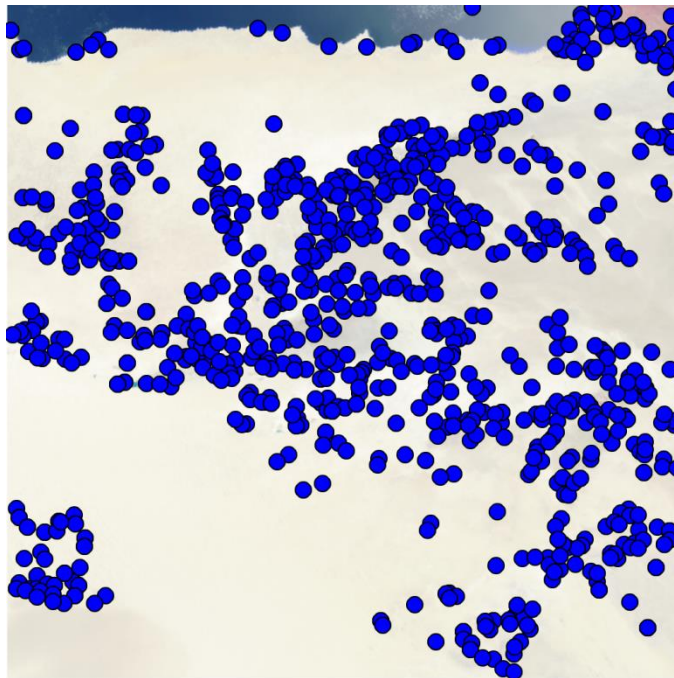


Figure 2-6. The first 1000 features detected by SURF from the nadir 0.55 μ m channel from the image presented in Figure 2-3.

Once a feature pixel has been detected in the image and its raster coordinates recorded, its local neighbourhood is used to generate a scale and illumination invariant descriptive feature vector of 64 bits in length, referred to hereafter as a descriptor. Scale and illumination invariance are important algorithmic qualities, as the ATSR instruments, due to the conical scanning geometry, have varying pixel resolution and viewing angles which lead to differing scale and illumination effects respectively.

Matching of the SURF descriptors leads to the generation of a tie-point set between the forward and nadir views. The matching is performed by first normalising the descriptors from each view to unit vectors. Once normalised the inverse cosine of the dot product is used as the distance measure. For each

¹ Scale can be thought of as being analogous to resolution

descriptor from the forward view a comparison is made to every descriptor from the nadir view in order to evaluate potential matches, which can be written as follows,

$$C_i = \cos^{-1}(\hat{P}_i \cdot \hat{R}^T) \quad \text{Eq. 2.1}$$

Where \hat{P}_i is the i^{th} descriptor in unit vector form from the forward view and \hat{R} is a matrix of descriptors in unit vector form for every feature detected in the nadir view. The vector C_i contains the costs between the i^{th} descriptor from the forward descriptor set and all nadir descriptors.

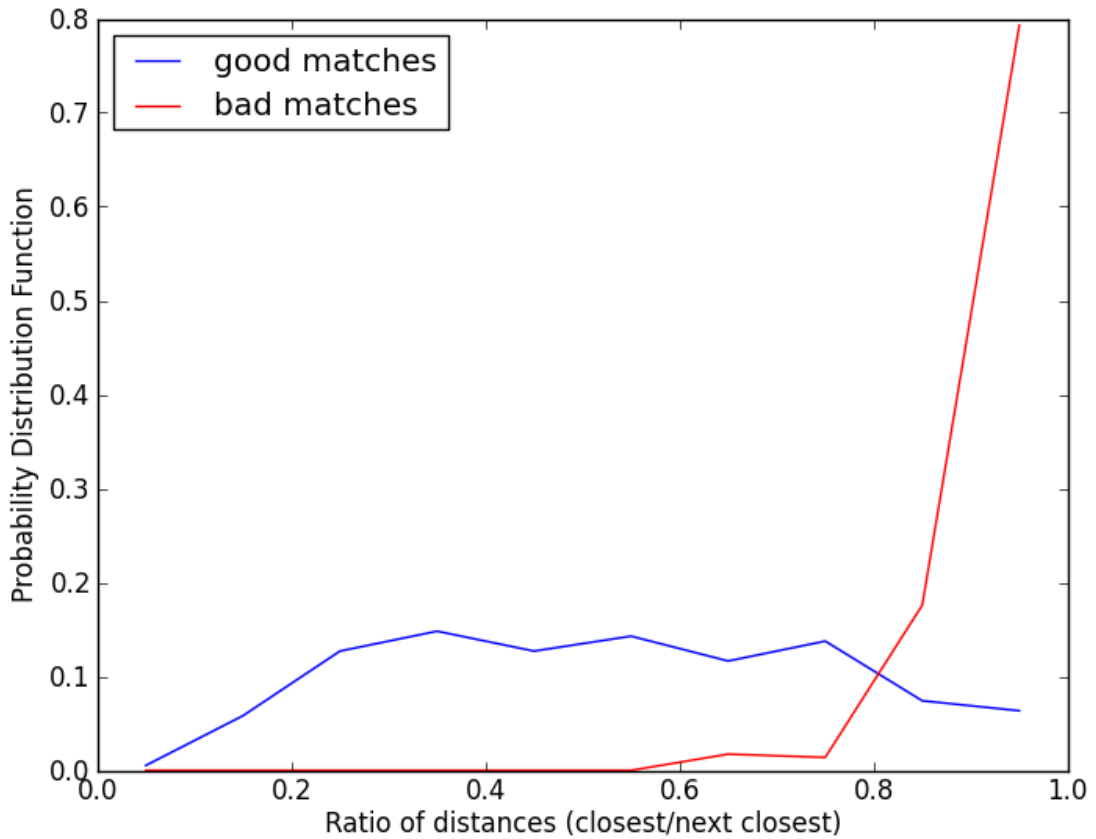


Figure 2-7. The effect of the matching threshold t and the good and bad SURF match probability distribution functions. This analysis was obtained from the scene presented in Figure 2-3. This plot indicates that a threshold of 0.8 retains around 80% of the good matches, whilst rejecting around 85% of the bad matches.

Once the costs for the i^{th} descriptor have been computed, two straightforward quality control checks are performed. This first check involves analysing the ratio of the two smallest costs contained within C_i . If this ratio is above a certain threshold then the match is rejected, as the minimum descriptive vector is deemed insufficiently unique enough to provide a reliable match. This process can be written as,

$$Q = \begin{cases} \text{if } \frac{C_i^j}{C_i^k} < t, & 1 \\ \text{else} & 0 \end{cases} \quad \text{Eq. 2.2}$$

Where the superscripts j and k denote indexes to the minimum and next smallest cost contained in C_i respectively, t is a threshold which in this study is set to 0.8, and Q is either 1 or 0 depending on the outcome of the test. If Q is 1 then the forward descriptor under evaluation is deemed to have suitable discriminability and the index j denotes the location of the matching descriptor contained within $\hat{\mathbf{R}}$. This screening process, with t set to 0.8, has been shown to be very reliable removing up to 90% of erroneous matches, whilst retaining 95% of those which are correct (Lowe, 2004). An analysis of the suitability of $t=0.8$ for the $0.67\mu\text{m}$ channel for the scene shown in Figure 2-3 is presented in Figure 2-7. The scene is of very low elevation terrain ($\ll 1\text{km}$) and has been corrected using the derived co-registration coefficients. Given these conditions correctly matched SURF features should have disparities in both axes of less than 1 pixel. This allows for the detection of incorrectly matched features (they will exhibit disparities of >1 pixel). The plot in Figure 2-7 demonstrates that a threshold of 0.8 excludes $\sim 85\%$ of bad matches (disparities of > 1 pixel), whilst retaining $\sim 80\%$ of good matches (disparities of < 1 pixel).



Figure 2-8. The first 100 matched tie-points between the forward and nadir AATSR views derived by SURF from the $0.55\mu\text{m}$ channels from the images presented in Figure 2-3. Incorrect matches are generally indicated by those tie-point pairings that do not conform to the general consensus (i.e. lines which are not parallel). It is evident from the above example that there no incorrect matches, i.e. all lines are parallel.

The second check involves performing a left-right assessment, where the matching is repeated with the nadir and forward descriptors switched, i.e. the nadir descriptors are defined by \hat{P}_i and the forward descriptors by $\hat{\mathbf{R}}$. If the same vector pairing is found matching in both directions then the match is deemed reliable and the raster coordinates associated with the descriptor are included as tie-points. The

end result of the feature detection and matching processes is a set of tie-points which can be used to derive a warp between the forward and nadir views to provide improved co-registration. An example output of this tie-point matching process is shown in Figure 2-8.

2.3.2.2 Transfer Function

The tie-points are used to generate a warping function which maps the forward tie-point locations onto those in the nadir image using a least squares minimisation. The first step in this process is to scale all tie-point coordinates to the range [-1, 1] in order to reduce the introduction of errors (NIMA, 2000). The next step in the process involves determination of the number of trials required to obtain an optimal warp. Within a typical set of tie-points there are a very large number of potential tie-point combinations. Some tie-point combinations will provide a better co-registration result than others; location of such combinations is therefore of great importance in ensuring the quality of the end result. Testing all potential combinations is unfeasible; therefore the binning method from Tao and Hu (2001) is employed, where the image is divided up into bins, or neighbourhoods. The number of tie-point combinations tested is dependent on the number of bins and the ratio of bins with tie-points to those without, it can be defined as,

$$T = \frac{\ln(1 - \lambda)}{\ln(1 - \frac{L}{w})} \quad \text{Eq. 2.3}$$

Where, T is the number of trials/tests, λ is the desired probability² of the optimal tie-point subset being selected, here 0.99, L is the total number of bins containing tie-point populations, and w is the total number of bins. Tie-point distribution is of critical importance, and this in turn is affected by the number of bins: too many bins and a poor tie-point distribution may lead to an infinite number of trials; too few bins may not effectively capture enough tie-point location variability to find the optimum warp. Therefore the number of trials is calculated for varying bin sizes (16^2 , 32^2 , 64^2 , 128^2 and 256^2 pixels), which in turn varies w and L . The smallest bin size which leads to $T < 100$ is used.

Once the number of trials has been determined, a pseudo-random subset of tie-points is selected, one tie-point from each bin with a tie-point population. Bins with a greater number of tie-points will have more likelihood of being selected, leading to non-random behaviour in the trialling process. Therefore the random bucketing technique of Zhang (1995) is employed. This method ensures random selection through a weighting process where tie-points from more densely populated bins are assigned a lesser weight and therefore are less likely to be selected. An example tie-point selection is shown in Figure 2-9. Once selected, the tie-points are used to determine a simple geometric relationship between the forward and nadir views in the form of a least squares derived polynomial transfer function. The order of the polynomials used to derive the transfer functions for all three instruments are determined from the results given in Casadio (2010). Linear behaviour in error was demonstrated in all cases in the study for ATSR-1 and ATSR-2. Therefore, first order polynomials of the form,

² For the effects of varying the probability see Table 1 in Tao and Hu (2001)

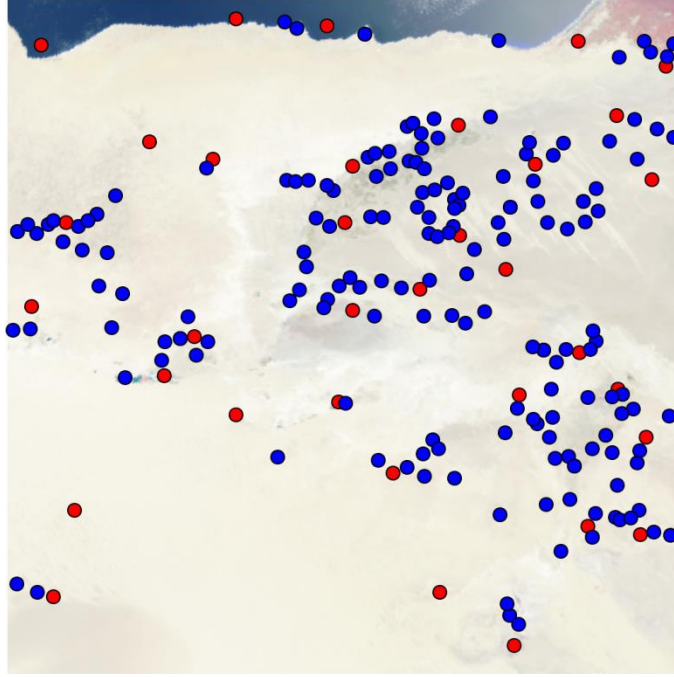


Figure 2-9. An output from the random tie-pointing process. The red dots are those features selected by the binning process as tie-points. The blue points are those left unselected, and they are used as check-points to assess the quality of the warp.

$$\begin{aligned} x_f &= a_0 + a_1 y_n + a_2 x_n \\ y_f &= b_0 + b_1 y_n + b_2 x_n \end{aligned} \quad \text{Eq. 2.4}$$

are used, where x and y with a subscripted n are the nadir raster coordinates, and with a subscripted f the forward raster coordinates. The coefficients a and b define the warp coefficients which relates any given nadir view pixel coordinate to its corresponding pixel coordinate in the forward view. For AATSR, across track errors were shown to vary non-linearly depending upon across track pixel position (Casadio 2010). The non-linear behaviour is corrected for using polynomials of the form,

$$\begin{aligned} x_f &= a_0 + a_1 y_n + a_2 x_n + a_3 x_n^2 \\ y_f &= b_0 + b_1 y_n + b_2 x_n + b_3 x_n^2 \end{aligned} \quad \text{Eq. 2.5}$$

where the extra term corrects for such behaviour in the across track axis. The coefficients are determined by way of least squares minimisation,

$$\mathbf{S} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad \text{Eq. 2.6}$$

where \mathbf{A} contains the nadir view tie-point raster coordinates, and \mathbf{b} is a vector of the corresponding tie-point coordinates from the forward view raster. The warping coefficients are contained in \mathbf{S} .

To improve the quality of the end result S is computed twice. In both iterations the coefficients in S are applied to both the nadir tie-points and the check-points (the unused tie-points in each bin) and compared to their respective locations in the forward image to assess the quality of the derived warp. The first iteration is for screening purposes and the mean and standard deviation of the Euclidean distances between the warped nadir tie-point raster coordinates and their corresponding forward raster coordinates are calculated. Any tie-point pair with a Euclidean distance of greater than 3 standard deviations from the mean is excluded from the tie-point set for the second iteration. In the second iteration the tie- and check-point RMSEs are calculated from the Euclidean distances and summed. The S associated with the trial which minimises the summed RMSEs is the retained warp.

Once the warp has been derived it can be used to transform the imagery from the forward view onto that of the nadir view. This is achieved by using a nearest neighbour interpolation routine,

$$\hat{F} = F(x_N^S, y_N^S) \quad \text{Eq. 2.7}$$

where \hat{F} is the warped ATSR forward image, F is the original forward image and x_N^S and y_N^S are the nadir image x and y raster coordinates which have been warped by the appropriate parameters contained within S .

2.3.3 Transformation Coefficients

In this section the application of the automated tie-pointing and warping algorithm described above is discussed, the main consideration being selection of suitable scenes for deriving the transfer functions. Furthermore, once derived, an independent evaluation of the transfer functions is carried out using a pyramidal version of the Gotcha algorithm (Day and Muller 1989) which in various tests (Heipke et al., 2007) has been shown to produce excellent, albeit very slow results.

2.3.3.1 The coefficients

Due to the temporal variability (increasing along and across track errors through time) in the AATSR and ATSR-2 co-registration errors, yearly transformation functions, whilst not strictly necessary, are derived. For ATSR-1, the temporal variability is insignificant; therefore only one scene of forward and nadir views is required for the calculation of transformation parameters.

To test its efficacy over different land surfaces, the SURF algorithm was applied to numerous scenes of differing land cover at low topographic elevations in the $0.55\mu\text{m}$ channel for AATSR and ATSR-2 and the $1.6\mu\text{m}$ channel for ATSR-1. Low elevation ($<800\text{m}$) scenes with limited terrain variability were used to avoid the introduction of parallax³ effects during the transfer function derivation. This led to a number of specific target regions being tested, namely the Pampas region in South America, the East European Plain, the West Siberian Plain and the coastlines of Libya and Egypt. From this evaluation, scenes of Libya and Egypt provided both the greatest density and the best distribution of features amid all the scenes analysed. Scenes from these areas were therefore employed for the calculation of the

³ The proximity dependent shift in a feature when viewed from different viewing angles. See Chapter 3.

transformation parameters, and the scene used for the derivation of the 2008 coefficients is shown in Figure 2-3.

Following scene selection, the method outlined in section 2.3.2 was applied to 512^2 pixel images to obtain transformation parameters for each defined correction epoch. This provided 20 sets of warping parameters for the correction of the co-registration errors present in the ATSR instruments, which are given in Table A1, contained in appendix A. The subset size of 512^2 pixels was chosen due to the processing requirements of other applications developed within this thesis – the stereo processing algorithm for example subsets each orbit into 512^2 pixel chunks prior to processing.

2.3.3.2 Independent Evaluation

To provide an independent and automated method for measuring the quality of each transformation a pyramidal version of Gotcha algorithm, due to its excellent potential for sub-pixel accuracy measurement is employed. Following application of the derived warp, pyramidal Gotcha, due to the underlying adaptive least squares correlation (ALSC) stereo matching algorithm (Gruen 1985), can be used to measure the co-registration quality in both axes at potentially every image pixel. Images, which are perfectly co-registered and have had potential terrain elevation induced parallax effects removed, will exhibit a shift of zero pixels in both axes.

For the 512^2 pixel ATSR images, four image pyramid levels are used with the smallest resampled image being 64^2 pixels. Initial tie-point locations are then automatically selected at regular intervals (every 8 pixels in this instance) across the image and matched using ALSC. The ALSC algorithm has an effective convergence radius of ± 3 pixels, i.e. if the shift is greater than 3 pixels then ALSC will not be able to measure it reliably. The image pyramiding has the effect of reducing the disparity, e.g. a disparity of 16 pixels at full ATSR resolution will be two pixels at the fourth pyramid level, allowing ALSC to converge to a solution effectively and with greater speed. In order to evaluate every image pixel from the initial set of ALSC matched tie-points, a seed-growing algorithm is employed. The seed-growing method supplies a priori disparity estimates to the ALSC algorithm from matched points in close proximity (typically distances of 1 pixel) to the pixel being matched. These processes are iterated down the pyramid until the shifts are measured at full image resolution.

Gotcha is slow, but due to ALSC provides sub-pixel accuracy (reportedly 0.01 pixels with well signalled points) and achieves good coverage. It potentially provides an assessment of the co-registration quality across the entire scene, and not just at sparse feature locations (Otto and Chau 1989). The results of the quality assessment for the warping parameters are given in Table 2-1 and Table 2-2.

	X					Y				
	RMSE	Mean	Std	Bias	Npts	RMSE	Mean	Std	Bias	Npts
AATSR										
2011	2.06	-1.65	0.93	-1.95	221729	1.86	-1.51	0.74	-1.70	221729
2010	1.95	-1.61	0.82	-1.88	224812	1.85	-1.53	0.77	-1.80	224812
2009	2.21	-1.63	1.19	-1.94	220426	1.83	-1.45	0.84	-1.73	220426
2008	1.91	-1.56	0.85	-1.82	224864	1.77	-1.47	0.77	-1.71	224864
2007	1.68	-1.25	0.86	-1.52	215678	1.71	-1.37	0.73	-1.66	215678
2006	1.63	-1.16	0.86	-1.48	206915	1.74	-1.33	0.79	-1.68	206915
2005	1.89	-1.20	1.17	-1.54	205172	1.69	-1.28	0.75	-1.64	205172
2004	1.52	-1.11	0.83	-1.33	218915	1.59	-1.29	0.68	-1.54	218915
2003	1.23	-0.94	0.64	-1.09	225420	1.72	-1.42	0.72	-1.66	225420
2002	1.27	-0.68	0.87	-0.89	198683	1.69	-1.25	0.76	-1.65	198683
ATSR2										
2003	2.77	-1.99	1.37	-2.61	200179	1.39	-1.00	0.68	-1.31	200179
2002	4.45	-3.66	1.84	-4.34	220998	2.36	-1.96	0.92	-2.33	220998
2001	3.66	-2.35	1.87	-3.48	176884	2.79	-1.81	1.41	-2.69	176884
2000	2.43	1.78	1.40	2.10	226771	3.76	-3.1	1.68	-3.55	226771
1999	0.80	-0.32	0.66	-0.38	226112	1.10	-0.81	0.57	-0.93	226112
1998	1.24	0.85	0.7	0.99	225656	0.96	-0.7	0.55	-0.82	225656
1997	0.65	0.47	0.38	-0.54	229246	1.19	-0.97	0.53	-1.11	229249
1996	0.43	0.21	0.32	0.25	209749	1.26	-0.99	0.57	-1.25	209749
1995	0.41	0.18	0.33	0.21	229941	1.43	-1.23	0.52	-1.40	229942
ATSR1										
1994	2.31	1.86	0.99	2.24	218021	0.86	-0.27	0.74	-0.34	218020

Table 2-1 - This table presents the results from the application of the P-Gotcha for each instrument and year for images of size 512 x 512 (262,144 pixels). The images have had no correction applied. The X values are the statistics for the across track displacements between the nadir and the corrected forward view, all units are in pixels. The Y values are the statistics for the along track displacements, again units are in pixels.

	X					Y				
	<i>RMSE</i>	<i>Mean</i>	<i>Std</i>	<i>Bias</i>	<i>Npts</i>	<i>RMSE</i>	<i>Mean</i>	<i>Std</i>	<i>Bias</i>	<i>Npts</i>
AATSR										
2011	0.63	-0.08	0.38	-0.21	98039	0.6	-0.13	0.35	-0.34	98040
2010	0.62	-0.09	0.38	-0.23	103330	0.5	-0.05	0.31	-0.15	103327
2009	0.51	-0.11	0.33	-0.22	123132	0.40	-0.05	0.27	-0.12	123131
2008	0.53	-0.07	0.34	-0.15	118889	0.37	0.02	0.24	0.04	118889
2007	0.52	-0.60	0.31	-0.17	98763	0.42	-0.06	0.25	-0.15	98765
2006	0.56	-0.10	0.35	-0.23	108316	0.45	-0.07	0.28	-0.17	108312
2005	0.35	-0.08	0.31	-0.09	224378	0.25	-0.12	0.19	-0.14	224378
2004	0.49	-0.12	0.32	-0.25	126365	0.38	-0.05	0.26	-0.11	126363
2003	0.44	-0.09	0.31	-0.17	142942	0.36	-0.03	0.26	-0.06	142944
2002	0.51	-0.08	0.33	-0.18	118135	0.37	-0.01	0.25	0.02	118135
ATSR2										
2003	0.65	-0.22	0.51	-0.30	194874	0.45	-0.03	0.38	-0.04	194869
2002	0.68	-0.22	0.54	-0.29	191935	0.30	0.01	0.26	0.01	191932
2001	0.59	-0.24	0.46	-0.32	198802	0.41	-0.05	0.35	-0.07	198799
2000	0.49	-0.11	0.39	-0.17	176679	0.37	-0.17	0.24	-0.27	176678
1999	0.45	-0.09	0.32	-0.17	143475	0.37	-0.03	0.27	-0.05	143474
1998	0.34	-0.07	0.27	-0.11	169619	0.22	0.01	0.18	0.01	169617
1997	0.47	-0.22	0.32	-0.32	179194	0.40	-0.19	0.28	-0.28	179191
1996	0.48	-0.09	0.32	-0.19	121422	0.38	-0.03	0.25	-0.06	121420
1995	0.42	-0.09	0.31	-0.15	156896	0.36	-0.04	0.27	-0.07	156895
ATSR1										
1994	0.38	-0.03	0.35	-0.03	215558	0.29	-0.09	0.25	-0.11	215560

Table 2-2- This table presents the results from the independent warping coefficient evaluation for each instrument and year for images of size 512 x 512 (262,144 pixels). For the evaluation each warp is applied to the scene from which it is derived, the statistics being generated using the P-Gotcha stereo matching algorithm. The X values are the statistics for the across track displacements between the nadir and the corrected forward view, all units are in pixels. The Y values are the statistics for the along track displacements, again units are in pixels.

2.3.3.3 Discussion

The statistical analysis presented in Table 2-2 demonstrates that when the warping parameters are applied to the scene from which they are derived, the co-registration error (RMSE) in both image axes is improved to pixel level of better. This is a significant improvement in most cases over the co-registration quality in the uncorrected images as shown in Table 2-1. The evaluations are performed on at least 37% of the pixels within the 512x512 scenes used to derive the coefficients, giving good overall coverage and significantly more points than the tens of points used during the warping coefficient derivation. The analyses show biases in both axes that are far less than the RMSE. The cause of the biases is not fully understood, the along track bias is most likely due to the effects of terrain parallax, the cause of the across track bias is more challenging to determine. This independent analysis of the warping parameters demonstrates that a suitable improvement in the co-registration between the forward and nadir images of the ATSR instruments to pixel level accuracy can be achieved using this automated correction approach.

Further analysis of the warping coefficients is required to determine whether each of the derived warping parameters is able to provide a universal correction for the co-registration problem, and not just on the scene from which it is derived.

2.3.4 Global Evaluation

In order to provide an assessment of the global applicability, the transformation coefficients are evaluated at numerous sites distributed globally to determine whether the improvement in the co-registration is consistent at disparate locations. Whilst complete global evaluation is unachievable due to the need for detectable features (e.g. the sea surface provides no features) to generate tie-points for evaluation, a number of orbits with a suitable spread of coverage in terms of longitude and with significant percentages of land mass are evaluated to assess the universality of the transformations. To ensure suitable coverage complete days' worth of orbits are selected giving up to 14 orbits of data per day, with good distribution globally. To further improve coverage and to account for the effects of cloud cover reducing observable landmass, four days' worth of orbits in each test year were selected. With two days from Northern hemisphere summer and two from Southern, giving up to 56 orbits in total for each assessment. The orbits were selected to give a good temporal distribution throughout each parameterisation period to check whether any decrease in the efficacy of a particular warp occurs through time and also to ensure that a suitable number of features are detectable (typically more features will be detected in Summer (i.e. snow free scenes)).

2.3.4.1 Processing Chain

The global evaluation processing chain, implemented in IDL, processed the 4 days' worth of orbits selected from each year of operation for each instrument to assess the performance of the associated transformation coefficients, the analysis being performed across all instrument channels.

The first step in the processing chain involves segmenting each orbit into 512x512 pixel subsets for each channel. Then, using the built in ATSR cloud flagging, the clear land percentage for each scene is then estimated; if at least 20% of the scene is clear land, processing is begun upon the scene. The main

computational effort involves the application of the automated tie pointing. Prior to the application of the tie-pointing, the forward image for each channel within the subset is warped using the appropriate transformation coefficients. Following detection of tie-points, a number of screening steps, described in the following sections, are employed to ensure the removal of tie-points associated with cloud features and to account for the poor geo-location accuracy of the ATSR instruments. Following screening, the nadir and forward raster coordinates of each retained tie point pair are stored in addition to the spectral channel, the shift in pixels, and the latitudes and longitudes for further analysis. A sample distribution of these points can be seen in Figure 2-10.

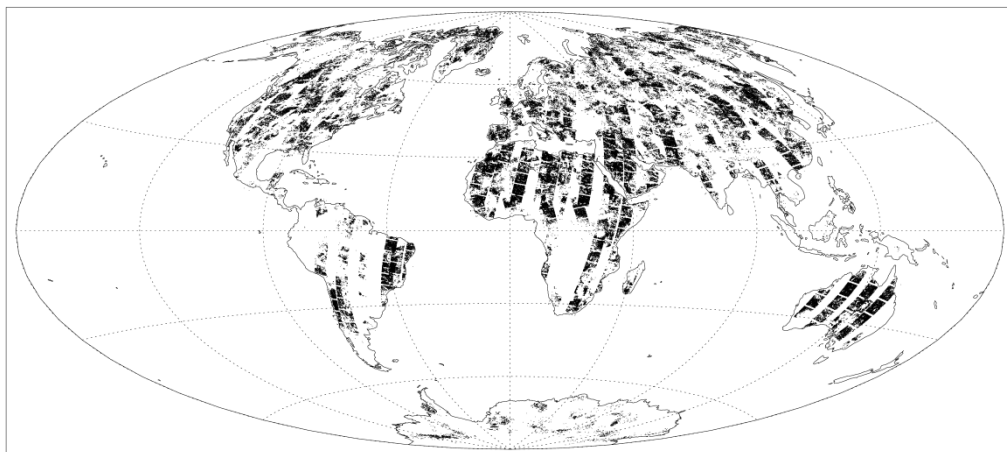


Figure 2-10. Example assessment locations, where a black point indicates an assessments, for the 2008 AATSR global applicability assessment. The tie-point distribution provides reasonable coverage over all major landmasses with the exception of Antarctica

2.3.4.2 Cloud Screening

Cloud screening is vital to effectively assess the quality of the transformation in the direction of parallax (in the along track direction for ATSR). Any cloud features, which are not screened lead to an overestimation of parallax and therefore introduce errors during statistical analysis of the improvement in co-registration. The effects of clouds in the across track direction are dependent on the local winds (Denis et al. 2007) and again, can be a source of error during statistical evaluation. To screen clouds the ATSR forward cloud-flag products are used. To ensure as many cloud tie-points are removed as possible a smoothing filter is applied to the ATSR cloud mask to provide a buffering effect and remove features on the edges of clouds, which are generally problematic. It was found that even after these cloud screening steps, clouds still contaminated some scenes leading to errors in the statistical analysis. Therefore a further screening step is carried out through elevation normalisation of the along track results. Once elevation normalised, any feature, which is more than 2 km above the surface, is discarded as cloud.

2.3.4.3 Elevation Normalisation

To effectively evaluate the improvement in along track co-registration the parallax effects introduced by terrain need to be removed. This is achieved by converting a digital elevation model (here GETASSE30)

height for each matched feature into a disparity using an inverse ATSR camera model⁴ (Denis et al., 2007). This value is then subtracted from the matched features' along track disparity, leading to an elevation normalised result.

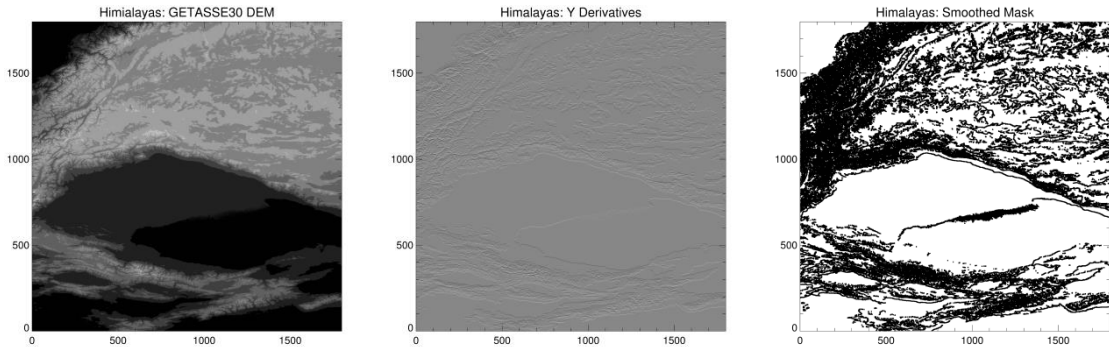


Figure 2-11. This figure shows the elevation screening process. The left most image is the input DEM scene of the Himalayas and Tibetan Plateau. The central image contains the y derivatives computed using Eq. 2.9. The right image is the buffered mask generated from both the x and y derivatives, where any pixel with a rate of change of greater than 0.5km in either axis is screened off (here represented as a black pixel)

Unfortunately, for the ATSR instruments, regions with rapidly and significantly varying terrain elevations lead to ambiguities in the elevation normalisation process. This is due to the quality of the instrument's geo-referencing (here assumed ± 2 pixels). Therefore, an approach is adopted which excludes from further evaluation all tie-points located on terrain that varies above a pre-set threshold. The threshold limit is defined as any pixel whose terrain gradient rate of change is outside the range ± 0.5 km. The gradient is determined by calculating the second derivatives of a GETASSE30 1km resolution digital elevation model (DEM) which has been geographically collocated with the image. The DEM image second derivatives in the x-axis are found using,

$$\Delta x_2 d(i, j) = d(i + 1, j) + d(i - 1, j) - 2d(i, j) \quad \text{Eq. 2.8}$$

and similarly for the y-axis using,

$$\Delta y_2 d(i, j) = d(i, j + 1) + d(i, j - 1) - 2d(i, j) \quad \text{Eq. 2.9}$$

where d is the DEM and (i, j) are the DEM pixel coordinates. A gradient outside ± 1 is deemed unacceptable, as it indicates a rate of gradient change over the pixel that would lead to a greater than 1 pixel displacement due to the effects of elevation parallax. This is greater than the given accuracy requirement of the transformation (≤ 1 pixel), and would introduce errors into the statistical analysis. Hence ± 0.5 was chosen as an acceptable limit, as it does not remove too many regions from the transformation parameter quality assessment and maintains confidence in the analysis. As the geo-

⁴ A camera model defines the instrument viewing geometry. See Chapter 3.

location accuracy of the ATSR instruments is ± 2 pixels, a buffer of the same dimension is applied around those pixels that demonstrate a gradient greater than the threshold. This ensures the robustness of the elevation normalisation process and hence, of the statistical evaluation. The masking process is presented visually in Figure 2-11.

2.3.5 Results and Analysis

The results output from the global evaluation for each warping coefficient provide a large number, in the hundreds of thousands, of worldwide distributed tie-points between the nadir and warped forward views, along with associated latitudes, longitudes, elevation and the channel from which the tie point is obtained. To analyse such a large number of data points, it is preferable to combine the results where possible to provide single statistical measures of the overall quality of the warp. The appropriate statistical measure for the task is the RMSE along with a measure of bias to check for introduction of potential trends into the warped data.

To generate a more manageable dataset, the channels are first analysed separately for each instrument and its year of operation to check for any large differences in the quality of the co-registration across bands. Assuming no large variations in co-registration quality and behaviour occur between each channel then the channels for each instrument and year of operation are combined for further analysis. Searching for trends in the errors as a function of longitude or latitude is the next step. The warping coefficients must be applicable with similar quality to all latitudes and longitudes. Therefore in this analysis, divisions of the points along lines of latitude and longitude are made to search for any behaviour which would indicate that the warping parameters are not applicable globally. The analysis could have been further split to assess the individual days, which make up each yearly test epoch, to check for higher resolution temporal variability in the co-registration quality. However, this is deemed unnecessary here as if one day's worth of orbits within the assessment are significantly different then the outlying points would have a strong impact on the overall RMSE leading to their detection

2.3.5.1 Analysis by Channel

Firstly, analysis by channel was performed for each instrument and year. This was done to check for any significantly different behaviour in the errors between channels for the corrected scenes. An RMSE and bias value is derived from all measurements for each epoch for each instrument. The results from this are presented in Figure 2-12. Looking first at the outputs from the AATSR analysis, the RMSE values for all channels and years are below 0.8 pixels and all biases are within the range -0.5 to 0.0 pixels for the along track (Y) error and within the range -0.5 to 0.5 pixels for the across track (X) error. The RMSE and bias for each year for AATSR tend to be clustered within 0.2 pixels, showing that there is limited variation in error across channels. This permits combined channel analysis for assessing the behaviour of the error with varying latitude and longitude.

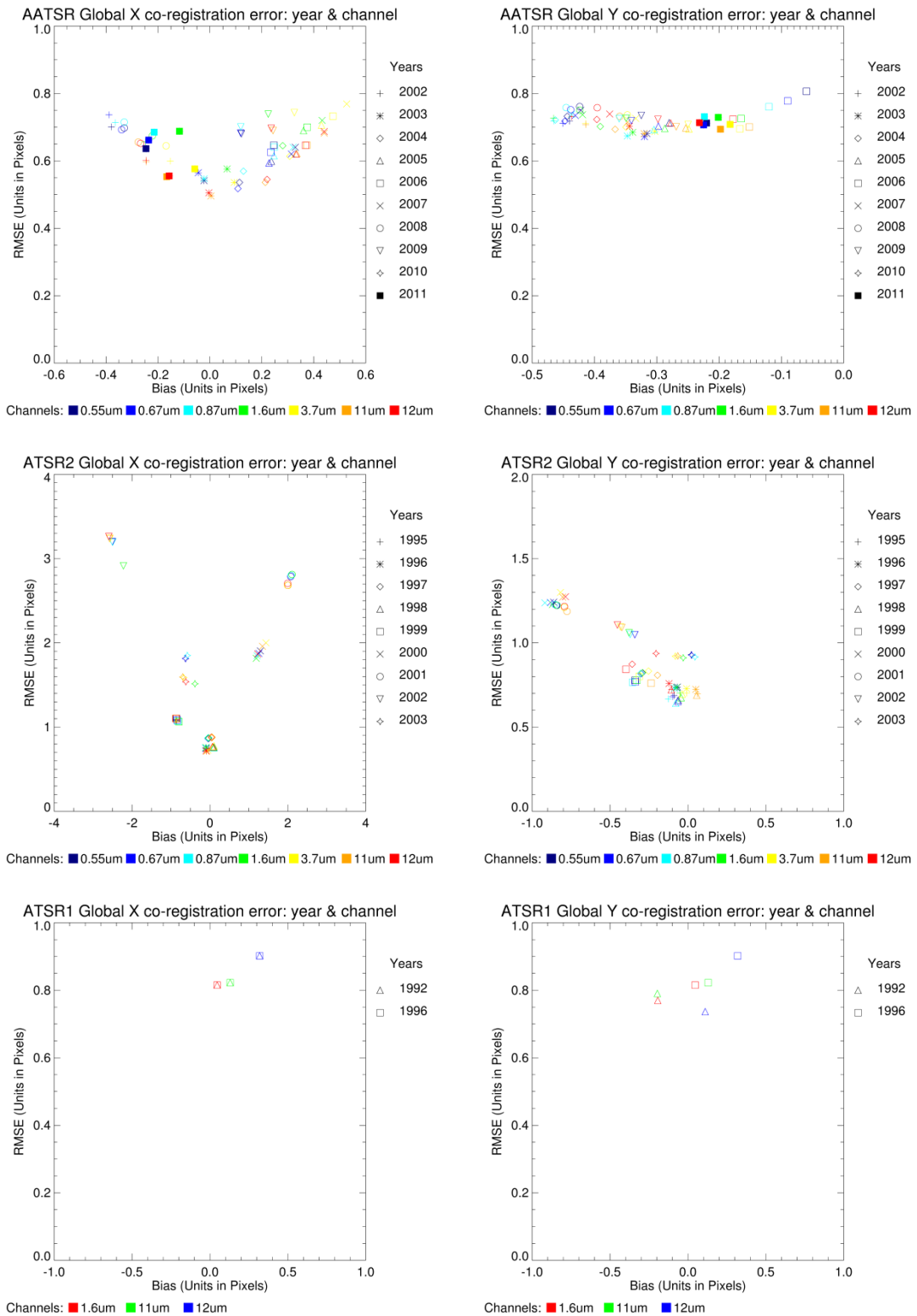


Figure 2-12. These six plots show the across (X) and along (Y) track co-registration errors by channel for each of the ATSR instruments as determined by the global assessment. The left column contains the across track errors and the right column the along track. The top two plots are for AATSR, the middle two for ATSR-2 and the bottom two for ATSR-1.

The middle two plots in Figure 2-12 present the channel results from ATSR-2. Looking firstly at the along track error it is evident that from 2000 the RMSE is greater than 1 pixel and the bias is approaching -1 pixels. This increase in RMSE and bias from 2000 is also evident at even greater magnitudes in the across track analysis, with the 2002 results presenting RMSE and bias values of around 3 pixels. However, the channel results are again generally well clustered in terms of the behaviour of the error for each year. Therefore combining the channels for analysis is permissible.

The bottom two plots are for ATSR-1 as only two years of analysis are presented due to the limited variability of the error through time. For the analysis, the first year of operation, 1992, and a later year, 1996, are assessed following application of warping parameters derived for scenes from 1994. The along track assessment shows RMSE values of less than 1 in both instruments and bias values between -0.5 to 0.5 pixels. For across track errors, similar results are presented. The clustering of the results for each year for ATSR-1 is not as apparent as for the other two instruments. However, the error does not vary by more than a fraction of a pixel so the channels were combined for further analysis.

2.3.5.2 Analysis by Latitude

For latitudinal analysis the channel results for each instrument and its year of operation were combined into 15° intervals from 90° S through 90° N. The results from this output are given in Figure 2-13; again with RMSE on the y-axis and bias on the x-axis, colour denotes latitudinal group and symbol the year of operation. The AATSR along track plot shows that in general the RMSE is below 1 pixel and the bias varies between ± 0.5 pixels. The main digression occurs at higher latitudes denoted by darker colours. The RMSE at higher latitudes approaches closer to, and in some cases exceeds 1 pixel. Whilst the bias tends to be positive, at lower latitudes, negative biases are exhibited. The across track error presents similar behaviour, with the biases having a more equal distribution with no noticeable latitudinal effect.

The ATSR-2 results show much more complex behaviour with a large proportion of the results having RMSE and bias values greater than 1 pixel. The along track RMSE and bias values do not show much latitudinal dependence, rather those points with more significant errors come from the instrument years post 2000. Prior to 2000, the darker coloured points indicating matches from higher latitudes do show larger deviations from zero in terms of RMSE and bias, as is also shown for AATSR. The across track errors show much more extreme behaviour with RMSE and bias approaching deviations from zero of ± 5 pixels for the instrument post 2000. Prior to 2000 the results are more clustered within the range ± 1 pixels for RMSE and bias. However, outputs from 1995 and 1999 do present some points with RMSE scores and biases of greater than ± 1 pixels with some of the points indicating origination from lower latitudes.

For ATSR-1 the results are more consistent with AATSR, with the majority of scores from lower latitudes for both across and along track analysis showing values within the range ± 1 pixels for both RMSE and bias.

2.3.5.3 Analysis by Longitude

For longitudinal analysis the channel results for each instrument and its year of operation were combined into 30° intervals from 180° E through 180° W. The results from this analysis are plotted in Figure 2-14

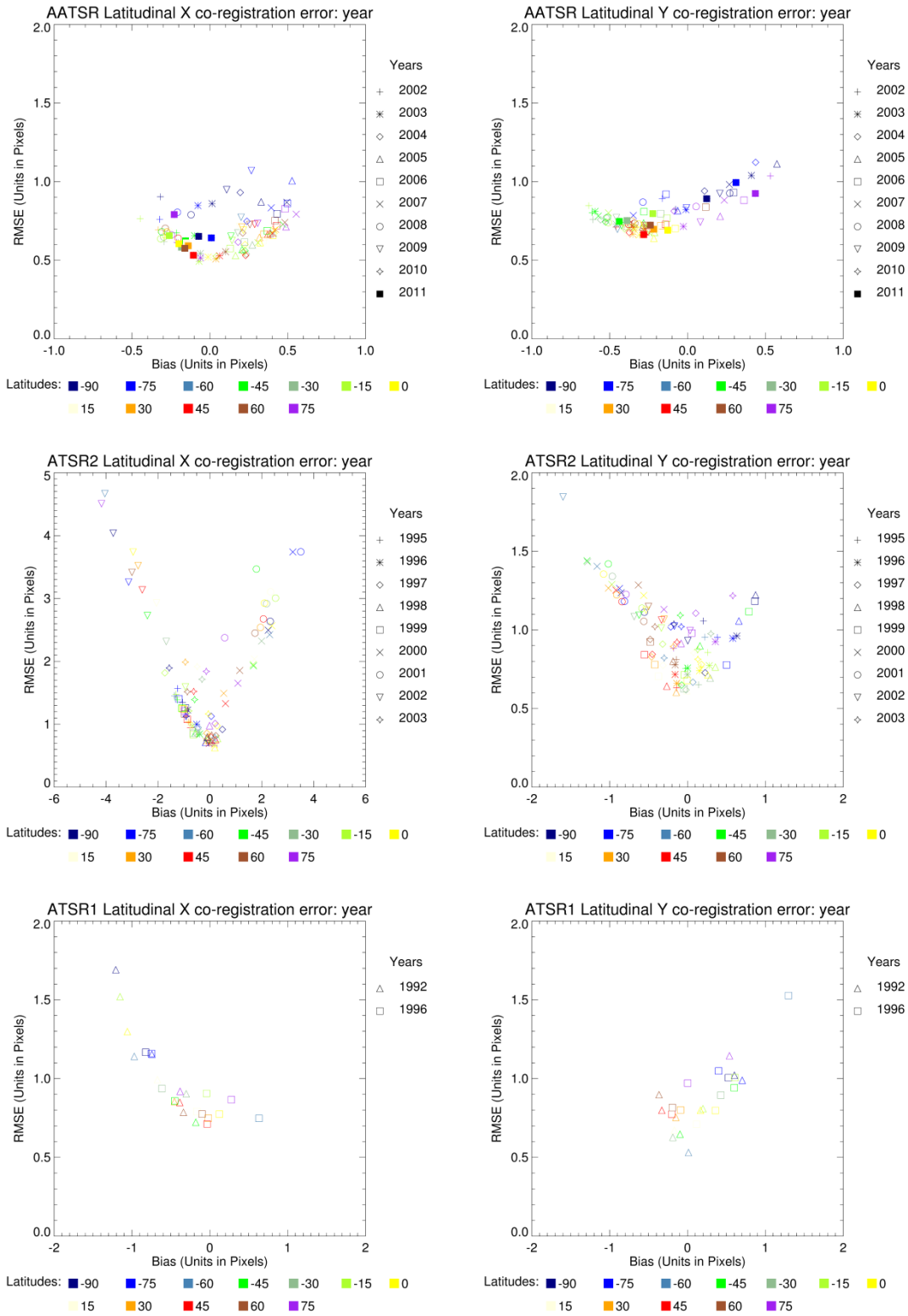


Figure 2-13. These six plots show the combined channel across (X) and along (Y) track co-registration errors by latitude for each of the ATSR instruments as determined by the global assessment. The left column contains the across track errors and the right column the along track. The top two plots are for AATSR, the middle two for ATSR-2 and the bottom two for ATSR-1.

with the same setup as for the latitudinal analysis but the colours represent longitude rather than latitude. Looking firstly at AATSR along track co-registration, the results are well distributed with no apparent longitudinally dependent trends exhibited. The RMSE is below 1 pixel for all measurements and the bias is generally in the range -0.5 to 0.0 pixels. For the across track co-registration, again all scores are below 1 pixel for the RMSE and the bias is distributed through the range ± 0.5 pixels.

ATSR-2 exhibits similar behaviour in its along track RMSE and bias scores as presented in the latitudinal along track analysis with the post 2000 scores being generally greater than 1 pixel in terms of RMSE. The same can be said for the across track analysis, with large deviations from zero being present in the score post 2000 and some results from 1995 and 1999 showing co-registration quality outside of that specified for sub-pixel accuracies.

For ATSR-1 the along track longitudinal results have bias scores which are typically less than 1 pixel and biases in the range ± 0.5 pixels. One score from 1996 for matched points located in the longitude range 150-180° W shows slightly greater RMSE and bias than the other longitudes/years. The across track RMSE and bias scores are generally below 1 pixel and in the range 0.5 to 0.0 pixels respectively.

2.3.6 Discussion

The results of the global analysis present a number of interesting findings with significant implications for the application of the warping coefficients globally. Overall, the AATSR warping coefficients perform very well for correcting the co-registration errors between the nadir and forward views in both the along and across track directions. The combined channel analysis shows limited variation between channels with each year clustered with similar bias values. The RMSE scores are quite consistent across years with values in the range 0.8 pixels (± 0.15 pixels) for the corrected scenes. The longitudinal analysis demonstrates an apparent drop in performance in the ability of the warping coefficients to improve the co-registration at higher latitudes, most notably in the southern hemisphere below 60°. The drop in performance can be attributed to a number of other factors rather than a decline in the quality of the co-registration. The most significant of these factors is the limited amount of landmass in the southern hemisphere below 60°. This leads to fewer points being matched allowing outliers to affect the statistical analysis to a greater degree. Further, the land mass is relatively spectrally homogenous especially in southern hemisphere winter. As the matching algorithm is based on image texture, homogenous regions lead to an increase in the number of mismatched points and points with low texture are challenging to remove through the ALSC quality assessment algorithms. Thus more mismatched points are likely to be present at these latitudes. Lastly, poor AATSR cloud masking in these regions further exacerbates the problem. However, considering these challenges the overall quality of the matched points at these latitudes is not significantly greater than 1 pixel indicating that the majority of the points are likely to be good matches. This demonstrates that the algorithm correction is still performing well in these regions.

The biases for high quality sets of measured locations (i.e. at latitudes below $\pm 60^\circ$ of latitude) in the across track direction is below ± 0.3 pixels for AATSR indicating high confidence in the quality of the results. The biases for the along track are shifted with the high quality measured locations in the range -

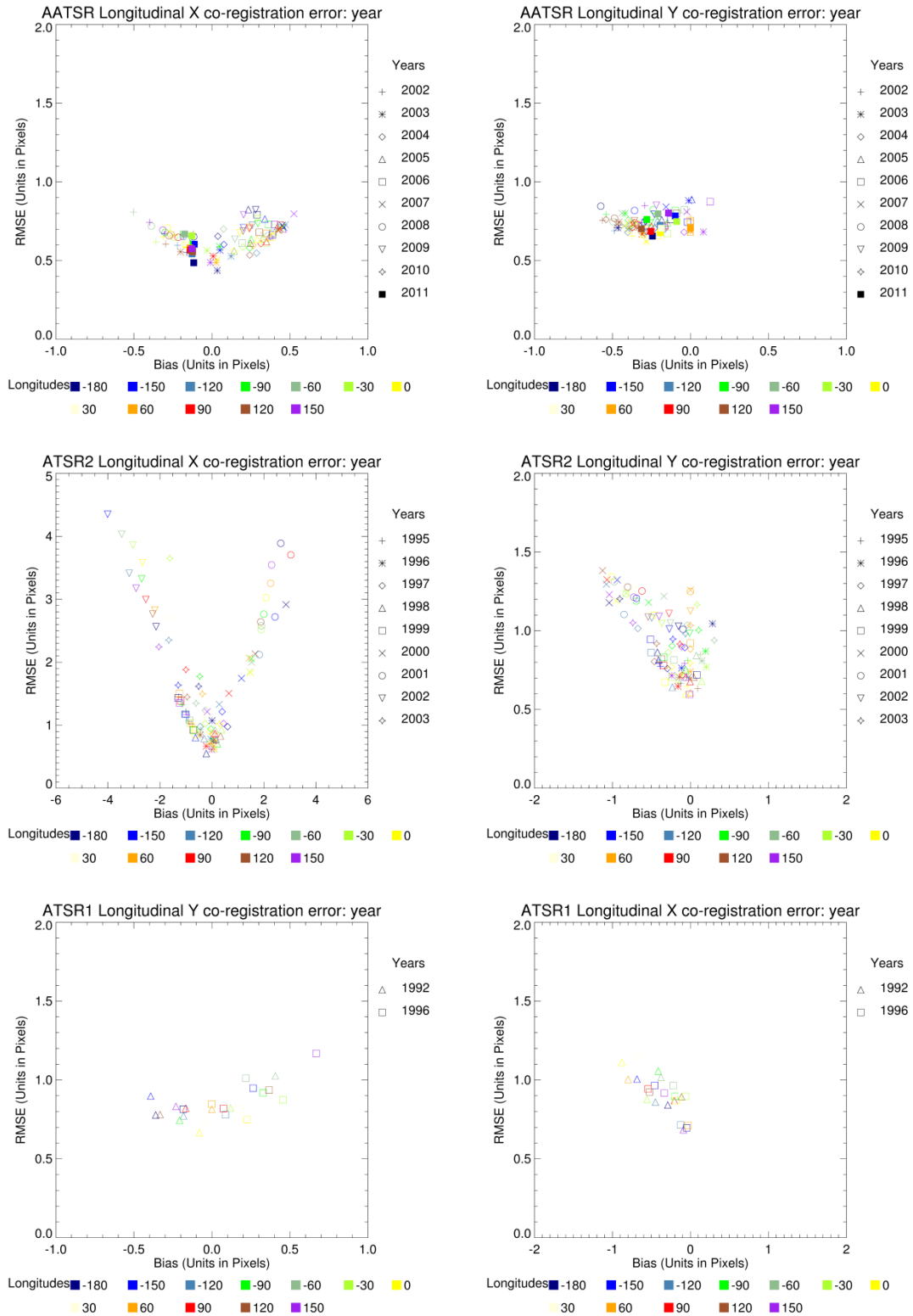


Figure 2-14. These six plots show the combined channel across (X) and along (Y) track co-registration errors by longitude for each of the ATSR instruments as determined by the global assessment. The left column contains the across track errors and the right column the along track. The top two plots are for AATSR, the middle two for ATSR-2 and the bottom two for ATSR-1.

0.1 to -0.7 pixels. These biases are attributable to the elevation normalisation process and the algorithms and parameters it employs. The main cause is likely to be the Mannstein camera model, which is used to convert the DEM height for each feature location in the nadir image into a disparity to provide the elevation normalisation. This camera model in another study (Denis et al. 2007) was shown to provide differing elevation values from a reference DEM, with a tendency to underestimate the disparity value that the elevation corresponds to, particularly for lower heights (< 3km). This leads to the slight underestimation of the elevation disparities used in the elevation normalisation process and hence leads to the introduction of the negative bias. In the independent study performed on low lying terrain (where no camera model was applied) presented in section 2.3.3.2, the biases were more in line with a suitable range of deviations from zero to -0.2 pixels. The larger negative biases in the along track direction can be attributed to the normalisation process and not any defect in the warping coefficients.

The improvement in co-registration for ATSR-2 is more variable, with certain years performing very poorly both across and along track. Focusing firstly on the years when the warping coefficients perform adequately, similar conclusions on quality can be made to those given for AATSR. Though the only years where the co-registration is improved to acceptable levels are 1996, 1997 and 1998. The across track individual channel results for 1995 and 1999 have an RMSE value of just greater than 1 pixel. However, the bias is also around -1 pixel the cause of this is not currently understood. For the years from 2000 onwards, there is a distinct drop in performance of the warping coefficients. This is not entirely unexpected and can be explained by ERS-2 operating in mono-gyro mode from January 2000 (Accica and Goryl, 2002). Seiz (2003) also showed that this has a significant effect on the co-registration between the nadir and forward views, although very few scenes were assessed. Here, we have shown that the co-registration for ATSR-2 for years following 2000 is not correctable using a single set of warping coefficients and corrections will have to be undertaken on an orbit by orbit basis.

For ATSR-1, an incomplete time series analysis was performed, however the years analysed and the findings from Casadio (2010) can be used to confirm that the warping coefficients will provide at worst pixel-level correction to the co-registration errors both along track and across track for the entire ATSR-1 data set. The along track channel analysis found the assessed channels (1.6 μ m, 11 μ m and 12 μ m) co-registration was improved to sub-pixel level accuracy, with RMSE scores below 0.9 pixels and biases in the range ± 0.4 pixels. Similar results were found in the across track analysis. In the latitudinal analysis similar results were found both along and across track as for AATSR, with latitudes below -45° exhibiting greater errors, leading to the same conclusion as given for AATSR. One unexplained outlier is present in the across track co-registration error for 1992 for the latitudinal range $0-15^\circ$, the cause of which is likely to be mismatched points in the global assessment. For longitudinal analysis, the across track errors have RMSE less than or very close to 1 pixel and the biases are in the range -0.5 to 0.0 in the majority of cases. For the along track longitudinal analysis, all of the points bar one are sub-pixel in terms of RMSE and the bias for these points range between ± 0.5 . The outlying point is for the longitudinal range $150-180^\circ$ W for 1996 and has an RMSE of near 1.2 pixels and a bias of 0.7. Whilst these values are not exceptional they can be used to explain the poor quality of the longitudinal co-registration assessment results in certain regions. Similar to the cause of errors in latitudes below 60° S, certain longitudes have limited land-mass leading to poorer results in certain geographic regions.

2.4 Chapter Summary

This chapter has introduced the ATSR instrument series. Furthermore, driven by the necessity to perform accurate stereo-photogrammetric measurements from ATSR, an automated method for correcting the co-registration errors present in the instrument datasets to, at worst, pixel level accuracy has been described (see also Fisher and Muller, 2013). The approach developed employs an automated tie-pointing algorithm based on the SURF algorithm, followed by application of polynomial derived warping coefficients. This technique has been employed to derive yearly warping coefficients from the 0.55 μ m channel for both AATSR and ATSR-2 and a single set of coefficients from the 1.6 μ m channel for ATSR-1. The quality of each warp has been evaluated independently on the scenes used for their derivation to provide an initial performance assessment. In this initial analysis, the automated tie pointing and warping methods were shown to perform well, with the imagery being effectively co-registered to pixel level accuracies. A global analysis on a subset of orbits has also been undertaken to assess whether the warping coefficients provide suitable corrections universally for each instrument and year or mission period. This analysis shows that for AATSR the derived coefficients work well in all instances and the method provides a suitable correction to the co-registration. For ATSR-2 the corrections work well in most cases pre-2000. For ATSR-1 the derived coefficients work effectively for both test years and provide a suitable correction for the co-registration.

3 STEREO-PHOTOGRAMMETRY

3.1 Introduction

Stereo-photogrammetry is the art of using stereo overlapping images with known imaging geometry to obtain 3D information about the objects contained within the scene under observation (Barnard and Fischler, 1982). It has wide ranging application. Since its inception it has been applied extensively in topographic mapping tasks (Panton, 1978; Day and Muller, 1989; Hirano et al., 2003) and more recently, with advances in computational power, automated navigation systems in robotics (Murray and Little, 2000), object recognition (Helmer and Lowe, 2010), and also in the generation of global environmental datasets from satellite imagery (Moroney et al., 2002; Muller et al., 2002; Nelson et al., 2008; Martin et al., 2010). It is the latter of these applications which is of interest in this thesis, with Chapters 4 through 6 examining the application of stereo-photogrammetric techniques to the ATSR instruments for various tasks within the field of atmospheric sciences.

Section 3.2 introduces the critical concepts of stereo-photogrammetry: the imaging system, the stereo matching algorithm, and the camera model. Of these three aspects, in many ways the most important is the matching algorithm and this, therefore, is where the majority of research occurs in stereo-photogrammetry: stereo matching algorithm development. Section 3.3 reviews the main stereo matching challenges encountered in this thesis, covering those caused by the ATSR instruments and those caused by the environment, in order to demonstrate why the majority of stereo matching algorithm development and assessment studies are often inappropriate in terms of assessing real-world applicability. A pertinent real-world assessment study by Hirschmuller and Scharstein (2009) is then reviewed in order to select an algorithm well suited to the stereo matching tasks undertaken in this thesis. The algorithm chosen, due to its excellent robustness to image radiometric variability and reduced propensity to introduce smoothing artifacts, is the census algorithm (Zabih and Woodfill, 1994).

Section 3.4 introduces the three algorithms applied in thesis, M4 (Muller et al., 2007), the census algorithm, and M6 (Fisher et al., 2013). The introductions provide algorithm design considerations and mathematical descriptions. Following their description two assessments are undertaken in section 3.5. The first looks at the implications of critical stereo matching parameters on the performance of M4 and census. The second provides a brief inter-comparison study between all three algorithms and a state-of-the-art DEM in order to provide an initial assessment of their capabilities when applied to AATSR data.

3.2 Stereo Photogrammetric Concepts

3.2.1 The Imaging System

The stereo photogrammetric technique relies on the principle of parallax in order to obtain three-dimensional measurements of an object. Parallax is the apparent change in relative position of a stationary object when viewed from different positions (Lillesand et al., 2004). Objects which are closer to the observation plane demonstrate a larger parallax than those more distant (see Figure 3-1), and it is this phenomenon which can be used to determine the distance of an object from the sensor, or conversely, above a reference datum such as an Earth ellipsoid.

FIGURE REMOVED DUE TO THIRD PARTY CONFLICTS

Figure 3-1. This figure, adapted from Lillesand et al. (2004), demonstrates the concept of parallax. The two images in the figure are obtained from different locations, L and L' . These images overlap and both observe features A and B , with B being nearer to the observation plane than A . Due to greater proximity and the parallax effect, the displacement between b and b' , the locations of B in each image, is greater than that observed between a and a' , the locations of A . These displacements, being elevation dependent in terms of magnitude, can be used to infer 3D location.

Satellite imaging systems, such as the ATSR instruments, are located in fixed orbits. This makes obtaining views from differing positions challenging. There are a number of ways of surmounting this problem; the most common approaches being imaging systems either with moveable optics or, alternatively, multi-view/multi-angle setups, preferably along-track. The ATSR instruments employ the latter, with a dual view conical scanning arrangement. To demonstrate ATSR's capability, a stereo image pair of the Himalayas from AATSR is provided in Figure 3-2 and the strong parallax effect caused by the extreme elevation of the mountains is evident between the forward and nadir views.

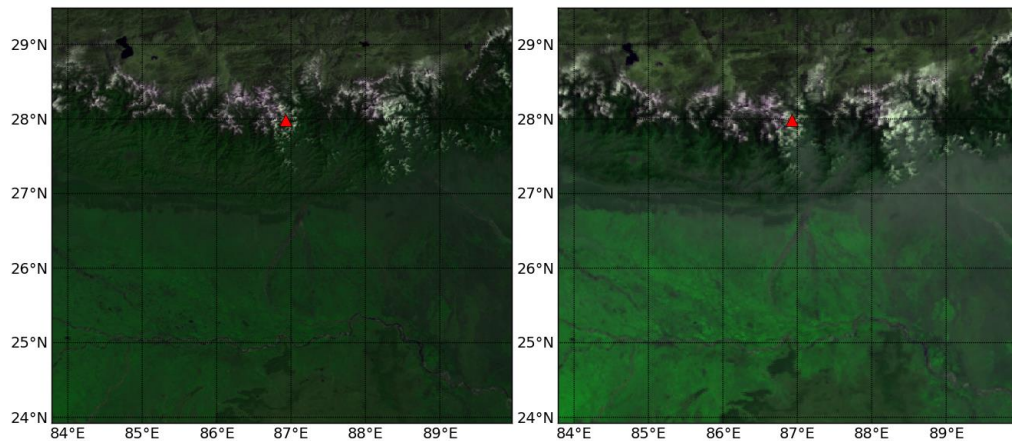


Figure 3-2. The left image is a false colour composite (R: $0.67\mu\text{m}$; G: $0.87\mu\text{m}$, B: $0.55\mu\text{m}$) derived from AATSR nadir view for orbit 36473. The red triangle within the image is Mount Everest. The right image is the same but observed from the forward view. The first observation that can be made is the displacement of the mountainous features in a Southerly direction due to the parallax effect. Another observation is the blurred appearance of the AATSR forward view; this is due to the reduced resolution of this view. The final observation is the increased haze in the lower Eastern half of the forward view; this is due to an increased atmospheric path length. The latter two effects can have significant impacts on matching accuracy and are considered in depth in this chapter.

One critical aspect of a stereo capable imaging system is that its design, specifically the instrument base-to-height ratio, impacts on the potential accuracy with which the measure of parallax can be obtained. The base-to-height ratio for a satellite-borne sensor is the distance between the stereo acquisitions, the base, divided by the elevation of the satellite above an Earth ellipsoid, the height. This ratio determines the amount of parallax features exhibit in the obtained stereo imagery (Lillesand et al., 2004). Small base-to-height ratios lead to limited parallax induced displacement between the images, reducing the potential accuracy with which the distances can be measured. Large base-to-height ratios will exhibit large parallaxes, which in turn may lead to excessive projective distortion of the imagery between acquisitions and therefore also lead to an accuracy reduction. Many stereo photogrammetric imaging systems (Hirano et al., 2003; Li and Liu, 2004) have base-to-height ratios of between 0.5 and 1, as this provides a suitable trade-off. The ATSR instruments have a base-to-height ratio which varies between 0.7 and 1.2. This variation is due to the conical scanning operation of the instruments which causes the base separation to vary between $\sim 560\text{km}$ and $\sim 960\text{km}$ depending on across track position.

In combination with the instrument spatial resolution and the accuracy of the stereo matching algorithm, the base-to-height ratio can be used to estimate the instruments stereo height retrieval accuracy as follows,

$$\Delta h = \frac{a}{(B/H)} \times r \quad \text{Eq. 3.1}$$

where B is the base distance between acquisitions in metres, H is the height of the instrument in metres, a is the accuracy of the stereo matching algorithm in pixels, r is the instrument resolution in meters, and Δh is the achievable height accuracy also in meters. For the ATSR instruments Δh varies between 830m and 1430m for pixel level stereo matching accuracy (Seiz, 2003).

3.2.2 The Stereo Matching Algorithm

Given a stereo capable imaging system, an image processing technique is required to obtain measures of parallax; this technique is referred to as stereo image matching. The aim of any stereo matching algorithm is to locate corresponding image locations (i.e. pixels) in the stereo imagery. Once correspondences have been found, the pixel-wise distance between them in the parallax direction is the measure of parallax. Assuming the direction of parallax is in y axis, in the along-track direction, as is the case for the ATSR instruments, then parallax can be defined as

$$v_x = y_x^R - y_x^C \quad \text{Eq. 3.2}$$

where v_x is the observed parallax at image location x , and is determined by differencing y_x^R , the y -coordinate at the location of x in R , the reference stereo image, and y_x^C the y -coordinate at the location of x in C , the comparison stereo image. The conical scanning geometry of the ATSR instruments, and the associated azimuthal angle variation, does not invalidate Eq. 3.2 for the determination of parallax with ATSR. This is due to parallax displacements occurring only parallel to the line of flight (Lillesand et al., 2004), that is the along-track direction for ATSR. To assess whether this assertion is accurate for ATSR, an inter-comparison has been made with collocated AATSR and CALIOP observations used in an accuracy assessment analyses undertaken in Chapter 4. In the analysis undertaken here the effect of AATSR across track pixel number on the stereo CTH retrieval is used as a proxy for assessing azimuth angle effects, as pixels nearer the swath edges are observed from greater azimuth angles. If azimuth angle has an effect on the stereo retrieval then it would be expected to result in some form of non-linear, parabolic, behaviour as a function of across track pixel when compared against the higher accuracy CALIOP observations. The plot shown in Figure 3-3 demonstrates for the measure of bias that no such behaviour occurs, rather there is a linear change in bias of $\sim 100\text{m}$ across the AATSR swath. Other statistical measures (RMSE and standard deviation) assessed also demonstrated linear behaviour as a function of across track pixel.

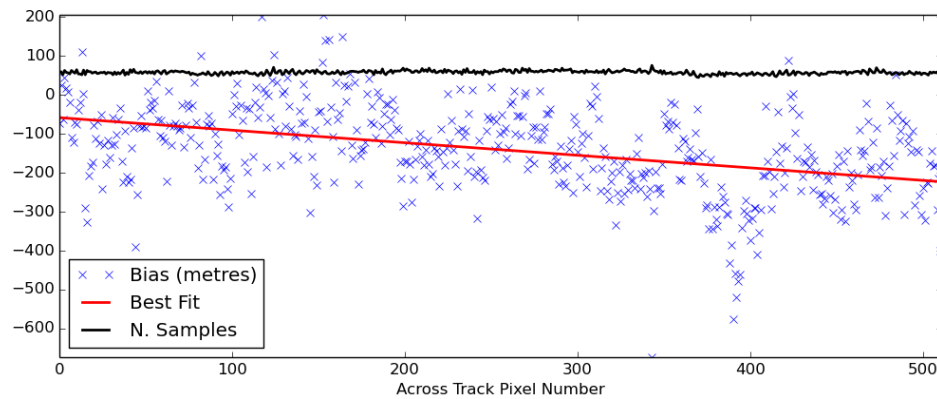


Figure 3-3. This plot shows the biases observed between collocated AATSR 11 micron channel stereo cloud top heights and CALIOP cloud top layer observations across the AATSR swath. The number of samples is the total AATSR to CALIOP collocations for the given across track pixel (on average 60 collocated observations per AATSR pixel). The behaviour of the bias between the AATSR and CALIOP observations is linear as a function of across track pixel, as demonstrated by the best fit line. The data employed comes from the AATSR inter-comparisons undertaken against CALIOP in chapter 4.

The measure of parallax is often referred to as disparity, especially when it relates to imaging arrays (Marr, 1982). The concept of a disparity can be broadened to include displacements not only in the direction of parallax, but also in the orthogonal axis. For ATSR, these x , or across-track, disparities are important, as they can be used to measure other atmospheric variables, such as across-track wind speed, as demonstrated in Chapter 6. For ATSR, two distinct disparity maps can therefore be generated, one for the along-track direction to measure elevation induced parallax, and another for the across-track direction to measure displacement effects, such as wind from cloud tracers.

Accurately measuring disparity is a challenging and very computationally expensive task, and as such, there is a voluminous body of work on the development of stereo matching algorithms that are able to achieve this feat effectively and efficiently (<http://vision.middlebury.edu/stereo/>; Scharstein and Szeliski, 2002). Stereo matching algorithms can be split into three main groups: feature based methods (Bolles et al., 1993; Hsieh et al., 1992), local (or window based) methods, and global methods (Scharstein and Szeliski, 2002).

Feature based methods locate and match features, such as corners, edges, or other points of “interest”, between stereo imagery using algorithms such as SURF introduced in Chapter 2. The matched features are used to generate measures of disparity. Once measures of disparity have been obtained at feature points, interpolation routine such as Delaunay triangulation are employed to fill in unobserved locations. Feature based methods do not perform effectively when applied to the atmospheric stereo tasks. This is due to the interpolation schemes employed assuming the presence of continuous surfaces, which is commonly not the case. For example, assume the simple situation of an along-track disparity observed for a cloud feature, and another for a nearby land feature. Assuming a linear interpolation routine is employed to generate the unobserved disparities between the two features, the disparities from the cloud

to the land feature would decrease linearly, which is unlikely to be an accurate physical representation. In order to generate an accurate interpolation result, a feature would be required on each side of the land/cloud boundary. This requirement may not always be fulfilled leading to poor interpolation results. For this reason feature based stereo algorithms are not considered for the atmospheric applications undertaken in this thesis.

The local methods which encompass the second group of algorithms, are able to more effectively represent the piecewise smooth (smooth, but not continuous) surfaces typically found in atmospheric stereo tasks, as they attempt to generate a disparity observation for every pixel, thus avoiding interpolation errors. To generate correspondences, local methods generally employ three distinct computational stages: cost computation, cost aggregation, and disparity assignment (Scharstein and Szeliski, 2002). Cost computation involves assessing some metric computed between a patch of pixel values taken from the reference image, N_R against a patch of pixel values from the comparison image, N_C . The most commonly employed metrics include the sum of absolute differences, $\sum |N_R - N_C|$, and the sum of the squared differences, $\sum (N_R - N_C)^2$. The chosen metric is computed at varying disparities, the range of which is defined by a search window (either in one or both axes) in the comparison image. For each disparity assessment a cost aggregation step is typically performed. This generally involves a summing or averaging of the costs over a finite window, effectively applying a smoothness constraint to the costs and therefore, the disparities (Scharstein and Szeliski, 2002). This smoothing helps reduce noise in the resultant disparity maps. The final step, disparity assignment, involves assessing all the disparities and associated aggregated cost results for a given pixel and defining the disparity with the optimal cost. There are numerous methods employed to achieve this, the most basic is winner takes all (WTA) where the disparity associated with the minimum cost is selected, based on the assumption that the minimum cost indicates the best match. More complex techniques can be employed to find sub-pixel disparity estimates using interpolation between cost values.

Global stereo algorithms, the final group, attempt to minimise some cost function over the entire image rather than for windowed subsets as with local algorithms. The cost function is generally comprised of two separate terms, one associated with the data and another associated with smoothness (Scharstein and Szeliski, 2002). The data cost term is computed using methods similar to those employed by local algorithms. The smoothness cost term is typically determined by comparing the gradients of either the disparities or pixel intensities within the local neighbourhood of each pixel. The challenge of global methods arises in finding a global solution which minimises both the data and smoothness cost terms in the cost function. There are a large number of techniques which can be employed to locate this solution, examples include: dynamic programming (Ohta and Kanade, 1985), belief propagation (Sun et al., 2003), and perhaps the most commonly employed, graph cuts (Boykov et al., 2001). The final disparity set is the result of the minimisation process.

There is no single best stereo matching algorithm. Rather, both global and local methods have their particular benefits and drawbacks. Global algorithms have the greatest potential for accuracy (fewest erroneous disparity measurements) as shown by their prevalence near the top of the rankings in the one hundred plus inter-comparison studies contained in the Middlebury stereo algorithm evaluation database

(<http://vision.middlebury.edu/stereo/>; Scharstein and Szeliski, 2002). Local algorithms achieve lesser accuracy as they make the assumption that the disparities within the cost computation and aggregation windows are constant. This is not the case at discontinuities⁵ and leads to the occurrence of blurring artifacts at such interfaces (Hirschmüller et al., 2002). The pixel-wise techniques of global algorithms are able to avoid these artifacts as they do not employ local smoothness assumptions. However, whilst local algorithms tend to be less accurate, they have a number of potential benefits over global algorithms, including: ease of implementation; low processing overheads; and effective operation without epipolar geometric constraints.

Now examining these three benefits in detail, firstly it is the case that local algorithms are generally easy to implement, and can be written to perform efficiently in high-level programming languages such as Python or IDL. Most global algorithms necessitate development in C or similar low-level languages to perform efficiently, which increases their complexity and therefore development time. The second aspect, low processing overheads, is a significant issue, as even when written in a low-level language, the majority of global algorithms tend to exhibit significant memory consumption and require a large number of computational clock cycles to complete processing (Mei et al., 2011). Recent advancements have been made on porting global algorithms to graphics processing units (GPU) and excellent performance has been demonstrated (Mei et al., 2011). However, GPU porting is complex, and the computational facilities available at MSSL, in general, are CPU based. Local methods can perform quickly on CPU based systems (see section 3.5) and this is important when processing large quantities of images, as is necessitated in this thesis (see Chapter 4). The third aspect is that of epipolar constraints, which nearly all stereo processing systems make extensive use of. Epipolar constraints employ knowledge of the epipolar geometry of the stereo imaging system. If the relative translation and rotation of the sensors used to generate the stereo pair are known, then it is possible to reduce the search for a corresponding pixel pair from one of a 2D problem to that of a 1D problem, where the disparity search is performed along an epipolar line (see Figure 3-4). This constraint cannot be employed in atmospheric stereo tasks as clouds and other atmospheric features are often in motion, making such constraints invalid. Furthermore there has been no work undertaken on understanding the application of epipolar constraints to conical scanning instruments such as ATSR and epipolar assumptions are made based on a conventional 2D frame camera system, and are therefore not applicable to conical scanners (D. Shin pers. comm.). Not being able to apply epipolar constraints generates a number of problems: it generally precludes application of open-source global stereo matching algorithms, such as the openCV semi-global matcher (Hirschmüller, 2005), which rely on such constraints; and it also significantly increases processing time, for example, given a disparity range of 50, with epipolar constraints 50 evaluations would have to be carried out, without epipolar constraints this potentially increases to 50^2 . CPU-based global algorithms are already slow, squaring the number of clock cycles and also, potentially, the memory overhead is unacceptable for the science tasks carried out in this thesis. Lastly, in real-world applications, it has been shown that local algorithms are not substantially less accurate (~5%, depending on the algorithm) than global algorithms (Hirschmüller and Scharstein, 2009). For these reasons local stereo matching algorithms are employed over global algorithms to obtain measures of disparity from the ATSR instruments in this thesis.

⁵ A discontinuity is a change in scene depth, e.g. moving from a cloud feature to the Earth's surface

FIGURE REMOVED DUE TO THIRD PARTY CONFLICTS

Figure 3-4. This figure demonstrates epipolar lines which can be used to improve stereo algorithm processing performance given a suitable imaging system. The epipolar processing allows the search for matching stereo pixels to be reduced from a 2D problem to that of a 1D problem – where the search is only performed along the epipolar line: the line running between points p and e . Taken from

<http://bibliothec.immateriel.fr/baw/9780596516130/httpatomoreillycomsourceoreillyimages213352.png>
(last accessed 10th August 2013)

3.2.3 The Camera Model

Following determination of the parallax disparity set by the stereo matching algorithm, the measured parallaxes must be related to a 3D coordinate system in order to obtain meaningful results. This is achieved with a camera model that replicates the geometry of the imaging system. For the ATSR instruments two separate camera models have been proposed, one by Prata and Turner (1997) for ATSR-1, and another more recently by Denis et al. (2007) for ATSR-2. These two camera models are referred to as the Prata camera model and the Mannstein camera model respectively. Both models have a similar form to calculate height,

$$H = \frac{v}{\tan(\theta_f) - \tan(\theta_n)} \quad \text{Eq. 3.3}$$

where H are the returned elevation in meters, v is the ATSR parallax disparity set in pixels, θ_f is the set of ATSR forward viewing zenith angles and θ_n is the set of ATSR nadir viewing zenith angles. The two algorithms differ in the viewing zenith angle terms, and an inter-comparison study comparing the two separate models is made in the paper which introduces the Mannstein model (Denis et al., 2007). The inter-comparison assessments presented show that the Mannstein model is on average closer to ground

truth measurements (taken from the GLOBE DTED-0 DEM (Hastings and Dunbar, 1998)) than the Prata model, with mean differences of 1.13 km and 1.35 km respectively, when assessed over seventeen 512 by 512 pixel cloud free scenes from the ATSR-2 11 μ m channel. Both camera models share similar standard deviations of around 1.2 km. From this analysis, both camera models are close to the achievable ATSR accuracy determined by the base-to-height ratio, as reported in section 3.2.1, which ranges from 0.8 to 1.2 km.

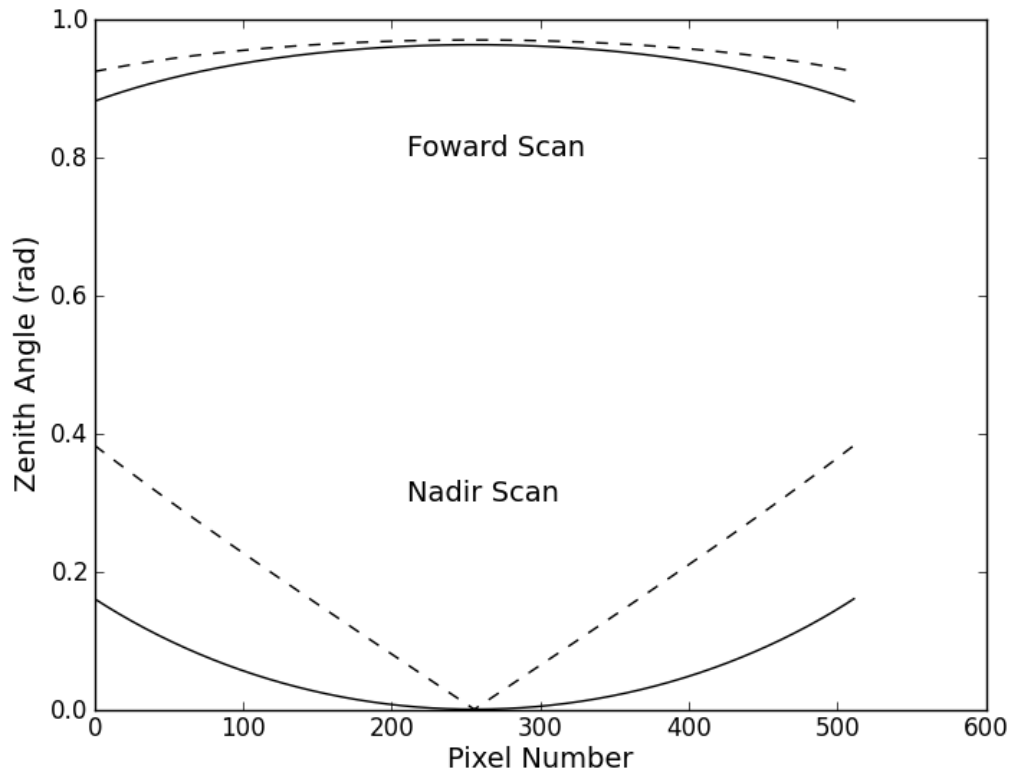


Figure 3-5. This figure show the zenith angles for the Prata and Mannstein camera models. The Prata model is represented by the dashed line, note the poor representation of the nadir view scan angles. The Mannstien model, represented by the solid line, has a much better representation of the ATSR conical scanning geometry leading to more accurate heights.

The reasons for the slightly improved accuracy of the Mannstein camera model is an improved representation of the nadir scan angles, leading to a parabolic curve similar to the instrument geometry. The Prata model on the other hand, has poor nadir scan angle representation, and represents the scan arc as two linear features centred on the nadir sub-satellite pixel (See Figure 3-5). This significantly underestimates the baseline, B , particularly toward the edges of the ATSR swath. Due to its improved accuracy, the Mannstein camera model⁶ is applied to convert disparities to elevations in all cases in this thesis.

⁶ Whilst the Mannstein camera model is applied in this thesis to compute the ATSR viewing angles, it should be noted that this is no longer necessary – the viewing angles are now contained within the Envisat products. Eq. 3.3 can be used with the provided angles to calculate heights.

3.3 Stereo Matching Challenges

The majority of studies undertaken in the field of stereo matching algorithm development tend to focus on achieving optimal performance on controlled laboratory generated stereo datasets such as the Middlebury dataset (<http://vision.middlebury.edu/stereo/>; Scharstein and Szeliski, 2002). The stereo images which comprise datasets such as this are controlled in a number of ways: the illumination conditions are regulated and consistent; the scenes observed tend to have the good spatial heterogeneity required for effective stereo matching; the cameras used generate consistent images and are well calibrated; and the images are post-processed with rectification algorithms, allowing application of epipolar constraints. Such conditions allow for excellent algorithmic evaluation, however, there is an inherent drawback. When selecting an algorithm for application to real-world tasks, such as those encountered in this thesis, these controlled algorithm assessments, due to their unrealistic assessment characteristics, do not provide reliable quantification of real-world algorithm performance.

Fortunately a recent study by Hirsh Müller and Scharstein (2009) has provided a robust analysis on the performance of algorithms under real-world conditions, allowing justifiable algorithm selection for real-world stereo matching tasks. This section, before reviewing this assessment study, firstly outlines the main challenges encountered in real-world stereo matching, focusing on the environmental and ATSR specific instrumental factors that need to be accounted for in order to generate reliable stereo matching outcomes.

3.3.1 Stereo Image Dissimilarity

One of the main challenges to real-world stereo matching performance is that of stereo image dissimilarity, which can be defined as any radiometric (in the case of ATSR, brightness temperature or reflectance values) difference between observed pixels values for corresponding stereo pixels. Such radiometric differences, without some form of correction, lead to ambiguous stereo cost evaluations, as an incorrect disparity may correspond to the optimal cost. There are various causes of radiometric differences, stemming from either the sensing system or the feature being observed. Some are present both in the laboratory environment and the real-world, such as the non-Lambertian qualities of the feature being observed. Others that are present in the real-world are eradicated from the laboratory environment through the regulation of the illumination conditions and the camera setup and calibration. Of the radiometric differences encountered in this thesis, three separate causes can be identified: those caused by Lambertian effects; those caused specifically by the ATSR sensing system; and those caused specifically by clouds and/or smoke plumes. Each cause is dealt with separately in sections 3.3.1.1, 3.3.1.2 and 3.3.1.3.

Furthermore, it should also be noted that stereo matching ambiguity can also be introduced from sources other than image radiometric differences, the prime example being that of occlusion, where a pixel is visible in one stereo image but not in the other(s). This leads to ambiguous results in the stereo outcome

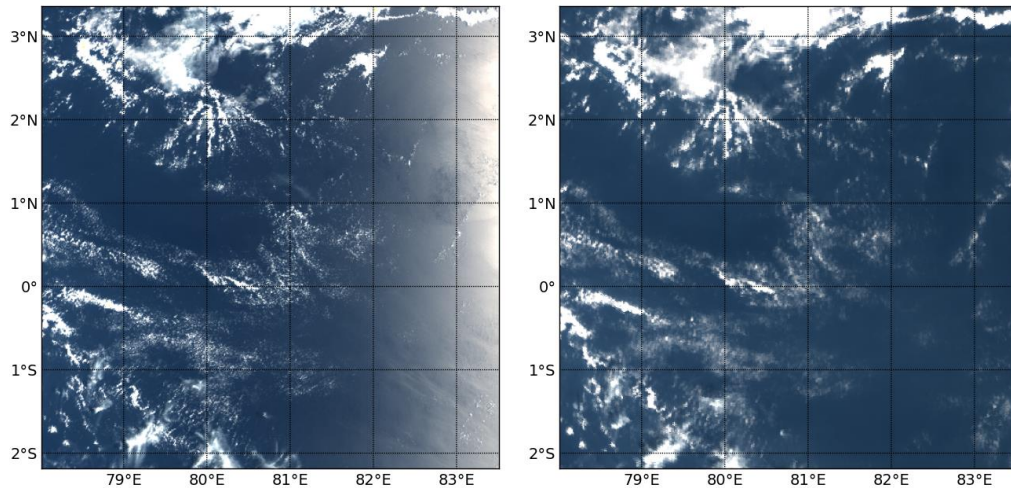


Figure 3-6. The left image is a histogram stretched false colour composite (R: $0.87\mu\text{m}$; G: $0.67\mu\text{m}$; B: $0.55\mu\text{m}$) from the AATSR nadir view (Orbit: 36473). The right image is the same from the forward view. Firstly, note the brightening effect associated with the sun glint on the Eastern edge of the nadir image. This effect is not apparent in the forward view, demonstrating non-Lambertian behaviour. Also, the differing atmospheric path lengths lead to differing cloud characteristics between the forward and nadir views.

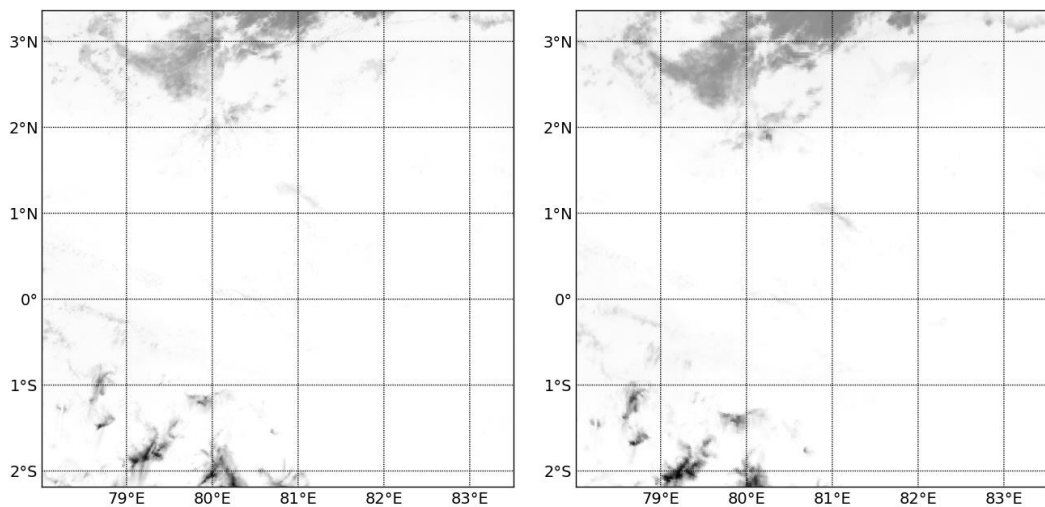


Figure 3-7. The left image is from the AATSR $11\mu\text{m}$ nadir view for the same scene presented in Figure 3-6. The right image is that of the forward view. Note the lack of scattering artifacts due to the long wavelength of observation. Also, note the darker appearance of cloud features in the forward view; this is due to reduced brightness temperature, which in turn is due to increased atmospheric path length.

as there are no true matching pixel pairs between the stereo imagery in occluded regions, so any match found is erroneous. Another important example is that of image homogeneity, i.e. image regions where there is no discernible structure. In such instances, local stereo matching, depending on the size of the

cost window, is ambiguous, as without identifiable image structures the local stereo matching algorithms fail to locate correct correspondences. Contrary to stereo image dissimilarity, which can be improved through image pre-processing steps, the errors introduced by occlusion and homogeneity can be dealt with effectively through application of post processing algorithms (Hirschmuller, 2008; Chapter 4) and are considered in this thesis on a task by task basis.

3.3.1.1 Lambertian Related Differences

Lambertian surfaces are those which scatter radiation equally in all directions. Due to this characteristic they look the same irrespective of viewing angle, and can be considered isotropic. There are no real-world surfaces that are truly Lambertian (Schaub et al., 2011), and therefore all surfaces look different to a certain degree when viewed from different viewing angles. Observing a feature from differing viewing angles is a key requirement of stereo photogrammetry for satellite imaging platforms, therefore the anisotropic surfaces of nature lead to radiometric differences between the stereo imagery and need to be accounted for. An example of anisotropy seen in ATSR imagery is presented in Figure 3-6.

Additionally, it must be understood that scattering effects predominate in the visible and near-infrared regions of the electromagnetic spectrum (c.f. the thermal region for example). For the ATSR thermal channels, which observe terrestrial emissions, scattering effects are of secondary importance to absorption and emissions effects and so, Lambertian image dissimilarities are limited to the visible channels (this is demonstrated in Figure 3-7).

3.3.1.2 ATSR Related Differences

The ATSR instruments introduce a specific challenge to the stereo matching task due to variable resolution between looks. As described in Chapter 2, the ATSR nadir view is comprised of 555 pixels resampled to 512 one kilometre pixels; the forward views is comprised of 371 overlapping pixels interpolated to the same size using nearest neighbour interpolation. This difference in sampling results in a significant reduction in the forward image resolution when compared to that of the nadir, as demonstrated in Figure 3-2. The different image resolution means that there are no direct correspondences (and in the context of this section, radiometric matches) between any of the forward and nadir pixels.

3.3.1.3 Cloud and Smoke Plume Related Differences

The forward view path length of the ATSR instruments can be up to 74% (800km vs. 1395km) longer than that of nadir depending upon scan position. Such differences in atmospheric path length can greatly vary the radiometric qualities of cloud and smoke plume features in the stereo imagery. These differences occur primarily due to changes in the apparent optical depth, where optical depth is the negative natural logarithm of the amount of radiation that is not scattered along a given path. Clouds (stratiform) and smoke plumes tend to have a greater horizontal extent than vertical; and as such, when viewed in the ATSR visible/NIR channels the amount of radiation scattered back to the forward looking view will be greater than that of the nadir view (see Figure 3-2 & Figure 3-6). Such path length effects are also present in the thermal channels (see Figure 3-7). It should be noted that these effects

predominate for optically thin clouds, as over a certain optical depth the scattering or radiance contribution is similar irrespective of the view.

Another factor which needs to be considered is that clouds and smoke plumes are often in motion. This may lead to changes in their structure over the two minutes time difference between the two ATSR views, and also lead to errors in the parallax measurements due to displacement effects. Errors in parallax are not considered here, as they do not impact on the stereo matching in terms of radiometric dissimilarity and performance. Instead the effects of such errors are considered in the specific science tasks undertaken in Chapters 4 and 5. Structural changes caused by motion can be problematic, as they likely result in changes in Lambertian properties and therefore introduce radiometric dissimilarity.

3.3.2 Real-world Algorithm Assessments

Quantitatively judging the performance of a stereo algorithm and its ability to deal with the radiometric distortions present in real-world imagery is problematic, as the ground truth comparison data may not be of suitable accuracy. This is another factor behind the appeal of laboratory testing with well validated data; high quality disparity maps are available for assessment purposes. One solution to this problem is taking a laboratory generated dataset and applying synthesised radiometric dissimilarities to the imagery. This approach has been applied with success by Hirschmuller and Scharstien (2009), where 15 different stereo matching metrics were assessed with three different stereo algorithms (local, semi-global, and global) in the presence of a number of radiometric distortions, including Gaussian noise, gain changes⁷ and gamma changes⁸ at varying degrees of severity. Whilst such distortions may not give an identical replication of the radiometric image differences outlined in Section 3.3.1, they do provide analysis of an algorithms ability to deal with radiometric variability.

The matching metrics evaluated in the Hirschmuller and Scharstien (2009) study followed a number of different approaches for dealing with radiometric variability. The commonest approach analysed involves applying some form of parametric transformation to the image prior to cost computation. In such transformations, the radiometric distortions between the stereo images are reduced through use of parametric statistical measures such as the mean and standard deviation. A prime example, encountered in Section 3.4 is NCC, where each image pixel is normalised by subtracting the mean and dividing by the standard deviation, both derived from the local neighbourhood. Following normalisation, a correlation function is used to determine costs between the images. The normalisation process applied by NCC, at least empirically, deals with any gain changes between the imagery reducing dissimilarity prior to cost computation.

An alternative approach analysed comes in the form of non-parametric transforms. Again, these transformations replace each pixel with a new value or set of values, the difference being that the statistical measures are derived from the ordering of the pixel values rather than the actual pixel values. Examples include the rank transform (Zabih and Woodfill, 1994), which replaces each pixel with the sum of the number of pixels within a local neighbourhood which are of a lesser value than the pixel of interest,

⁷ A multiplicative change between stereo images.

⁸ A non-linear change between stereo images.

and the census transform (Section 3.4, Zabih and Woodfill, 1994) which replaces each pixel with a bit vector encoding the pixel ordering and the structure of the local neighbourhood. Once transformed, cost computation is performed using SAD in the case of rank and the Hamming distance in the case of census. These transformations are very powerful, as they are robust to all radiometric distortions which preserve the ordering of the pixels within the transformation window.

Of the cost metrics analysed in the Hirsh Muller and Scharstein (2009) study, the census transform was found to be best able to deal with all the radiometric distortions encountered for all stereo algorithm types. Previous real-world studies, whilst not specifically assessing robustness to radiometric distortions, also identified non-parametric transforms, particularly the census transform, as being more effective than parametric stereo matching techniques. It is therefore proposed that the census transform is likely to provide the most effective stereo matching technique in the presence of the radiometric differences described in Section 3.3.1 and as such the algorithm is described in Section 3.4 and compared against other stereo algorithms applied to the ATSR instruments in Section 3.5.

3.4 ATSR Stereo Matching Algorithms

The potential for stereo-derived products from the ATSR instruments was first proposed prior to the launch of ATSR-1 by Lorenz (1985), who suggested its potential for the stereo determination of cloud top height (CTH). After many years, this idea was first implemented on ATSR-1 by Shin and Pollard (1996) followed in quick succession by Prata and Turner (1997). More recently, Seiz (2003) applied stereo matching techniques to the ATSR-2 instrument and Muller et al., (2007) developed M4, an advanced stereo processing scheme for the same instrument based on previous work undertaken for MISR stereo processing tasks (Muller et al., 2002). This section, after describing M4 in detail, builds upon this history of stereo matching with ATSR by describing two algorithms new to the instruments. The first being the census algorithm, its selection justified in section 3.3.2. The second algorithm to be described is M6, developed ‘in house’ at MSSL for application within the ESA ALANIS project.

3.4.1 M4 Algorithm

3.4.1.1 Development

The M4 stereo matching algorithm (Muller et al., 2007) was developed for application to ATSR-2 data under the CloudMap2 project (Muller and Fischer, 2007) and built upon previous algorithms developed for MISR (Muller et al., 2002). M4 exhibits a substantial reduction in processing time compared to its predecessors, as it pre-calculates the disparity set using phase correlation and employs image convolution extensively. The matching metric employed in M4 is Normalised Cross Correlation (NCC), which is statistically the optimal method for dealing with Gaussian noise and is also able to account for gain differences between the stereo matching windows (Hirschmuller and Scharstein, 2009). The cost evaluation metric employed by M4 is winner takes all (WTA). The algorithm operates at pixel level accuracy, that is, integer disparities are the output from its application. M4 has been shown to be accurate

(Naud et al., 2007) to the accuracy limit defined by the ATSR geometric configuration as presented in Section 3.2.

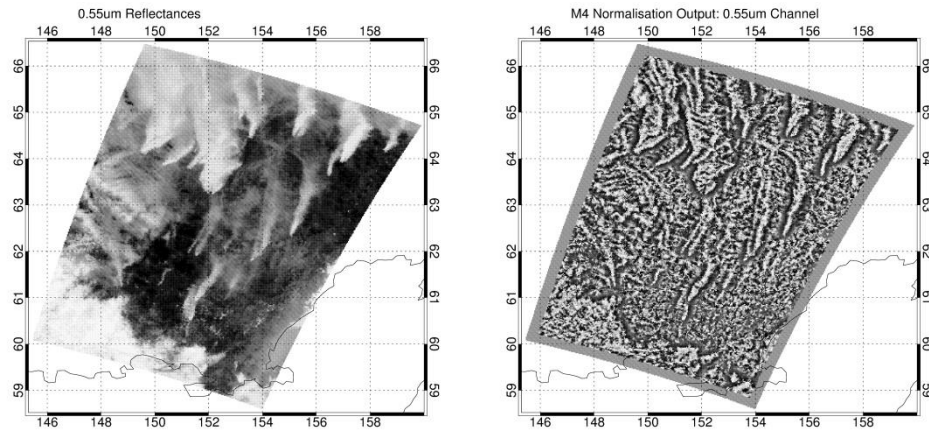


Figure 3-8. This figure demonstrates the loss of image texture caused by the parametric normalisation employed in the M4 stereo matching algorithm. The left image is an AATSR 0.55 μm nadir view from orbit 43986 of wildfire events and associated smoke plumes in the Eurasian boreal region, the right image is the outcome from the M4 normalisation process. Note the loss of image detail at the interfaces between features of differing intensity.

However, M4 is not without its limitations. The underlying NCC metric has a tendency to blur depth discontinuities more than other matching costs, as extreme pixel values within the NCC cost computation window lead to high errors in its calculation (Hirsh Muller et al., 2002). An example of this effect is shown in Figure 3-8, where there is a distinct loss of texture around smoke plume features in the normalised image. This effect occurs as the mean and standard deviation used in the image normalisation process in NCC (see section 3.4.1) assume the intensity values observed within the normalisation window have a unimodal distribution. In the image shown in Figure 3-8 this is regularly not the case, as the smoke plume and land intensity values form two distinct probability distributions. When the normalisation window straddles the two separate distributions, the calculated mean and standard deviation are unrepresentative of either distribution (Zabih and Woodfill, 1994), leading to ineffective normalisation and the loss of image texture. When the matching algorithm within NCC is applied to the normalised images, matching in the textureless regions is ambiguous which leads to the occurrence of blurring artifacts in the resultant disparity maps.

Another issue associated with M4 are non-local disparity errors introduced by its application of phase correlation to determine the disparity set for assessment (L. Brinkmann, Pers. Comm.). The cause of this error is unknown, however, they are substantial and therefore an M4 variant, referred to as M5, which does not use a phase correlation derived disparity subset, is employed in this thesis and is described in the following section.

3.4.1.2 M5 Description

In its image normalisation stage, which accounts for radiometric variability between the stereo imagery, the M5 algorithm makes extensive use of a Gaussian kernel, K , with the size (p, q) and the standard deviation, σ set to $(21, 21)$ ⁹ and $\frac{21}{4}$ respectively. The kernel is defined as follows,

$$K(k, l) = \frac{\frac{1}{2\pi\sigma^2} e^{-\frac{k^2+l^2}{2\sigma^2}}}{\sum_k \sum_l \frac{1}{2\pi\sigma^2} e^{-\frac{k^2+l^2}{2\sigma^2}}} \quad \text{Eq. 3.4}$$

where $k \in [-\frac{p-1}{2}, \frac{p-1}{2}]$ and $l \in [-\frac{q-1}{2}, \frac{q-1}{2}]$. Following definition of K it is possible to normalise the reference image, R , and the comparison image, C , using locally computed means and standard deviations derived through its application. For every pixel x in R there is an associated line coordinate i and sample coordinate j . The local mean $\bar{x}_{i,j}$ for each pixel $x_{i,j}$ is computed as follows,

$$\bar{x}_{i,j} = \sum_{k=0}^{p-1} \sum_{l=0}^{q-1} R(i+k-\frac{p}{2}, j+l-\frac{q}{2}) \times K(k, l). \quad \text{Eq. 3.5}$$

The local standard deviation, $x_{i,j}^\sigma$, is defined by the formula,

$$x_{i,j}^\sigma = \sqrt{\sum_{k=0}^{p-1} \sum_{l=0}^{q-1} [R(i+k-\frac{p}{2}, j+l-\frac{q}{2}) - \bar{x}_{i,j}]^2}. \quad \text{Eq. 3.6}$$

After using equations 3.5 and 3.6 to obtain the local mean and standard deviation for every pixel x in R , giving the means of the reference image, $R_{\bar{x}}$, and the standard deviations of the reference image, R_σ , the normalisation of the reference image can be performed,

$$R_{norm} = \frac{R - R_{\bar{x}}}{R_\sigma + \varepsilon}. \quad \text{Eq. 3.7}$$

The small positive number ε is set here to avoid zero divisions and to limit the amplification of image noise. Its default size is set equal to 10^{-3} . The same methods are also applied to normalize the comparison image,

$$C_{norm} = \frac{C - C_{\bar{x}}}{C_\sigma + \varepsilon}. \quad \text{Eq. 3.8}$$

The last step in the normalisation process is to reject from further consideration the edges of the normalised images. Edge artefacts occur in the normalisation process due to image padding required for convolution of the image with the Gaussian filter. Any pixels which are within $\frac{p}{2}$ pixels of the image boundary are set to null values, and not considered in the matching process.

Following image normalisation, the correlation metric can be computed for r different across track displacements, u , and along track displacements, v . For each displacement, the matching cost at each

⁹ As assessment of the normalization and aggregation window sizes is provided for M5 and Census in section 3.5.1.

pixel is aggregated from the local neighbourhood using a smaller Gaussian kernel, K_s , derived from equation 3.4 with the size (p, q) and the standard deviation, σ , set to $(11, 11)$ and $\frac{11}{4}$ respectively. The M5 cost is computed as follows,

$$M5(x_{i,j}, u_r, v_r) = \sum_{k=0}^{p-1} \sum_{l=0}^{q-1} \left| R_{norm}\left(i+k-\frac{p}{2}, j+l-\frac{q}{2}\right) - C_{norm}\left(i+u_r+k-\frac{p}{2}, j+v_r+l-\frac{q}{2}\right) \right| \times K_s(k, l). \quad Eq. 3.9$$

The M5 correlation costs are stored at each pixel $x_{i,j}$ for all displacements in the range defined by u_r and v_r giving a total of $u_r \times v_r$ different costs. The across track and along track disparities for a given pixel are the u_r and v_r displacement pair which minimise the M5 cost.

3.4.2 Census Algorithm

3.4.2.1 Development

The main weakness of M5 and other such parametric algorithms is their poor performance in the presence of multi-modal intensity distributions common at discontinuities, such as the cloud/land boundary example given in Section 3.2.2. The motivation behind the development of the census (Zabih and Woodfill, 1994) and other such non-parametric algorithms (Banks et al., 1997; Bhat and Nayar, 1998) is to provide improved stereo matching performance in such regions and a reduction in the smoothing artifacts associated with parametric transforms. The census algorithm achieves this through application of the census transform, which encodes local image information into informative bit vectors. A bit is set in the vector if the corresponding pixel in the local neighbourhood is of a lesser intensity than the pixel of interest. This transformation replaces the image normalisation stages employed in parametric transform such as NCC.

The strength of the transformation lies in the fact that the use of bits effectively limits the influence of statistically outlying pixels (pixels from a different probability distribution) on the pixel of interest during correction for radiometric dissimilarity prior to stereo cost computation. To illustrate this limiting effect, an example from Zabih and Woodfill (1994) is employed: assume a 3 by 3 region of pixels taken from an 8 bit image with values of $[127, 127, 129, 126, 128, 129, 177, 131, A]$ and a value of A that ranges from $0 \leq A < 255$. Firstly looking at the case of a parametric transform, such as NCC, which employs the mean and the standard deviation to account for radiometric variability prior to cost computation, as A varies through its range of values, the mean changes from 114 to 142 and the variance from 2 to 1823, indicating the substantial susceptibility of the normalisation outcome on the value of A . The census transform, on the other hand, is far more stable in its reaction to variability in A . The central value of the 3 by 3 region is 128, therefore the bit which represents A in the bit vector is 1 when $A < 128$ and 0 when $A \geq 128$. If A is an outlying value then the effect in the cost computation stage is minimised by the census transform to the size of the minority, in this instance, at most one bit.

From the above demonstration it is apparent that census transform is more robust to outliers within the local image neighbourhood than parametric transforms. Furthermore, as mentioned in Section 3.3.2, it is unaffected by all radiometric distortions as long as they do not alter the pixel ordering. The main

weakness of the census algorithm is that it is susceptible to mismatches in regions with repetitive image structures, though in the natural environment such occurrences are limited.

3.4.2.2 Algorithm Description

For any pixel x we can define the census transformation as,

$$\Gamma(x) = \otimes f(x, n) \quad \text{Eq. 3.10}$$

where \otimes is a concatenation operator which concatenates the results of the comparison function, f , to the bit vector, Γ . The comparison function returns a 1 if $n < x$ and a 0 if $n \geq x$, where $n \in N$, the local neighbourhood centred on x . For the tasks undertaken in this thesis N is set to a radius of 5 pixels as this provides suitable discriminative power, see section 3.5.1, whilst not increasing computational cost significantly.

Applying Eq. 3.10 to every pixel in both the reference and comparison images yields two 3D arrays of bits vectors. In order to locate the correspondences the Hamming distance metric is used to compare the bit vectors as follows,

$$S(x_{i,j}, u_r, v_r) = \sum \Gamma_{i,j}^R \vee \Gamma_{i+u_r, j+v_r}^C \quad \text{Eq. 3.11}$$

where, the cost S for a given pixel x at the reference image location i, j is determined for r different across and along track displacements, u and v . This is achieved by summing the exclusive or comparisons, as determined by the exclusive or operator \vee , between the reference bit vector, Γ^R , at location i, j and comparison image bit vector, Γ^C , at the displaced location $i + u_r, j + v_r$. The costs for all disparity assessments are aggregated by a 7 pixel radius mean filter to reduce noise. Following cost aggregation a simple spline interpolation routine is employed to estimate sub-pixel disparities in the along-track direction from the along-track disparities associated with the five smallest costs. Across track disparities are returned at integer level accuracy.

3.4.3 M6 Algorithm

3.4.3.1 Development

The M6 stereo matching algorithm (Fisher et al., 2013) was developed by Vladimir Yershov at MSSL to effectively determine smoke plumes injection heights for the ESA ALANIS smoke plumes project which is presented in Chapter 5. Although named after the other M-series algorithms (M2-M5) and employing local based stereo matching techniques, M6 approaches the stereo matching task with a number of significant modifications. These modifications occur in the image normalisation (see Figure 3-9) and cost aggregation stages, where, instead of using the entire normalisation or cost computation window, only those neighbourhood pixels with radiometric similarity to the pixel of interest are employed. The rationale being that the pixels employed in the normalisation or cost aggregation stages then come from similar intensity distributions (and likely therefore disparities), meaning firstly that the statistical measures used in the normalisation are more robust leading to a reduction in smoothing artifacts at

discontinuities, and secondly that the pixels used in the cost aggregation come from similar disparities leading to a reduction in potential projective distortion errors. This need for a reduction of smoothing artifacts in the ALANIS project is due to the fact that the smoke plume features under observation are often small, encompassing tens of pixels. With M5 many of the smaller smoke plume features are significantly eroded (see Chapter 5) leading to a reduced detection rate.

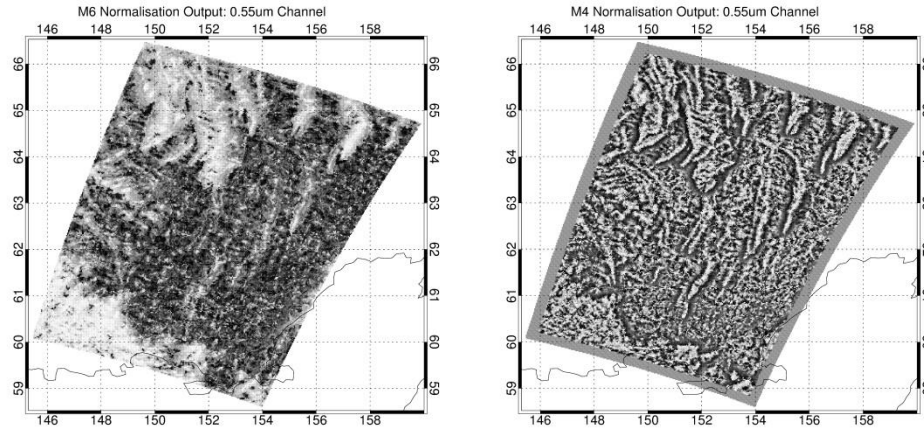


Figure 3-9. This figure demonstrates the alternative normalisation achieved by the M6 algorithm. The left image is the M6 normalisation output of the AATSR nadir view presented in Figure 3-8. The right view is the M5 normalisation output from the same figure. Note the significant reduction in low texture artifacts in the M6 normalisation output.

3.4.3.2 Algorithm Description

Surrounding each pixel x in the reference image R is a local neighbourhood N . From N a subset of pixels ρ is defined by similarity to x , giving a reduced neighbourhood N^ρ . This is achieved through,

$$N^\rho = |x - N| < s, \quad \text{Eq. 3.12}$$

where s is a threshold. Once N^ρ has been defined it is possible to use it to normalise x in place of the Gaussian filter employed by M5. Similar to M5, the size of N is (21,21) pixels. The threshold s is set dynamically to retain those pixels ρ whose absolute differences from x are within the 25% percentile¹⁰.

Letting ξ define the subset of pixels which comprises N^ρ , i.e. $\xi \in N^\rho$, and assuming that these pixels are sorted by value so that $\xi_0 < \xi_1 < \dots < \xi_{n-1}$ where n is the total number of pixels in N^ρ , it is possible to calculate the median of the subset as follows,

¹⁰ According to V. Yershov this threshold appeared to give the best result for the determination of smoke plume elevation across a number of scenes evaluated by hand measurement of the retrieved disparities.

$$\tilde{x} = \begin{cases} \xi_k & \text{if } n = 2k - 1 \\ (\xi_k + \xi_{k+1})^{0.5} & \text{if } n = 2k \end{cases}. \quad \text{Eq. 3.13}$$

With the median defined it is possible to compute the standard deviation of the subset N^ρ relative to the median,

$$x^\sigma = \sqrt{\frac{\sum_{m=0}^{n-1} (\xi_m - \tilde{x})^2}{n} - \left(\frac{\sum_{m=0}^{n-1} \xi_m - \tilde{x}}{n} \right)^2}. \quad \text{Eq. 3.14}$$

Eqs. 3.12, 3.13 and 3.14 can be used to define the subset medians and standard deviations for every pixel for the reference image R , giving the median reference image $R_{\tilde{x}}$ and the standard deviation reference image R_σ . These can be used to normalise the reference image as,

$$R_{norm} = \frac{R - R_{\tilde{x}}}{R_\sigma + \varepsilon}, \quad \text{Eq. 3.15}$$

where, like M5 the small positive number ε is set to 0.001 to avoid division by zero errors and to limit the amplification of image noise, an example of the M6 normalisation is shown in Figure 3-9. The same methods can also be employed to normalise the comparison image, C ,

$$C_{norm} = \frac{C - C_{\tilde{x}}}{C_\sigma + \varepsilon}. \quad \text{Eq. 3.16}$$

Once the reference and comparison images are normalised it is possible to carry out the along track stereo cost computations. For M6 a modified sum of the absolute difference (SAD) metric is employed and only along track disparity estimates are made. Across track assessments were found to add only increased computational cost whilst not improving the along-track disparity estimates substantially (Vladimir Yershov Pers. Comm.). The SAD modification is that the sum is computed using only a subset of pixels from the cost aggregation window, similar to variable window size algorithms (Veksler, 2003). In order to determine the pixel subset Eq. 3.12 is applied to each pixel in R_{norm} . Here, however, the indices of the pixels $\xi \in N^\rho$ are used to define the size N neighbourhood, K , over which the cost aggregation is performed, that is, $K(\xi) = 1$, and all indices that are not of the set N^ρ , are set to 0, that is, $\psi \notin N^\rho$. Following determination of the pattern kernels it is possible to evaluate the matching costs. Letting i and j define the line and sample coordinates for all pixels x in the normalised reference and comparison images R_{norm} and C_{norm} , then the SAD metric for r different along track displacements v_r can be computed as follows.

$$SAD(x_{i,j}, v_r) = \frac{\sum_{k=0}^{p-1} \sum_{l=0}^{q-1} \left| R_{norm}(i+k-\frac{p}{2}, j+l-\frac{q}{2}) - C_{norm}(i+k-\frac{p}{2}, j+v_r+l-\frac{q}{2}) \right| \times K(k,l)}{\sum_{k=0}^{p-1} \sum_{l=0}^{q-1} K(k,l)}. \quad \text{Eq. 3.17}$$

Where p is the number of samples and q is the number of lines in N . While working through the list of along track displacement vectors v_r , the value of the metric and the associated displacement are stored

for $x_{i,j}$. Then, to determine the final disparity, a spline interpolation routine is applied to the set of costs to find a sub-pixel disparity value.

3.5 Algorithm Assessment

Two assessments are undertaken in this section. The first evaluates the impact of window size, for both normalisation and cost aggregation windows, on the performance of the Census and M5 stereo matching algorithms. The second assesses the three stereo algorithms introduced in Section 3.4 in terms of their ability to replicate a collocated digital elevation model (DEM) from a region of the Himalayas taken from the GMTED 2010 DEM (Danielson and Gesch, 2010) and shown in Figure 3-10.

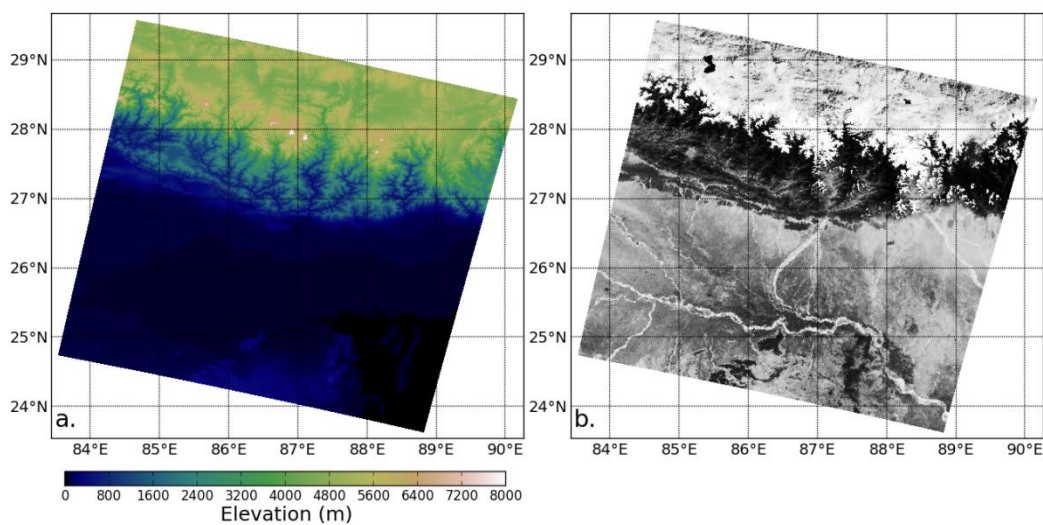


Figure 3-10. The left image is the GMTED2010 DEM subset of the AATSR image shown on the right. The AATSR image is taken from the 0.55μm channel which has been histogram equalised. The image comes from the same AATSR orbit as used in Figure 3-2.

3.5.1 Window Size Assessment

Window size plays a critical role in the performance of local stereo matching algorithms in two key areas: accounting for radiometric dissimilarities during the normalisation phase; and defining the number of local costs to take into account when assigning a disparity estimate. Here an evaluation on the effect of window size is undertaken for both M5 and census, enabling justification of the window sizes defined for each algorithm in the mathematical descriptions given in the previous section.

The evaluation assesses three different normalisation window radii (5, 10 and 15 pixels) and five different cost aggregation window radii (1, 3, 5, 7 and 9 pixel), giving a total of 15 window combinations, for both M5 and census. A cloud-free 512 by 512 pixel 11μm AATSR stereo pair (taken from orbit: 36473) collocated with the GMTED 2010 DEM shown in Figure 3-10 is used to generate the stereo derived terrain heights. A disparity search radius of 12 pixels is used in the direction of parallax, and the co-

registration correction coefficients defined in Chapter 2 are applied to the derived disparities prior to conversion into heights. For each normalisation-aggregation window combination the 11 μ m stereo derived terrain heights are differenced with the collocated DEM and summary statistics (mean absolute difference and root mean square error) are calculated. All image edges are excluded from the analyses, with exclusion extent determined by the largest of the normalisation or aggregation window size. The results from this analysis are presented in Figure 3-11.

3.5.1.1 Window Size Results

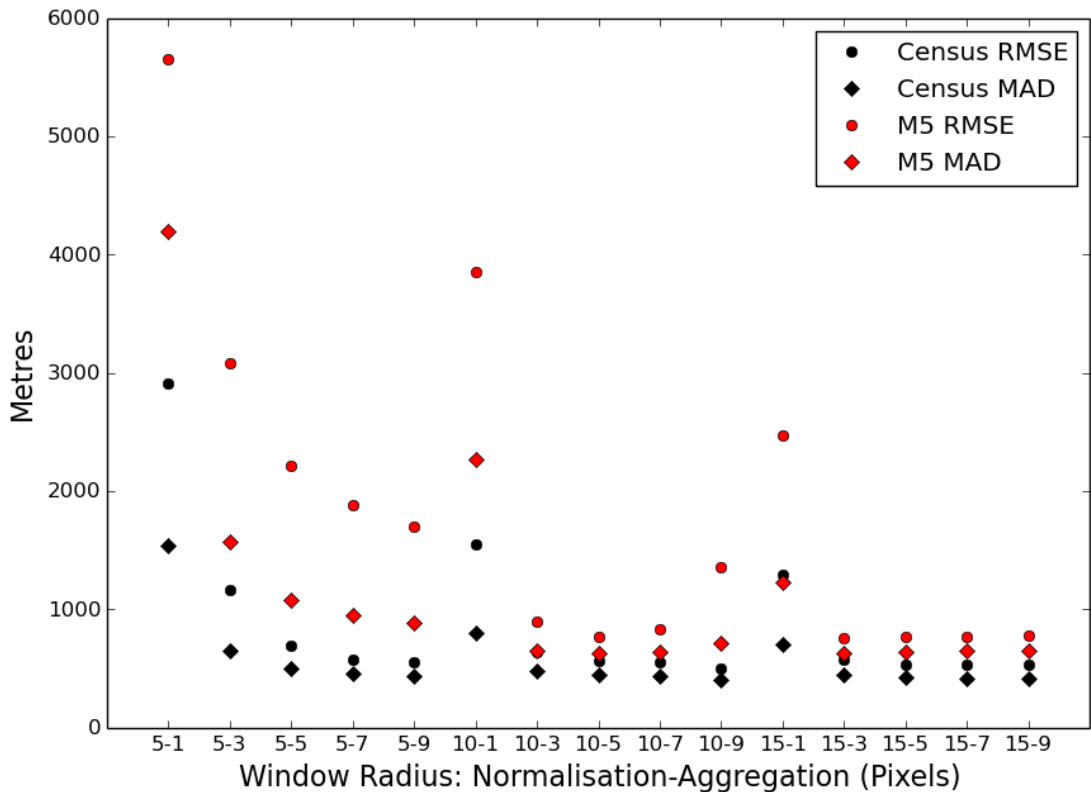


Figure 3-11. This plot summarises the effects of window size on the census and M5 stereo matching algorithms. The x-axis presents various normalisation and aggregation window size radii, ranging from a 5 pixel normalisation window and a 1 pixel aggregation window, through to a 15 pixel normalisation window and a 9 pixel aggregation window. The black diamonds show the census MAD from the collocated reference DEM. The black circles show the census RMSE. The red diamonds show the M5 MAD, the red circles the M5 RMSE.

3.5.1.2 Window Size Discussion

The statistical measures employed in this window size analysis are the mean absolute difference (MAD) and the root mean square error (RMSE). The first of these measures provides the measure of average height difference between the stereo retrievals and the collocated DEM over the given sample. The second measure returns the average magnitude of the height difference, giving more weight to larger height differences due to the squaring term. In combination the two measures enable better insight into

the height difference variations observed between the stereo terrain elevation retrievals and the DEM – the lesser the difference between the two measures the lesser the variance in the individual height differences across the set of collocated observations and the more precise the matching algorithm.

Looking at the results shown in the plot in Figure 3-11, census outperforms M5 for every window size combination. The optimal outcome is achieved by the census stereo algorithm when the normalisation window radius is set to 10 pixels (giving bit strings of length 441) and the aggregation window radius is set to 9 pixels. With these settings an average height error of ~400m between the stereo retrievals and the DEM is returned. The variance in the height differences across the inter-comparison is also small when using this combination of window sizes, as shown by the RMSE of ~500m. The one drawback using large normalisation and aggregation window sizes is that there is a significant smoothing of the output disparities. This is problematic when stereo matching cloud and smoke plume features as their extent is exaggerated. By using a smaller normalisation and aggregation windows set to 5 and 7 pixel radii respectively the smoothing effects are reduced, and as seen from the plot, the change in retrieval accuracy is ~100m across both measures. The processing time is also reduced by using the smaller aggregation window. These results provide justification for the normalisation and aggregation window radii of 5 and 7 pixels respectively as defined in the previous section and applied in the census stereo processing throughout this thesis.

For M5 the optimal performance in the analysis is achieved through application of the 15 pixel radius normalisation window in combination with the 3 pixel radius aggregation window, with a MAD of ~600m and an RMSE of ~700m. The difference between the best window combination outcome for M5 in this analysis and the window combination employed in the paper which introduces the M5 algorithm (Muller et al., 2007), 10 pixel normalisation window radius and 5 pixel cost aggregation window radius, is minimal. The retrieved statistics for both window size combinations show differences of <50m, yet the processing time is more than doubled with the larger normalisation window. For this reason, the window radii defined in the paper that introduced M5 are employed in all cases throughout this thesis.

3.5.2 DEM Inter-comparison

To provide an initial assessment on the quality of the outputs from the three different stereo matching algorithms the technique in Muller et al. (2002) is employed, where the retrieved terrain heights are compared against elevations from an independent ground surface terrain model. To perform this assessment, the first requirement is to obtain cloud free imagery so that the retrieved heights represent the terrain. In this assessment a cloud free 512 by 512 pixel image of a region in the Himalayas from AATSR orbit 36473 is employed and is shown in Figure 3-10. The Himalayan region makes for good comparison datasets due to the close proximity of extremely high and also low terrain features ranging from 0km to more than 8km, enabling evaluation of stereo matching performance over a wide range of elevations. The inter-comparison data employed is from the GMTED 2010 DEM (Danielson and Gesch, 2010) and is also shown in Figure 3-10. This DEM is comprised of the best current available global elevation data, and has RMSE values that range between 25 and 42 metres when compared against a global set of stereo derived height observation from high resolution stereo imagery. This accuracy is significantly better than that achievable by the ATSR instruments, so provides a valid inter-comparison

test dataset. The same evaluation process employed in section 3.5.1 is employed here. However, in this instance all channels are processed rather than just the 11 μ m channel and the coefficient of determination statistic is calculated in addition.

3.5.2.1 Stereo Assessment Results

	RMSE	MAD	R**2
12 μ m	754, <i>1133</i> , 473	610, <i>842</i> , 351	0.94, <i>0.77</i> , 0.95
11 μ m	768, <i>1150</i> , 471	630, <i>860</i> , 347	0.94, <i>0.76</i> , 0.96
3.7 μ m	1784, <i>3510</i> , 564	946, <i>2857</i> , 362	0.58, <i>0.04</i> , 0.93
1.6 μ m	918, <i>972</i> , 748	644, <i>744</i> , 398	0.88, <i>0.82</i> , 0.89
0.87 μ m	820, <i>1276</i> , 563	649, <i>1035</i> , 375	0.93, <i>0.67</i> , 0.93
0.67 μ m	815, <i>1063</i> , 494	660, <i>849</i> , 366	0.93, <i>0.78</i> , 0.95
0.55 μ m	820, <i>1276</i> , 563	649, <i>1035</i> , 375	0.93, <i>0.67</i> , 0.94

Table 3-1. The statistical analysis of the three stereo matching algorithms applied to the AATSR scene in Figure 3-10 and differenced from the DEM in the same figure. The M5 results are in plain text, the M6 results in italic text, and the census results in bold text. The number of points analysed for each channel and matcher is >200,000. RMSE and MAD units are in metres.

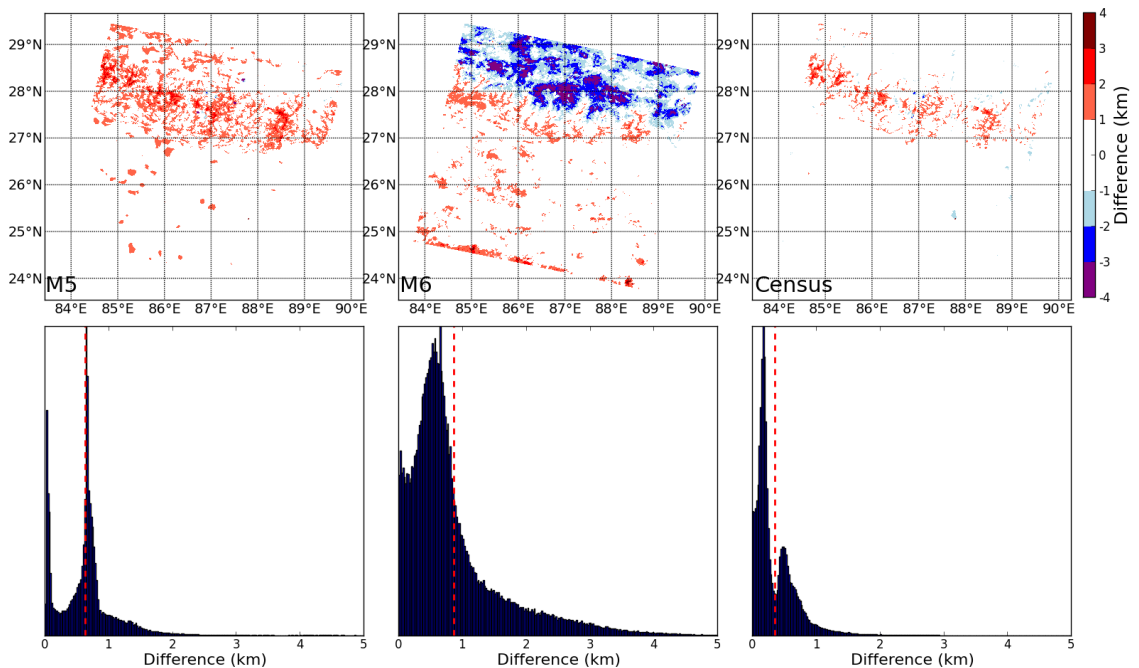


Figure 3-12. The top three plots show the differences between the stereo outputs from the 11 μ m channel for M5, M6 and census respectively and the collocated GTMED 2010 digital elevation model. Beneath each spatial plot is the associated histogram of the absolute differences (250 bins). red dashed line in each histogram shows the location of the MAD.

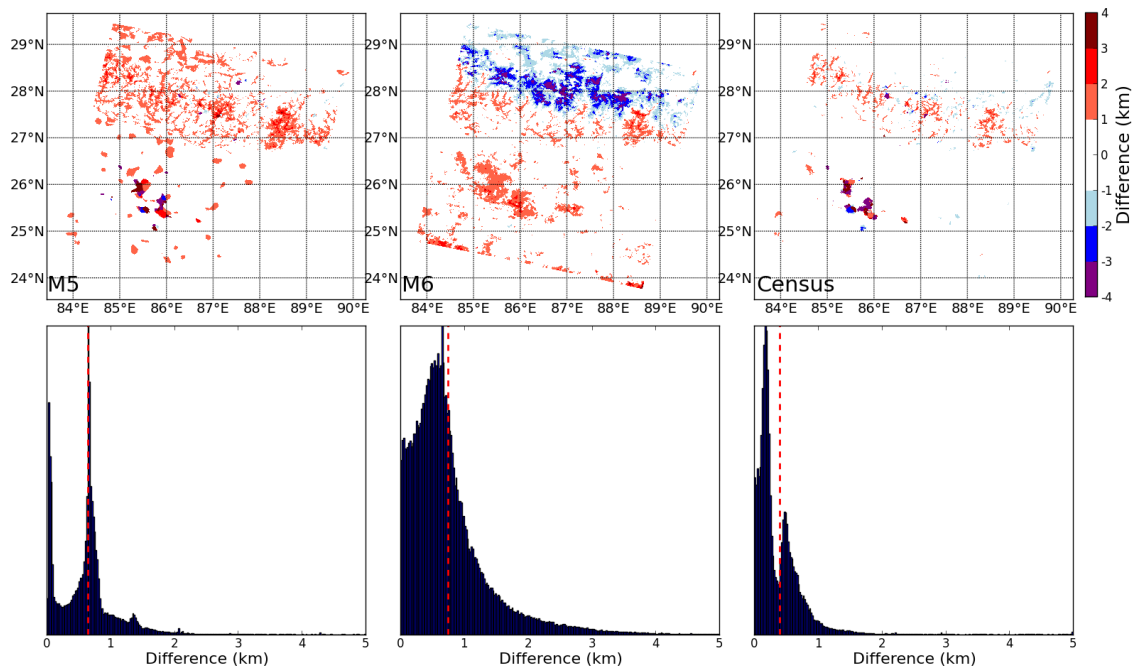


Figure 3-13. The top three plots show the differences between the stereo outputs from the $1.6\mu\text{m}$ channel for M5, M6 and census respectively and the collocated GTMED 2010 digital elevation model. Beneath each spatial plot is the associated histogram of the absolute differences (250 bins). The red dashed line in each histogram shows the location of the MAD.

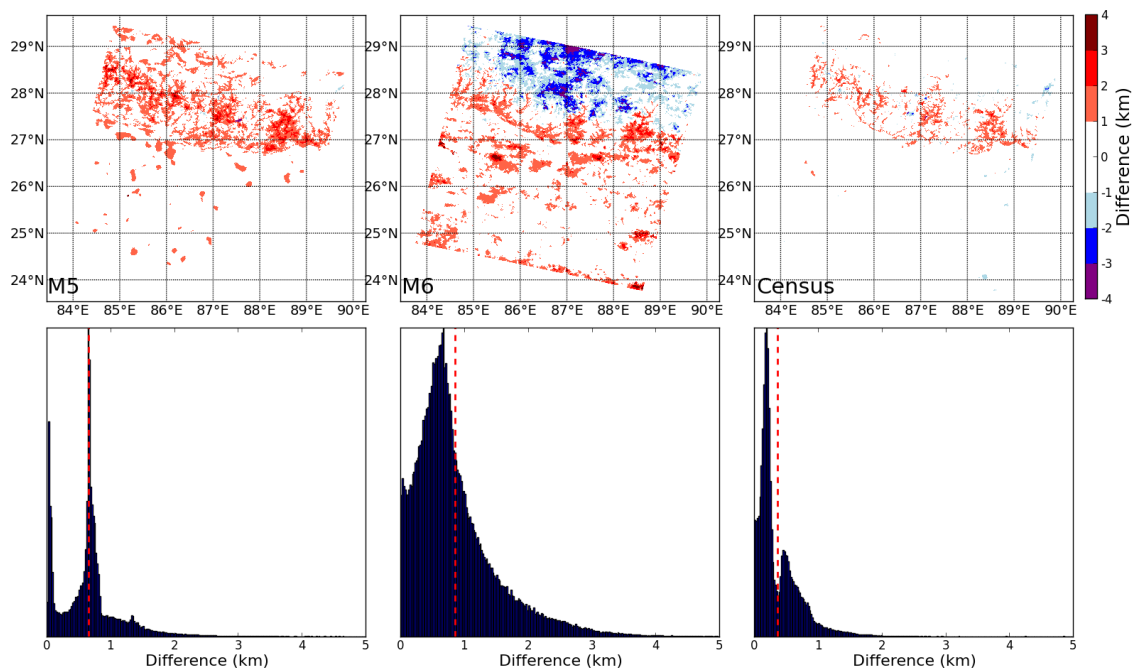


Figure 3-14. The top three plots show the differences between the stereo outputs from the $0.55\mu\text{m}$ channel for M5, M6 and census respectively and the collocated GTMED 2010 digital elevation model. Beneath each spatial plot is the associated histogram of the absolute differences (250 bins). The red dashed line in each histogram shows the location of the MAD.

3.5.2.2 M5 Discussion

Table 3-1 shows that for the M5 algorithm the AATSR channel with the best performance is 11 μ m with an RMSE 754m, a MAD of 610m, and a coefficient of determination of 0.94. The RMSE is within the achievable accuracy of the ATSR instruments. The difference between the RMSE and MAD results is small demonstrating little variance across the sample. The coefficient of determination value indicates that there is good agreement between the stereo results and the DEM across the more than 200,000 individual points evaluated.

The statistical performance is similar across all AATSR channels, with a slight decline in quality at shorter wavelengths. This is likely due to the fact that at shorter wavelengths there is increased atmospheric scattering by aerosols, leading to increased noise in the imagery, and therefore poorer stereo outcomes. Looking at visible channels assessed in this analysis, which are shown in the false colour composite in Figure 3-2, large amounts of what appear to be aerosols are apparent. It should also be noted that the poor ATSR instrument geo-referencing is likely leading to a slight decline in the quality of the returned statistics. However, this error is common for all three assessments, so does not significantly impact on the inter-comparison outcome.

Looking at the M5 11 μ m DEM difference map presented in Figure 3-12, there are a significant number of pixels with differences of greater than ± 1 km in regions of rapidly varying terrain. These errors are likely due to the increased smoothing caused by the parametric normalisation measures and the large size requirement of the cost aggregation window to generate robust matches.

In terms of processing time, a Python version of the M5 algorithm is able to process the 512 by 512 pixel scene in less than 20 seconds (~30 minutes an orbit) on an Intel based i5 processor running at 2.66 GHz. This is entirely acceptable for processing large quantities of EO data.

3.5.2.3 M6 Discussion

The M6 algorithm is the worst performing out of the three algorithms in this analysis. Looking at Table 3-1 the best performing channel is 1.6 μ m with an RMSE 972m, a MAD of 744m, and a coefficient of determination of 0.82. The substantial decline in performance in the 3.7 μ m channel is currently undetermined (V. Yershov, pers. comm.).

Looking at the DEM difference maps presented in Figure 3-12 through Figure 3-14 it is apparent that M6 performs well in regions of low and also rapidly varying terrain. However, for terrain above 5km the performance declines substantially. The cause of this decline is also unknown (V. Yershov pers. comm.). However, for the smoke plumes the algorithm is designed to measure this is not entirely problematic, as fire events are rarely intense enough to inject smoke into the atmosphere at elevations of more than 5km (see Chapter 5).

The M6 algorithm is written in Java to facilitate incorporation in BEAM and to provide suitable processing speed for the intensive window computations. Despite this the algorithm is still slower than M5 or census, taking approximately 3 minutes to process the 512 by 512 pixel scene on the same computing system.

3.5.2.4 Census Discussion

The census algorithm outperforms both M5 and M6 in this assessment. The optimal channel is 11 μ m, with an RMSE of 471m, a MAD of 347m, and a coefficient of determination of 0.96. The RMSE and MAD have a difference of 124m indicating little variance across the inter-comparison sample and that the algorithm is precise. There is a slight decline in the quality of the RMSE statistic with wavelength; again this is likely due to atmospheric aerosols confusing the algorithm.

Looking at the DEM difference maps presented in Figure 3-12 through Figure 3-14 it is apparent that the census algorithm performs well in regions of rapidly varying terrain. The improvement in performance is due to the non-parametric nature of the transform and the reduction in disparity smoothing effect in this class of algorithm. In terms of processing performance, a Python implementation of the census algorithm is slower than M5, taking approximately 1 minute (90 minutes per orbit) to process the 512 by 512 pixel scene on the same computer system. This time could however be reduced, through application of multi-threading techniques, to approximately 45 seconds.

3.6 Chapter Summary

To summarise, this chapter has introduced the main concepts of stereo-photogrammetry: the imaging system, the stereo matching algorithm, and the camera model. The imaging system, comprised of two or more cameras separated by a distance referred to as the baseline, obtains the stereo imagery. The stereo matching algorithm determines, typically through a cost minimisation process, the displacements of objects within the imagery caused by the parallax effect. The measured displacements are converted into real-world measures of distance through application of the camera model, which attempts to replicate the imaging systems geometry. Following definition of these key concepts, the stereo image matching process has been identified as the most important consideration, in terms of ease of manipulation and impact on the final outcome quality, in the ATSR stereo-photogrammetric processing chain. To that end the main stereo photogrammetric challenges encountered in real-world stereo image matching tasks, arising from both the environment and the sensing system, in this instance AATSR, have been described. An algorithm well suited to coping with radiometric distortions common in the AATSR imagery, the census algorithm, has been identified, described, and evaluated using AATSR data against two other stereo matching techniques, M4 and M6. The M4 and M6 algorithms, also recently applied to the AATSR instrument, have been described in detail. In the evaluation the heights retrieved by the census stereo matching algorithm when compared with a collocated DEM taken from the GMTED2010 digital elevation model demonstrate for the 11 μ m channel an RMSE of 471m, a MAD of 347m and a coefficient of determination of 0.96. M4 in comparison returns an RMSE of 768m, a MAD of 630m and a coefficient of determination of 0.94 for the 11 μ m channel. M6 returns an RMSE of 972m, a MAD of 744m and a coefficient of determination of 0.82 for the 1.6 μ m the channel, the channel on which it performs best.

4 GREENLAND CLOUD STUDY

4.1 Introduction

4.1.1 Clouds and the Earth's Climate System

4.1.1.1 Earth's Changing Climate

According to observational records, the Earth's climate is experiencing unprecedented and accelerating warming (Solomon et al., 2007; Stocker et al., 2013). The effects of this warming manifest in numerous different climate variables, in particular: rapid sea-level rise, increasing sea and land surface temperatures, changing precipitation patterns, and reductions in sea-ice extent (Solomon et al., 2007; Stocker et al., 2013). The root cause of these changes is generally regarded as being an anthropogenic forcing of the climate, linked to industrialisation processes and the associated emissions of greenhouse gases (Solomon et al., 2007; Stocker et al., 2013). Significant effort is being undertaken to determine and understand the future implications of such a forcing on the climate system by projecting its future state using global models of the Earth's climate system, General Circulation Models (GCMs). This task, due to the complexity of the climate system and its internal processes known as feedbacks is inherently challenging. There is consensus between GCMs that the Earth will continue to warm, however, agreement on the magnitude of this warming is some way from realisation (Solomon et al., 2007; Stocker et al., 2013). This section discusses the processes of climate forcing and feedback and defines where the main challenges in determining the future state of the climate exist.

4.1.1.2 Climate Forcing – The Earth's Radiation Balance

Prior to discussing how the climate system is forced and how feedbacks affect this process, it is necessary to review the Earth's radiation balance. To begin with, it must be understood that the climate system is comprised of a highly complex set of interactions between the atmosphere, the ocean and the land surface (see Figure 4-1), and that the system maintains radiative equilibrium – the energy output from the system is the same as the amount input. Annually there is an approximate global top of the atmosphere (TOA) average of 340 W m^{-2} of incident radiation, S , originating from the Sun, and this comprises the vast majority (99.7%) of the energy input into the climate system. Of this incident radiation, around 100 W m^{-2} is reflected directly back into space due to the Earth's albedo¹¹, α , which is approximately 0.3 (Wielicki et al., 1995). The remaining 240 W m^{-2} is absorbed into the climate system, and, to maintain radiative

¹¹ Albedo is the ratio of reflected radiation from the surface to the incident radiation upon it.

equilibrium, is output to space through longwave (LW) emission, L . (Wielicki et al., 1995). This process can be defined as,

$$R = S(1 - \alpha) - L. \quad \text{Eq. 4.1}$$

When the climate is in a state of equilibrium, the net radiance at the TOA, R , is equal to zero, however, as the climate system is dynamic, there is continual variation in the equilibrium as it strives to maintain radiative balance. From this definition it is possible to introduce the concept of a climate forcing as any process by which the equilibrium is disturbed, or in the terminology of climate studies, perturbed.

FIGURE REMOVED DUE TO THIRD PARTY CONFLICTS

Figure 4-1. Radiative interactions between the atmosphere, the ocean and the land surface. Adapted from http://spacefellowship.com/wp-content/uploads/2009/12/412218main_radiation-budget2923.jpg (Last accessed 1st September 2013).

4.1.1.3 Climate Forcing – Radiative Forcing

As implied above, radiative forcing of the climate is caused by changes in the net irradiance/exitance balance of the climate radiative system. That is to say, perturbations in S , α , and/or L , cause a state of radiative imbalance and, therefore, the equilibrium requirement is no longer fulfilled, i.e. $R \neq 0$. Radiative forcing processes are considered to be external processes, in that they are independent from the climate system. This can be illustrated by looking at typical forcing agents, which can be either natural or anthropogenic. Natural agents include changes in the behaviour of the sun (e.g. Lean and Rind, 1998), reducing or increasing S , or extreme volcanic events (e.g. Minnis et al., 1993) that tend to significantly increase α . Anthropogenic agents are primarily atmospheric in nature (see Figure 4-2) and generally relate to increased emission of aerosols and radiatively active gases. These gases, mainly CO_2 and CH_4 , strongly absorb in the infrared part of the electromagnetic spectrum (see Figure 4-3), reducing L . Following a forcing on the climate, there must be a follow on response by the climate system to restore radiative equilibrium. There are many techniques to measure this response, but one of the most commonly applied is that of climate sensitivity.

FIGURE REMOVED DUE TO THIRD PARTY CONFLICTS

Figure 4-2. Climate radiative forcing agents (adapted from IPCC AR5, Stocker et al., 2013)

4.1.1.4 Climate Forcing – Climate Sensitivity

Following the terminology from Andrews et al. (2012), regard G (W m^{-2}) as the forcing on the climate, caused by variation in the aforementioned parameters on the right hand side of Eq. 4.1, that cause a change in R of ΔR (W m^{-2}). The magnitude of ΔR is approximately linearly dependent on the global-mean surface air temperature change ΔT (K),

$$\Delta R = G - \lambda \Delta T, \quad \text{Eq. 4.2}$$

where $-\lambda$ ($\text{W m}^{-2} \text{K}^{-1}$) is the climate feedback parameter. If G and λ are constant then ΔR is a linear function of ΔT , with a slope of λ and an intercept of G ; and from this it is possible to obtain both the forcing and the feedback parameters by way of linear regression. The ΔT required to restore the climate to a state of equilibrium, ΔT_{eqm} , is the measure of climate sensitivity, and can be estimated through extrapolation of Eq. 4.2 to a state of equilibrium, i.e. $\Delta R = 0$ and $\Delta T = G/\lambda$.

Climate sensitivity assessments have been applied extensively in the IPCC climate projection analyses (Houghton et al., 2001; Solomon et al., 2007; Stocker et al., 2013). In the IPCC studies, a standardised climate sensitivity measure referred to as the equilibrium climate sensitivity is employed. The standardisation occurs in the forcing, which is set as a doubling of atmospheric CO_2 , a particularly pertinent forcing agent given the currently understood causes of climate change. In a climate system only perturbed by a doubling of CO_2 the ΔT_{eqm} required to restore equilibrium is estimated to be around 1 K. However, in the most recent IPCC analysis (at time of writing, Stocker et al., 2013) ΔT_{eqm} was found to range (at a high level of confidence) from 1.5 K to 4.5 K across GCMs, demonstrating greatly increased

and strongly disagreeing magnitudes of warming required to restore radiative equilibrium in the climate system. The majority of this uncertainty arises in the climate feedback parameter λ , which contains the internal processes of the climate system that either amplify or dampen the effect of the external forcing (Randall et al., 2007). Indeed, it has been shown that differences in feedbacks contribute almost three times more to the range in ECS estimates than differences in the radiative forcings applied in GCMs (Webb et al., 2006). Improved understanding of climate feedback processes contained in the feedback parameter is therefore of critical importance in gaining better agreement on the future state of the Earth's climate and in turn better quantification of anthropogenic forcing impacts.

FIGURE REMOVED DUE TO THIRD PARTY CONFLICTS

Figure 4-3. Radiation and its transmission through the atmosphere. Note the strong absorption features over many wavelengths by water vapour – the strongest greenhouse gas. Also note the strong absorption bands for CO₂ and CH₄ in the thermal regions of the spectrum. Lastly note the atmospheric window around 10µm which is employed in the IR-window CTH assignment approach. Figure adapted from www.researchgate.net (last accessed 2nd September 2013). Originally prepared by Robert A. Rhode for the Global Warming Art Project.

4.1.1.5 Climate Feedbacks – Introduction to Feedbacks

As temperature changes within a forced climate other climate variables internal to the climate system, and intrinsically linked to temperature, also change. Examples include atmospheric water vapour content, cloud characteristics, sea-ice extent, and so on (Bony et al., 2006). These internal climate system processes, if they amplify or dampen the effect of a forcing agent, are referred to as feedbacks. In the

IPCC assessments (Houghton et al., 2001; Solomon et al., 2007; Stocker et al., 2013) a feedback is deemed to occur when an initial process (such as a forcing) triggers a change in a second process (a feedback) that in turn influences the initial one. The uncertainty associated with the feedback processes at work within the Earth's climate system stems from their complex interlinked and internal nature which leads to great difficulty in their detection and quantification.

The four most important physical feedback process, given their potential radiative impacts following a doubling of CO₂ are: the water vapour feedback, the most important positive feedback within the climate system; the lapse rate feedback, a strong negative feedback which offsets the water vapour feedback; the albedo feedback, which goes some way to explaining why Polar Regions are more sensitive to climate change; and cloud feedbacks, which lead to the greatest uncertainty in GCM projections.

4.1.1.6 Climate Feedbacks – Key Feedback Mechanisms

Water Vapour Feedback - The most important positive feedback process within the climate system is thought to be that of water vapour, the gaseous phase of water (Randall et al., 2007). Water vapour, due to its total atmospheric mass (~4% of the atmosphere), is the most prevalent of the greenhouse gases, and as such contributes a warming effect approximately eight times greater than that of CO₂ (Hartmann 1994). Furthermore, as defined by the Clausius-Clapeyron equation, there is a near exponential increase in the moisture holding capacity of the atmosphere with temperature (Randall et al., 2007). Therefore, climate warming leads to an increase in the amount of water vapour in the atmosphere and increased water vapour leads to increased warming - constituting a feedback effect. Recent studies have estimated a water vapour related increase of $1.80 \pm 0.18 \text{ W m}^{-2} \text{ K}^{-1}$ in the climate feedback parameter (Soden and Held, 2006).

Lapse Rate Feedback - Heterogeneous changes in column atmospheric temperature profiles cause the lapse rate feedback. The sign of the lapse rate feedback (i.e. whether it is positive or negative) is dependent upon where in the atmospheric column the warming occurs. In the tropics, GCMs generally predict enhanced warming in the upper troposphere due to increased concentration of greenhouse gases. Enhanced warming at higher altitudes leads to greater LW emission to space, leading to a negative feedback. In middle and high latitudes greater warming tends to occur in the lower troposphere due to radiative heating, and this has the reverse effect leading to a positive feedback. Overall the effect in the tropics dominates, leading to a strong negative feedback, reducing the climate feedback parameter by $-0.84 \pm 0.26 \text{ W m}^{-2} \text{ K}^{-1}$ (Soden and Held, 2006).

The water vapour and lapse rate feedbacks are inherently linked - as temperature changes in the atmosphere so does water vapour content. Due to this close link between the two processes, they are typically considered together in GCMs. The main benefit of considering the two processes together is a reduction in the overall uncertainty when compared to a separate consideration. The global mean combined water vapour-lapse rate feedback has been estimated to have a positive effect on the climate feedback parameter of $0.97 \pm 0.11 \text{ W m}^{-2} \text{ K}^{-1}$ (Soden and Held, 2006).

Albedo Feedback - Polar Regions are generally ascribed to be both extremely sensitive to changes in global climate, and also, to present the greatest potential to change global climate (Solomon et al., 2007;

Stocker et al., 2013). Numerous recent studies have continued to add strength to the sensitivity argument, demonstrating various effects resulting from changes in the Polar climate regimes. Observed phenomena include: decreasing ice-sheet mass balance (Shepherd et al., 2012); rapidly decreasing North Pole sea-ice extent (Comiso et al., 2008; Parkinson and Comiso, 2013); loss of permafrost (Grosse et al., 2011); and increasing glacier velocities (Moon et al., 2012). The observed changes in Polar Regions are thought to be mainly caused by declining surface albedo, due to declining high albedo sea/terrestrial ice and increasing low albedo open water/rock extents, leading to further warming and melt. Reduced albedo leads to increased absorption of insolation, increasing local temperatures. This process, referred to as the ice-albedo feedback mechanism (Curry et al., 1993), is generally considered to be one of the key feedback mechanisms behind the above average (1.5 to 4.5 times above the global average) warming trend observed in the Arctic in recent decades (Holland and Bitz, 2003). When looking at global albedo feedbacks on forcing agents the estimated amplification of the climate feedback parameter is on the order of $0.26 \pm 0.08 \text{ W m}^{-2} \text{ K}^{-1}$ (Soden and Held, 2006), and three quarters of this feedback is thought to arise in the Northern Hemisphere from changes in the cryosphere (Winton, 2006).

Cloud Feedbacks - Clouds strongly modulate the incoming SW and outgoing LW radiation, as shown in Figure 4-1. Global fractional cloud occurrence is on the order of 0.6 to 0.7 (Stubenrauch et al., 2013), leading to a doubling of the Earth's average albedo from 0.15 to 0.3 (Cess, 1976). Furthermore clouds also have a greenhouse effect and trap outgoing LW radiation. Clouds are therefore of great importance to the radiation balance, and any changes in cloud properties may constitute feedbacks that could have significant impacts on the future state of the climate. The specific details of cloud radiative and feedback processes that may affect the climate system are omitted here as they are discussed in the following section. The point to be made is that the physical feedback processes introduced above are all relatively well constrained, with small uncertainties in comparison to overall magnitude, and that this is in stark contrast to the uncertainty associated with cloud feedbacks. Cloud feedback processes have an estimated amplification effect on the climate feedback parameter that ranges from 0.31 to $1.07 \text{ W m}^{-2} \text{ K}^{-1}$ across GCMs (Soden and Held, 2006). In the 4th IPCC analysis (Solomon et al., 2007) it was demonstrated that in the presence of water vapour, lapse rate and surface albedo feedbacks, but excluding cloud feedbacks, the GCMs analysed predict a ΔT_{eqm} of $1.9 \text{ K} \pm 0.15 \text{ K}$. When cloud feedbacks are incorporated into the same models the mean and standard deviation of ΔT_{eqm} significantly increase to $3.2 \text{ K} \pm 0.7 \text{ K}$ indicating strong disagreement on the magnitude of the cloud feedback effect. Reducing this uncertainty has been identified by the IPCC as the key requirement for improving consensus between climate projections and therefore, gaining a better understanding of the future state of the climate (Randall et al., 2007).

4.1.1.7 Climate Forcing and Feedback Overview

The above sections have introduced the external and internal climate change concepts of radiative forcing and feedback. In the case of climate forcing, the Earth's radiation balance and the concept of radiative equilibrium have been introduced. Some of the natural and anthropogenic forcing agents that can perturb the climate equilibrium have been briefly addressed. GCMs, one of the main tools applied to understand climate change processes, have been introduced, as has the measure of climate sensitivity, which is used to highlight disagreements between GCM projections. The cause of these disagreements has been

attributed to internal climate feedback processes, the most significant of which have been reviewed. Of the feedback processes discussed, clouds, due to significant uncertainties in their radiative effects, have been identified as the most important for furthering understanding on the condition of the Earth's future climate. Given their importance, cloud radiative effects are discussed along with a more in-depth review of the uncertainty they cause in models in the following sections to highlight why clouds are radiatively complex and why this leads to uncertainty in models.

4.1.1.8 Cloud Radiative Interactions

The radiative effect of clouds is typically determined by differencing clear-sky and all-sky radiances at the TOA, and is referred to as cloud radiative forcing (*CRF*; Ramanathan et al., 1989). The net *CRF* can be defined as

$$CRF_{net} = (1 - N)F_{clr} - NF_{cld} \quad Eq. 4.3$$

The fractional cloud cover (referred to hereafter as CF) of the scene is, in this instance, defined by N and modulates the clear and cloudy sky top of atmosphere fluxes, F_{clr} and F_{cld} , respectively. The CRF_{net} can be further reduced into SW and LW components, as follows,

$$CRF_{sw} = NS(\alpha_{clr} - \alpha_{cld}) \quad Eq. 4.4$$

and,

$$CRF_{lw} = N[F_{clr} - F_{cld}(\varepsilon, T_c)] \quad Eq. 4.5$$

Looking at the CRF_{sw} component, the amount of reflected radiation is dependent on the cloud fraction N , the insolation S , and the albedo of the cloud, α_{cld} , and the surface, α_{clr} . The size and state of the cloud particles, the thickness of the cloud and the solar zenith angle determine the cloud albedo. The CRF_{sw} effect is one of cooling, with the largest effect occurring in the presence of high albedo clouds over dark surfaces, such as bodies of water, and the smallest in the case of low albedo clouds over bright surfaces, such as snow or deserts (Bony et al., 2006; Stephens, 2005).

The CRF_{lw} component is dependent on the cloud fraction, and the surface and cloud emittances. Cloud emittance is a function its temperature T_c , which for clouds is height dependent, and emissivity¹² ε , which is dependent on cloud thickness, and particle size and state. The CRF_{lw} effect is one of warming, with the greatest magnitudes in the presence of cold clouds over warm surfaces and the smallest in the case of warm clouds over cold surfaces (Bony et al., 2006; Stephens, 2005).

The main assessments of CRF_{sw} , CRF_{lw} , and CRF_{net} are from the Earth Radiation Budget Experiment (ERBE, Ramanathan et al., 1989) and its follow on, the Clouds and Earth's Radiant Energy System (CERES, Wielicki et al., 1996) experiment. From the CERES analysis, shown in Figure 4-4, it can be seen that, in accordance with the reasoning above, the strongest CRF_{sw} effects generally occur over tropical oceanic regions, where there is low surface albedo, high cloud albedo, and generally substantial

¹² Emissivity in this instance is the ratio of the energy emitted by the cloud compared to the amount that would be emitted by a blackbody at the same temperature.

FIGURE REMOVED DUE TO THIRD PARTY CONFLICTS

Figure 4-4. These figures show for the period 2001-2007 the various radiative effects of clouds in terms of mean SW, LW and net forcing. The left image displays the SW forcing, the central the LW and the right the net forcing. Note the complex radiative behaviour associated with clouds. Figure adapted from figure 3 in Allan, 2011.

cloud coverage. The weakest CRF_{sw} effects generally occur over high albedo surfaces with low cloud cover, such as deserts. The strongest CRF_{lw} effects also occur in the tropics in the presence of the coldest cloud tops over the warmest land or ocean surfaces. The weakest CRF_{lw} effects occur at higher latitudes, where the surfaces are cooler, and therefore the difference with the cloud top temperature is reduced.

The CERES CRF results demonstrate the complexity of cloud radiative effects, with the opposing effects in the tropics, whilst having a significant impact on the CRF_{sw} and CRF_{lw} fluxes, nearly negating each other. In the regions pole-ward of 30° , CRF_{net} is found generally to be negative; in particular a very strong effect is demonstrated over the Pacific and Atlantic oceans with values approaching -50 W m^{-2} . In Polar Regions over high albedo surfaces (i.e. ice sheets) the CRF_{net} is generally positive due to the 'radiation paradox' (Ambach 1974). When globally averaged, CRF_{net} demonstrates a cooling effect of clouds on the climate, with an estimated value of -18.2 W m^{-2} (Allan, 2011). This finding corroborates outcomes from other assessments such as the ERBE with a global mean CRF_{net} of -19 W m^{-2} (Kiehl et al., 1994), a Nimbus-7 assessment, with -27 W m^{-2} (Ardanuy et al., 1991), an ERBE re-assessment with -20 W m^{-2} (Kiehl and Trenberth, 1997), and a CERES re-assessment using additional data sources with -21 W m^{-2} (Allen, 2011).

To summarise, cloud radiative behaviour is a complex function of the negative SW and positive LW radiative effects. The magnitudes of the conflicting radiative components are in turn dependent on a number of microphysical (particle size, phase) and macrophysical (fraction, elevation) cloud parameters. The observational analyses presented demonstrate that when combined, the current overall radiative effect of clouds is one of SW cooling, but there is significant spatial heterogeneity. Due to the conflicting radiative effects and strong spatial variability, effective incorporation of cloud radiative behaviour into climate models is tremendously challenging and leads to great uncertainty in projecting the future state of the climate. In the next section this uncertainty is further demonstrated through a more detailed review of GCM inter-comparison studies where cloud radiative processes (feedbacks) in a forced climate are assessed.

4.1.1.9 Cloud and Models – GCM Inter-Comparison Studies

One of the first GCM inter-comparison studies to assess the effects of clouds on climate projections was undertaken by Cess et al. (1989). In their analysis 14 GCMs were assessed using an inverse climate change approach. In that, rather than apply a forcing to the system and record the modelled climate change response, the GCMs were instead perturbed by a ± 2 K change in sea surface temperature as a surrogate for climate change, and the forcing leading to such a change recorded. Using this design, for each model the forced clear sky fluxes, F_{clr} , were globally averaged separately from the all sky fluxes, F_{cld} . Analysing the climate feedback parameter under F_{clr} , excellent agreement between models was demonstrated, whilst for F_{cld} a three-fold variation was observed. The authors argue that since the uncertainty is much greater in F_{cld} (i.e. in the presence of clouds) than in F_{clr} (i.e. no clouds) then the uncertainty must be mainly attributable to cloud feedback processes. Evaluation of cloud feedback was achieved by taking the ratio of the forcing induced change in cloud radiative forcing to the overall forcing, that is $\Delta CRF_{net}/G$, and was found to range between $-0.4 \text{ W m}^{-2} \text{ K}^{-1}$ and $1.5 \text{ W m}^{-2} \text{ K}^{-1}$ across models. Other studies led by the same author (Cess et al., 1990; Cess et al., 1996) demonstrated similar findings reinforcing the above conclusion.

A study by Colman (2003) questioned whether the large cloud feedback variability found between models in the Cess studies was perhaps attributable to other non-cloud feedback processes contained within CRF_{net} (aliasing) and therefore not related to cloud feedbacks at all. To test this hypothesis 10 climate models ranging from full 3D GCMs to more simplistic 2D and 1D radiative convective models were assessed using a proscribed forcing of $2 \times \text{CO}_2$. Using the approach first employed by Wetherald and Manabe (1988), each model was run in offline mode (i.e. no interaction with other model fields allowed) and the TOA flux assessed on a field by field basis to evaluate the associated feedbacks, allowing their separation. This approach is completely different to the simplistic one employed in the Cess et al. studies. However, despite this dissimilarity, the modelled cloud feedback processes of the offline approach produced a range of $-0.1 \text{ W m}^{-2} \text{ K}^{-1}$ to $1.4 \text{ W m}^{-2} \text{ K}^{-1}$, similar in terms of the positive range to that found in the Cess et al. (1990) study ($-0.6 \text{ W m}^{-2} \text{ K}^{-1}$ to $1.5 \text{ W m}^{-2} \text{ K}^{-1}$) but with a much reduced negative feedback.

A study undertaken by Soden and Held (2006) again expressed doubts regarding certain aspects of the Cess analyses, including: the unrealism of globally uniform perturbation in temperature as the climate change; the intentional suppression of surface albedo feedbacks; and the use of the F_{cld} to assess cloud feedback, which, as stated above, aliases non-cloud feedback processes, so the feedback variability observed may not be due to cloud processes. In their assessment, 14 GCMs were evaluated under a forcing of $2 \times \text{CO}_2$, with cloud feedback being calculated by differencing the climate feedback parameter (derived from Eq. 4.2, where all parameters excluding the climate feedback parameter are obtained *a priori* or from the model and used for its calculation) with the sum of the water vapour, lapse rate and albedo feedbacks (obtained using a field by field analysis). The findings from this assessment are presented in the feedback summary in section 4.1.1.6. The derived cloud feedback results contained within the climate feedback parameter ranged from 0.31 to $1.07 \text{ W m}^{-2} \text{ K}^{-1}$, the largest spread of all the main climate feedbacks, and dissimilar to the Cess et al., studies in that no negative cloud feedbacks are obtained and the range is reduced. A more recent study led by Soden (Soden et al., 2008), using the same

models applied in the 2006 study, separated out cloud feedbacks from CRF_{net} using radiative kernels. This avoids the need to compute cloud feedbacks as a residual term (as in the method applied above) and is less sensitive to uncertainties in external radiative forcing. From this evaluation similar outcomes to those presented above were obtained, with neutral to positive cloud feedbacks.

Under the auspices of the upcoming IPCC 5th report on climate change (due for publication in 2014), Andrews et al. (2012) evaluated G , $-\lambda$ and equilibrium climate sensitivity in response to an abrupt quadrupling of CO_2 across the 15 GCMs employed in the Coupled Model Intercomparison Project Phase 5 (Taylor et al., 2012). The first finding, in accordance with previous studies (Webb et al., 2006), demonstrated that the climate feedback parameter, $-\lambda$, conferred greater uncertainty to equilibrium climate sensitivity (2.3 K to 5.5 K) than forcing, G (2.4 K to 4 K). Assessment of the feedback processes contained within the climate feedback parameter was achieved through its decomposition into LW clear-sky, SW clear-sky, and the CRF_{lw} and CRF_{sw} components, which were combined to form CRF_{net} . For the clear-sky components good agreement was found across models, with strongly negative (mainly due to temperature negative feedback) LW clear-sky climate feedback parameters ranging from $1.6 \text{ W m}^{-2} \text{ K}^{-1}$ to $2.0 \text{ W m}^{-2} \text{ K}^{-1}$. The SW clear-sky feedback parameters were also shown to agree well with a positive effect ranging from $0.5 \text{ W m}^{-2} \text{ K}^{-1}$ to $0.9 \text{ W m}^{-2} \text{ K}^{-1}$ (positive mainly due to declining albedo and water vapour SW effects). Greater differences were found in the CRF_{net} analysis, ranging from $-0.5 \text{ W m}^{-2} \text{ K}^{-1}$ to $0.7 \text{ W m}^{-2} \text{ K}^{-1}$, which explains the majority of the range in the climate feedback parameter. The observed CRF_{net} range is similar to that observed by Soden et al., (2008) prior to adjustment for other aliased feedback processes, therefore the authors conclude that similar (i.e. neutral or positive) cloud feedbacks occur within the models assessed in their study.

From this chronological summary it appears that more recent GCMs demonstrate greater cloud feedback consensus, with a neutral to positive cloud feedback. Whether this is due to improved model physics, improved model analysis, or model convergence is uncertain. The studies presented do suggest that the analysis method employed plays a large role in accurate determination of cloud feedback processes, and from this it can be said that analysis by way of CRF_{net} , without correction for various aliased non-cloud feedback processes, appears to be an unreliable indicator of cloud feedbacks. What is definitive is that cloud feedbacks remain the greatest source of uncertainty irrespective of model maturity. In the next section, the cause of this continued uncertainty, the cloud parameterisation problem, is introduced along with methods that can be applied in order to reduce the uncertainty associated cloud feedbacks in models.

4.1.1.10 Cloud and Models – Reducing Uncertainty

GCMs, due to computational and other limitations, typically operate on mesoscale (>100km) resolution grids. This operating resolution is suitable for resolving the large-scale atmospheric motions that drive the climate system. However, the majority of the microphysical and macrophysical cloud processes relevant to climate feedback occur at sub-grid resolution. These processes must therefore be represented through physical approximations, and these approximations are referred to as the cloud parameterisation problem (NRC, 2003; Stephens, 2005). There are two main types of cloud parameterisation - convective and large-scale (Stephens, 2005). Convective parameterisations represent the effects of convection in drying and warming the large-scale atmospheric environment. Large-scale parameterisations define the

cloud properties as functions of the thermodynamic and dynamical processes that are resolved by the model. The parameterisations employed within a GCM (Stephens, 2005) define the modelled cloud characteristics and are the primary cause of the inter-model differences seen in the assessments presented above (Stephens, 2005).

In order to improve cloud parameterisations and in turn reduce feedback uncertainty, proposed methods include more definitive use of observations and the development of techniques that provide more rigorous assessment of models (Stephens 2005). For the latter of these proposals progress has been made and is documented in the previous section. However, much of the focus on understanding feedback uncertainty has been placed on these inter-model comparisons. Whilst these comparisons have proved useful for determining that differences in cloud feedbacks are the primary cause behind the spread in climate sensitivity, they do not provide detailed information on how the various model parameterisations lead to the observed discrepancies (Stephens 2005). Furthermore, as suggested above, model-to-model comparisons may lead to model convergence, but there is no evidence that this convergence indicates that the projections are nearer to the truth (NRC 2003). A more detailed and informative test of the cloud parameterisation employed in a climate model can be achieved through assessing the model's ability to replicate cloudy conditions of the present day, necessitating the use of observational data.

There are a number of ways in which observational data are employed to evaluate the modelled cloud conditions in GCMs, including comparisons of: cloud occurrence (Miller et al., 1999; Hogan et al., 2001); cloud regime properties (Jakob and Tselioudis, 2003); cloud-radiation interaction (Webb et al., 2001); and the relationship between clouds and atmospheric dynamics (Norris and Weaver, 2001). Of these, cloud occurrence evaluation is the most straightforward, and involves quantitatively assessing a model's ability in replicating cloud observations, in terms of fraction and elevation (i.e. macrophysical cloud features). If the model well replicates the macrophysical cloud characteristics seen in observations, then this denotes that the model has good potential for the study of feedback processes as the cloud parameterisation schemes within the model lead to realistic outcomes (Stephens, 2005).

In order to undertake such inter-comparisons, global observational datasets of macrophysical cloud parameters are required, and it is the creation of such datasets from AATSR using stereo photogrammetry which is this chapter's primary aim. There are, however, numerous different techniques and instruments which can be employed to derive observations of macrophysical cloud properties, and prior to discussing the generation of a novel macrophysical climatological dataset from AATSR, it is useful to review the alternative methods of operation and their particular strengths and weaknesses to better place AATSR in the context of observational cloud climatologies. Therefore, in the following section, section 4.1.2, the main approaches for the retrieval of macrophysical cloud parameters, and their particular strengths and weaknesses, are described. In Section 4.1.3 the main instruments which employ these techniques are outlined through a brief review of the GEWEX Cloud experiment (Stubenrauch et al., 2013), four of which, sharing similar orbital characteristics to Envisat, are reviewed in more detail due to their application for climatological inter-comparison purposes in Section 4.4. The need for multiple observational cloud climatologies is also defined, as are the potential unique contributions from AATSR. Finally, given the current strengths and weaknesses of the majority of the current macrophysical cloud

climatology datasets, the decision to generate a cloud climatology over Greenland from AATSR, in the first instance, is justified

4.1.2 Macrophysical Cloud Properties from EO

4.1.2.1 Cloud Top Height Determination – IR Window Technique

The gases and aerosols that compose the Earth's atmosphere, particularly water vapour, CO₂, and ozone, limit the regions within the thermal infrared (IR) part of the EMS where observations of terrestrial exitance can be made. This is demonstrated in the atmospheric transmission profile shown in Figure 4-3. In this profile, a number of atmospheric windows, regions where the atmospheric constituents have limited effect on the outgoing thermal IR radiation (between 3µm and 14µm), are evident. Of these atmospheric windows, the thermal IR window (hereafter referred to as IR-Window), centred on 11µm can be used to derive measurements (e.g. Rossow and Schiffer, 1999) of cloud top pressure (CTP), which defines a clouds vertical location by the mass of the atmosphere above the cloud.

To estimate CTP, the cloud top temperature must be estimated by converting at sensor radiance observations into brightness temperatures¹³. If the cloud is suitably opaque, then the measured BT at the sensor will be a combined contribution of the cloud radiance and any absorption and emission processes by atmospheric constituents along the atmospheric path between the cloud and the sensor. To enable separation of the cloud top temperature (CTT) from extraneous contributions, radiative transfer models, which describe the absorption and emission processes along the atmospheric path length, are employed. Following atmospheric correction, the CTT is converted into CTP through reference to a temperature-pressure profile obtained from a numerical weather prediction model.

The accuracy of the IR-window method is dependent upon a number of factors: primarily the physical representativeness of the radiative transfer model used to retrieve the parameters, and the accuracy of the NWP derived temperature profile applied to infer pressure. The radiative transfer model maps from state to measurement, where the state contains the retrieved cloud parameters, and the measurement the observations. With a single layer cloud radiative transfer model, there is no physically sound solution for mapping the state vector onto the measurement vector given a multilayer cloud system. This shortcoming manifests in the presence of optically thin (< 5 optical depths) ice clouds e.g. cirrus, overlying an optically thick cloud e.g. cumulus (Poulsen et al., 2012). The retrieved CTP is an intermediate function of the CTT from the upper and lower cloud layers. This is particularly problematic as a large proportion of cloud systems are multi-layered (see Wind et al., 2010 Figure 10). Recently, a two-layer radiative transfer scheme applied to IR measurements has shown some usefulness in detecting multi-layer clouds, demonstrating that this limitation is partially surmountable (Watts et al., 2011).

To effectively minimise the cost between the state vector and observations the radiative transfer model also relies on ancillary information that describes the local atmospheric and surface conditions (e.g. water vapour, ozone, albedo, etc.). Errors and shortcomings in these external datasets also lead to degradation in the accuracy of the output parameters. In the case of CTP assignment, a thermal profile is used to assign a CTP to the CTT observation. As an example, the temperature profile in the polar troposphere is

¹³ BT is the temperature a blackbody would have to be to reproduce the observed at sensor radiance.

typically isothermal during summer (Karlsson and Dybbroe, 2010). In these conditions there are potentially many solutions for the CTP, and the retrievals therefore are unreliable (Schweiger and Key, 1992).

Other points which need to be considered are that the IR-window method is less effective for the detection of low clouds due to a lower signal to noise ratio, leading to inaccuracies in the measurement of CTT, and that the method is also susceptible to instrument trending, in particular any degradation in radiometric calibration. Therefore, for IR-window systems, it is necessary to regularly assess the calibration of the instruments to ensure that introduction of spurious CTH trends are kept to a minimum.

4.1.2.2 Cloud Top Height Determination – CO₂ Slicing

The CO₂ slicing technique (Menzel et al., 1983; Wylie and Menzel, 1999), in contrast to IR-window, exploits the partial absorption of outgoing LW radiation within the broad 15µm CO₂ absorption waveband (see Figure 4-3) to infer CTP. This inference is achieved by calculating the ratio of differences between observed cloudy and expected clear sky radiances at two nearby wavelengths from within the waveband. The expected clear sky radiances are obtained from ancillary data sources. The sensors employed often contain multiple bands within the CO₂ absorption region, so various CTP estimates can be obtained. To select the most representative CTP, an alternative CTP is estimated using a forward radiative transfer model, and the observed CTP closest to that of the model is retained.

As CO₂ slicing also relies upon a radiative transfer model the same problems as discussed for the IR-window method, that is the single layer cloud assumption and reliance upon ancillary data, apply here. Other similarities with the IR window technique, in terms of limitations can be seen in the need for accurate instrument calibration.

Problems specific to CO₂ slicing occur in the presence of high-level, optically thin cirrus (<1 optical depths) and more generally to lower level clouds. In the case of an optically thin cloud, there will be significant transmission through the cloud of surface radiation, and therefore the difference between the observed cloud radiance and the expected clear sky radiance will be small, leading to a low signal-to-noise ratio. In these instances the cloud may well be missed. However, for high-level cloud with optical depths of greater than 1, the CO₂ slicing method is one of the most effective height assignment approaches (Naud et al., 2007a; Marchand et al., 2010). Signal-to-noise considerations also preclude the detection of low clouds (> 700 hPA; Menzel et al., 2008). This is due to lower clouds typically being warmer, and therefore emitting a greater proportion of their radiation in shorter wavelengths (i.e. close to 11µm), as defined by Wien's displacement law¹⁴.

4.1.2.3 Cloud Top Height Determination – O₂A Band

The O₂A band method (Fischer and Grassl, 1991), similar in some respects to CO₂ slicing, relies on absorption of radiation, in this case by O₂, centred on the O₂ absorption band at 761nm. The strength of the signal recorded at the sensor is relative to the mass of atmosphere above the cloud, the CTP. Lower

¹⁴ The inverse relationship between the wavelength of peak emission of a black body and its temperature.

clouds have a greater mass of atmosphere above them leading to a reduction in the transmission of scattered insolation radiation in the O₂A band, for high clouds the converse is true. CTP cannot be inferred to a high degree of accuracy from O₂A band observations without consideration of the multiple scattering effects within clouds, which is in turn a function of cloud optical thickness (Preusker and Lindstrot, 2009). Multiple scattering is corrected for using a proximal window channel, to estimate optical thickness. By comparing the ratio of the absorption channel and the window channel to radiative transfer simulations the CTP can be accurately inferred.

The O₂A band method, as with the IR-window and CO₂ slicing methods, employs radiative transfer modelling, and as such shares the same radiative transfer specific limitations as those techniques. The most significant unique limitation of this technique is its inability to detect optically thin and even certain optically thick clouds over bright surfaces (albedo > 0.6). To counter this weakness global albedo models are required to define the local surface albedo. The algorithm's accuracy, is therefore, implicitly dependent upon the accuracy of the albedo model.

4.1.2.4 Cloud Top Height Determination – Stereo-Photogrammetry

An entirely different approach for the retrieval of cloud elevation from EO imaging systems is offered by stereo-photogrammetry. Refer to Chapter 3 for a detailed review of the techniques involved and accuracy limitations. The stereo approach is able to overcome a number of the more significant problems associated with the absolute radiance based techniques discussed previously in this section. Stereo methods perform similarly over all land cover types (assuming sufficient image texture), they can detect clouds at all levels within the atmosphere with similar effectiveness, and depending upon the sensor channel configuration, can detect multiple cloud layers. Importantly, stereo methods work with relative rather than absolute radiances so are not affected by instrument calibration trending.

4.1.2.5 Cloud Top Height Determination – Active Techniques

The observational methods discussed so far all rely upon passive imaging systems. That is, the source of the observed radiation is, either directly or indirectly, solar in origin. Active sensors present an alternative approach, where the observed radiation is emitted by the imaging system itself. The most prevalent active systems are those that operate in the microwave portion of the EMS, and they are referred to as Radio Detection and Ranging systems, or radars. More recently, systems employing lasers, referred to as Light Detection and Ranging systems, or lidars, have been developed for space-borne remote sensing application. The principle of both techniques is to obtain time-of-flight measurements – the time between the emission of the radio wave or light and its return to the sensor – to discern an objects location. Cloud observation specific space-borne radar and lidar instruments are limited to CloudSat (Stephens et al., 2002) and the CALIOP (Winker et al., 2003) instruments respectively.

Active methods provide potentially the most accurate tool for the retrieval of CTH, CALIOP for example can retrieve CTH to accuracy of 30-60m depending on the altitude of the cloud (Vaughan et al., 2009). The main weakness of high vertical resolution cloud radars and lidars are the pencil sampling characteristics giving very small spatial coverage. This is a significant limitation in terms of cloud climatology development, as much meteorological variability is missed.

4.1.2.6 Cloud Fraction Determination

There are a multitude of different techniques employed in the determination of CF. The approaches employed in the major observational cloud climatology datasets generally revolve around some combined temporal and/or spatial radiometric consistency test. ISCCP (see section 4.1.3) for example employs a five step process: firstly a spatial contrast test applied to an IR channel, where cloudiness is assessed using a single scene by assessing thermal variability over a small image region; the second step looks at thermal variability of a given pixel through time, with changes in radiances used to determine the presence of clouds; the third step performs a similar task as the first two steps, but over larger spatial and temporal domains; the fourth step combines the information from the first three steps in an attempt to define local clear sky conditions for a given pixel; given an estimated clear sky radiance, cloudiness is inferred in the final step by setting any pixel to cloudy when it differs from the clear sky radiance by more than estimated uncertainty of the clear sky values (Rossow and Gardner 1993a). Such radiometric cloud detection techniques are very commonly employed (e.g. Menzel et al., 2008; Minnis et al., 2011). More modern radiometric techniques apply machine learning algorithms trained over various imaging channels in order to provide more robust cloud detection as well as classification (Mazzoni et al., 2007a). However in the current major cloud climatology datasets only the more traditional RT approaches are employed.

Alternative CF determination can be achieved through imagers with variable imaging geometry, with MISR (see section 4.1.3.3) being the main proponent (Diner et al., 1999; Di Girolamo and Davies, 1994). The first geometrical technique is based on the fact that cloudiness increases with viewing obliquity due to greater atmospheric path length, and, by employing certain statistical checks the presence of cloud can be inferred across the range of viewing angles. The other geometrical CF determination method is provided by stereo-photogrammetry in combination with a local DEM. The cloud mask is retrieved by comparing the stereo retrieved elevation against that of the collocated DEM, any difference which is greater than the retrieval accuracy of the employed stereo algorithm, is, assuming a noise free case, a cloud feature. The benefit of the geometric approach over the radiometric is that it performs effectively irrespective of the underlying surface; this is typically not the case for radiometric cloud masking algorithms.

4.1.3 EO Macrophysical Cloud Climatologies

4.1.3.1 Macrophysical Cloud Property Datasets – Overview

The most comprehensive review of cloud climatologies is given by Stubenrauch et al., (2013) under the remit of the GEWEX cloud project. In this study all of the major climatological cloud datasets, in terms of macrophysical and microphysical cloud parameters, are reviewed and inter-compared extensively. The climatologies employed in the GEWEX study are as follows: ISCCP (Rossow and Schiffer, 1999), PATMOS-x (Heidinger et al. 2012), AATSR-GRAPPE (Poulsen et al., 2012), MODIS-Science Team (Menzel et al., 2008), MODIS-CERES (Minnis et al., 2011), HIRS-NOAA (Wylie et al. 2005), TOVS Path-B (Stubenrauch et al. 2006), AIRS-LMD (Stubenrauch et al. 2010), CALIPSO Science Team (Winker et al. 2009), CALIPSO-GOCCP (Chepfer et al. 2010), POLDER (Ferlay et al. 2010), and MISR

(Di Girolamo et al. 2010). The instruments use varying techniques to derive the macrophysical cloud parameters, and in terms of CTH or CTP the following divisions can be made: ISCCP, PATMOS-x, AATSR-GRAPE, MODIS-CERES use approaches similar to the IR-window method described in section 4.1.2.1 to assign CTP. MODIS-Science Team, HIRS-NOAA, TOVS Path-B and AIRS-LMD all employ some form of CO₂ slicing (in combination with IR-Window to detect low clouds), described in section 4.1.2.2. The CALIPSO groups use the active techniques described in section 4.1.2.5. The POLDER team employs the O₂A band method described in section 4.1.2.3, and finally the MISR team use the stereo-photogrammetric approach outlined in section 4.1.2.4. Cloud fraction for the climatologies is estimated using the various approaches described in section 4.1.2.6, depending upon the particular instrument characteristics. In this chapter only the climatological macrophysical datasets derived from MODIS, MISR and AATSR are employed for purposes of comparison due to similar or shared observations times (see section 4.4.1), and the techniques they employ to derive macrophysical cloud observations are detailed in sections 4.1.3.2 through 4.1.3.4.

4.1.3.2 MODIS Macrophysical Cloud Property Determination

The NASA MODIS instrument has been operational in sun synchronous orbits since 1999 on Terra (10:30 a.m. equatorial overpass) and 2002 on Aqua (1:30 p.m. equatorial overpass). It is a multi-spectral whiskbroom ($\pm 67^\circ$ from nadir) scanning radiometer with 36 channels between 0.415 μ m and 14.235 μ m at 12-bit radiometric sensitivity. The large scanning angle generates a swath of 2,330km giving near complete daily coverage of the globe, with nadir spatial resolution ranging between 250m and 1km depending on channel. MODIS also has excellent geo-referencing, with pixels geo-located to an accuracy of 150m. There are two separate science teams deriving cloud parameters from the MODIS instruments and they are referred to in the Stubenrauch et al., (2013) paper as MODIS – Science Team (Menzel et al. 2008), and MODIS – CERES (Minnis et al. 2011). In this section the approaches of both teams for CTH and CF determination are reviewed.

MODIS Science Team CTH Determination – The MODIS science team derive CTP by applying either the CO₂ slicing approach (see section 4.1.2.2) or the IR-window method (see section 4.1.2.1) to the appropriate MODIS bands (from channels 31 through 36). The differences in approach are due to the fact that CO₂ slicing is more effective for determining the elevation of mid and high level clouds. As mentioned in section 4.1.2.2, the CO₂ slicing method begins to perform poorly when the cloud signal is similar to the instrument noise levels as tends to occur when either the clouds are lower in the atmosphere or when the clouds are optically thin (<1 optical depth). In the case of low clouds the MODIS science team switch to the IR-window method, which, due to its region of operation within the EMS, is more sensitive to lower clouds than CO₂ slicing. The MODIS CTP measurements are provided at 5km resolution through aggregation of the retrieved CTP values of the cloudy part of a 5 by 5 pixel array. This aggregation is performed to reduce radiometric noise levels. The retrieved CTP is converted into CTH through application of the National Centre for Environmental Prediction (NCEP) Global Forecast System (Derber et al., 1991).

MODIS Science Team CF Determination – CF is assigned at 1km resolution using a radiometric confidence-based system, with four different classifications (confident clear; probably clear; probably

cloudy; cloudy) derived using up to 20 of the 36 possible bands (Ackerman, 1998). The cloud masking process begins by defining a land type (e.g. desert, ocean, snow, etc.) for the pixel under assessment. Following land type assignment, a series of five groups of radiometric thresholds are then applied to estimate cloud coverage: The first group is employed to detect thick high clouds using BT thresholds; the second group of tests is for determining the presence of thin clouds and uses BT difference tests; the third group is used to check for low clouds and employs a reflectance ratio test alongside and a BT difference test; Upper tropospheric thin clouds are detected using an additional reflectance threshold test; the final test is for detecting thin cirrus clouds using a number of BT difference tests. Each cloud detection test returns a confidence level that the pixel is clear ranging in value from 1(high) to 0 (low). The minimum confidence from all tests within a given group is taken to be representative of that group.

MODIS CERES CTH Determination – The MODIS CERES team employ two different daytime retrieval methods, depending on whether the underlying surface is comprised of snow/ice or not. As in this chapter only daytime CTHs are effectively retrieved with the current AATSR L2 stereo processing algorithm (see section 4.3.1.2), and the study is over the Greenland ice sheet, only the Shortwave-infrared Infrared Near-infrared height assignment technique, which is employed over snow and ice, is reviewed here. This approach is essentially the IR-window technique and uses MODIS channel 31 to determine a CTT, which is then assigned a height using a look up table derived from a NWP model. As in most instances the CTT retrieved corresponds more closely to the centre of the cloud/some depth below the cloud top, a further correction is applied to the retrieved cloud temperature to provide better estimate of the true CTT, and in turn CTH and CTP, using retrieved optical depth and cloud emissivity values. The CTT values are converted into CTHs using reanalysis data from the GMAO reanalysis (Stubenrauch et al., 2013).

MODIS CERES CF Determination – Cloud detection by the CERES team is again dependent on whether the underlying surface is snow/ice or not. The cloud masking over snow/ice algorithm (Trepte et al., 2002) employs reflectance and emission RTMs for a modelled snow surface at 0.65, 1.6, 3.75 and 11 microns. The modelled radiances are then used to discriminate between clear and cloudy surfaces.

4.1.3.3 MISR Macrophysical Cloud Property Determination

The Multiangle Imaging SpectroRadiometer instrument, MISR (Diner et al., 1998), has been in operation since its launch alongside MODIS on the NASA Terra satellite in 1999 and so shares the same equatorial overpass time of 10:30 a.m. MISR obtains imagery of the Earth from nine pushbroom cameras arranged in a bank of different along track viewing angles, covering nadir, 26.1°, 45.6°, 60.0° and 70.5°, both fore and aft of nadir. These cameras are referred to as AN at nadir, AA, BA, CA, and DA for the aft camera group, and AF, BF, CF, and DF, for the forward camera group. It takes approximately seven minutes for all nine cameras to view the same point on the Earth. The swath width of the sensor is 360 km and it has a 16 day orbit repeat cycle. MISR obtains imagery in the visible and near-IR region of the EMS (443, 555, 670 and 865 nm) and has two different modes of operation – global and local. The general operation mode, Global mode, provides imagery at 275 m resolution across all channels in the nadir view and in the red channel across all cameras, all other channels and cameras are resampled to 1.1 km to reduce data

load. In the alternative local mode, no data resampling is performed giving a product resolution of 275 m across all channels.

MISR is the only imaging system that employs stereo methods to derive CTH operationally (Moroney et al. 2002) and this data is available in the form of the MISR level 2 TOA Cloud data set (L2TC). The stereo matching algorithm employed for MISR is known as the multi-point matching algorithm. There are two versions employed operationally on MISR, the Multi-point Matcher using means, referred to as M2, and the Multi-point matcher using Medians, referred to as M3 (Muller et al., 2002). These algorithms are the precursors to M4 as employed on ATSR-2 (Muller et al., 2007), introduced in the form of M5 in Chapter 3. The base-to-height ratio of the MISR instrument is dependent upon the combination of views used to derive the stereo observation. The operational algorithms use the three views centred on nadir (AA, AN, AF), giving a base to height ratio of 0.45. This preserves cloud similarity between scenes and allow CTH to be resolved to an accuracy of ~560m (Muller et al., 2002). A significant strength of the MISR stereo, due to it having more than three observations of the same cloud feature, is the ability to separate wind motion induced parallax from CTH parallax (Horvath and Davies, 2001). An assessment of wind induced parallax was made in Seiz (2003) across a limited number of scenes, and found to lead to errors in heights far greater in magnitude than those introduced by the imaging geometry; the ability of MISR to inherently correct for these errors is therefore of great benefit.

For CF determination, MISR employs three separate cloud different masks in its operational product: a radiometric camera-by-camera cloud mask (Zhao and Di Girolamo, 2004; Yang et al., 2007); a stereoscopically derived cloud mask (Diner et al., 1999); and an angular signature cloud mask (Di Girolamo and Davies, 1994; Di Girolamo and Wilson, 2003; Zhao and Di Girolamo, 2004). The radiometric cloud mask generates a per-camera cloud mask using one spectral and one spatial metric which differ whether the observation is over land or ocean. Due to the observational channels available with MISR the radiometric cloud mask performs inadequately over snow and ice, and over these land surface types the two other cloud masks are employed. The stereo cloud mask is straight forward and sets any pixel to cloudy if it is more than 562m above the underlying terrain. The angular cloud mask operates by looking at the change in value of the difference between the 446nm and 886nm bands across the MISR view angles. If the slope between the differences is large and positive then the feature is cloud. Small positive or negative slopes indicate snow, ice, or land, and large negative slopes indicate water features.

4.1.3.4 ATSR Grape Macrophysical Cloud Property Determination

The ATSR instruments (see Chapter 2), due to the spectral configuration and viewing geometry, are capable of both radiometric CTP and geometric CTH retrieval. The main radiometric approach currently employed on ATSR-2 and AATSR is the ORAC algorithm (Sayer et al. 2011, Poulsen et a. 2012), which employs an optimal estimation algorithm originally developed for the Spinning Enhanced Visible and InfraRed Imager carried onboard the current Meteosat Second Generation weather satellites (Watts et al., 1998). The optimal estimation method, in addition to retrieving microphysical and macrophysical cloud parameters, also provides estimates on the quality of the retrieval, which is vital for effective assimilation of the data into climate models. The optimal estimation method employs all ATSR channels in its

retrieval algorithm. However, the main contribution to the observed CTP comes from the 11 μ m and 12 μ m BT channels. As these channels are within the infrared IR-window waveband, and height assignment is achieved in a similar manner to that as the IR-window method, therefore the same benefits and limitations apply. The AATSR CTP values are converted into CTHs using reanalysis information from ECMWF (Stubenrauch et al., 2013).

The ORAC authors state that the CF output from the optimal estimation retrieval is generally of a low quality and not reliable and this is reportedly due to the fact that the cloud fraction retrieved states tend to change erroneously to compensate for other inadequacies within the retrieval. So the current ORAC CF outputs are not expected to perform well.

4.1.3.5 The Need for Multiple Cloud Climatologies

The ATSR, MODIS and MISR instruments introduced above and their respective macrophysical cloud property determination algorithms, and indeed all other instrument/algorithm combinations reviewed in the GEWEX cloud study, all return differing cloud characteristics. All radiance based algorithms (IR-window, CO₂ Slicing, O₂A band), discussed in section 4.1.1.2, assign CTP or CTH through application of ancillary, reanalysis derived, atmospheric temperature, pressure, or geopotential height profiles. Differences in the employed ancillary datasets lead to differences in the returned CTP or CTH estimates. Furthermore, these techniques tend to experience poor performance in the presence of temperature inversions, such as with stratocumulus clouds (Wang et al., 1999; Menzel et al., 2008; Garay et al., 2008; Marchand et al., 2010) or in the polar troposphere during summer months (Karlsson and Dybbroe, 2010). Such problems are not encountered by active instruments or those which employ geometric approaches for height assignment. MISR for example, can provide reliable CTH estimates, with small biases, in isothermal conditions due to the stereo techniques employed (Marchand et al., 2010).

All passive algorithms generally retrieve a radiometric height which lies some way between the cloud top and the cloud base; in some cases a number of kilometres below the true cloud top (Stubenrauch et al. 2013). The stereo algorithms employed on MISR and AATSR require sufficient image texture to perform effectively. Therefore, stereo retrievals tend to relate to where the cloud optical depth reaches one to two (Naud et al., 2004) rather than the true cloud top identified by lidar or radar. In the case of MISR, for low clouds, which tend to be optically thicker, differences of less than 0.5 km (near to the potential accuracy of the MISR product, which is ~560m) are typically found against comparisons with active instruments (Naud et al., 2004; Marchand et al. 2007). In the case of mid- to high- level clouds, which tend to have lower optical thicknesses, larger biases of up to 3km are found, with the retrieved CTH tending to correspond to the cloud base (Naud et al., 2004; Marchand et al. 2007). The CO₂ slicing algorithm, employed by MODIS Science Team, shares similar characteristics for high cloud retrieval to MISR, tending to retrieve the CTH at an elevation corresponding to a cloud optical depth of one to two (Menzel et al., 2008), with biases in the region of 3km (Naud et al., 2004). In the presence of low or semi-transparent clouds (optical depth of <1, Naud et al., 2007a) the CO₂ slicing method no longer functions effectively due to signal to noise considerations (the cloudy-sky signal is very close in magnitude to the clear-sky signal) and the algorithm reverts to the IR-Window method, which is also employed by the MODIS-CERES, AATSR-GRAPES and a number of other science teams in the GEWEX cloud study. The

outputs from the IR-Window method tend to provide better CTH retrievals for clouds at low-levels than mid- to high- levels (Liao et al. 1995; Wang et al. 1999; Naud et al., 2004; Minnis et al. 2008; Stubenrauch et al. 2010), however the results are often biased low (i.e. associated with the CBH rather than CTH) when the optical depth is less than 3 (Stubenrauch et al., 2013), and can be some way below the true CBH when the optical depth is low enough to allow for radiative contributions from the underlying surface (Naud et al., 2004; Marchand et al., 2010; Stubenrauch et al., 2013). Correction for these radiative CTH biases can be made with knowledge of cloud emissivity and/or optical depth, however in the case of multiple cloud layers such corrections are often underestimated (Jin and Rossow, 1997).

Multi-layer cloud systems are a common occurrence (as shown in Wind et al., 2010, Figure 10) and present a significant problem to the majority of CTH assignment algorithms. The algorithm most susceptible is the IR-window method, which often exhibits CTH retrievals laying at an intermediary height between the true layers (Marchand et al., 2010; Stubenrauch et al., 2013). This occurs due to both the tendency of upper clouds types (e.g. cirrus) to be optically thin (typical cirrus optical depth <5; Poulsen et al., 2012) allowing radiative contributions from warmer, underlying cloud features to bias the retrievals high (high in terms of radiance/CTP, low in terms of CTH), and the general application of single cloud layer RTMs, which have no physically sound solution for multi-layer clouds (Wind et al., 2010; Watts et al., 2011). The CO₂ slicing method, as long as the cloud has an optical depth of greater than one and is at an elevation of more than 700 hPa, detects only the upper cloud layer CTPs (Naud et al., 2004; Naud et al., 2007). However, if the optical depth falls below one, thermal influence from the lower cloud layers is observed by the sensor, also leading to an over estimation of CTP (Baum and Wielicki, 1994; Rossow et al., 2005). In the case of MISR, the CTHs tend to relate to the most reflective (highest contrast) cloud feature in the field of view, which is typically the lowest cloud feature observed (Naud et al., 2004). Therefore MISR generally “sees through” optically thin high clouds, such as cirrus, and effectively retrieves the lower cloud layers (Marchand et al., 2010; Stubenrauch et al., 2013).

CALIOP, the sole active system employed in the GEWEX study, is the only sensor that can provide accurate CTH observations (to within 30-60m of the cloud top) and effective multiple cloud layer retrievals (Stubenrauch et al., 2013). The multi-layer cloud retrieval potential is again limited by the optical thickness of the overlying cloud; with clouds of an optical depth of greater than 3 being effectively opaque (i.e. no further underlying cloud layers are retrieved due to complete attenuation of the lidar signal, Vaughan et al., 2009). CALIOP also brings to light sampling issues, which, given its pencil sampling (see section 4.2.3.4) samples only a small percentage of the Earth atmosphere, leading to noisy climatologies (Stubenrauch et al., 2013). Sampling differences in the form of overpass time can lead to differences of a few percent across different datasets obtained using shared approaches (Stubenrauch et al., 2013). Furthermore, the general top down observation approach of the majority of passive EO instruments can also lead to differences in the retrieved cloud characteristics. Low- and mid-level clouds tend to only be observable in single layer cloud conditions, i.e. when there are no overlying clouds at high altitudes, potentially leading to underestimation of low- and mid- cloud amounts (Stubenrauch et al., 2013).

In terms of CF, retrieval performance depends greatly on instrumental capabilities (e.g.: look angles; channels of observation; imaging resolution). For example observations from higher resolution instruments will provide higher estimates of cloud amount due to increased sensitivity to sub-pixel clouds, imagers with channels observing close to 1.6 microns will be better able to discriminate between cloud and snow/ice features, and more oblique look angles will tend to increase cloud amount due to perspective distortions (Wielicki and Parker, 1992; Di Girolamo and Davies, 1997; Rossow et al, 1993b; Cawkwell et al., 2002; Maddux et al. 2010; Stubenrauch et al., 2013). According to the GEWEX cloud study, global mean CF for clouds of an optical depth of greater than 0.1 is ~0.68. Taking into account sub-visible cirrus, only observable by CALIOP, the mean global CF increases to 0.73, clearly indicating the importance of approach.

Observational cloud macrophysical cloud climatologies, due to the various reasons outlined above, tend to be complimentary, with more observations allowing for more effective insights to be made about cloud characteristics (Marchand et al., 2010). This is particularly the case in Polar Regions where current observational climatologies show substantial disagreement, contrary to all other regions where agreement between climatologies is high (see Stubenrauch et al., 2013 Figure 2.). It is the case that the combination of AATSR and stereo photogrammetry has the potential to provide an important contribution to the current suite of observational cloud climatologies, providing useful complimentary global observations, and much needed observations where a number of the current main algorithm/instrument pairings tend to perform poorly (e.g. Polar Regions), and this potential is discussed in the next section.

4.1.3.6 The Potential Contribution of AATSR Stereo

There are two current limitations of the observational climatologies reviewed in GEWEX where AATSR in combination with stereo photogrammetry is likely to provide unique or strongly beneficial insights. The first is a general inability of all current single instrument/algorithm setups to effectively resolve multiple cloud layers at a suitable spatial sampling (whilst CALIOP is capable, its spatial sampling is prohibitive). AATSR has the potential, as shown in previous studies (Naud et al., 2007b), to effectively resolve both high and low cloud features through application of stereo photogrammetric algorithms to its different channels of observation. The second is a general lack of consensus between cloud climatologies in Polar Regions (particularly the North pole) as demonstrated in Stubenrauch et al., (2013, see Figure 2). This lack of consensus probably stems from the general inability of radiometric techniques to effectively resolve cloud parameters in Polar Regions due to small thermal gradients (Karlsson and Dybbroe, 2010). This, combined with the fact that only two of the twelve datasets studied in GEWEX do not rely on radiometric methods and associated thermal gradients, means that there is a gap in the knowledge of cloud parameters in Polar Regions (Stubenrauch et al., 2013). AATSR stereo would provide a much needed non-radiometric based contribution to the current observational climatologies for improving polar cloud observations. Furthermore, previous work has demonstrated the potential for employing ATSR in combination with stereo for effective CF and CTH determination at high latitudes (Cawkwell et al., 2001; Cawkwell et al., 2002; Griggs and Bamber, 2008).

In terms of general complimentary contributions to the current set of cloud climatologies, the following can be identified for an AATSR stereo climatology: the potential to retrieve CTH to an accuracy of 1km

(Cawkwell et al., 2001; Naud et al., 2007; Muller et al., 2007; see section 4.3.2 for census accuracy); the potential for effectively retrieving CF for all clouds at elevations of >1km using a stereo derived cloud mask (Cawkwell et al., 2001; Cawkwell et al., 2002); diurnal stereo CTH and CF observations; three separate ATSR instruments, with potentially 20+ years of stereo CTH/CF observations available; the potential for synergy with the ATSR GRAPE retrievals; and a geometric retrieval method which is unaffected by instrument trending.

4.1.3.7 A Case Study – Greenland

Processing an entire AATSR orbit with the census algorithm introduced in Chapter 3 takes approximately 1.5 hours. In parallel across the 160 cores available on the MSSL Imaging Group Linux server, processing the ~50000 AATSR orbits would take around 20 days of processor time. Therefore an initial case study over a reduced geographical region is preferred for assessing the potential of the AATSR stereo contributions identified in the previous section. The geographical subset chosen is that which comprises Greenland [latitudinal range: 50°N-85°N; longitudinal range: 80°W-5°E], which, due to its extensive ice sheets, makes it a particularly challenging region for the majority of current cloud property algorithms (Stubenrauch et al., 2013).

The remainder of this chapter documents the generation of the Greenland cloud climatology from AATSR, beginning with the generation of the L2 census stereo CTH and radiometric CF (see section 4.3.1.2 for justification) products in section 4.3. Validation of these products is undertaken against higher resolution (in terms of CTH) observations from MISR and CALIOP; Section 4.3 also contains the algorithm description of the climatological aggregation process for generating the stratified L3 CFbA product. In Section 4.4 the AATSR monthly and yearly L3 CTH and CF climatologies for 2008 derived from the stratified products are inter-compared against collocated climatologies from the MODIS, MISR and AATSR-GRAPE datasets introduced in Section 4.1.2. From this inter-comparison the additional insights into cloud characteristics over Greenland which AATSR provides are further clarified. Finally section 4.5 provides a summary, addresses the scientific objectives and provides recommendations for future work.

4.2 Scientific Aims and Objectives

- Generate L2 CTH and CF products for the AATSR data time series over Greenland.
- Determine the accuracy of AATSR stereo CTH retrieval against higher accuracy observations of CTH from MISR and CALIOP.
- Determine any gain in performance achieved through application of the census stereo matching algorithm over that of the M4/M5 stereo matching algorithm.
- Aggregate the L2 AATSR products into height stratified macrophysical cloud climatologies containing CF on a by height basis.

- Inter-compare the stereo derived AATSR cloud climatology against the GEWEX CTH and CF climatologies from MODIS (Science Team and CERES), MISR, and AATSR-GRAPE to determine any new data contributions provided by AATSR stereo.

4.3 Cloud Climatology Development

4.3.1 L2 Processing

The first stage in generating a stereo-derived cloud climatology is to process the L1 geo-rectified AATSR radiance and reflectance imagery into meaningful stereo derived CTH and CF observations. The section deals with this process, detailing the AATSR data employed, the stereo processing chain and the validation of the output products.

4.3.1.1 AATSR Data and Stereo Processing

The entire AATSR time series for the geographical subset [latitudinal range: 50°N-85°N; longitudinal range: 80°W-5°E], was made available by the NEODC at RAL. A Python based stereo processing chain incorporating the census transform introduced in Chapter 3 was implemented and used to process the AATSR data into CTH on the MSSSL Linux server. The processing chain is split into the following stages:

1. Ingest the AATSR data and select the channel to be processed. If the channel observes in the visible region of the spectrum the SZN angles are assessed to determine those image pixels with SZNs of greater than one degree (the census algorithm was found to be robust even to very low illumination angles). Any illuminated regions are then marked for further processing. If a mid- or thermal-IR channel is being processed then the SZN check is not employed.
2. The nadir and forward views of the channel under processing and then subset into overlapping x, y subsets of 512 by 512+overlap pixels in order to reduce memory overhead and improve processing performance. The amount of overlap is determined by the sum of the census transform radius and the expected CTH induced displacement in pixels. The census transform radius is set at five pixels, as this provides suitable discriminability for the matching algorithm, whilst limiting stereo smoothing effects and processing costs. The y -search radius for the matching algorithm is set to 17 pixels, as in Polar Regions the Tropopause is generally far below the ~17km elevation such a search radius will allow for, and therefore all cloud features of interest are detectable with this search limit.
3. The census stereo matching algorithm (see Chapter 3) is applied to each forward and nadir subset. Following processing the central 512 by 512 pixel region (i.e. non-overlapping region) from each subset is extracted, the along track disparities corrected using the AATSR co-registration correction coefficients (see Chapter 2) and converted into above ellipsoid heights using the Mannstein camera model (see Chapter 3). The extracted heights and across track disparities are then placed into the correct location in output arrays. Once processing is

complete the filled output arrays are written out in geotiff format along with associated georeferencing information.

This processing chain was used in parallel on the MSSSL Linux server to process the AATSR 11 μ m, 1.6 μ m and 0.67 μ m channels for the entire AATSR time-series for the previously given geographic region. These channels are employed as they each give unique information on CTH (Muller et al., 2007, Naud et al., 2007b; see Figure 4-5). The unprocessed channels do not give any new CTH information and tend to perform with lesser efficacy in CTH determination (e.g. the 0.67 μ m and 0.55 μ m channels tend to observe the same cloud features, however, shorter wavelengths more effectively detect atmospheric aerosols, this leads to corruption in the CTH signal and a reduction in the stereo matching algorithm performance, therefore the 0.55 μ m channel tends to provide less accurate CTH observations than the 0.67 μ m channel). An example output from the processing chain for each channel is shown in Figure 4-5.

4.3.1.2 Post-Processing

Stereo matching algorithms, due to a reliance on suitable image texture to locate distinct matches, fail to perform effectively in low texture regions. In the presence of low texture, matching is ambiguous, leading to noisy stereo outputs and, in this instance, erroneous CTH retrievals. Low texture in EO imagery can be caused either by low illumination levels (in the visible channels), or particular terrestrial features with extensive homogenous regions, such as ice sheets. This presents a particular problem over Greenland, as the majority of its surface is covered in ice and low illumination angles are common throughout polar winter. The effects of low illumination angles are not particularly problematic for the census stereo matching algorithm, as it is robust to even very low SZNs. Low texture regions are far more problematic, and lead to significant errors in the census outputs as shown in Figure 4-5. The previous ATSR study on cloud properties over Greenland by Cawkwell et al. (2002), overcomes this problem by applying a radiance threshold to the 1.6 μ m channel to effectively delineate low texture snow and ice features. This is possible due to the low reflectance of solar radiation of ice and snow in that region of the EMS as shown in Figure 4-5. However, the Cawkwell method is quite simplistic and there is no information in the paper on either the value of threshold or how it is derived.

More recently, Istomina et al (2010) developed an automated technique for the detection of cloud-free snow-covered areas using AATSR. The approach employs all seven AATSR channels across five different RTM derived thresholds: $(3.7\mu\text{m} - 11\mu\text{m}) / 3.7\mu\text{m} < 3\%$; $(3.7\mu\text{m} - 12\mu\text{m}) / 3.7\mu\text{m} < 3\%$; $(0.87\mu\text{m} - 1.6\mu\text{m}) / 0.87\mu\text{m} > 80\%$; $(0.87\mu\text{m} - 0.67\mu\text{m}) / 0.87\mu\text{m} < 10\%$; $|(0.67\mu\text{m} - 0.55\mu\text{m}) / 0.67\mu\text{m}| < 40\%$. Any check which fulfils all of these thresholds is regarded as being ice or snow. The performance of the above threshold combination was assessed against lidar observations over 100 AATSR scenes and was found to be 95% accurate in detecting cloud free regions over snow. The missed cloud detections by AATSR tended to be high, optically thin, cirrus clouds. The approach can therefore be used to screen the problematic snow and ice regions and also provide an effective cloud mask over the Greenland ice sheet.

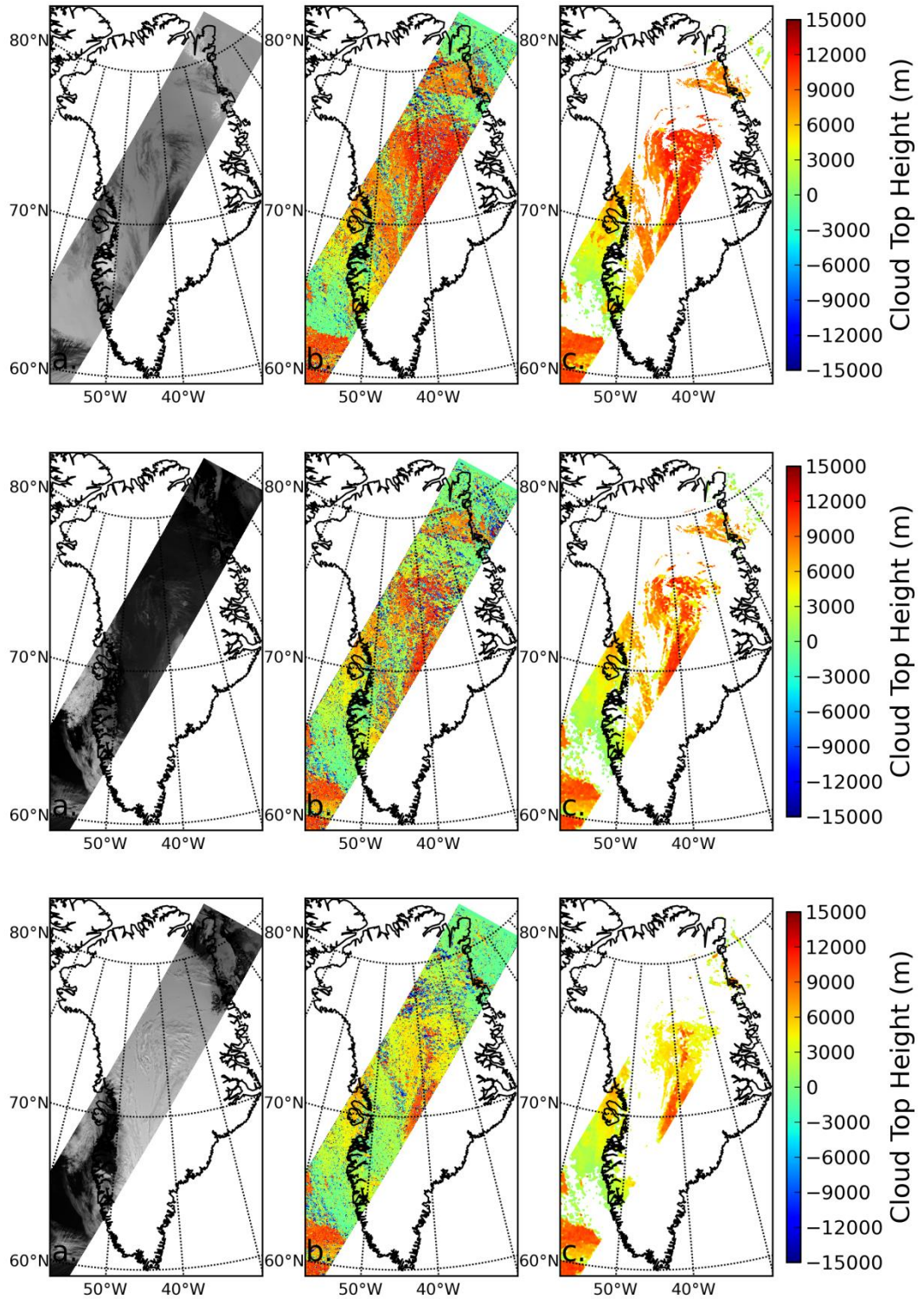


Figure 4-5. These three figures present the stereo outputs from the $11\mu\text{m}$ (upper plot), $1.6\mu\text{m}$ (middle plot), and $0.67\mu\text{m}$ (lower plots) channels. The left column is the input data, the middle column shows the raw output stereo heights, and the right column the heights after post-processing with the Istomina cloud mask and a 500m height threshold applied. Note the differing detection abilities, with the $11\mu\text{m}$ channel detecting far more high cloud than the other channels. Also note the very low reflectance of snow and ice

in the $1.6\mu\text{m}$ channels. Finally observe the poor performance of the stereo matching algorithm of low texture ice regions, resulting in negative and erroneous positive stereo outputs, which lead to false positive cloud retrievals.

One significant drawback, however, is that the threshold process requires the visible channels in its operation; precluding year round screening of low texture features/cloud masking. Once the low texture regions have been screened, effectively generating a cloud mask, a 7×7 median filter is applied to the raw CTHs to reduce the effects of stereo matching noise within the detected clouds. The results of this post-processing can be seen in Figure 4-5.

4.3.1.3 Wind Correction

A further step, which is often needed and applied in other studies using stereo observations for CTH determination, is correction for wind induced along track disparities (Horvath and Davies, 2001; Seiz, 2003). Along track wind, depending upon the node of the satellite and the prevailing wind direction, can lead to an under- or over-estimation of cloud top height. Correction for such wind displacements can be achieved through either the instrument having three or more views of the scene, as with MISR (Horvath and Davies, 2001), dual stereo satellite operation (see Chapter 6), or through the use of ancillary data, such as wind products from geo-stationary or polar orbiting instruments, or numerical weather prediction outputs. For this study wind corrections are not applied to the AATSR stereo output, as tropospheric wind speeds over Greenland tend to be less than $\pm 2\text{ms}^{-1}$ according to both a 10-year analysis of latitudinally and longitudinally averaged observationally derived zonal and meridional wind components from MISR presented in Figure 4-6. In order to bias the retrievals the local wind speeds must exceed at least 8ms^{-1} , as this equates to approximately a one pixel shift in cloud position between the stereo acquisitions. It can be seen from the MISR analysis, that above 9km the wind speed is on average greater than 8ms^{-1} , so biases of a few hundred meters may occur for higher clouds.

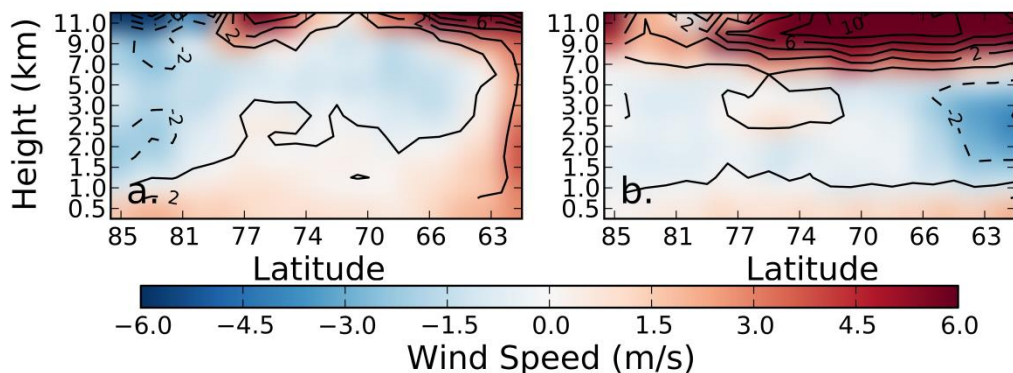


Figure 4-6. This figure presents the latitudinal 10 year average (2002-2011) wind speeds by elevation over Greenland from the MISR L3 AM1_CGCL yearly wind component product. The left plot shows the zonal (positive, west-to-east; negative, east-to-west) wind component averages, the right the meridional component averages (positive, south-to-north; negative north-to-south). The contours plot wind components from $\pm 12\text{ms}^{-1}$ at 2ms^{-1} quantisation, with filled lines representing

positive wind components and the dashed lines negative components. The applied colour look up table has been scaled to $\pm 6\text{ms}^{-1}$ to improve the visualization. The high meridional winds over the Southern portion of Greenland blowing in the south-to-north direction above 7km are likely caused by the annual variations in the polar jet (see Figure 4-13).

4.3.2 Level 2 Validation

The post-processed census stereo L2 outputs are extensively validated against alternative, higher resolution observations of CTH from both the MISR and the CALIOP instruments in this section. Furthermore an additional inter-comparison is made using the M5 algorithm applied to the same AATSR data in order to provide an insight into any gains in performance granted through the application of the census stereo algorithm.

4.3.2.1 Validation against MISR

For the validation against MISR, a search was made using the COVE tool (Chander et al., 2010) for the month of July 2008 for all collocated AATSR-MISR overpasses with a time difference of less than 10 minutes between acquisitions. With this threshold in place a total of 17 orbits pairs were located. The most recent MISR TC_CLOUD products (Mueller et al., 2013), which provide superior coverage and accuracy compared to the previous TC_STEREO products (Moroney et al., 2002), associated with the defined orbits were extracted from the MISR data server. The TC_CLOUD dataset is comprised of a number of different variables, the variable selected for this inter-comparison is the *CloudTopHeight_WithoutWindCorrection* product, as this provides a more direct inter-comparison (as no wind correction is applied to AATSR). The AATSR L2 data employed is that derived from the $0.67\mu\text{m}$ channel, chosen for its similarity to the $0.67\mu\text{m}$ channel which is employed on MISR to derive stereo observations of CTH.

To collocate the datasets the geographic grids associated with each swath are employed to resample both datasets to a shared grid defined by the Bamber et al., (2001) Greenland elevation model (which is employed in the GMTED2010 DEM dataset (Danielson and Gesch, 2011)). Non-common areas (i.e. regions with no shared AATSR-MISR orbit) are screened from further analysis. Prior to re-projection to the Greenland grid, the MISR data is smoothed with a 5x5 pixel median filter. This image processing is performed due in an attempt to account for the poor geo-referencing of the AATSR data (see Chapter 5 for more details), which is thought to be good to $\pm 2\text{km}$. Further screening is performed using the collocated DEM so that only points over Greenland are employed in the analysis (this is done as this current study is interested only in those clouds over Greenland). Lastly, any AATSR or MISR clouds which are less than 500m above the terrain are excluded from the analysis. This threshold is chosen as neither the census algorithm nor the MISR instrument (Muller et al., 2002), due to stereo accuracy limitations, are able to provide an accurate height assessment on clouds which are within a 500m proximity to the surface. The results of collocation methodology from the 17 collocated orbits are presented for the census and M5 stereo algorithm outcomes in scatterplot form in Figure 4-7.

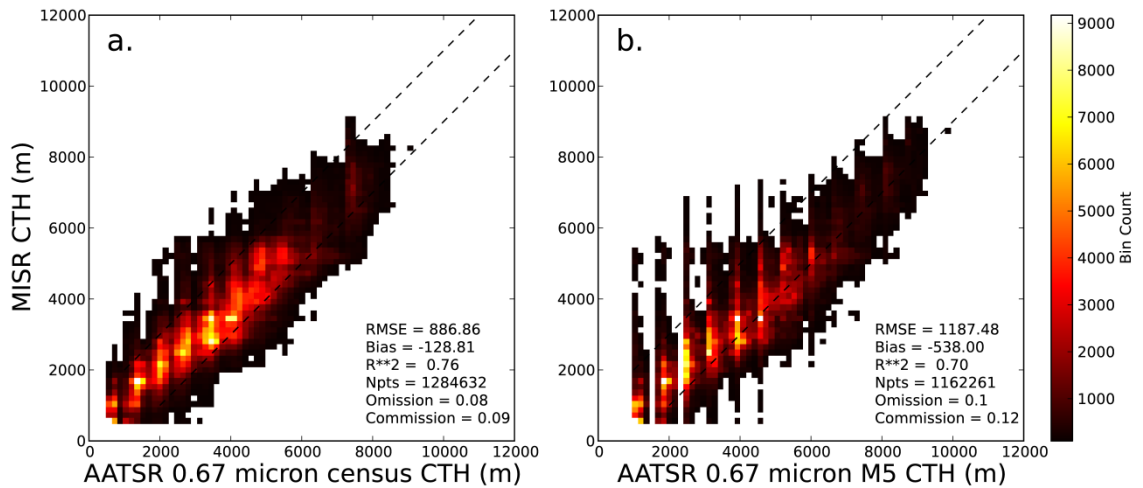


Figure 4-7. This figure presents the inter-comparison of AATSR 0.67 μ m census stereo outputs against collocated MISR stereo observations in the left histogram, and the same inter-comparison but employing M5 to process the AATSR data in the right histogram. There is generally very good agreement between AATSR census and MISR, with the majority of height observations within 1km of each other, this agreement is represented in the statistics. The M5 outputs demonstrate less agreement with MISR than census and appear to be noisier.

4.3.2.2 MISR Discussion

The inter-comparisons between AATSR census, and AATSR M5, CTH estimates and those from MISR compare over ~1 million samples of CTH obtained from the 17 orbit pairs. The statistical analysis (inset in the scatterplots in Figure 4-7) for the census algorithm, following an outlier removal process (all points greater than 3 sigma from the mean excluded), gives an R^2 value of 0.79, demonstrating very good agreement with MISR, a small bias of -129m, and a RMSE of 887m, which is very near the limit of AATSR's theoretical accuracy for pixel-level accuracy stereo matching, ~830m (Seiz, 2003). It should be made clear that the accuracy of the census stereo output, even with the application of the sub-pixel interpolation routine, is still at best (i.e. taking the sub-satellite nadir and forward pixels, that give the optimal base to height ratio) ~800 meters. This is due to that fact that the co-registration correction coefficients (Fisher and Muller, 2013) in the along track direction improve co-registration accuracy to ~0.8 pixel (see Chapter 2). The M5 inter-comparison with MISR results in an RMSE of 1187m, a bias of -538m, and an R^2 value of 0.7, demonstrating that census retrieves CTH to a greater degree of accuracy with reference to MISR.

Statistics of omission and commission are also generated between the AATSR and MISR observations to provide further insights into the census and M5 stereo algorithm characteristics. If the MISR cloud observations are assumed to be the truth, then an error of omission would be an instance where MISR detects a cloud feature and AATSR does not, and an error of commission the converse of this. This form of analysis enables determination of errors of representation, i.e. is one instrument or algorithm seeing far more cloud than the other? To calculate the omission statistic the ratio of the number of cloudy pixels

that both MISR and AATSR detect to the total number of MISR cloudy pixels is calculated. The commission statistic is the ratio of the number of cloudy pixels that both MISR and AATSR detect to the total number of AATSR cloudy pixels. For census, there is an omission percentage of 8%, that is AATSR misses 8% of the clouds that MISR detects. The commission statistic for census is 9%, indicating that MISR misses 9% of the clouds that AATSR observes. The difference in total cloud amount between the datasets is therefore ~1%. In the case of the M5 algorithm, the omission statistic is 1%, whilst the commission statistic is 12%, equating to an 11% difference in the amount of cloud detected by the two datasets.

The variability between the census and M5 cloud amounts is due to the cloud mask being comprised of two separate components in this instance: one radiometric, the Istomina cloud/ice mask; and one geometric, exclusion of all CTHs which are less than 500m above the surface terrain. The source of the differences is the increased smoothing characteristics of the M5 algorithm. The radiometric cloud mask is effective at detecting cloud features below 500m, whilst the MISR and AATSR stereo masks are not. Therefore the radiometric mask extent tends to be greater than that of the stereo mask. Due to the smoothing effect of the M5 algorithm (see Chapter 3) the extent of the stereo cloud mask (i.e. regions with a CTH elevation of >500m) may be increased, and dilation of stereo cloud features into the radiometric mask will occur, leading to the observed bias in the M5 algorithm outputs. This analysis provides further insight into the problems of the M5 algorithm and the distortion of cloud features, as whilst the total extent of the cloud mask in the L2 product is defined by the radiometric mask boundaries, the smoothing errors of M5 may potentially lead to biasing of cloud amounts at certain elevations. Such biases, in this analysis, appear to be reduced through application of the census algorithm due to its reduced smoothing effects (see Chapter 3).

This omission-commission analysis also demonstrates that AATSR census geometric masking alone provides a similar CF estimate to that of the MISR masks employed over snow and ice (stereo and angular masks, see section 4.1.3.3), which is very encouraging, and implies in regions with high-texture (i.e. that do not confuse the matcher leading to erroneous heights) a simple height threshold mask may well be suitable for AATSR census stereo CF determination.

4.3.2.3 Validation against CALIOP

CALIOP has been making measurements of clouds and aerosols since 2006. The instrument is located in the NASA A-Train satellite constellation and therefore has an equatorial overpass time of approximately 1:30 p.m. and a 16-day orbital repeat cycle. The lidar has a ground footprint on the ellipsoid of 100m and pulses every 333m along track. It receives backscattered radiation in three channels, two at 532nm with sensitivity to the backscattered intensity at orthogonal polarisations and one at 1064nm. The vertical resolution is between 30-60 metres depending on the altitude of the cloud, with 30m resolution achievable in the troposphere (Vaughan et al., 2009). If the uppermost cloud layer has an optical depth of less than three (Vaughan et al., 2009), then CALIOP is able to detect the presence of lower cloud layers.

For the period April to October 2008, all collocated AATSR-CALIPSO orbits over Greenland with less than 10 minutes between overpasses (to reduce the effects of clouds being displaced by winds) are extracted for this analysis, giving a total of 70 CALIPSO orbits split between the months of April, July,

August, September and October (no collocations fulfilling the requirements were found for May or June). The CALIOP L1 data is processed into various L2 products, of which the 1km cloud product, CAL_LID_L2_01kmCLay-ValStage1-V3-01, is applied for validation here due to its similar resolution to the AATSR instrument. The lower limit for cloud detection for this product is a backscattered signal of greater than $1 \times 10^{-3} \text{ km}^{-1}\text{sr}^{-1}$ (equivalent to an optical depth of 0.01 for cirrus clouds; McGill et al, 2007, Kahn et al., 2008; Vaughan et al., 2009). The products are similarly masked as with the MISR analysis, with all AATSR and CALIOP cloud features below 500m excluded from the analysis, and the data is resampled onto the Bamber Greenland DEM grid for analysis. To account for poor AATSR georeferencing in this analysis the bounding box method applied for AATSR validation against CALIOP in the ALANIS study is employed (see Chapter 5). The results of the inter-comparison over Greenland for the $0.67\mu\text{m}$, $1.6\mu\text{m}$ and $11\mu\text{m}$ channels are presented in Figure 4-8. A further analysis which assesses how often a given AATSR channel detects the uppermost cloud layer detected by CALIOP is presented on a monthly basis in Table 4-1. This table also presents the monthly mean height and optical depth of the uppermost cloud layer detected, and the monthly mean height and optical depth of the uppermost CALIOP cloud layer when it is missed by AATSR.

4.3.2.4 CALIOP Discussion

The CALIOP inter-comparison compares over 15000 samples of CTH data, with comparisons against the visible, NIR and thermal AATSR channels for stereo outputs from both the census and M5 algorithms. Statistics have been generated from the inter-comparison, with, in all cases, points greater than three standard deviations from the mean difference being rejected from the analysis. The best comparison, see Figure 4-8, is achieved by the $11\mu\text{m}$ channel with the census stereo algorithm, with an R^2 value of 0.78, a root mean square error (RMSE) of $\sim 1200\text{m}$, and a bias of 45m (AATSR heights are on average 45m below those of CALIOP). This small bias value is in contrast to the census results from the NIR and visible channels, with AATSR returning heights 348m below those of CALIOP for the $1.6\mu\text{m}$ channel and 678m below CALIOP for the $0.67\mu\text{m}$ channel. These biases, whilst large for the NIR and visible channels are substantially less than those exhibited when CALIOP is compared to other passive sensing techniques. Studies examining collocated MODIS and ISCCP observations demonstrate underestimations typically to the order of 2.5km (Holz et al. 2008; Vaughan et al., 2009).

The inter-comparison between CALIOP and M5 (also shown in Figure 4-8) demonstrates that census performs more effectively for this comparison. M5's RMSE values are not as good as those from census, ranging from 1400m-1700m for M5, versus 1200m-1500m for census. So the census CTH estimates typically show a tighter distribution, i.e. less noise, than the M5 results. This assertion of improved performance is further backed up by the R^2 values, which range from 0.70-0.79 for census, and 0.58-0.63 for M5, showing that census better describes the CALIOP data. The biases for the thermal and NIR channels are larger for M5 than for census, though in both cases, the difference is less than 100m: 45m against 177m for the thermal channel, and 375m against 413m for the NIR. In the visible channel the bias for the census retrieval is greater than that of M5: 695m versus 424m. The cause of this reduced bias for M5 in the visible channel is not currently understood.

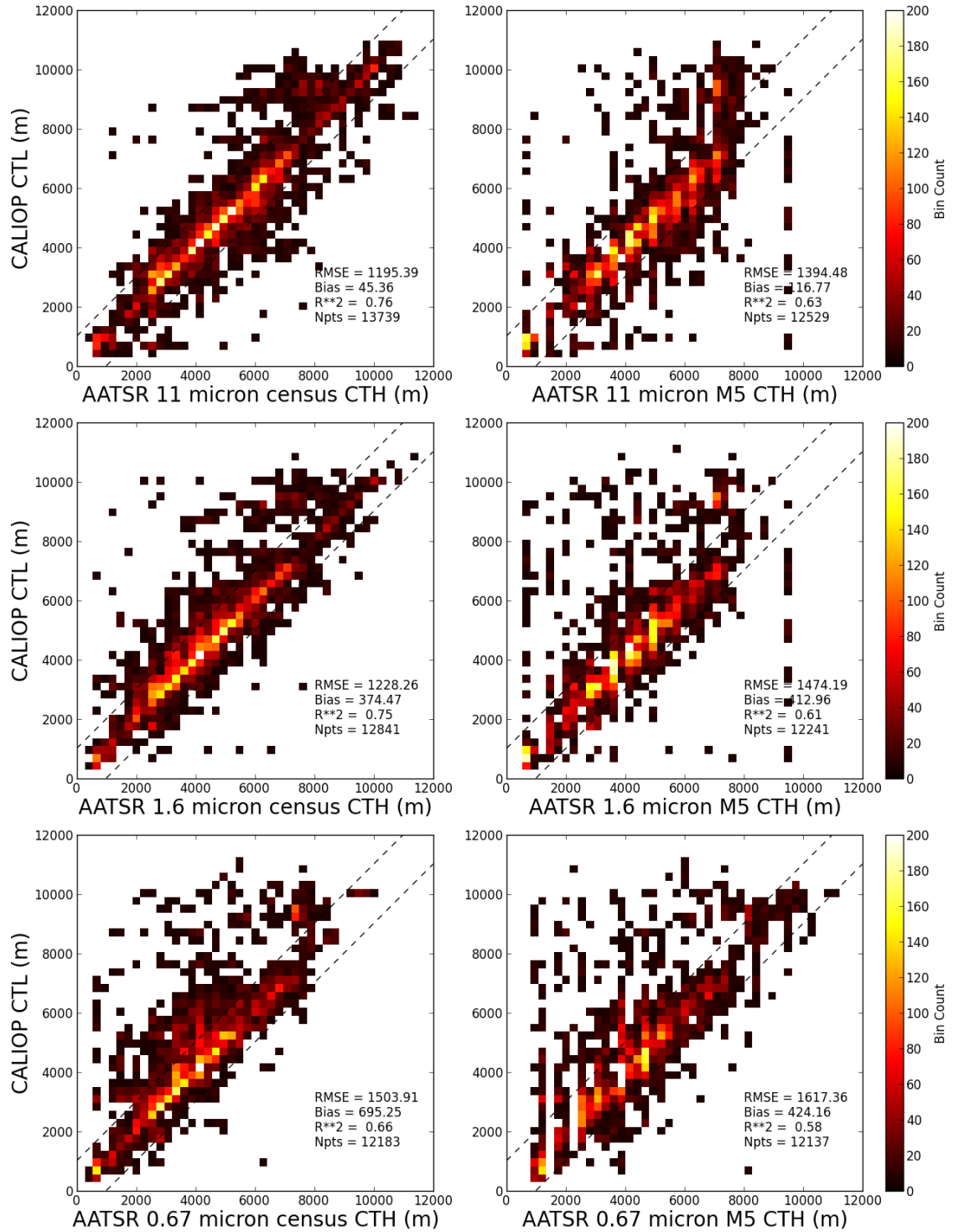


Figure 4-8. The figure contains the histograms of the AATSR census and M5 stereo output inter-comparisons against CALIOP lidar observations. The upper row contains the 11 μ m results, the middle row the 1.6 μ m results, and the bottom row the 0.67 μ m results. The left column comprises the census inter-comparisons, the right column the M5. Note, that in general the census results are far less noisy than M5 and compare far better with CALIOP, particularly for high cloud observations (above 8km) in the 11 and 1.6 μ m channels.

	11μm D	11μm M	1.6μm D	1.6μm M	0.67μm D	0.67μm M
Apr						
%	96	6	85	15	75	25
H(m)	6033(\pm 2563)	8368(\pm 1780)	5767(\pm 2469)	8246(\pm 1758)	5779(\pm 2316)	8406(\pm 1716)
OD	0.39(\pm 0.68)	0.16(\pm 0.27)	0.41(\pm 0.69)	0.16(\pm 0.28)	0.45(\pm 0.76)	0.19(\pm 0.29)
Jul						
%	90	10	81	19	78	22
H(m)	5075(\pm 2220)	7169(\pm 2042)	4800(\pm 2079)	6947(\pm 1962)	4448(\pm 2305)	6689(\pm 1769)
OD	1.43(\pm 1.42)	0.41(\pm 0.42)	1.51(\pm 1.44)	0.54(\pm 0.61)	1.74(\pm 1.49)	0.62(\pm 0.61)
Aug						
%	90	10	85	15	83	17
H (m)	5710(\pm 2438)	7633(\pm 2235)	5706(\pm 2425)	7364(\pm 2147)	5755(\pm 2407)	6977(\pm 2221)
OD	1.62(\pm 1.43)	0.43(\pm 0.51)	1.68(\pm 1.43)	0.41(\pm 0.53)	1.69(\pm 1.43)	0.44(\pm 0.55)
Sep						
%	88	12	83	17	81	19
H (m)	4837(\pm 2440)	8311(\pm 2214)	4878(\pm 2249)	8230(\pm 2414)	4691(\pm 2163)	8121(\pm 2480)
OD	1.04(\pm 1.21)	0.31(\pm 0.36)	(0.99 \pm 1.17)	0.31(\pm 0.18)	1.04(\pm 1.21)	0.32(\pm 0.38)
Oct						
%	90	10	86	14	83	17
H (m)	5625(\pm 2783)	6570(\pm 2706)	5446(\pm 2702)	7101(\pm 2802)	5199(\pm 2750)	7334(\pm 2556)
OD	1.03(\pm 1.32)	0.48(\pm 0.69)	1.08(\pm 1.34)	0.32(\pm 0.43)	1.19(\pm 1.42)	0.38(\pm 0.47)

Table 4-1. Here the results of the detection efficiencies of the AATSR channels in detecting the uppermost cloud layer detected by CALIOP are presented. Each row in the table represents a month from the AATSR CALIOP inter-comparison dataset, and each column an assessed AATSR channel. For each channel a number of statistics are generated for both when the AATSR channel observes the uppermost CALIOP cloud layer (denoted by D), and for when the AATSR channel misses the uppermost CALIOP cloud layer (denoted by M). The statistics are the percentage of detected and missed upper cloud layers(%), the monthly mean CTH of the detected and missed CALIOP cloud top layers along (H), and the monthly mean optical depth of the detected and missed CALIOP cloud top layers (OD). The bracketed values within the table are the associated monthly standard deviations.

The cause of the different biases across the AATSR channels is difficult to ascertain. A previous study comparing ATSR-2 stereo retrievals against the ground based Chilbolton radar undertaken by Naud et al., (2007b) did not demonstrate such significant bias differences between the channels. With biases of -0.9km for the 0.67 μ m channel, -0.8km for the 1.6 μ m channel, and -1.1km for the 11 μ m channel. However, the study was only comprised of 11 observations (versus ~15000 in the study presented in this chapter), and therefore may not be very representative in terms of the actual performance of stereo algorithms when applied to the ATSR instruments.

A possible reason for the biases observed is the requirement of suitable image texture for reliable stereo matching. One potential hypothesis is that the 11 μ m channel, due to AATSR's excellent thermal sensitivity, reaches sufficient image texture for effective stereo matching at lower cloud optical depths, than in the NIR/visible channels. Regions of the cloud with lower optical depth will tend to be closer to the true cloud top, and this would explain the current observed reduction in CTH bias in the thermal channel when compared to CALIOP (which can effectively observe the CTH to an accuracy of between 30m to 60m). In order to facilitate this assessment, firstly the ability of the 11 μ m channel to detect the top-most CALIOP cloud layer is compared against the ability of the NIR and visible channels. The results of this assessment are presented in Table 4-1. Looking at combined mean results, the 11 μ m channel detects the first CALIOP cloud containing layer 90% of the time, this drops to 84% for the 1.6 μ m channel, and to 81% for the 0.67 μ m channel. The second evaluation looks at the mean optical depths (as determined by CALIOP¹⁵) when first layer clouds are detected by each of the AATSR channels. Again the results are presented in Table 4-1. The combined means demonstrate that the 11 μ m channel can on average detect clouds with an optical depth of 1.22 (\pm 1.35), this drops to 1.26(\pm 1.35) for the 1.6 μ m channel, and to 1.34 (\pm 1.39) for the 0.67 μ m channel. The same result, albeit with lower optical depths, is found when evaluating the medians, with an optical depth of 0.68 for 11 μ m, 0.75 for 1.6 μ m, and 0.83 for 0.67 μ m. Though further work is required, this initial finding implies that the thermal channel is more sensitive to lower optical depths, meaning firstly that clouds with a lower optical depth can be effectively retrieved, and secondly that the retrieved heights are nearer to the cloud top as determined by the CALIOP lidar.

Table 4-1 also offers insights into the detection limits of the AATSR stereo retrieval system. When AATSR misses the first CALIOP layer the elevation of the cloud detected by the lidar is typically above 7km in elevation (vs. ~5km when AATSR detects the first cloud layer as determined by CALIOP). AATSR therefore has a tendency to miss the very high clouds over Greenland, but generally only 10% of the time in the thermal channel which has been shown to be the most sensitive. Of the first layer clouds that are missed the optical depths tend to be substantially smaller than those layers that are detected. Furthermore, the mean behaviour of all of the missed cloud optical depth is very similar across all AATSR channels. The 11 μ m channel generally misses clouds with an optical depth of 0.37 (\pm 0.46), the 1.6 μ m exhibits the same mean detection limit with a slightly larger standard deviation, 0.37 (\pm 0.48), and the 0.67 μ m has a slightly increased mean and standard deviation 0.39 (\pm 0.49). The median optical depth detection limit is 0.21 across all channels. This detection behaviour is similar to the detection limits

¹⁵ It should be noted that the CALIOP optical depth observations can be biased by up to 30% (Josset et al., 2012)

achievable by ISCCP, MISR and MODIS which range from between about 0.1 and 0.4 (Rossow and Schiffer, 1999; Dessler and Yang, 2003; Marchand et al., 2007; Ackerman et al., 2008; Marchand et al., 2010). Importantly, these results show that the AATSR 11 μ m, in combination stereo, likely has better potential for retrieving optically thin high cloud than MODIS with the CO₂ slicing method, which fails to operate effectively at optical depths of less than 1 (Naud et al., 2007a).

4.3.3 L2 Summary

The entire AATSR dataset for Greenland has been processed using the census stereo matching algorithm, providing ~10 years of CTH data. Due to problematic low texture regions caused by the Greenland ice sheet there is substantial noise within the stereo matching outcomes, this noise is problematic as it leads to false positive cloud features. In order to remove these false positive clouds a RT derived set of radiometric thresholds which effectively delineates ice from other spectral features has been employed. This threshold application also provides an effective cloud mask, as any non-ice features over the ice sheet are typically clouds. However, it precludes year round and diurnal stereo operation, which is a significant drawback. The census derived CTH observations have been validated against higher resolution observations of CTH from both MISR stereo retrievals and CALIOP lidar returns. In both validations the census CTH outputs are shown to agree well, and in almost all cases demonstrate better agreement than the M5 algorithm. Furthermore, the AATSR 11 μ m stereo results, in comparison against CALIOP, have been shown to be more sensitive than the stereo retrievals from both the 1.6 μ m and the 0.67 μ m channels, leading to a greater likelihood of clouds being detected and also a reduction in the bias of the retrieved CTH.

4.3.4 L3 Processing

The AATSR L3 CFbA climatology is derived from the AATSR L2 data using the same general principles as those employed by the MISR science team to derive the MISR CFbA product (Di Girolamo et al., 2010). This product is in a stratified format, that is, CF is recorded for multiple height levels. This approach is beneficial as it allows retention of both CF and CTH data in one dataset, and removes the ambiguities that occur when attempting to create mean CTH values from bimodal cloud distributions, which occur with regularity in the tropics (Stubenrauch et al., 2013). The CFbA processing chain is as follows:

1. The first step in the process is to establish a climatological grid into which the AATSR data is aggregated. The grid employed is of 0.5° resolution and has 17 vertical levels. The first sixteen levels contain the CFbA data at 1km quantisation above the WGS84 reference ellipsoid, apart from the first level (0.5km-1km), and the sixteenth level (which contains all clouds detected by the Istomina cloud mask, but having no valid height, i.e. <500m CTH). The final level contains the total cloud fraction.
2. Following generation of the climatological grid, a L2 processed AATSR orbit for a given day is ingested into the processing chain and divided into 16 x 16 km² regions. Within each region the CF is computed by calculating the ratio of cloud to non-cloudy pixels and the CTH determined by taking the median above ellipsoid height of the cloudy pixels.

3. The CF for each region is then projected into the climatological grid. The horizontal location is determined by taking the central latitude and longitude values for the region under consideration and selecting the closest 0.5° grid cell within the climatological grid. The vertical position is determined by the median height value. If the median above ellipsoid CTH value is less than 500m above the median local terrain height taken from the GMTED2010 DEM data, then the CF data is placed into the 16th grid level.
4. Once all regions from the orbit have been placed into their respective height bins the mean cloud fraction at each height bin is calculated. This is achieved by dividing the summed cloud fraction of a given height bin by the number of observations which comprises it.
5. The next step is to normalise the CF for a given horizontal location so that when summed over its vertical extent it is equal to 1. This is achieved by dividing each bin by the total number of samples over its vertical extent (from bins 1 through 16).
6. Once each orbit grid has been through this normalisation process, it can be combined into a daily grid. This is achieved by summing all orbit grids together and dividing each grid cell by the total number of orbits which observe it. The total CF bin for each climatological grid cell is also computed at this point by summing over bins 1 through 16.
7. Step 6 is then applied to combine the daily data into monthly data. All daily grids are summed together and each grid cell is divided by the number of observation days. Yearly climatologies are similarly constructed by the summing the monthly climatologies and dividing by the number of observation months. This daily, monthly, yearly aggregation approach ensures that the resulting CFbA values are unbiased, versus aggregating by orbit where the CFbA would be biased towards the observations with the most samples (Di Girolamo et al., 2010).

This processing chain was used to process the AATSR L2 CTH 11 μ m, 1.6 μ m and 0.67 μ m datasets for the entire AATSR time-series for the previously defined geographic region over Greenland. Three separate datasets were generated for each channel. One derived from the AATSR descending node (a 10:30 a.m. dataset), one from the ascending node (10:30 p.m. dataset), and combined ascending and descending node dataset. Example climatological outputs from this processing chain are shown in Figure 4-9.

4.4 L3 Inter-Comparison Study

4.4.1 GEWEX Inter-comparison Datasets

The climatologies employed for inter-comparison against the new AATSR-stereo L3 CFbA climatology are taken from the GEWEX cloud climatology database (<http://climserv.ipsl.polytechnique.fr/gewexca/index-2.html>, last accessed 28th Sep 2013). The climatological products employed from the database are the MODIS-Science Team, MODIS-CERES, MISR and AATSR-GRAPe datasets, with the respective methods behind their L2 product derivation discussed in section 4.1.3. The climatological aggregation process differs between the datasets, though

all the GEWEX climatological products are available on a 1° grid, with monthly time stamps. The MISR science team developed and employ the CFbA climatological aggregation approach that is also employed for developing the AATSR CFbA product in this chapter. The only difference between the MISR and the AATSR approaches is that the MISR L2 products are averaged onto a 17.6km grid (rather than a 16km

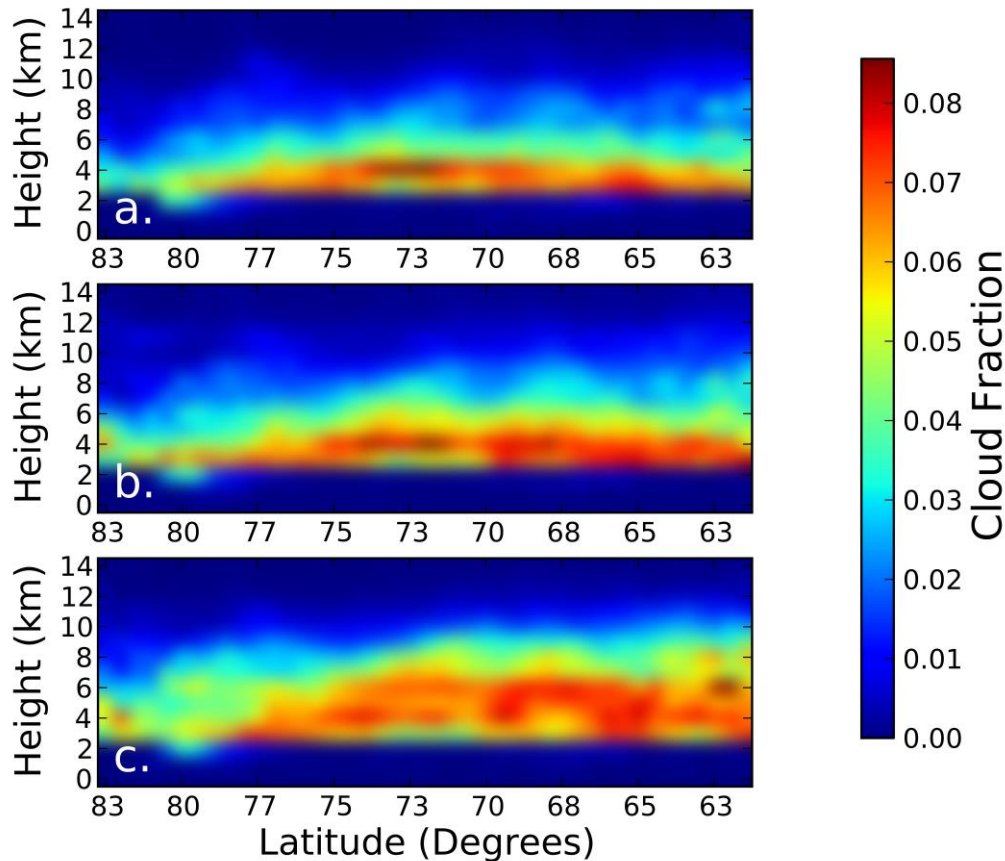


Figure 4-9. This figure displays the latitudinal 2008 descending node mean CFbA for the $0.67\mu\text{m}$ (upper plot), $1.6\mu\text{m}$ (middle plot), and $11\mu\text{m}$ (lower plot). There is a general increase in high cloud amount as the wavelength of observation increased, particularly over the southern half of Greenland. The cloud fraction below 2km is limited due to the elevation of Greenland itself, which is on average greater than 2km. The majority of clouds are below 8km, this is to be expected as in Polar Regions the troposphere rarely exceeds 8km in elevation.

grid as employed with AATSR) prior to climatological aggregation. The MISR CFbA product is typically available at a 0.5° resolution, however as mentioned, within the GEWEX database, the MISR CFbA is available only on a 1° grid. All of the MODIS monthly climatological products are formed through aggregating daily L3 products (aggregated from the L2 products) into a 1° equal angle grid. In the inter-comparisons used in this section only the daytime MODIS Terra products are used (10:30 a.m.) for their similarity to the overpass time of Envisat. The AATSR-GRAPPE climatology is aggregated in a method similar to that of the MODIS datasets.

Two different comparisons are made in the section, one evaluating CF, the other CTH. Both comparisons employ data from 2008. As the GEWEX climatologies are available at 1° resolution they are resampled using bilinear interpolation to the 0.5° grid employed by the AATSR stereo (hereafter referred to as AATSR-Stereo) CFbA grid. In the CF analysis, undertaken in section 4.4.2, the total CF layer is taken from the 2008 0.67µm descending AATSR CFbA product. The dataset used from AATSR-Stereo only depends on the node, as the total CF is the same across CFbA products due to consistent application of the Istomina cloud mask to define CF. The descending AATSR node climatology is derived from data closest in time to the other climatologies used in this comparison, hence its selection. As the GEWEX CF products are available as monthly datasets they are combined into a yearly product by averaging across each month.

The CTH inter-comparison, undertaken in section 4.4.3, requires conversion of the AATSR-Stereo CFbA products into mean CTHs. This is achieved by summing scaled heights over all bins for a given location. The heights are scaled by the ratio of CF in a given bin to the total CF across all bins for the given location, e.g. $5 * (0.6/0.64) + 4 * (0.04/0.64)$, where there are CF observations in the bins corresponding to 4km and 5km and the total CF across all bins is 64%, 60% of which is at 5km and 4% of which is at 4km, giving a CTH for the location of 4.94km. In the CTH inter-comparison two different AATSR-stereo products are evaluated, the 2008 0.67µm descending AATSR-Stereo CFbA product and the 2008 11µm descending AATSR-Stereo CFbA product. The 0.67µm and 11µm products, whilst sharing total CF, are expected to exhibit different mean CTHs due to the previously demonstrated differing detection capabilities, with the 11µm product expected to return greater heights than that of the 0.67µm. Two comparisons are made for the CTH analysis. The first, similar to the CF analysis, assesses the general large scale CTH variation over Greenland through comparison of the yearly averages. The second looks at the variation of the monthly total mean CTH over Greenland across the datasets.

4.4.2 CF Inter-comparison

4.4.2.1 Yearly Comparison

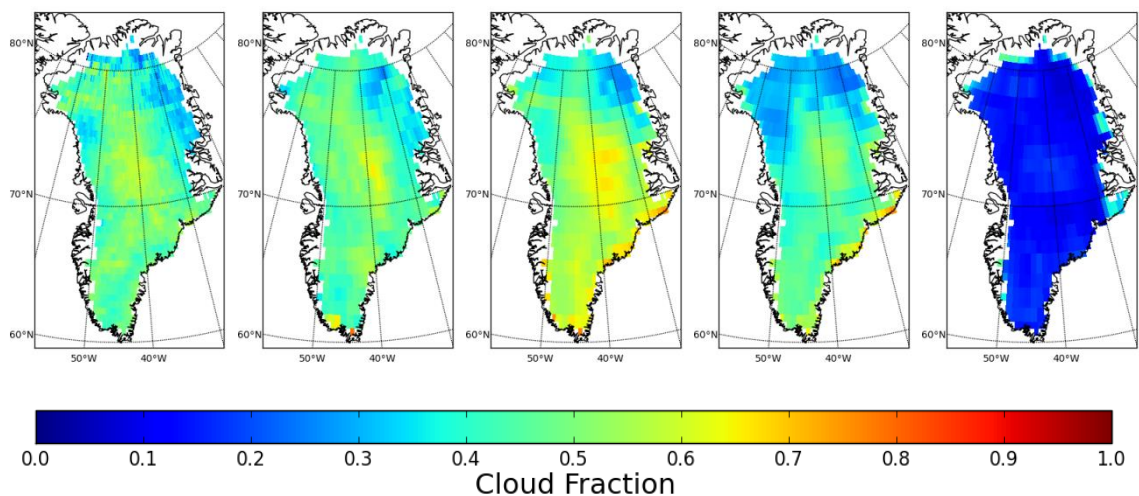


Figure 4-10. These plots show the mean CF for 2008 from left to right: AATSR-Stereo, MISR,

MODIS-CERES, MODIS-Science Team, AATSR-GRAPe. In general there is good agreement across the first four datasets for large scale cloud structure. The MODIS-Science team product underestimates CF in the north west of Greenland compared to the other three reliable datasets. The AATSR-GRAPe dataset appears to be ineffective in CF determination.

4.4.2.2 CF: Yearly Discussion

Of the yearly CF climatology plots shown in Figure 4-10 AATSR-Stereo and MISR demonstrate the best agreement in terms of large-scale CF organisation. This is somewhat surprising given the fact that the AATSR-Stereo CF is derived using the Istomina radiometric masking method, whilst the MISR CF is determined using angular and stereo thresholds. This demonstration further validates the quality of the Istomina snow/ice mask for effective cloud discrimination over snow and ice. Looking at the mean 2008 CF for AATSR-Stereo a value of $0.45 (\pm 0.08)$ is returned, for MISR a value of $0.45 (\pm 0.07)$ is returned. For the other datasets some disagreement is found. For AATSR-GRAPe the retrieved CF appears to be far too low, however, this is not too surprising as the AATSR-GRAPe team states that the CF retrieved by the optimal estimation algorithm is generally unreliable (see section 4.1.3.4). The mean CF for AATSR-GRAPe for 2008 is 0.15 ± 0.06 . The two MODIS teams also retrieve quite dissimilar results given the fact that the same instrument is used in their derivation; this demonstrates the importance of algorithm in defining the final outcome. The MODIS Science Team product agrees well with MISR and AATSR-Stereo over the central region of Greenland, however in the Northern regions it substantially underestimates CF. The MODIS CERES CF climatology for 2008 has a very similar CF organisation to MISR and AATSR, but in many locations appears to be detecting slightly more cloud, as evidenced by the stronger yellows and oranges. The MODIS Science Team and CERES return 0.39 ± 0.09 and 0.49 ± 0.1 as their mean CFs for 2008 respectively.

4.4.3 CTH Inter-comparison

4.4.3.1 Yearly Comparison

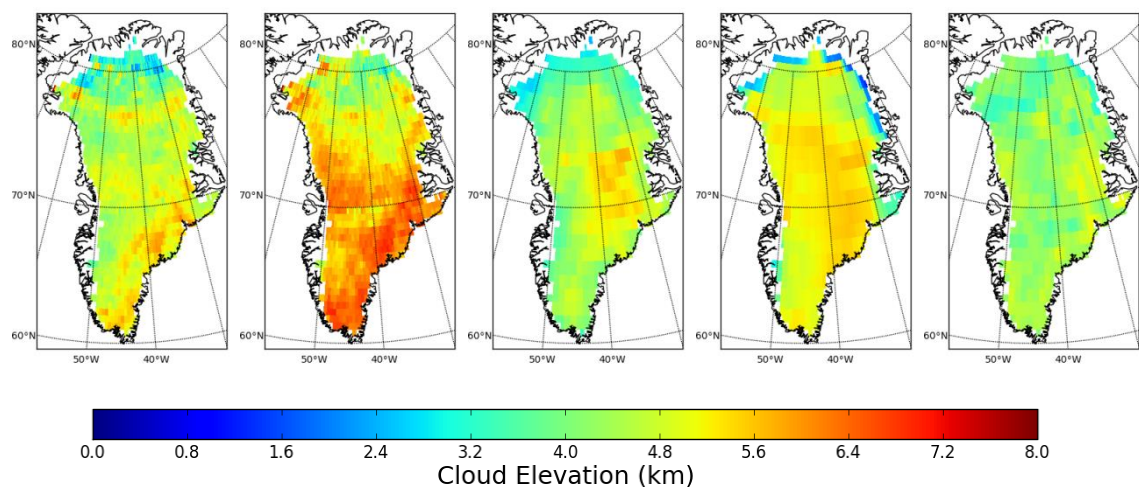


Figure 4-11. These plots show the mean CTH for 2008 from left to right: AATSR-Stereo $0.67\mu\text{m}$,

AATSR-Stereo 11 μ m, MISR, MODIS-CERES, AATSR-GRAPe. All of the climatologies demonstrate quite different large scale CTH distributions. Of most interest in this figure is the difference between the two AATSR-Stereo plots, demonstrating the effective multilayer cloud detection capabilities of AATSR in combination with stereo. This plot also demonstrates that that the other datasets under consideration tend to miss high clouds.

4.4.3.2 Monthly Comparison

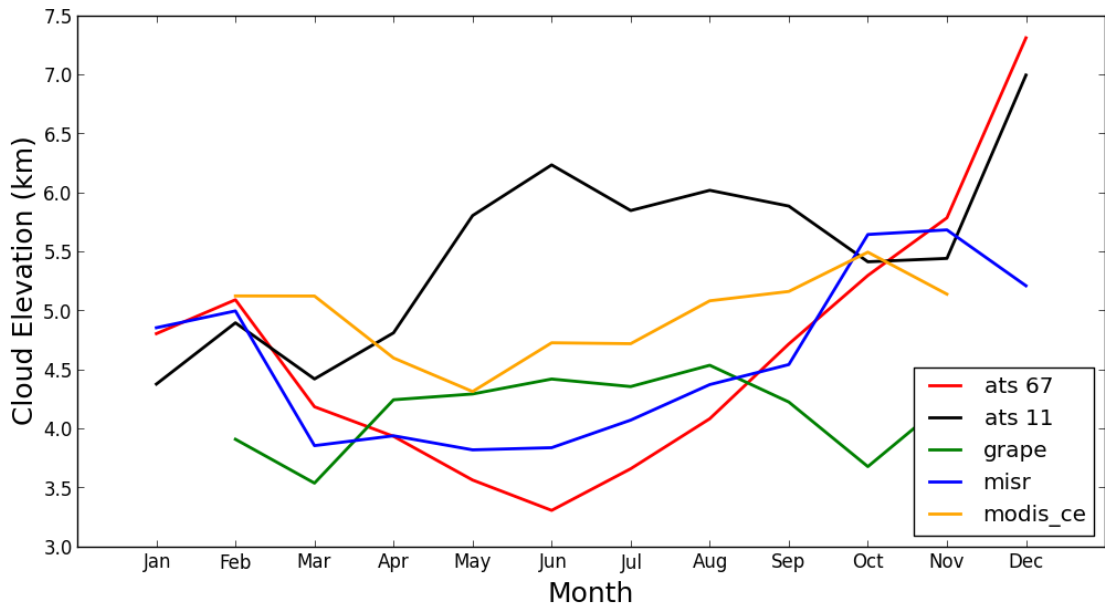


Figure 4-12. The above figure plots the monthly mean CTHs over Greenland for each month of 2008 for the AATSR-Stereo 0.67 μ m, AATSR-Stereo 11 μ m, AATSR-GRAPe, MISR and MODIS-CERES climatologies. Only months March through October have complete coverage of Greenland due to illumination conditions. Of most interest in this plot is the high cloud detection of the AATSR-Stereo thermal channel, as also demonstrated in Figure 4-11, and also the intermediate CTHs returned by AATSR-GRAPe and MODIS-CERES. The dashed lines in the plot represent ± 1 standard deviation.

4.4.3.3 CTH: Yearly Discussion

The yearly averages of CTH in Figure 4-11 all show quite different large-scale structures. The AATSR-Stereo 0.67 μ m product, MISR and AATSR-GRAPe climatologies show the most similarity but the agreement is limited. The MODIS-CERES CTH climatology appears to lose all spatial heterogeneity present in the other datasets, setting the majority of Greenland to heights close to 5km. The MODIS-CERES CTHs also appear to be generally higher than the other datasets with the exception of the AATSR-Stereo 11 μ m product. The thermal AATSR-Stereo climatology returns greater heights than the other climatologies, particularly in the Southern half of Greenland, implying the presence of multi-layer clouds. Looking at the statistics for Greenland the mean CTH for the AATSR-Stereo 0.67 μ m is 4.5 ± 0.7 km, for MISR it is 4.3 ± 0.8 km, for AATSR-GRAPe it is 4.2 ± 0.4 km, for MODIS-CERES it is

4.9km±0.8km, for AATSR-Stereo 11µm it is 5.3±0.8km. The MODIS-Science Team have no CTH data for 2008 so are not considered.

4.4.3.4 CTH: Monthly Discussion

Of the monthly mean CTH amounts for Greenland shown in Figure 4-12, only the months of March through October have complete coverage of Greenland and are considered here. The remaining 4 months have limited coverage due to polar winter, and the requirement of illumination is present for the MISR, AATSR-Stereo and AATSR-Grape climatologies. Looking firstly at the AATSR-Stereo 0.67µm, in the summer months (May-July) it is shown that on average the AATSR-Stereo dataset underestimates the MISR observations by ~500m, with a peak of this behaviour June. The AATSR-Stereo 11µm detects CTHs 2-3km above the other datasets for the period March through September. This result implies firstly that there is an increase in high cloud in summer months and that in the spring and autumn it appears that the high cloud is reduced as both the assessed AATSR-stereo channels return similar CTH statistics. Secondly this result implies that there is a substantial amount of multilayer cloud over Greenland in the summer months which is effectively detectable with AATSR-Stereo. This reconfirms the findings of differing cloud layer detection capabilities of the AATSR channels when processed using stereo presented in the CALIOP inter-comparison and also in Naud et al., (2007). The MODIS-CERES and AATSR-GRAPe climatologies appear to detect inter-mediate heights between the AATSR-stereo 11µm and 0.67µm results in Greenland summer; this result is not unexpected, in multi-layer cloud situations the IR-window method employed by these instruments tends to return an intermediary CTH value somewhere between the upper and lower cloud decks (Stubenrauch et al., 2013). In the Spring and Autumn months, with an apparent decline in multilayer cloud conditions, the MODIS-CERES and AATSR-GRAPe mean CTH estimates fall more in line with those from MISR and AATSR-Stereo.

4.5 Summary

4.5.1 Objective Review

The scientific objectives outlined in section 4.2 are reviewed here:

‘Generate L2 CTH and CF products for the AATSR data time series over Greenland.’

This objective has been achieved successfully. The entire AATSR time series over Greenland has been processed into above ellipsoid CTHs using the census stereo algorithm in combination with the Mannstein camera model. The Istomina snow/ice mask has been implemented to screen low texture snow and ice regions which were found to lead to erroneous stereo CTH retrievals. Furthermore, this method provides an effective cloud mask, and therefore CF observations, as any non-snow/ice regions over a snow/ice surface generally correspond to cloud features. One side-effect of the application of the Istomina cloud mask is that it requires all AATSR channels to effectively identify snow/ice regions. This precludes the generation of effective and accurate stereo CTH observations over Greenland during non-

illuminated periods (e.g. night time/polar winter) from the AATSR thermal channels, which is a substantial drawback.

‘Determine the accuracy of AATSR stereo CTH retrieval against higher accuracy observations of CTH from MISR and CALIOP.’

The L2 AATSR census stereo outputs have been compared extensively against observations from both MISR and CALIOP. Census has been shown to perform very well. In the 0.67 μ m comparison against MISR an RMSE of ~900m is returned along with a bias of approximately -100m. These values demonstrate that the algorithm leads to performances very close to the optimum accuracy achievable by the AATSR instruments, which is ~800m (a limit introduced by the co-registration correction coefficients). A further comparison against MISR was undertaken to estimate biases in CF. This omission-commission analysis demonstrated that a ~1% difference exists between the CF detected for MISR and for AATSR for CTHs above 500m, demonstrating excellent agreement between the two cloud masks.

In the comparison against CALIOP all three AATSR channels processed with census have been compared. The 11 μ m stereo results show the best agreement with an RMSE of 1200m and a bias of only 45m. At shorter wavelengths the biases of the AATSR stereo retrievals are shown to increase, from 375m at 1.6 μ m to 695m at 0.67 μ m. These bias differences have been attributed to differing detection efficiencies across the AATSR channels, and the 11 μ m channel has been shown to be more sensitive to cloud features, requiring lower optical depths to effectively perform stereo matching, resulting in heights retrieved closer to the true cloud top as detected by the lidar instrument and fewer missed high clouds, which tend to be optically thin (Poulsen et al., 2012).

Additionally, the CTH determination efficacy of the AATSR instrument in combination with stereo has been assessed by looking at the typical characteristics of the clouds, in terms of optical depth and elevation, which it commonly misses. This assessment, performed against the CALIOP lidar data, has demonstrated that the majority of clouds missed by the AATSR observations over Greenland are above 7km in elevation and on average have an optical depth of approximately 0.4. This detection limit of 0.4 is in line with other CTH retrieval systems. One area where AATSR in combination with stereo may provide performance benefits over the current passive CTH assignment systems is for the detection of high optically thin cirrus clouds. Currently only the CO₂ slicing algorithm employed by the MODIS Science Team provides accurate CTH assignment for high optically thin cirrus clouds. However, this algorithm fails on clouds with an optical depth of less than one. AATSR stereo observations from the 11 μ m channel could therefore provide improved coverage of high cirrus clouds over the current passive CTH assignment systems.

‘Determine any gain in performance achieved through application of the census stereo matching algorithm over that of the M5 stereo matching algorithm.’

The M5 algorithm outputs were compared against the same MISR and CALIOP datasets as used for the census validation. In nearly all cases census was shown to outperform M5, indicating that it is more effective for the task of CTH determination.

‘Aggregate the L2 AATSR products into height stratified macrophysical cloud climatologies containing CF on a by height basis.’

The AATSR L2 data has been effectively aggregated into a CFbA product almost identical to that employed by the MISR Science Team. The resultant L3 products are available on a daily, monthly and yearly basis for the AATSR 11 μ m, 1.6 μ m and 0.67 μ m channels. The climatological datasets are further split into descending, ascending and combined nodes of observation.

‘Inter-compare the stereo derived AATSR cloud climatology against the GEWEX CTH and CF climatologies from MODIS (Science Team and CERES), MISR, and AATSR-GRAPe to determine any new data contributions provided by AATSR stereo.’

Two separate inter-comparisons have been made against cloud climatologies from the GEWEX cloud database. Firstly a CF inter-comparison was made using the AATSR-Stereo L3 datasets for the descending node of 2008 over Greenland against collocated CF climatologies from MODIS-Science Team, MODIS-CERES, MISR, and AATSR-GRAPe. In this comparison the AATSR-Stereo CF climatology was demonstrated to agree very well with that of MISR, returning the same overall mean CF and have a very similar large scale structure despite differences in the employed masking approach. The MISR instrument does not rely on radiometric CF assignment, so is therefore assumed robust in terms of CF assignment over snow and ice covered regions. This agreement between MISR and AATSR therefore provides further encouragement on the performance of the current cloud masking employed in the AATSR-Stereo cloud climatology derivation.

The second comparison involved comparing CTH climatologies derived from the 2008 AATSR-Stereo descending node 11 μ m and 0.67 μ m CFbA products against collocated CTH climatologies from the MODIS-CERES, MISR and AATSR-GRAPe products. Two separate comparisons were made, one employing yearly mean maps of CTH over Greenland, and another looking at the monthly mean CTH over Greenland. In the yearly analysis the general organisation of CTH differs quite substantially between the datasets. The differences between AATSR-Stereo 0.67 μ m and MISR are somewhat unexpected due to the excellent agreement demonstrated in the validation. Of the yearly comparisons the most important result is the differing cloud detection between the AATSR-Stereo 11 μ m and 0.67 μ m datasets, which well demonstrates the effective multi-layer cloud detection that is achievable.

The monthly mean analysis also returns some interesting results. In terms of monthly Greenland averages the MISR and AATSR-Stereo 0.67 μ m results tend to agree very well, and are more in line with the findings in the validation. Mean differences between the two datasets of up to 500m do occur in the summer, with AATSR-Stereo 0.67 μ m appearing to underestimate CTH compared to MISR. The AATSR-Stereo 11 μ m 2008 monthly averages again demonstrate a substantial amount of high cloud particularly in the summer months. This increase in high cloud in the summer may be caused by shifting patterns in the polar jet. Assessing the ERA interim monthly mean wind speeds at 300 hPa (approximately the pressure of the jet stream, and also the high clouds detected by AATSR-Stereo 11 μ m channels) for March to October 2008 (May through August shown in Figure 4-13), the polar jet appears

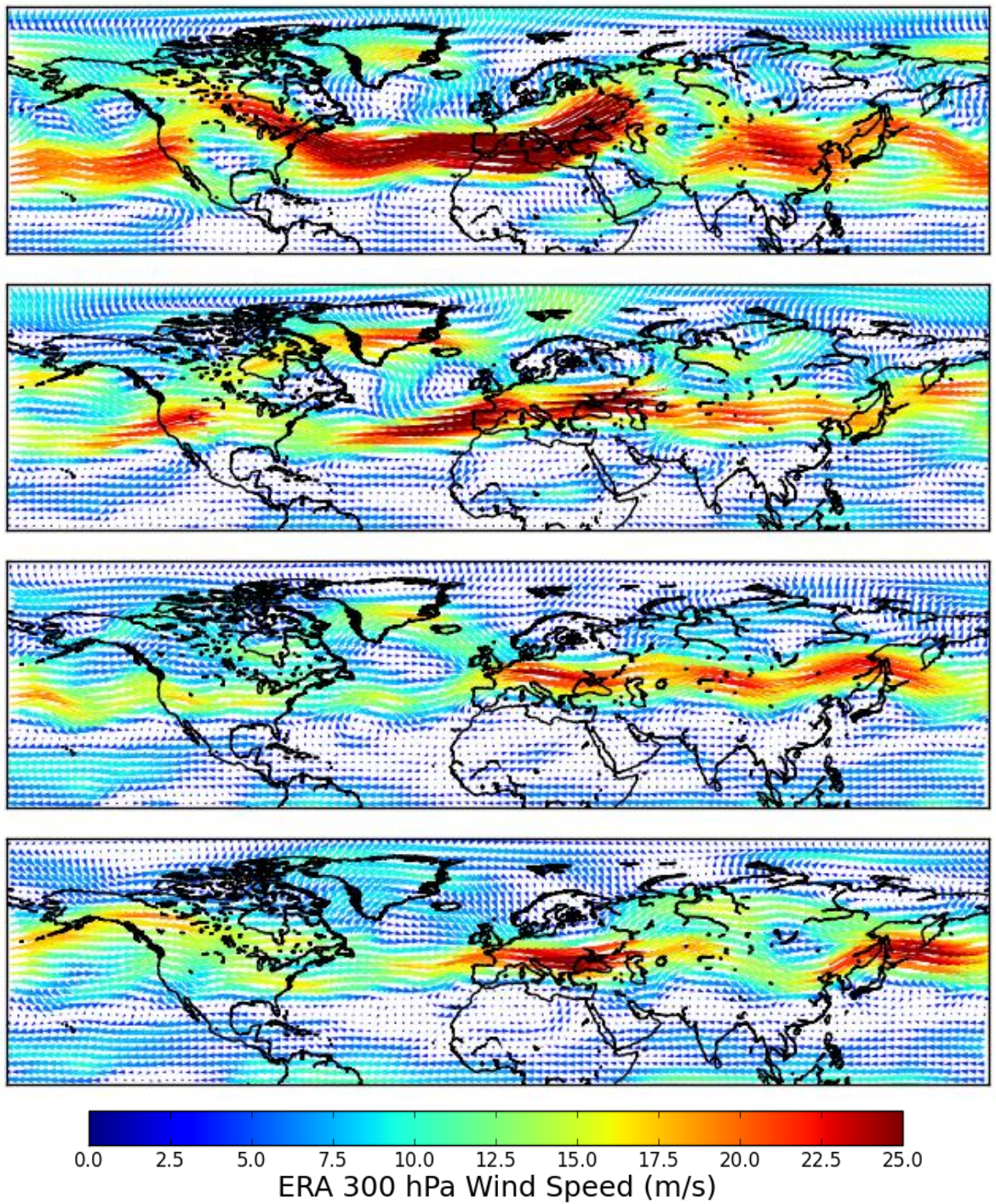


Figure 4-13. These four plots show the polar jet position for May, June, July and August 2008 at 300 hPa as defined by the zonal and meridional wind profiles taken from the ERA-Interim reanalysis. In May, June and July, the position of the polar jet is clearly over Southern Greenland, and is perhaps the cause of the increased high cloud amount in the summer months.

to oscillate over southern Greenland through the summer months, potentially pushing high cirrus clouds contained within the jet (Dowling and Radke, 1990) over the ice sheet as a result. The MODIS-CERES and AATSR-GRAPE datasets are shown in the evaluation to return an intermediary height value between the upper and lower cloud decks during the summer months – an inherent weakness of the IR-window method employed by these instruments. In summary it appears that when compared to the other climatological datasets, AATSR-Stereo has the potential to provide valuable new insights into the high cloud characteristics over Greenland and important complimentary low and middle cloud information.

4.5.2 Future Work

From the objective review a number of areas for future work can be defined. Firstly an image processing technique needs to be located or developed which is able to effectively screen off low texture areas that tend to confuse the matching algorithm. The image processing technique would have to be channel independent in its operation (i.e. only requiring the input channel and applicable across all channels). This would allow year round retrieval of macrophysical cloud parameters from the AATSR thermal channels. In the case of snow and ice, the masking approach would likely exploit the difference in texture between snow/ice and cloud features, as shown in Cawkwell et al., (2001, Figure 3). Following this development the next logical step is to process the complete AATSR stereo time series into the climatological CFbA product to provide another complimentary data set for climate studies. This would also require further validation of the stereo products in other regions globally to better quantify the stereo CTH and CF errors at disparate locations. This further validation would likely also lead to greater insights into the current biases and detection characteristics exhibited in the AATSR stereo data when compared against CALIOP. Locating triple collocations between the AATSR stereo, MISR and CALIOP data would also likely provide very performance insights across the three sensors.

Another avenue would be to apply the census transform to other stereo capable instruments such as MISR. Census has been shown in this chapter to significantly outperform the M4/M5 algorithm, which is very similar in approach to the one currently employed operationally by the MISR team.

4.5.3 Conclusions

This chapter has outlined the need for observational cloud climatology datasets within the study of climate change for validation efforts. The principle study which assesses these observational datasets, from the GEWEX cloud group, has been outlined, and a number of key climatologies from the assessment, namely MODIS-Science Team, MODIS-CERES, MISR, and AATSR-GRAPE, have been described in detail, as have the particular benefits of having multiple climatologies. The main undertaking in this chapter has been the development of a new, stereo derived, macrophysical cloud climatology from the AATSR instrument over Greenland. This aim has been successfully achieved through application of the census stereo algorithm in combination with the Istomina cloud mask to generate L2 CF and CTH products. These two products have been validated against observations from MISR and CALIOP and been shown to perform very well against the higher resolution observations with height accuracies in the region of ~1km and detection limits down to a cloud optical depth of 0.4. The derived L2 products have been aggregated into a number of different climatological CFbA products

inspired by the MISR product of the same name. The derived climatologies are split by channel (0.67 μ m, 1.6 μ m, 11 μ m) and orbital node (descending, ascending and total). The L3 products have been inter-compared against the aforementioned observational climatologies over Greenland for the year of 2008. This inter-comparison has demonstrated the worth of climatological processing of the entire ATSR time series with the census stereo algorithm; in particular for the effective determination of CTH in Polar Regions; the determination of multi-layer CTHs; and the effective determination of high cloud features with an optical depth of less than one.

5 ALANIS: ATMOSPHERE LAND INTERACTION STUDY

5.1 Introduction

The environmental and ecological importance of the boreal forest ecosystem is without doubt. Globally, a total of 1.2 billion hectares is classified as boreal forest, with 900 million hectares of this being closed forest. (Conard and Davidenko, 1998). These forests, which represent 30% of the world's forested area (Conard and Davidenko, 1998), store billions of tons of carbon, sequestered since the last glacial maximum (Adams et al., 1998; Dong et al., 2003; Prentice et al., 1993). Biomass burning events can potentially convert these ecosystems from carbon sinks to net sources (Dong et al., 2003; Turquety et al. 2007), releasing significant amounts of greenhouse gases, aerosols and also pollutants (Crutzen et al. 1979; Martin et al. 2010), possibly contributing to the current global warming trend and directly impacting on human health (Greenough et al., 2001; Künzli et al., 2006). Biomass burning events are not uncommon; one study assessing wildfire events in the Russia Boreal region for the period 1974-93 returned a mean of 17,000 fires per annum (Nyejluhto, 1994). Furthermore, due to the boreal region's Northerly latitude, located between 45° and 75° N, it has greater susceptibility to changes in climate (Stocks et al., 1998). With the current warming trend, some authors contend that this will lead to increased frequency and extremity of biomass burning events (Gillett et al., 2004; Kasischke et al., 2006; Turetsky et al., 2010). This is significant, as in extreme burning events with large energy outputs or under strongly convective atmospheric conditions, the fire emissions can be injected into the free troposphere (Martin et al., 2010). Under such circumstances, the atmospheric lifetime and distribution of most of the released trace gases and aerosols are significantly enhanced, potentially leading to greater climate and health impacts over increased geographical extents (Colarco et al. 2004).

In order to quantify the impacts of the emissions from biomass burning events on air quality and climate, chemical transport models (CTMs) are the tool of choice, e.g. (Krol et al., 2005). In order to generate effective predictions and better understanding, these CTMs must be initialised with reliable estimates of emissions including their vertical extent, and also validated against plume dispersion measurements over time. Due to the lack of available in situ or satellite data, particularly at higher latitudes (>60°N), which are not visible to geostationary satellites, often rather arbitrary assumptions are used for initialisation,

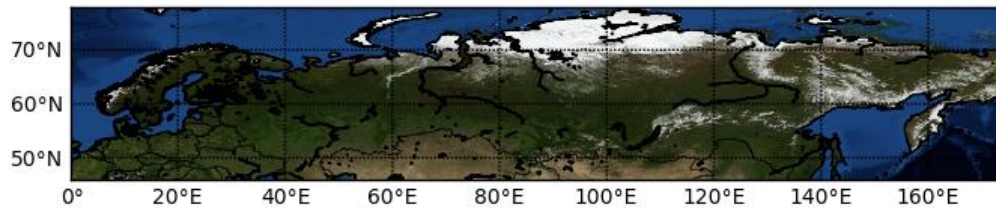


Figure 5-1. The map of the ALANIS study region (Source: Python Basemap)

such as fixed vertical injection levels (Turquety et al., 2007). This simplification is likely to lead to a reduction in the accuracy of the emission distribution outputs from such models (Kahn et al., 2008). Seeing this limitation, and, given the current, and possibly enhanced future potential of biomass burning events in Boreal forest ecosystems to impact both on the climate, in the longer term, and human health, in the shorter term, ESA along with the IGBP iLeaps (Marconcini et al., 2010) established a consortium to better understand and predict the potential impacts of such events in the Eurasian boreal region (see Figure 5-1) through the ALANIS smoke plumes project (Fisher et al., 2013; Krol et al., 2013).

The project aimed to improve the prediction accuracy of the TM5 CTM (Krol et al., 2005), through assimilation of a number of novel observational datasets, including EO derived smoke plume injection height (SPIH) estimates, over the Eurasian boreal regions (latitudinal range: 46°-78°, longitudinal range: 0°-174°) for the study period August 2008 through to August 2011. As the study region is vast the use of EO satellite data is necessitated to provide suitable coverage. However, currently there are few suitable systems for determining SPIH using EO. The current principle EO SPIH dataset is provided by the MISR Smoke Plumes project (Nelson et al., 2008; Nelson et al., 2013). This project employs the MISR instrument and the MISR-MINX visualisation tool (Nelson et al., 2008; Nelson et al., 2013) to generate accurate ($\pm 200\text{m}$) wind corrected stereo observations of SPIH. This project relies on manual stereo digitisation of smoke plumes by trained undergraduate student operators. This has a number of benefits and drawbacks: the main benefit being that smoke plumes can be accurately identified and digitised by human operators; the main drawback being that the dataset has limited spatial and temporal coverage, as the human cost of digitisation is high. The CALIOP instrument demonstrates another method; however, whilst the lidar SPIH measurements are very accurate (30-60m), its spatial sampling is extremely limited due to the employed technology. Indeed, the MISR instrument is between 40 and a few hundred times more likely to observe a smoke plume feature than CALIOP (Kahn et al., 2008). CALIOP's main strength is its increased sensitivity, being able to sample features to an optical depth of ~ 0.02 , whilst MISR is typically a factor of 10 less sensitive (Kahn et al., 2008). The AATSR instrument presents another alternative to measure SPIHs, with potential benefits over both MISR and CALIOP. In comparison to MISR, the main benefit of AATSR is the potential for simple automated determination of SPMs through spectral and elevation thresholds (see Section 5.5). Automated SPM determination with MISR can be achieved through machine learning techniques (Mazzoni et al, 2007b), but is not employed in the MISR plume height project owing to its high failure rate ($>15\%$). Against CALIOP, the main benefit of AATSR is the similar to that of MISR, vastly improved spatial and temporal sampling

potential. The potential detriment of AATSR versus MISR and CALIOP is a reduced SPIH retrieval accuracy due to reduced quantisation (1000m vs. 200m and 30-60m respectively).

This chapter (see also Fisher et al., 2013) examines the potential of AATSR for determining SPIHs and SPMs for application within the ALANIS plumes project in detail. Following an outline of the scientific objectives in Section 5.2, SPIH extraction from AATSR using the smoke plume tuned M6 algorithm which is described in Chapter 3 is presented in Section 5.3. This discussion also covers the processing chain employed to process the AATSR ALANIS dataset in detail. In Section 5.4 extensive validation of the AATSR SPIHs is undertaken against both MISR and CALIOP in order to evaluate the accuracy achieved through combination of M6 and AATSR for the task at hand. In Section 5.5, the SPM method is outlined and compared against manually derived masks from the MISR instrument. Finally Section 5.6 provides a summary and recommendations for future work.

5.2 Scientific Aims and Objectives

- Assess the performance of the M6 stereo matching algorithm applied to AATSR for determination of SPIH against collocated higher resolution observations derived from alternative satellite instruments.
- Develop and apply an automatic method for locating smoke plumes in the AATSR imagery.
- Inter-compare the AATSR smoke plume masks against collocated observations from other EO datasets.

5.3 SPIH Determination

5.3.1 SPIH Product Generation

The SPIH products are generated through application of the M6 stereo matching algorithm, introduced in Chapter 3, to the AATSR 0.55 μ m forward and nadir channels. The M6 algorithm was developed specifically for SPIH determination, with particular modifications to the normalisation and matching processes in an attempt to preserve smaller smoke features in the final stereo outputs. The 0.55 μ m channel was selected for processing due to its increased sensitivity to smoke plume particles as given by Rayleigh scattering theory¹⁶. Prior to application of M6 to the 0.55 μ m channels, a cloud mask is pre-computed using thermal thresholds applied to the 11 μ m forward channel. The derived cloud mask is then buffered by 2 pixels to ensure removal of all cloudy pixels and then applied to the 0.55 μ m forward and nadir scenes. Pre-application of the cloud mask is carried out for two reasons, firstly, to reduce the computational cost of the M6 stereo matching algorithm, and secondly, to avoid cloud disparities contaminating the M6 disparity map. It was found that with post application of the cloud mask, the

¹⁶ The intensity of the scattering varies inversely with the fourth power of the wavelength.

smoothing effects inherent in local stereo matching algorithms cause a bleeding of cloud disparities and therefore errors in the final SPIH outcomes. Following the application of the cloud mask and the stereo processing with M6, the resultant disparity map is corrected using the along track co-registration correction coefficients presented in Chapter 2 to ensure accurate disparity observations. Following correction, the disparities are converted into heights using the Mannstein camera model. The resultant heights are passed to the smoke plume masking (SPM) algorithm which is defined below in Section 5.5, leading to the determination of smoke plume locations and the final SPIH dataset. The processing chain is first summarised in the next section.

5.3.2 Processing Chain

The SPIH processing chain was implemented in Java to take advantage both of the BEAM API and to allow for integration in the BEAM visualisation software for distribution to other users within the scientific community. As indicated in the previous section, the processing chain has two main functions: the generation of pixel-level accurate SPIHs (pixel-level accuracy due to the co-registration correction coefficients) and, using AATSR's multispectral imaging capabilities, SPM generation. The SPIHs are generated using the M6 matcher and the Mannstein camera model both described in Chapter 3. The SPMs are then derived from the M6 output through inter-comparison against a local digital elevation model taken from the USGS GMTED2010 30 arc-seconds dataset (Danielson and Gesch, 2010). The key stages of the entire processing chain can be summarised as follows:

1. Ingest the AATSR Envisat style .N1 products and select for further processing observations from the two required spectral channels, 0.55 μ m Forward and Nadir and 11 μ m Nadir, as well as the relevant geo-referencing information, camera model, digital elevation model and co-registration correction coefficients.
2. Using the 11 μ m channel pre-compute the cloud mask using through a thermal thresholding process (see Section 5.5). Once computed, dilate the cloud mask by 2 pixels using a binary dilation operation and apply to the 0.55 μ m channel prior to stereo processing.
3. Apply the M6 stereo algorithm to the cloud masked 0.55 μ m channel stereo pair to generate a set of along tracks disparities. Once derived, correct the along track disparities using the co-registration correction coefficients.
4. Apply the Mannstein camera model to triangulate the disparities and return elevation estimates above a reference ellipsoid.
5. Generate the SPM by thresholding the elevations against a DEM. Apply the derived mask to generate the final masked SPIH product.
6. The resultant masked SPIH product is written out in geoTIFF format using the geo-referencing information and then converted into NetCDFv4. In addition to the SPIH layer, the following additional layers are incorporated into the product: MODIS fire radiative energy pixels of the nearest overpass; a false colour composite browse product of the forward 0.55 μ m, 1.6 μ m and 11 μ m channels; and a red-cyan stereo anaglyph of the 0.55 μ m (green) channel. However, these other aspects are not considered in this chapter.

The above processing chain was applied to the entire ALANIS dataset to locate and describe smoke plumes and smoke cloud features for ingestion into the TM5 CTM (see Krol et al., 2012). The results of the processing chain are available in the ALANIS smoke plumes database at <http://alanis.noveltis.com/>.

5.4 SPIH Inter-comparison

An initial evaluation of the M6 algorithm was provided in Chapter 3 through inter-comparison of the output stereo heights against a collocated DEM from the GMTED2010 model. This evaluation, whilst providing some insights into the particular performance characteristics of the M6 algorithm, is in many ways incomplete, as it does not provide an evaluation for the task for which it was developed, SPIH determination. In this section, this shortcoming is addressed through extensive inter-comparison of M6 SPIH outputs against collocated SPIH observations from both the CALIOP and the MISR instruments. Both of these instruments are able to generate SPIH measurements at an improved accuracy (30-60m for CALIOP; 560m for MISR) than compared to AATSR (~1km), so provide excellent comparison datasets. The section firstly describes the CALIOP inter-comparison approach and results, before moving on to that of MISR. Both inter-comparison efforts share a number of similarities in their approach, so there is significant overlap in the methods employed. In addition, a brief inter-comparison is made in this section between the performance of M6 and of M5 to determine whether or not M6 leads to improved SPIH estimates.

5.4.1 CALIOP Inter-comparison Approach

5.4.1.1 CALIOP Characteristics and Dataset

The CALIOP instrument, introduced in Chapter 4, provides an excellent SPIH comparison dataset. Here, the 1km CAL_LID_L2_01kmCLay-ValStage1-V3-01 product is employed due to its similar resolution to the AATSR instrument, and its excellent detection capability (down to optical depths of 0.02). As mentioned previously, the vertical resolution of the lidar is between 30-60 metres, with the accuracy depending on the altitude of the backscattering surface, with features at lower elevations retrieved with greater accuracy. This excellent vertical resolution makes for an ideal inter-comparison data set against which to compare the AATSR retrievals that, considering the analysis performed in Chapter 3, are expected to be accurate to around 1km for M6.

As the CALIOP instrument is on-board the CALIPSO satellite in the NASA A-Train satellite constellation it has an equatorial overpass local time of approximately 13:30, is in an ascending node, and has a 16 day repeat cycle. There are 233 fixed orbits over this 16-day repeat. This is in contrast to Envisat, which has an equatorial overpass time of 10:00 a.m., is in a descending node, and has a 35 day repeat cycle. Due to the different orbital cycles, nodes and equatorial overpass times, temporally collocated measurements between AATSR and CALIOP only occur with regularity at Polar latitudes (above 60°). As scene acquisitions descend towards the equator, the likelihood of finding co-located measurements decreases; furthermore, the time difference between any possible co-locations increases.

Due to its relatively high latitude, the ALANIS study region enables the location of collocated orbits. However, due to the different orbital configurations, there is still an approximate temporal dislocation of some two hours at the latitudes over the majority of the ALANIS study region. Furthermore, the ATSR instruments have a swath width of 512km, whereas CALIOP has a pencil sampling of $\leq 100\text{m}$, therefore even if the orbits are coincident and AATSR observes a smoke plume feature, CALIOP may still not observe the same feature.

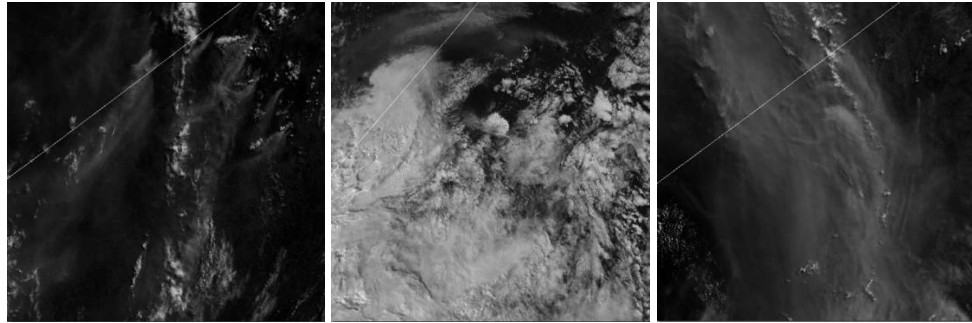


Figure 5-2. This figure presents the three AATSR scenes containing smoke plumes and collocated CALIOP observations of the plumes. The leftmost figure is from the $0.55\ \mu\text{m}$ forward channel from AATSR orbit 43976 (29th July 2010), the CALIOP path is the white line which tracks across the AATSR image. The central image is from the same channel/view from AATSR orbit 44015 (1st August 2010) and the rightmost image is again from the $0.55\ \mu\text{m}$ channel/view from AATSR 44033 (2nd August 2010). Again the white lines track out the collocated CALIOP observations.

To generate the inter-comparison dataset, firstly all AATSR observations of the study region for the 2010 study period were screened using the NEODC AATSR browse dataset. For any AATSR scenes that were found to contain smoke plumes a second search was undertaken using the COVE tool (Chander et al., 2010), which plots various instrument swaths, including AATSR and CALIOP, for a given time period onto the Google Earth globe. This tool enabled checking for coincidences between the instruments, and, through its application a total of three scenes with collocated observations were identified. These scenes are shown in Figure 5-2.

5.4.1.2 Accounting for Wind Displacements

Given the ~ 2 hour time delay between the instrument overpasses, a technique was developed to try to correct for any potential wind induced displacement of smoke plume features between observations in order to determine whether it would lead to improved collocations for the inter-comparison.

The method firstly involved locating the geographically closest AATSR pixel to each CALIOP smoke plume measurement using the associated geographical grids. The nearest geographically matched AATSR pixel was then used as an *a priori* estimate for an updated pixel position in a temporally consistent location. The updated pixel location was determined using the closest in time MERRA zonal and meridional surface wind profiles (Rienecker et al., 2011), by back projecting from the *a priori* location to the pixel location in the AATSR scene, which provides the most likely match to the CALIPSO

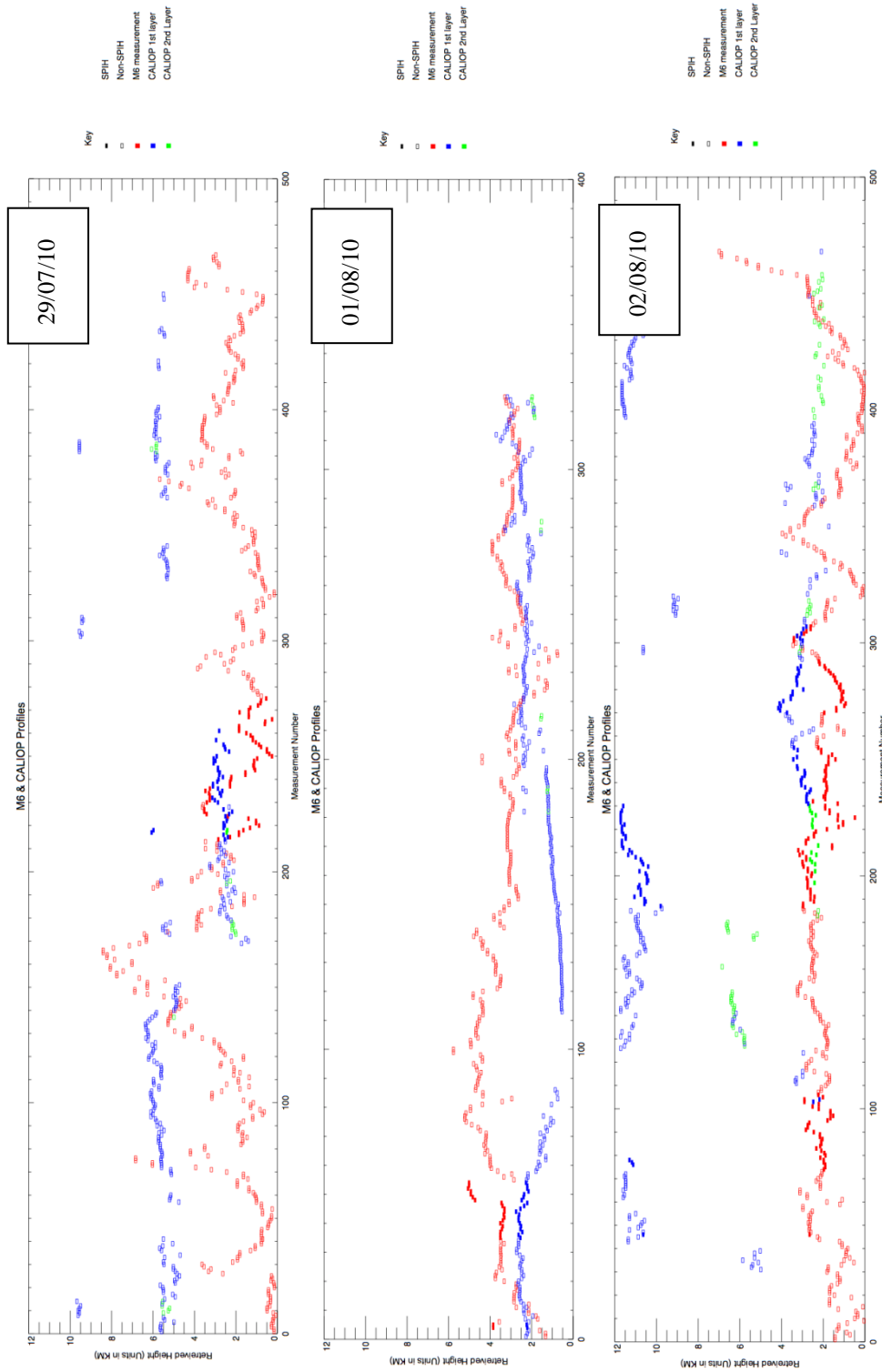


Figure 5-3. This figure shows the AATSR and CALIOP heights extracted along the CALIOP profile for each of the images presented in Figure 5-2. The filled blocks represent smoke plumes features, with blue representing the first (i.e. lowest) CALIOP layer detected, green representing the second CALIOP layer, and red representing the AATSR retrievals. The unfilled blocks represent non smoke retrievals, in this case clouds. Please rotate the page 90° in a clockwise fashion to view.

overpass when considering the wind speed. The method is straightforward and determines the shift in pixels using the following formula for both wind profile directions,

$$x_0 = x_1 - \frac{w_x \times t}{d} \quad \text{Eq. 5.1}$$

Where x_0 is the wind corrected x pixel location, x_1 is the *a priori* x pixel location, w_x is the across track wind component in the sensor coordinate system determined through a rotation of the meridional and zonal wind components by the satellite heading, t is the time between observations and d is a unit conversion factor. The y pixel locations can also be corrected for in a similar manner using the meridional wind components.

5.4.1.3 Accounting for Collocation Errors

Another issue is associated with the geo-location accuracy of the AATSR instrument. ATSR-2 has geo-location accuracy in the region of ± 2 km (Seiz, 2003), and similar accuracy is assumed for AATSR. Due to this poor geo-location accuracy the assumption that the AATSR and CALIOP observations are measuring the same point is unlikely to be true. To account for this a ± 2 pixel bounding box is used in the evaluation, and the AATSR pixel that has the closest SPIH observations to that of CALIOP (i.e. minimum distance) is returned as the AATSR pixel that is correctly collocated with CALIOP. The effect of using the minimum distance to define the collocated measurement is further assessed in the MISR inter-comparison and is shown not to significantly influence the outcome of the results.

5.4.1.4 Accounting for Multiple CALIOP Layers

A further issue occurs due to CALIOP's ability to retrieve multiple layers of cloud and aerosols. From the SPIH analysis shown in Figure 5-3, it can be seen that the top layer detected by CALIOP does not always correspond to the uppermost layer observed by AATSR. This is most likely due to the increased sensitivity of the lidar instrument compared to the passive imaging system. To account for this all layers detected by CALIOP are compared to the SPIH elevations within the AATSR bounding box and the closest height retained.

5.4.1.5 CALIOP Inter-comparison Summary

The issues outlined above, with the exception of the wind correction, are corrected for in the inter-comparisons shown in Figure 5-3, which show profiles, and Figure 5-4 which shows the scatter plot. The statistics generated from this analysis (presented in Section 1.4.2) are derived from all SPIH pixels detected by AATSR and CALIOP as defined by the AATSR SPM (see section 5.5), along with an outlier removal process. Here, outliers are defined as being two standard deviations from the mean difference. The outlier removal process applied to CALIOP in some instances observes wildly different SPIH values than AATSR (see Figure 5-4), likely due to the differences in spatial resolution and sensitivity in the AATSR SPM masking process. Furthermore, the wind induced shift correction to improve collocation was not applied in the analysis as it led to a decline in the quality of the comparison. The reasoning for this is thought to be that wildfires events generally have a point source and that smoke is being generated continuously from the fire event. Therefore under the assumption of near consistent wind speed/direction

and fire intensity between the AATSR and CALIOP observations the plume characteristics at the point of collocation are likely to be similar despite the temporal dislocation. Application of the wind induced shift correction likely introduces errors caused by further plume rise between observations as is typical of buoyant smoke plumes.

5.4.2 CALIOP Inter-comparison Results

The inter-comparison results for the profiles shown in Figure 5-3 are presented in Figure 5-4. The statistical outcomes of this statistical analysis, following the outlier removal process described in the previous section give a RMSE of 480m with a bias of 240m and a coefficient of determination of 0.5. For a discussion of these outcomes please refer to Section 5.4.5.

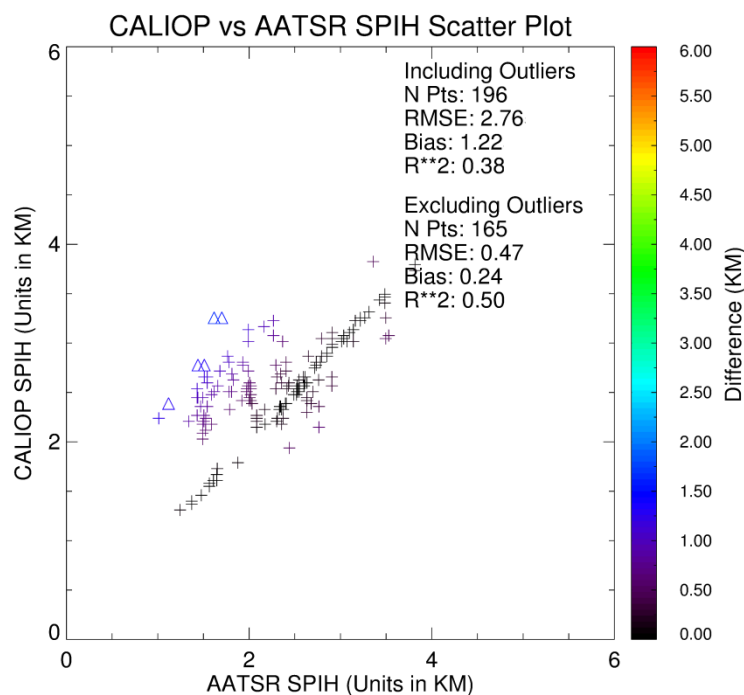


Figure 5-4. The inter-comparison scatter plot of the SPIH inter-comparison observations as shown for the three dates with collocations in Figure 5-3. Outliers are defined as unfilled triangles, of which a number are outside the plot axes and are instances of where CALIOP detects a cloud feature but AATSR detects smoke.

5.4.3 MISR Inter-comparison Approach

5.4.3.1 MISR Characteristics and Dataset

The AATSR and CALIOP inter-comparison does not provide convincing results (see Section 5.4.5), and the number of points inter-compared is very limited, necessitating further validation. This shortcoming is overcome, through a second analysis using the MISR instrument, also introduced in Chapter 4. Among

the important instrument characteristics are that the Terra satellite (on which MISR is carried) is in a descending node and has an equatorial overpass time of 10:30am, hence being much closer in time to Envisat (10:00am). This allows for a more direct inter-comparison when coincidences occur and therefore improved validation reliability. It also means that the effects of advective winds are similar in both datasets, and that the plumes are in similar locations.

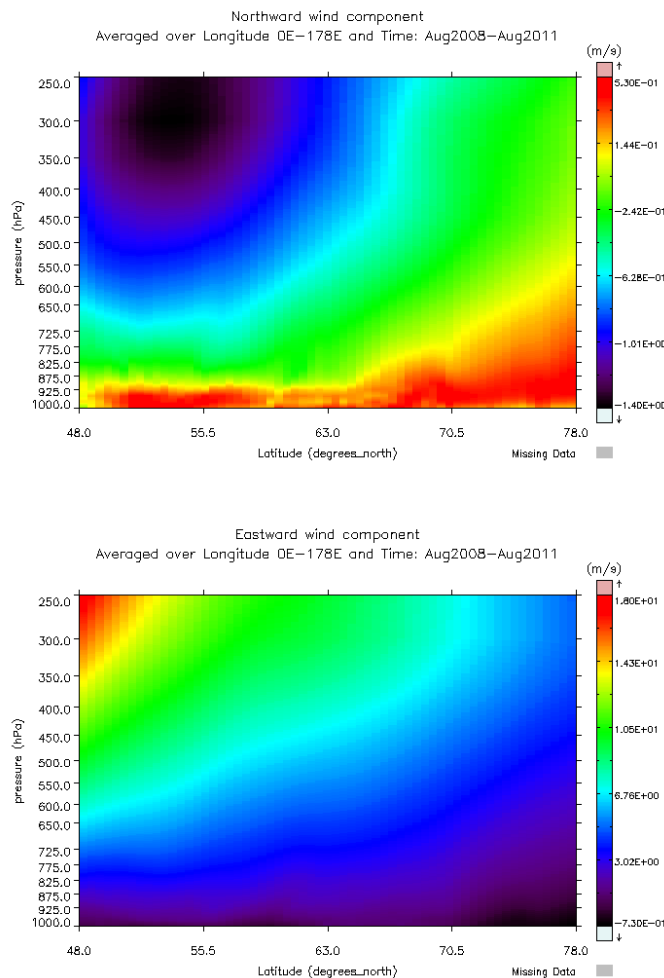


Figure 5-5. The upper plot shows the latitudinal average meridional (N-S) northward wind speeds over the ALANIS study regions for the period August 2008 through August 2011 as returned by the MERRA reanalysis. The lower plot shows latitudinal average zonal (E-W) eastward wind speeds for the same region and period.

The MISR instrument has been used to generate manually measured, stereo derived, SPIH datasets for a number of regions globally within the MISR Plume Height Project using the MINX toolkit. The MISR-MINX tool, is an IDL based visualisation and digitising toolkit for making manual stereo measurements of clouds, smoke/volcanic plumes, dust clouds, etc., from MISR imagery. There are a number of algorithms within MINX; however of most importance is the stereo matching algorithm. This algorithm uses Pearson's correlation coefficient to match an image patch extracted from the nadir camera to five other views (CF, BF, AF, AA, BA, CA) at 275 metre resolution. Once the patch has been matched in all

five views, a minimum curvature surface is fitted to the correlation matrix and this is used to interpolate the disparity to a sub-pixel location. Following matching, a wind induced disparity correction algorithm (based on Horvath and Davies, 2001; Davies et al., 2007) is applied leading to an assumed final height accuracy of approximately ± 200 (Davies et al., 2007) metres. In some cases wind correction is not applied due to failed wind retrieval. It should be noted that the AATSR disparities are not corrected for wind induced parallaxes in either the CALIOP or MISR inter-comparisons, or in the final ALANIS stereo SPIH product, and such wind induced disparity errors can, in some instances, be significant (Seiz, 2003). The justification for not applying wind induced disparity correction is reached through analysing the latitudinal average zonal and meridional mean wind speeds for the ALANIS study period and region from MERRA reanalyses as shown in Figure 5-5. Assuming the reanalyses are accurate (see Chapter 6 for observation/reanalysis inter-comparison), the wind speeds over the ALANIS study on average region do not exceed $\pm 8 \text{ ms}^{-1}$, the minimum wind speed required in order to produce a 1 pixel disparity (and in turn $\sim 1\text{km}$) error in the AATSR stereo outputs. This is particularly the case for the meridional winds which have the greatest potential impact, when mean wind speeds below 250 hPa rarely exceed 2 ms^{-1} . Given the above analysis, and the fact that the M6 algorithm is accurate to $\sim 1\text{km}$, points which failed in the MISR-MINX wind correction algorithm (i.e. no wind correction applied) are still employed in the analysis.

Of the numerous datasets generated with the MISR-MINX tool for the MISR plume height project one has been generated over Siberia for the 2008 fire seasons and an overview is shown in Figure 5-6. In this dataset, all MISR orbits from March through August were assessed for smoke plumes. Again a collocation assessment was performed using the COVE tool; this time however, the MISR Plume Height dataset was used as a reference rather than AATSR (as with the CALIOP analysis). From this analysis nine AATSR orbits were collocated with the MISR smoke plume measurements with overpass times differing by < 30 minutes.

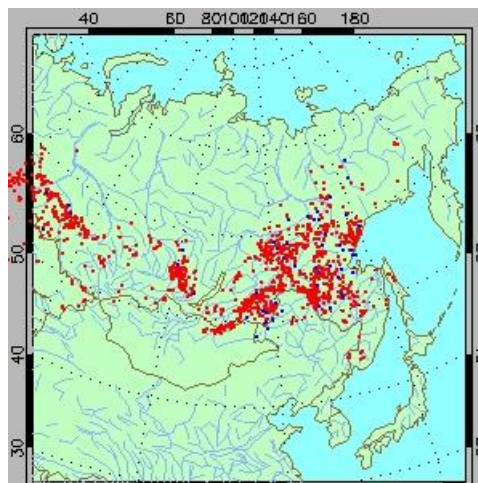


Figure 5-6. This image, taken from the MISR plume height project webpage (<http://www-misr.jpl.nasa.gov/getData/accessData/MisrMinxPlumes/>, last accessed 4th August 2013), shows all digitised MISR plumes for the 2008 MISR study over Siberia.

5.4.3.2 MISR Inter-comparison Summary

For the analysis, the bounding-box method (see Section 5.4.1.3) was again applied to account for the poor AATSR geo-referencing accuracy, with the closest heights between the AATSR M6 and MISR MINX SPIH observations being retained for the analysis. SPM masking was achieved by using the manually digitised MISR-MINX SPIH observations as the mask. In addition to the minimum, the mean, median and maximum differences from the bounding box are also analysed to get a better idea of the effects of choosing the minimum on the output statistics and also to better assess the overall quality. In total, 23,882 SPIH comparisons between AATSR and MISR-MINX SPIH observations were made across the 9 orbits, a significant increase over the tens of observation available in the CALIOP analysis.

5.4.4 MISR Inter-comparison Results

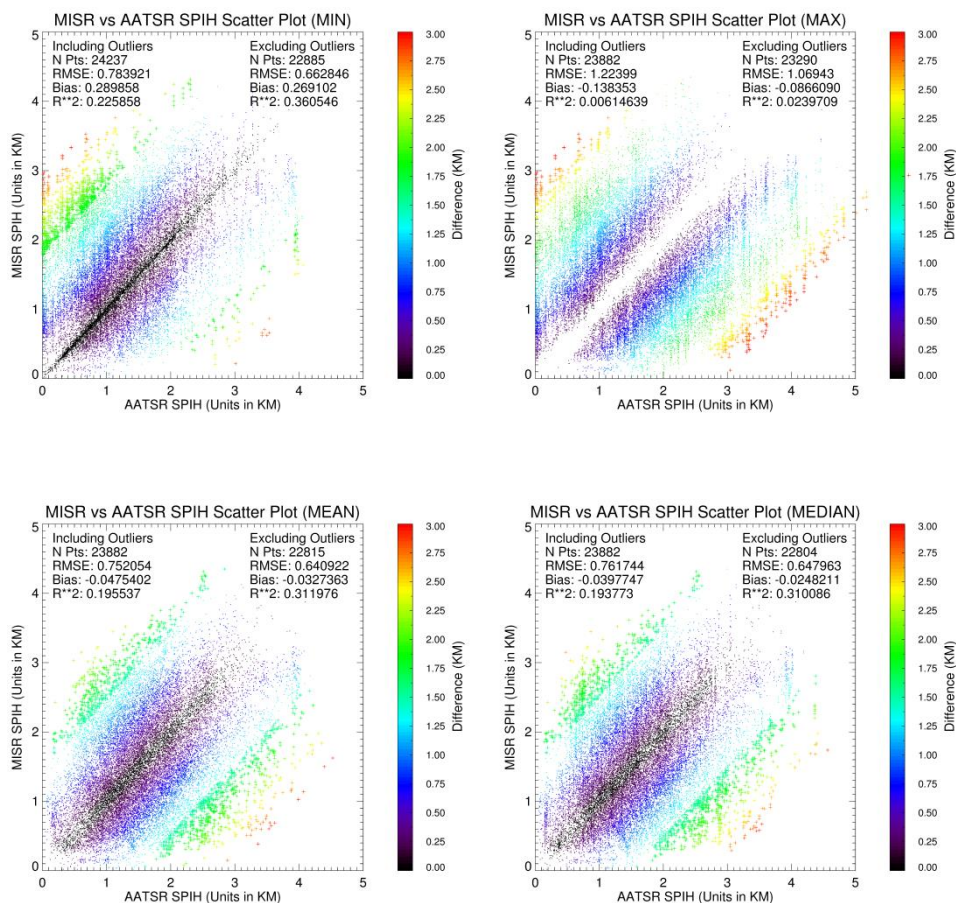


Figure 5-7. The plots in this figure show the AATSR inter-comparisons against MISR. The top left plot shows the results when using the minimum difference from within the bounding box, the top right the maximum difference from within the bounding box, the bottom left the mean difference with the bounding box, and the bottom right the median difference with bounding box. Note the similar outcomes with the minimum, mean and median. Outliers are plotted as crosses.

5.4.5 Inter-comparison Discussion

The CALIOP inter-comparison comprises a relatively small number of data points. This is due to the limited number of collocated measurements found, as would be expected at the study region's latitude range given the instrument characteristics. Following outlier removal, where an outlier is here defined as being more than two standard deviations from the mean, the inter-compared points result in an R^2 value of 0.5, which was found to be statistically significant to the 95% level using the Student t-test. The AATSR observations are biased low compared to CALIOP by 240m. This is probably due to increased lidar sensitivity to aerosol particles, causing CALIOP to retrieve heights nearer to the top of the plume. AATSR retrieves the SPIH at the point where the plume reaches a suitably high optical thickness to provide enough texture to be detected in the passive sensor array.

The MISR analysis provides a far more comprehensive inter-comparison of SPIH, with the number of individual pixels inter-compared numbering in the tens of thousands. Focusing solely on the results generated using the minimum difference shown in Figure 5-7 it is found that the initial correlation statistics computed for all points result in an R^2 of ~ 0.3 following outlier removal, with anything ≥ 2 standard deviations from the mean is classed as an outlier. Looking at the minimum plot in Figure 5-7, this appears to be due to the MISR data being used as the SPM and the fact that MISR detects many more low elevation smoke plumes which AATSR detects as the land surface, i.e. $< 1\text{km}$. This is due to MISR's higher resolution and therefore higher stereo accuracy, giving the ability to resolve smaller smoke plumes at lower elevations. AATSR due to its lower resolution, especially in the forward channel, fails to detect these plumes, leading to the poor R^2 result. Those pixels that AATSR detects that are potentially the land surface, when excluded as outliers, lead to a better evaluation of the quality of the M6 measurements, as the evaluation is constrained to be within the operational limits of AATSR and M6. To exclude the pixels from the statistical analysis the assumption is made that any AATSR SPIH measurement which is less than 1km above a collocated subset of the GMTED2010 DEM is potentially the land surface. With this threshold applied the statistics are greatly improved with an R^2 score of ~ 0.69 .

The use of the minimum height difference from the bounding box may be skewing the results and leading to an ineffectual and misleading assessment of the quality of M6. Selecting the minimum gives the optimum outcome; this best match from within the bounding box may not provide an accurate representation due to the noise effects common to stereo matching algorithms. Therefore further assessment was undertaken to better assess the effect of choosing the minimum and to determine whether this decision creates an unrealistically optimistic quality assessment of M6. This further assessment used the mean, median and maximum height differences between the AATSR and MISR observations within the bounding box. The results for this can be seen in Figure 5-7. The median and mean scatterplots show very similar distributions and statistics to the minimum plot. This is indicative that the majority of the heights within the bounding boxes are similar to the minimum height, leading to good confidence in using the minimum as the quality indicator. The maximum scatterplot shows that in most cases the largest height deviation within the bounding box between the AATSR and MISR is around 1km. Such deviations are likely due to boundary crossings within the bounding box, i.e. moving from a plume feature to a land feature. This further assessment using the median, mean and max verifies that the use of

the minimum height difference for validation is not skewing the results significantly and adds robustness to the evaluation.

5.5 SPM Determination

5.5.1 SPM Technique

SPMs are required for the accurate delineation of smoke from other features such as cloud and the land surface. Previous attempts at SPM generation have tended to focus upon radiometric threshold techniques. The thresholds are typically derived from statistical or machine learning methods (Mazzoni et al. 2007b; Xie et al. 2007). Whilst generally effective, these SPMs are often susceptible to misclassification due to the lack of a stable reflectance curve for smoke, with large overlaps existing between the spectral profile of smoke and non-smoke features (Xie et al. 2007). Through a combination of an elevation and radiometric thresholding, AATSR has a potentially enhanced capacity to effectively differentiate smoke for other atmospheric constituents.

The AATSR SPM is generated in a number of steps as follows. Firstly, a cloud mask is generated from the 11 μm channels using one of two radiometric thresholds,

$$M_c(x, y) = \begin{cases} 0 & \text{if } [F_{11}(x, y) < \bar{F} + F_\sigma] \text{ or } [F_{11}(x, y) < 280K] \\ 1 & \text{otherwise} \end{cases} \quad \text{Eq. 5.2}$$

Where M_c is the cloud mask, F is the forward scene for the subscripted channel, \bar{F} is the forward image mean and F_σ is the image standard deviation. The second step is to buffer the cloud mask. The buffering is required to try to remove as many cloudy pixels as possible from the SPM. As stereo results can vary between channels, there is not always a direct cloud correspondence between them, so buffering removes the errors associated with this. The buffering is achieved through application of a boxcar averaging filter to the cloud mask. This effectively dilates the cloud mask leading to the required buffering effect. Once derived, the cloud mask is applied to the forward and nadir scenes 0.55 μm scenes prior to stereo matching to eliminate these pixels from the matching process. This is required as area based methods (even M6 with its specific modifications) have a tendency to smooth features. Due to these smoothing effects, *a posteriori* application of the cloud mask rarely removes all cloud features.

The last step is to create the SPM, M_{sp} . This is achieved by removing land features by comparing the cloud masked 0.55 μm stereo results to the GMTED2010 DEM. Any pixel whose proximity to the DEM elevation is within the algorithms height accuracy ($\approx 1\text{km}$ for AATSR and M6), is flagged as land, and set to zero (i.e. non smoke),

$$M_{sp}(x, y) = \begin{cases} 1 & \text{if } SPIH(x, y) - DEM(x, y) > a \\ 0 & \text{otherwise} \end{cases} \quad \text{Eq. 5.3}$$

A number of example SPM outputs from this process are shown in Figure 5-8.

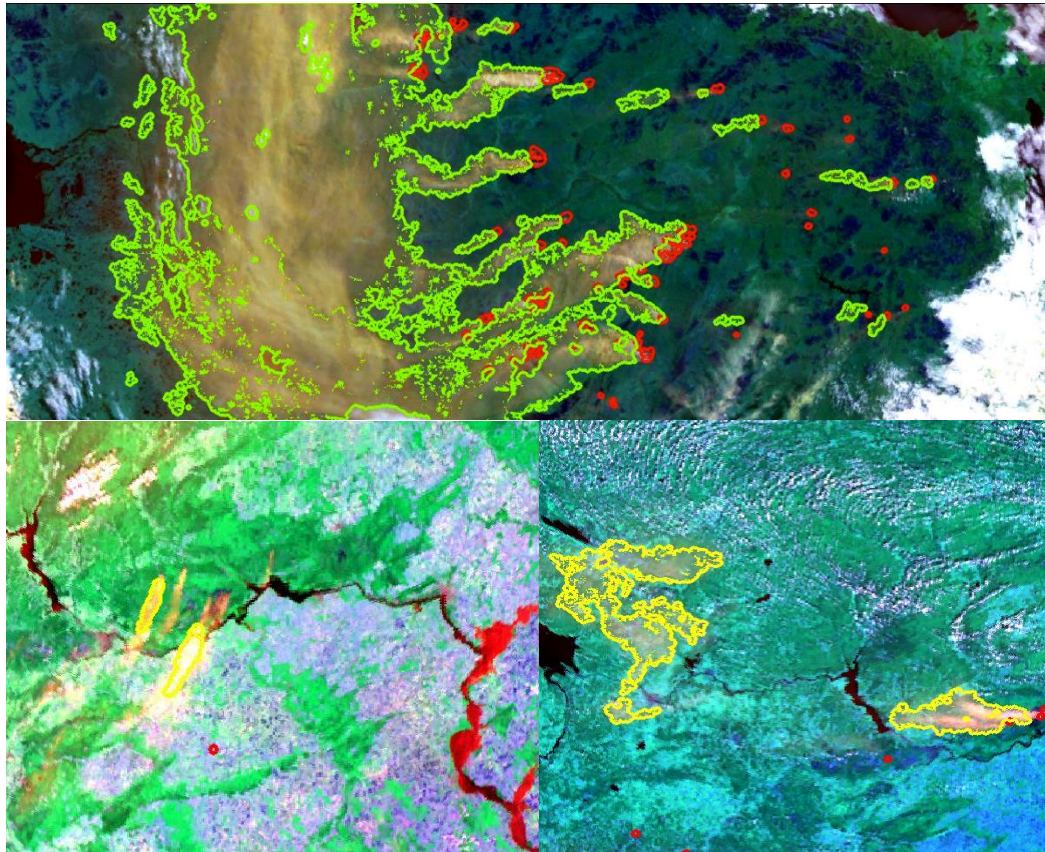


Figure 5-8. A selection of scenes showing the capabilities of the algorithm for smoke detection. The red points on the scene are fire hot spots (FRP) detected by the MODIS instrument. The yellow/green bounded regions represent where smoke plumes/clouds have been detected. The upper most scene was taken on the 30th July 2010, the bottom left image is from the 14th July 2010 and the bottom right from the 7th July 2010. Note in the bottom left scene that some of the smaller smoke plume features are being missed. This is due to the fact that they are less than 1km above the terrain.

5.5.2 SPM Evaluation Approach and Results

The SPM validation is carried out on three AATSR scenes co-incident with MISR in 2008. The scenes are taken from the same AATSR data set used in the SPIH validation against MISR. For each scene, the SPM is applied and then inter-compared with the MISR-MINX tool hand digitised MISR smoke plumes, which are assumed to be the reference “smoke truth” SPMs. The validation is carried out by statistical and visual assessment. The AATSR SPM is overlaid onto the scene of its derivation, as are the human operator detected SPMs from MISR that occur within the AATSR scene boundaries (see Figure 5-9 through Figure 5-14). Statistics are generated using a confusion matrix, typically used in classification quality assessment. The two classes here are smoke and non-smoke and the MISR SPMs are set as the ground truth to which AATSR is being compared. Omission and commission percentages are generated for the smoke class from the confusion matrix. When the omission percentage is very high, this represents MISR plumes that AATSR has not detected. A high omission percentage indicates a lack of

common regions between the masks. Commission occurs in the regions where AATSR detects smoke and MISR does not. A low commission percentage is ideal as it indicates that AATSR is detecting the same smoke as MISR, i.e. few false positives with regard to the MISR data. All statistics in the following sections are computed from smoke plumes above 1km in height (justified through the MISR SPIH validation in Section 5.4.4).

5.5.2.1 SPM Inter-comparison: 23rd April 2010

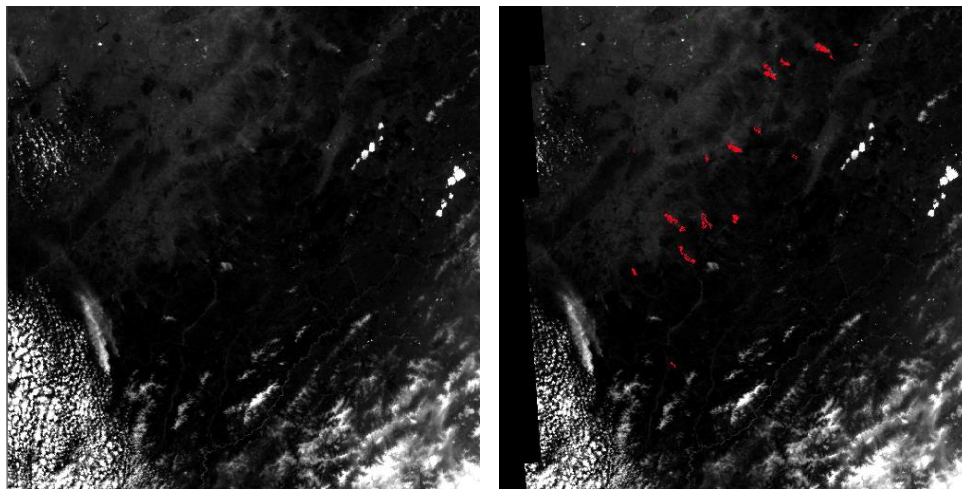


Figure 5-9. The left image is nadir 0.55 μ m nadir view from AATSR. The right image is the same view with MISR SPMs (Red) and AATSR SPMs (green) overlaid. Common areas are in yellow.

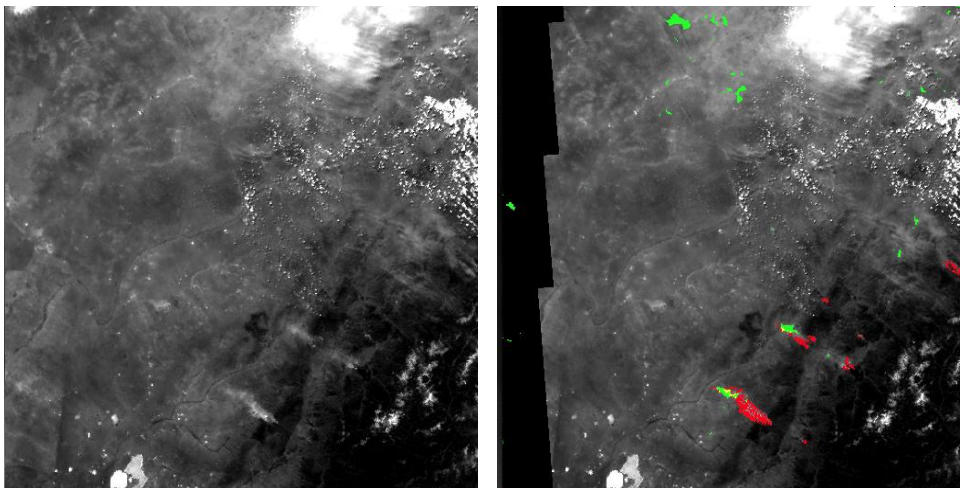


Figure 5-10. The left image is nadir 0.55 μ m nadir view from AATSR. The right image is the same view with MISR SPMs (Red) and AATSR SPMs (green) overlaid. Common areas are in yellow.

AATSR detects no smoke in Figure 5-9, all the red MISR plume regions are greater than 1km in height so they should be detectable by the mask. However, the MISR plumes are not substantially greater than 1km, and it is likely that the inherent noise in the M6 algorithm leads to their non-detection. Detection may also be problematic in these smaller plumes due to the averaging effect of the Gaussian window used

for convolution during the stereo matching (which is reduced in M5 but not removed entirely). The land regions surrounding the small plumes will tend to dominate leading to suppression of the plume feature, especially in a situation where the plume only comprises a very small region within the Gaussian window. In this instance the omission errors are 100%, as AATSR detects none of the MISR features. In Figure 5-10 larger plumes are present and AATSR and MISR detect plumes in similar locations. The omission statistic is quite poor, at 93.6% indicating that the AATSR plumes only detect 6.4% of the MISR plumes. The commission percentage is 77%, meaning that 77% of the smoke plumes detected by AATSR are not seen by MISR

5.5.2.2 SPM Inter-comparison: 17th May 2010

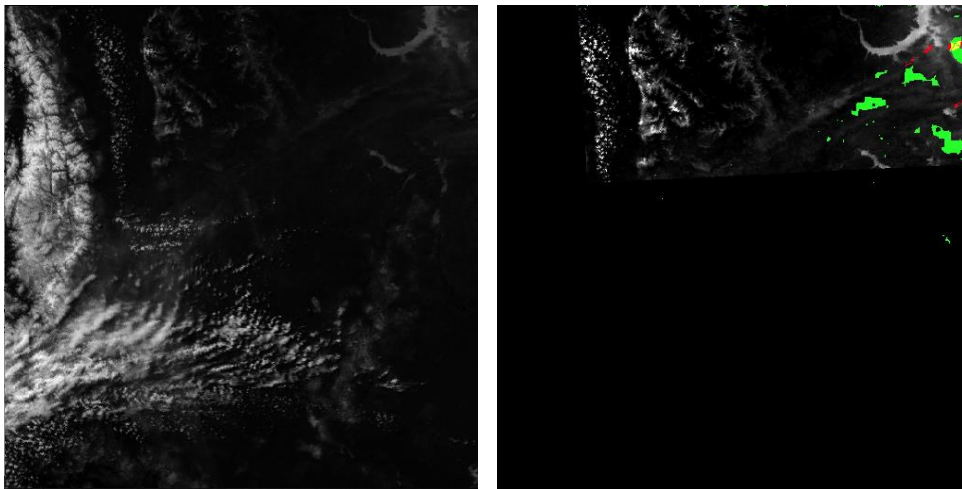


Figure 5-11. The left image is nadir 0.55 μ m nadir view from AATSR. The right image is the same view with MISR SPMs (Red) and AATSR SPMs (green) overlaid. Common areas are in yellow.

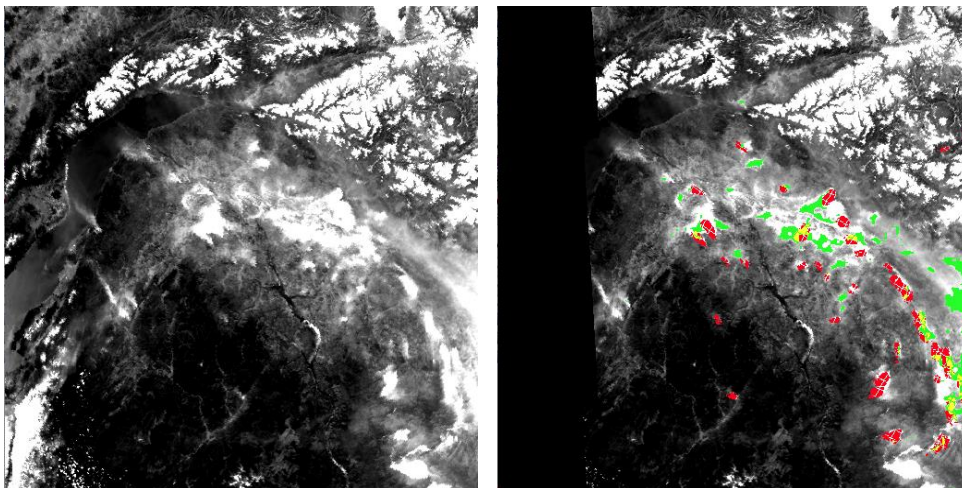


Figure 5-12. The left image is nadir 0.55 μ m nadir view from AATSR. The right image is the same view with MISR SPMs (Red) and AATSR SPMs (green) overlaid. Common areas are in yellow.

In Figure 5-11 the omission percentage is 65%, meaning that 35% of the plumes detected are in common. The commission percentage is again very high at 96%. In Figure 5-12 there are a significant amount of smoke clouds and smoke plumes, AATSR detects a number of the same plumes as MISR and also the smoke clouds. However, the smaller plumes are also missed here. The omission percentage for this figure 70% meaning that AATSR shares 30% similar plume coverage with MISR. The commission percentage is 79%, showing that 79% of the plumes detected by AATSR are not seen by MISR.

5.5.2.3 SPM Inter-comparison: 17th July 2010

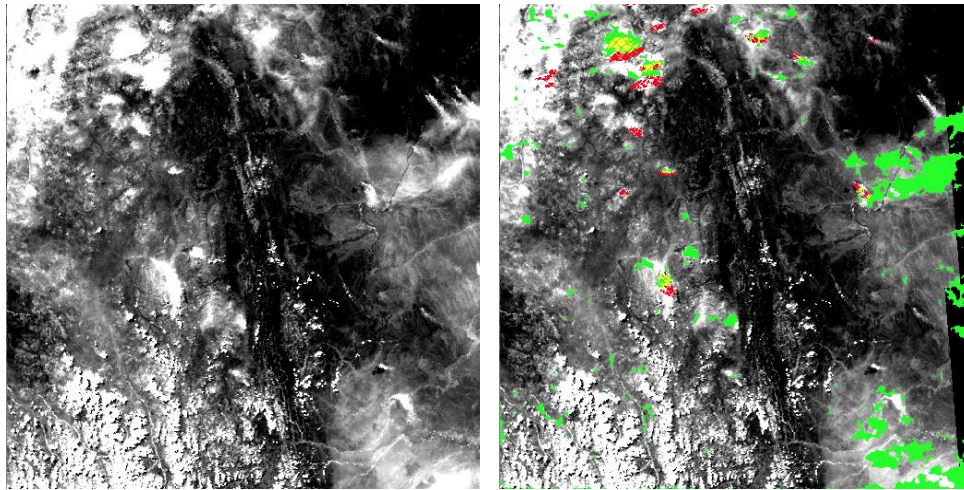


Figure 5-13. The left image is nadir 0.55 μ m nadir view from AATSR. The right image is the same view with MISR SPMs (Red) and AATSR SPMs (green) overlaid. Common areas are in yellow.

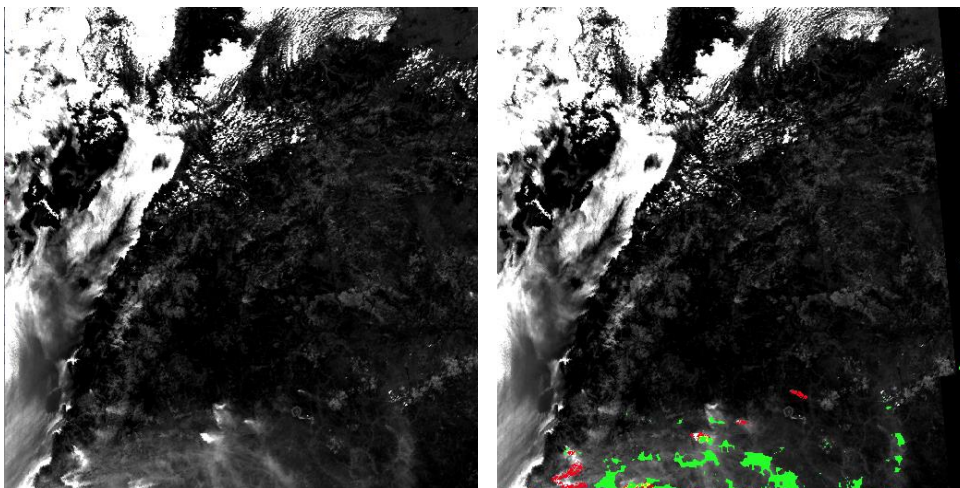


Figure 5-14. The left image is nadir 0.55 μ m nadir view from AATSR. The right image is the same view with MISR SPMs (Red) and AATSR SPMs (green) overlaid. Common areas are in yellow.

In Figure 5-13 AATSR detects a large number of smoke plumes and significant amounts of smoke clouds. The omission percentage is 67% with 33% of the MISR plume area being detected by AATSR. In Figure 5-14 the SPM also the omission percentage is 81% indicating a shared detection of 19%. In both instances high commission percentages of 95% are returned.

5.5.3 SPM Discussion

For large smoke plume features, such as those presented in Figure 5-8, the SPM algorithm appears to be performing well, effectively delineating the boundaries of the smoke plume features. However, compared to the MISR manually digitised features, the AATSR SPM performs poorly for the detection of smoke features (see Figure 5-9 through Figure 5-14), as demonstrated by both the errors of omission and commission. There are a number of reasons for these differences attributable to both to instrumental and algorithmic differences. Looking firstly at instrumental differences, MISR has more looks over a greater range of angles at the plume and the stereo matching is performed across looks with a reduced angular difference than achievable with AATSR. This means that there is greater similarity between plumes in the stereo imagery and therefore there is a greater likelihood of an effective stereo matching process. Furthermore, MISR has a higher resolution than that of AATSR (275m vs 1000m), MISR is therefore more able to resolve smaller plume features effectively. For these two instrumental features MISR is better able to detect smaller smoke plume features that AATSR routinely misses. In terms of algorithmic differences, the main benefit of MISR is the human operator, which allows for very reliable plume definition. M6 provides a far less reliable indicator which leads to a large number of false positive smoke features being detected (as shown by very high commission percentages), again particularly for lower elevation smoke plume features where noise in the M6 algorithm leads to poor SPM classification. In summary, whilst the current AATSR SPM method appears to be effective for large extensive smoke plume features (in the hundreds of pixels in area, rather than the tens of pixels), it is very much lacking in its ability to effectively detect smaller smoke plumes features.

5.6 Summary and Future Work

5.6.1 Objective Review

As with the previous science chapter, this section addresses the scientific objectives outlined in Section 5.2 in terms of their successful completion or insights gained.

‘Assess the performance of the M6 stereo matching algorithm in combination with AATSR for determination of SPIH against collocated higher resolution observations derived from alternative satellite instruments.’

The smoke plume tuned M6 stereo matching algorithm has been compared against both CALIOP and MISR. Against a sample of 165 CALIOP observations (following outlier removal with anything greater than 2 standard deviations from the mean excluded) an RMSE of ~500m and an R^2 of 0.5 were obtained. Furthermore, biases of ~250m were demonstrated. The MISR dataset provides a far larger inter-comparison set with ~20,000 collocated observations being compared. An RMSE of ~660m and bias ~270m were returned, along with a coefficient of determination of 0.69 following an outlier removal process which excluded all collocated observations greater than 2 standard deviations from the mean and less than 1km above the local terrain.

‘Develop and apply an automatic method for locating smoke plumes in the AATSR imagery.’

A simplistic spectral cloud masking and elevation threshold based SPM has been developed to automatically detect smoke features in the AATSR imagery. This approach firstly attempts to exclude all cloud features from the imagery through a threshold process applied to the ATSR 11 μ m channel. The mask derived from this process is applied to the 0.55 μ m channel prior to stereo processing. The SPM is then generated by comparison the stereo processed 0.55 μ m channel against a collocated DEM and rejecting any stereo height which is less than 1km above the local terrain height.

‘Inter-compare the AATSR smoke plume masks against collocated observations from other EO datasets.’

An inter-comparison of the AATSR SPM has been made against collocated SPMs from the MISR plume height project. In this inter-comparison, the AATSR and MISR masks show a mean commission percentage of ~90% and a mean omission percentage of ~80%. This indicates that the current AATSR SPM masking method does not perform effectively in the task of smoke plume detection, in particular for small plumes.

5.6.2 Future Work

From the review presented here, it is possible to outline a number of avenues to pursue in future studies. Perhaps the most critical improvement required is in the SPM generation. Whilst it is currently effective for large plume features, the noisy nature of M6 (see Chapter 3 Figure 3-12 to Figure 3-14) leads to much confusion with smaller smoke plume features, or those features which are not much greater than 1km above the land surface. In general, M6 appears to be accurate to ~1km for smoke plumes above 1km in elevation, as shown in the SPIH inter-comparison studies, but is imprecise and suffers from substantial matching noise. The census algorithm tends to be more precise (again see Chapter 3 Figure 3-12 to Figure 3-14), with less matching noise, so may provide a more robust SPM output with the elevation threshold approach. In additions census may lead to more accurate SPIH retrievals. Further enhancements are likely to be achieved through application of machine learning techniques which are able to take advantage of AATSR multi-spectral imaging capabilities to effectively differentiate between smoke and non-smoke plume features, such as that employed previously on MISR (see Mazzoni et al., 2007b). AATSR has the distinct advantage over MISR, in that its thermal channels will enable the machine learning method employed to more robustly separate cloud features from smoke features. This is likely to be the best direction for improving SPM outputs, and a similar technique has been shown to perform well for application with AATSR in other recent studies (e.g. Brockmann et al., 2013).

Another avenue for future work is to employ some feature recognition technique that can effectively differentiate between smoke plume and smoke cloud features. This task would be challenging as smoke plumes and smoke clouds rarely occur with regular structures. However, differentiating between smoke plumes and smoke clouds is of critical importance in order to effectively and automatically define new and old smoke features. This differentiation is important, as smoke plume features (i.e. new smoke) are important in model initialisation stages, whereas smoke clouds (i.e. old smoke) are only useful for model

validation. Effective separation is therefore critical, as employing smoke clouds for model initialisation will not result in reliable model distribution predictions.

5.6.3 Chapter Summary

This chapter has presented the application of AATSR for the determination of SPIH and SPMs for use within the ESA ALANIS project. The M6 algorithm, introduced in Chapter 3, is employed here for the determination of SPIH. In a statistical assessment against higher accuracy collocated SPIH observations from CALIOP and MISR, the AATSR M6 retrievals return an RMSE of ~500m and an R^2 of 0.5 against CALIOP and an RMSE of ~660m and an R^2 of 0.69 against MISR. The current SPM method has described and evaluated against SPMs from MISR. For large plumes the SPM masking method appears to perform effectively, however it has been shown to be ineffective in a comparison against MISR (mean commission percentage of ~90%; a mean omission percentage of ~80%), with particular difficulties in extracting smaller smoke plume features.

6 ATSR TANDEM ATMOSPHERIC MOTION VECTORS

6.1 Introduction

Observations of tropospheric winds are of critical importance in defining the current and predicting the future state of the atmosphere. Surface observation networks, whilst able to provide adequate mesoscale descriptions of atmospheric motion over continental regions (mainly in the Northern Hemisphere), fall short over the oceans (and over continental regions with low population density and/or low economic development), providing only very sparse coverage due to a lack of observing stations. Remotely sensed data from space-borne systems provides a suitable alternative data source, which not only has the potential to provide regular global sampling of atmospheric motion, but can also generate observations at microscale resolutions (1km or less). This potential has been developed through various techniques applied to both active (Browning and Wexler, 1968; Horstmann et al., 2000; Naderi et al., 1991; Stoffelen and Anderson., 1993) and passive remote sensing systems (Davies et al., 2007; Horváth and Davies, 2001a; Horváth and Davies, 2001b; Key et al., 2003; Moroney et al., 2002; Mueller et al., 2013; Muller et al., 2002; Velden et al., 1997; Velden et al., 2005). Of these approaches, the most popular is the tracking of atmospheric features (e.g. clouds) across temporally separated satellite images to infer atmospheric motion. This approach, first employed in the 1960s (Menzel, 2001) on the TIROS polar orbiter, was further developed for application to geostationary imaging systems in the 1970s and 1980s using both manual and automated techniques and, since 1996, has been employed operationally on the GOES geostationary imagers (Key et al., 2003). The assimilation of the AMVs derived from these initiatives into NWP reanalyses began in the 1980s, and is now commonplace across all major NWP systems (Kalnay et al., 1996; Rienecker et al., 2011; Simmons et al., 2007).

The geostationary imaging systems employed, such as those in the GOES program, provide atmospheric motion vectors (AMV) from tracking features across imagery in either the mid-infrared (water vapour) or thermal regions (clouds) of the EMS and are capable of providing excellent temporal coverage (every 15-30 minutes) at good spatial resolutions (1-8km, channel dependent). Furthermore, the tracking algorithms employed can achieve AMV accuracies on the order of $\sim 7.5 \text{ ms}^{-1}$ (Velden et al., 1997). However, geostationary imaging systems suffer from the substantial drawback of incomplete global coverage. This inherent weakness stems from the curvature of the Earth and geostationary imaging geometry limitations,

leading to such imagers having observational footprints covering only latitudinal regions inside the bounds of $\pm 60^\circ$. In order to provide AMV observations in those regions outside of geostationary footprints, polar orbiting satellite systems can be employed. Polar orbiting systems, whilst having reduced temporal sampling, typically three times per day at best, provide significantly improved spatial coverage, including complete polar observation given a suitable swath width. The importance of AMV observations from polar orbiting instruments on improving the robustness of NWP models is now well known (Bormann and Thépaut, 2004), with MODIS polar AMVs (Key et al., 2003) being assimilated in a number of current generation models and reanalyses (e.g. Rienecker et al., 2011).

AMVs have been derived operationally from MODIS (for instrument characteristics, see Chapter 4) through application of atmospheric feature tracking techniques since July 2002 (Key et al., 2003). Water vapour and cloud features are tracked across overlapping polar stereographic projected image triplets with a temporal separation of 100 minutes from the first acquisition to the last. The tracking algorithms are based on those employed on the GOES geostationary systems (Key et al., 2003), where the $11\ \mu\text{m}$ channel (MODIS band 31) is used to track cloud features and a $6.7\ \mu\text{m}$ channel (MODIS band 27) is used to track water vapour features. Cloud features to be tracked are identified by searching for the coldest BT within a $26\ \text{by}\ 26\ \text{km}^2$ region and then computing the local radiance gradients. Any observed features with gradients which exceed a pre-set threshold are employed for tracking. Water vapour feature determination follows a similar gradient/threshold process. Once the features are identified in the initial image they are located in the two subsequent images through use of the sum of the squared difference metric and a specified search window based on meteorological constraints derived from a forecast model. The two AMVs derived are then inter-compared to ensure consistency between the estimates and screen any observations outside empirically defined displacement constraints. For each defined AMV a wind height assignment is performed using one of either of the techniques employed on MODIS as discussed in Chapter 4. The AMVs derived from MODIS have been extensively assessed against $\sim 27\ 000$ rawinsonde observations, and in statistical analysis demonstrated an RMSD of approximately $8\ \text{ms}^{-1}$ and a bias of $-0.58\ \text{ms}^{-1}$.

In addition to MODIS, AMVs are also derived near operationally from the MISR satellite (for instrument characteristics see Chapter 4) based on the initial work undertaken by Horvath and Davies (2001a; 2001b). Since then a number of further enhancements have been implemented to improve the AMV quality (Davies et al., 2007; Mueller et al., 2013). The original AMV retrieval algorithm operated using a sparse feature based matching algorithm referred to as nested-maxima (Moroney et al., 2002; Muller et al., 2002) applied to the nadir, 46° and 70° views separately for both forward and aft camera triplets. The use of camera triplets enables the effective separation of cloud induced parallax and wind induced (advective) motion in the stereo axis. The nested-maxima algorithm operates with excellent efficiency, but has a low retrieval percentage, with observations typically generated for only 1-2% of the image data, and a low accuracy, with matches good to only ± 2 pixels. These performance characteristics necessitated the use of large aggregation regions in order to generate reliable AMV estimates, leading to a retrieval domain of $70.4\ \text{by}\ 70.4\ \text{km}^2$. The current AMV retrieval algorithm (Mueller et al., 2013) has switched to an area based stereo matcher for motion tracking which employs a pyramidal scheme and the SAD metric. The dense coverage returned by the area based stereo matcher has enabled a significant

improvement in the retrieval resolution, improving from 70.4 to 17.6 km. The accuracy achieved has remained relatively consistent across algorithms, with a slight improvement in the most recent approach (Horváth, 2013). The algorithm evaluation procedure employed (see Horváth, 2013) checks non-cloud features (i.e. the ground) for shifts, where a shift of 0 ms^{-1} indicates a perfect retrieval (as the ground should demonstrate no motion between retrievals). The most recent analysis returned RMSD values of 0.5 ms^{-1} for the meridional wind component and 1.7 ms^{-1} for the zonal wind component. The meridional ground shift observations were found to be practically unbiased, whilst a bias of 0.3 was found for the zonal analysis.

The MISR and MODIS instruments have some inherent drawbacks in their AMV retrieval schemes. Due to instrumental characteristics and the need for overlapping imagery MODIS can only provide AMV measurements at latitudes above 70° , it therefore provides incomplete global coverage. A far more critical consequence of the MODIS instrumental characteristics is that in order to obtain AMV estimates it must observe the cloud features from differing orbital paths. Observations from different positions have the potential to introduce parallax induced displacements (see Chapter 3). These parallax effects are currently unaccounted for in the operational retrieval (Key et al., 2003) and are likely introducing substantial, height dependent, biases into the AMVs. Another problematic aspect with MODIS is ineffectual wind height assignment caused by the RT based CTH algorithms employed, which often perform poorly in Polar Regions (Stubenrauch et al., 2013). This is a significant issue as the MODIS AMVs are generated for these regions in particular. The MISR-derived measurements, whilst far more robust and accurate in terms of speed and direction assignment due to the stereo methods employed, are only available from visible channels. This prevents both diurnal observations and the retrieval of winds over Polar Regions during winter months. Some problems with MISR occurred due to aliasing of cloud height and cloud motion in the instrument along track direction (direction of parallax), leading to a reduction in retrieval accuracy (Marchand et al., 2007; Hinkelman et al., 2009; Lonitz and Horváth, 2011). However, these appear to have been resolved in the latest version of the product (Mueller et al., 2013; Horváth, 2013).

The ATSR instruments represent an alternative that do not suffer from some or all of the limitations of MODIS and MISR. It is the case that alone (i.e. single operation) the instruments achieve wind measurement accuracy in the region of $8\text{-}10\text{ms}^{-1}$, and can only observe AMVs in the instrument across track direction (see section 6.3). These limitations occur due to the short time separation between the views (~ 2 minutes) and the dual view stereo arrangement. However, these confines are surmountable through a tandem ATSR sensor configuration, as is demonstrable for the period June 2002 through July 2003, when the ATSR2 and AATSR instruments shared near coincident following orbits. The tandem configuration enabled observations of the same location with the same viewing geometry at a temporal separation of ~ 30 minutes. The shared viewing geometry is critical, as any displacements measured between acquisitions are caused by motion within the scene, removing any parallax effects (as experienced with MODIS) or aliasing between parallaxes from different sources, such as motion and elevation (as with MISR). Furthermore, the height assignment can be achieved using stereo photogrammetric methods applied to the stereo pair from one of the imaging systems, enabling effective height assignment irrespective of the radiometric conditions under observations. The ATSR instruments

are also able to observe in the thermal region of the EMS, enabling observations of diurnal cycles and year round sampling in Polar Regions. The ATSR tandem operation therefore provides an excellent demonstration of the benefits of a tandem stereo capable satellite system for deriving AMVs from cloud tracking.

This chapter explores these benefits in detail. Following an outline of the scientific objectives in Section 6.2, AMV extraction from the ATSR tandem system is discussed in Section 6.3. This discussion includes accuracy considerations, cloud tracking algorithm selection, and conversion of the retrieved AMVs into meteorological zonal and meridional wind components. In Section 6.4, a ground based accuracy assessment of the derived AMVs is undertaken in order to evaluate the performance potential of the cloud tracking algorithm when applied to the ATSR tandem imaging system. In Section 6.5 an inter-comparison is made between the ATSR tandem meteorological wind components against collocated components from the state of the art MERRA reanalysis. Finally Section 6.6 provides a summary and recommendations for future work.

6.2 Scientific Objectives

The following scientific objectives have been identified for exploration in this chapter:

- *Explore the benefits of the ATSR tandem satellite imaging system for the determination of tropospheric AMVs through cloud tracking.*
- *Locate an algorithm that can effectively generate AMVs from the tandem imagery, preferably at pixel level resolution.*
- *Assess the accuracy of the AMV retrieval algorithm when applied to the tandem imagery through a ground displacement analysis.*
- *Compare the derived ATSR-2 AATSR meteorological wind components against reanalysis outputs from a state-of-the-art reanalysis, in the case MERRA, to provide an inter-comparison between observational and reanalysis winds.*

6.3 ATSR-2 AATSR AMV Extraction

6.3.1 Benefits of a Tandem System

In this section, prior to reviewing the benefits of a tandem satellite AMV retrieval system, the drawbacks of a single system will be reviewed. A single ATSR instrument has a number of weaknesses when it comes to determination of AMVs. The dual view configuration enables displacement determination in the instrument across track direction only. In the along track direction any observed displacements in the cloud features between the image acquisitions are a scrambled combination of parallax effects (caused by the cloud elevation and the different viewing positions) and atmospheric motion. Separation of these two

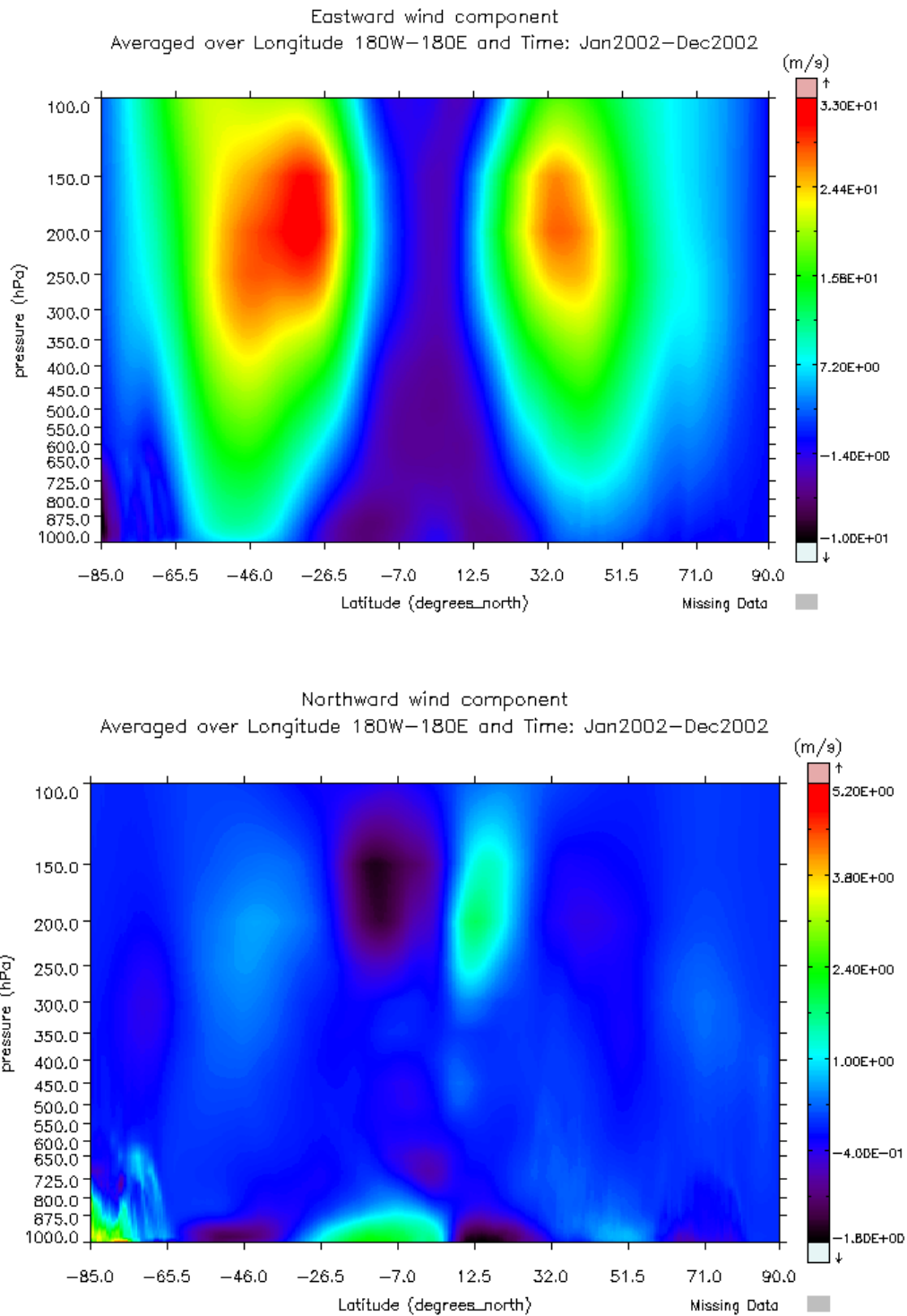


Figure 6-1. These plots present the MERRA zonal and meridional latitudinally averaged wind speeds between 1000 and 100 hPa for the period January to December 2002. Note that in both plots the mean wind speeds do not exceed $\pm 40 \text{ ms}^{-1}$. The plots were generated using the Giovanni tool (<http://disc.sci.gsfc.nasa.gov/giovanni>) and are cropped to only the data above 85°S due to missing data in MERRA over Antarctica.

components is impossible without prior knowledge of the CTH. The other main limitation associated with a single ATSR instrument is the short time lapse between image acquisitions, which is at most two minutes. Given the imaging resolution, nominally 1km, and the acquisition time lapse, it is possible to work out the achievable AMV quantisation with a pixel level cloud tracking algorithm as $a = (p * r)/t$, where a is the obtainable quantisation, p is the accuracy of the tracking algorithm in pixels, r is the resolution of the instrument in metres and t is the time separation between pixels in seconds. At the sub-satellite point, which has the longest time separation for ATSR, the hypothetical observational quantisation achievable for a single instrument is $\sim 8\text{ms}^{-1}$, towards the edges of the swath, which have a reduced time separation between acquisitions, the quantisation declines to $\sim 10\text{ms}^{-1}$. A study on mean global zonal and meridional wind speeds from the MERRA reanalysis, shown in Figure 6-1, demonstrates that given the achievable quantisation for across track winds, a single ATSR instrument provides little useful information on AMVs.

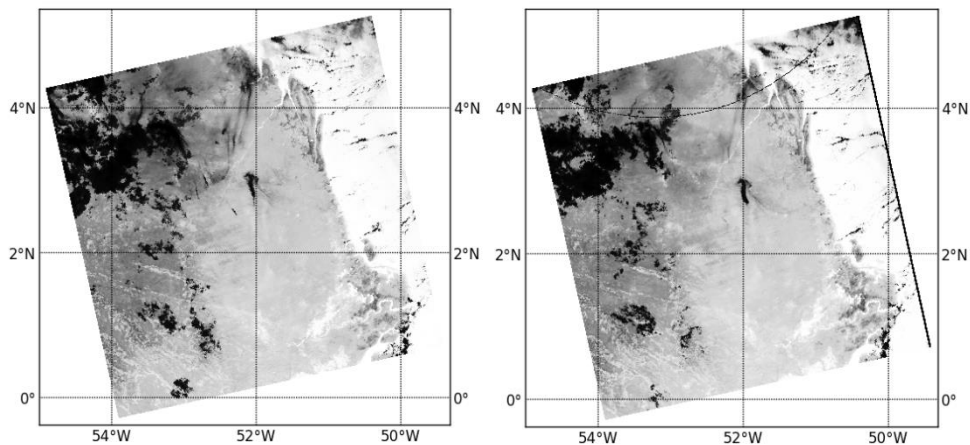


Figure 6-2. The left image is taken from the AATSR 11µm nadir channel, the right image is the same scene obtained from ATSR-2 11µm nadir channel 30 minutes later. Note the substantial changes in the cloud features (dark regions) between acquisitions, particularly in the western features.

The limitations encountered with a single ATSR instrument can be effectively overcome through tandem satellite operation, as suggested by Lorenz (1983). For the ATSR-2 AATSR tandem operation, employing either the two nadir or two forward views, the same location is observed from the same viewing angles with a time separation of approximately 30 minutes. As the same viewing angles are employed in either acquisition pair, any parallax effects are removed and all observed displacements observed are caused solely by the motion of features within the scene (Baker et al., 2011). This enables effective determination of both across and along track displacement measurements, and in turn AMVs. Furthermore, the increased time separation between the observations and the replication of the imaging geometry allows for a significant improvement in measurement quantisation and also leads to consistent observation conditions across the swath. Assuming a pixel level cloud tracking algorithm the 30 minutes between acquisitions improves the achievable quantisation to 0.5 ms^{-1} , allowing significant potential for improvement in accuracy and the usefulness of the AMVs for application in scientific study.

However, the increased time separation between acquisitions is not without drawbacks. As Figure 6-2 demonstrates, during the 30 minute time lapse between acquisitions there can be extensive changes to the structure of the cloud features within the scene. Such structural changes can confuse the cloud tracking algorithm employed, degrading the quality of the derived AMVs. Effectively accounting for such structural changes is critical for reliable AMV estimation as will become apparent in the next section.

6.3.2 Tandem AMV Retrieval Method

6.3.2.1 Cloud Tracking

The ATSR-2 and AATSR 11 μ m (selected to provide diurnal coverage) nadir views are used to derive the AMVs. The nadir view is chosen for its reduced sensitivity to optically thin cloud features. In the forward view, due to the longer atmospheric path length, these features are more apparent. This can be problematic as such features have a greater tendency to dissipate between observations leading to confusion in the applied tracking algorithm and erroneous measurements being returned.

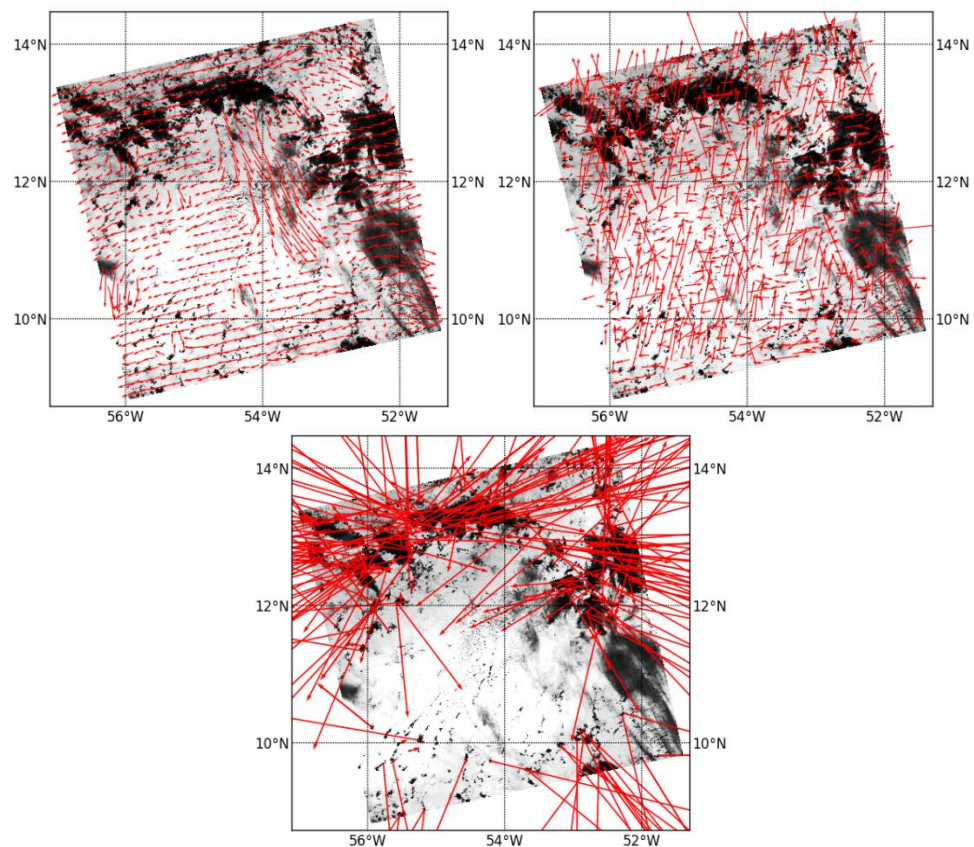


Figure 6-3. The top left image contains the AMV outputs from the Farneback algorithm applied to the ATSR tandem 11 μ m nadir views. The algorithm appears to better capture the atmospheric motion in the scene. In contrast the Census stereo algorithm outputs, the top right image, and the SURF algorithm outputs, the bottom image, are doing a very poor job of capturing the atmospheric motion within the scene.

Three different algorithms were tested for cloud tracking efficacy, the census stereo matcher, introduced in Chapter 3, the SIFT algorithm introduced in Chapter 2, and the Farneback optical flow algorithm which estimates displacement fields through analysis of polynomial expansion coefficients (Farneback, 2003) available in the OpenCV computer vision library (Bradski and Kaehler, 2008). Of the three algorithms evaluated the tracking achieved by the Farneback algorithm was found to correspond best both in terms of realistic AMV magnitude and direction. The Census stereo algorithm, whilst effective in scenes with limited distortion was unable to deal with the significant changes in cloud structure that occurs in the 30 minute delay between the ATSR-2 and AATSR acquisitions (similar outcomes were obtained when applying SAD, SSD and rank area based stereo matching algorithms). The same performance issues occurred when applying the SIFT feature detector, it being unable to deal with the structural changes and failing to locate matches robustly. The vector outcomes for all three algorithms are shown in Figure 6-3.

6.3.2.2 Component Conversion

The outputs from the Farneback routine are sub-pixel across and along track displacements. These displacements are in the instrument coordinate system and as such are not representative of meridional and zonal wind components; rather each displacement is a mixed function of both meteorological components. This mixing is due to the orbital inclination of the satellite, which for the ERS-2 and Envisat satellites was 98.5 degrees (Diekmann et al., 2010). To transform the observed displacements into meteorological wind components the underlying AATSR geographic grid is employed. The latitude/longitude position associated with each AATSR pixel is assumed as the cloud location at the initial observation time (i.e. the AATSR observation time). The across and along track displacements output by the optical flow algorithm are used to determine updated latitude / longitude position, with the geographic values again being taken from the AATSR geographic grid (any displacement which projects to outside the AATSR grid subset is rejected). The updated latitude / longitude positions represent the cloud positions at the second observation time (i.e. the ATSR-2 acquisition time). This process can be described as follows,

$$\begin{aligned}\phi_{i,j}^2 &= \phi^1[x_{i,j}, y_{i,j}] \\ \lambda_{i,j}^2 &= \lambda^1[x_{i,j}, y_{i,j}]\end{aligned}\tag{Eq. 6.1}$$

where, ϕ^1 and λ^1 are the AATSR latitude and longitude grids, x and y are the across and along track displacements, i and j are the scene pixel coordinates and ϕ^2 and λ^2 are the shifted latitude and longitude positions. Once derived the geographic coordinate pairings can be used to determine distance and bearing observations. The great circle distance between two geographic points is determined using the Haversine equation,

$$d = 2r \sin^{-1} \left(\sqrt{\sin^2 \left(\frac{\phi_{i,j}^2 - \phi_{i,j}^1}{2} \right) + \cos(\phi_{i,j}^2) \cos(\phi_{i,j}^1) \sin^2 \left(\frac{\lambda_{i,j}^2 - \lambda_{i,j}^1}{2} \right)} \right) \quad \text{Eq. 6.2}$$

where all angular measurements are in radians, r is 6371 km, and d is the distance between the two points in kilometres. The bearing between the two points is determined as follows,

$$\theta = \text{atan2} \left(\sin(\lambda_{i,j}^2 - \lambda_{i,j}^1) \times \cos(\phi_{i,j}^2), \cos(\phi_{i,j}^1) \times \sin(\phi_{i,j}^2) - \sin(\phi_{i,j}^1) \times \cos(\phi_{i,j}^2) \times \cos(\lambda_{i,j}^2 - \lambda_{i,j}^1) \right) \quad \text{Eq. 6.1}$$

where θ is the counter clockwise angle in radians between the heading at the initial location and geographic North. The distance is converted into speed in ms^{-1} by converting the distance between the points into metres and then dividing by the approximate time between acquisitions (1800 seconds). The speed and bearing effectively describe the AMV for each pixel. The AMVs can be converted to the zonal and meridional wind components typically output by NWP systems, with the following equations,

$$\begin{aligned} u &= -s \times \sin \theta \\ v &= -s \times \cos \theta \end{aligned} \quad \text{Eq. 6.1}$$

where s is the speed in metres and u and v are the zonal and meridional wind components.

6.3.2.3 Cloud Top Height Determination

In addition to zonal and meridional wind speed estimation, the processing chain also determines the CTH from the AATSR $11\mu\text{m}$ channel using the sub-pixel census algorithm described in Chapter 3. The $11\mu\text{m}$ AATSR channel is selected as it has been analysed extensively in previous chapters and has been shown to provide the best census stereo outcome. In addition it also provides observations irrespective of the illumination conditions. A search radius of 20 pixels is employed in the along track direction and 5 pixels in the across track direction. The along track radius is selected as the Tropopause in the tropics can often reach elevations approaching $\sim 17\text{km}$ (Seidel et al., 2001), and wind speeds rarely exceed 40ms^{-1} (see Figure 6-1). The along-track disparities are smoothed with a 7 by 7 pixel median filter to reduce noise and then converted to above ellipsoid CTH using the Mannstein camera model.

6.3.2.4 Other Considerations

Other considerations include the poor geo-referencing accuracy of the ATSR-2 instrument post 2000, when an on-board gyroscope failed leading to a decline in navigation accuracy (Accica and Goryl, 2002). In some instances this can lead to large geo-location errors (Seiz, 2003). Such errors have to be checked for in the outputs, as they may introduce measurement biases into the AMVs. A suitable checking process would be achieved through application of the census stereo cloud mask to the wind components. From this masked dataset a histogram of the cloud free wind components with bins of 1 pixel size could be created. It is then a simple process of rejecting any tandem orbit pairs which do not have the majority of cloud free wind vectors in the ± 1 pixel bins. Any scenes found to be poorly co-registered could be corrected for through application of the SURF algorithm presented in Chapter 2 (see also Fisher and

Muller, 2013). However, due to the single orbit pair processed in this study, the geo-referencing quality was assessed manually by visually examining displacements vectors for cloud free regions of all processed scenes. This histogram/SURF method would be employed for more extensive processing.

6.3.3 ATSR Tandem Wind Dataset

To test the AMV extraction algorithm ATSR-2 orbit 38977 and AATSR orbit 3105 from the 4th of October 2002 were selected. In the processing chain the 11 μ m nadir view for each orbit is extracted and separated into a series of 512 by 512 pixel chunks. Each chunk pairing is processed with the Farneback optical flow algorithm. The resultant displacements are converted into AMVs and then zonal and meridional components. The CTHs are computed from the AATSR nadir and forward view for each chunk using the census stereo algorithm. The u , v and CTH tiles are then stitched back into a full orbit and written out to netCDF along with a collocated DEM subset from GMTED2010.

6.4 AMV Accuracy Assessment

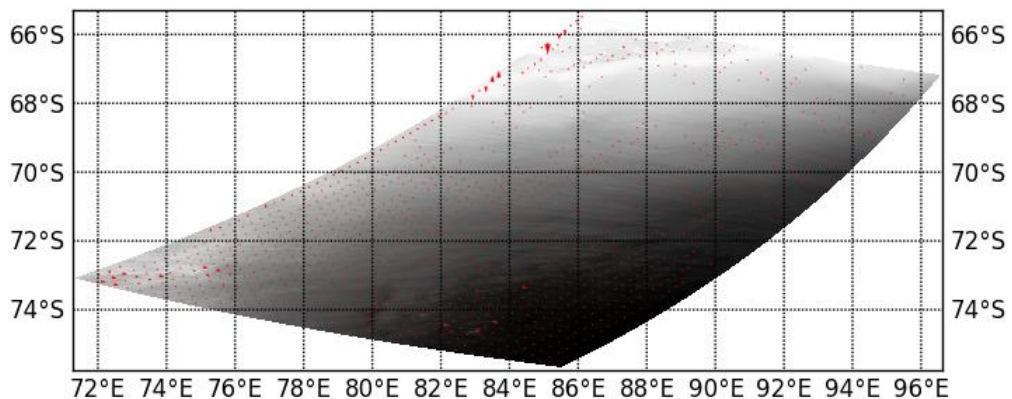


Figure 6-4. Extracted AMVs for a cloud free scene over Antarctica assessing the zero wind accuracy of the Farneback algorithm.

In order to assess the zero wind accuracy achievable by the Farneback algorithm a suitable approach is to evaluate the displacements returned for features where the expected displacement value is zero, i.e. the ground. In the 30 minutes between the tandem acquisitions, there should be no appreciable motion in the ground features, thereby providing this evaluable limit. A cloud free region over Antarctica from the processed tandem orbits was selected for evaluation and is shown in Figure 6-4. The histograms in Figure 6-5 present the across and along track shifts determined by the Farneback algorithm, in nearly all cases the observed shifts are less than a pixel. Statistical assessment of the displacements converted into wind components returns a RMSD of 0.37 ms^{-1} and a bias of 0.24 ms^{-1} for the zonal component. Similar quality is shown for the meridional component with an RMSD of 0.2 ms^{-1} and a bias of -0.14 ms^{-1} . These accuracies are below the expected accuracy achievable by the ATSR tandem operation due to the sub-

pixel displacement estimation capability of the Farneback algorithm. Furthermore, the algorithm returns the wind component estimations at the full instrument spatial resolution. In summary the performance of the Farneback algorithm in terms of accuracy and coverage has been demonstrated to outperform the algorithms currently applied operationally to the MISR and MODIS instruments.

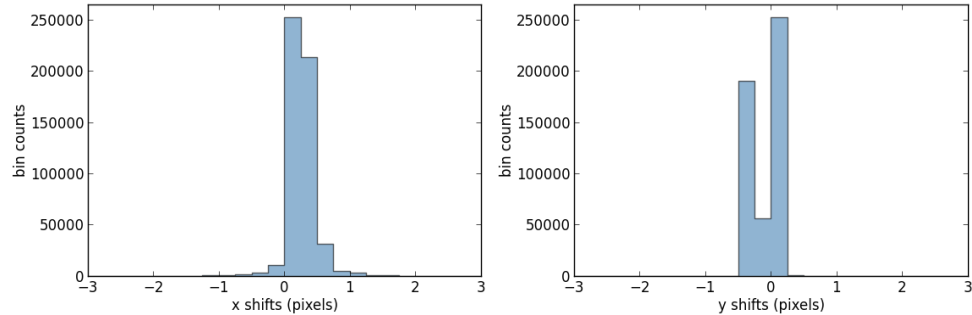


Figure 6-5. The left figure shows the histogram of the across track displacements obtained from the Farneback AMVs derived in Figure 6-4. The right figure shows the histogram of the along track displacements for the same figure. The bin size for both histograms is 0.25 pixels. Note nearly all observed displacements are accurate to less than 1 pixel, indicating an AMV accuracy of better 0.5 ms^{-1} .

6.5 Reanalysis Inter-comparison

6.5.1 MERRA Dataset

The MERRA reanalysis (Rienecker et al., 2011), is employed here to provide a suitable comparison dataset for the wind components. The MERRA reanalysis assimilates a large number of conventional and satellite based wind observations, including cloud and water vapour tracked winds from both geostationary observations and also polar observations from MODIS from 2002 onwards. For the inter-comparison the 12 a.m. UTC timestamp dataset from the 4th of October 2002 is used. This time stamp is selected as it is closest to that of the AATSR and ATSR-2 orbits, ~ 2 a.m. UTC.

6.5.2 Inter-comparison Method

6.5.2.1 Collocation

The MERRA reanalysis datasets employed in the inter-comparison are the zonal and meridional wind components, and the geo-potential heights. To facilitate inter-comparison the MERRA data is spatially collocated with that of the AATSR data using the BEAM tool. Following spatial collocation, collocation of pressure levels is also required. Making the assumption that the observed wind components are at the elevation defined by the stereo matched AATSR CTH output, it is possible to extract the reanalysis wind components at the collocated pressure level using the reanalysis geo-potential height data, i.e. the model

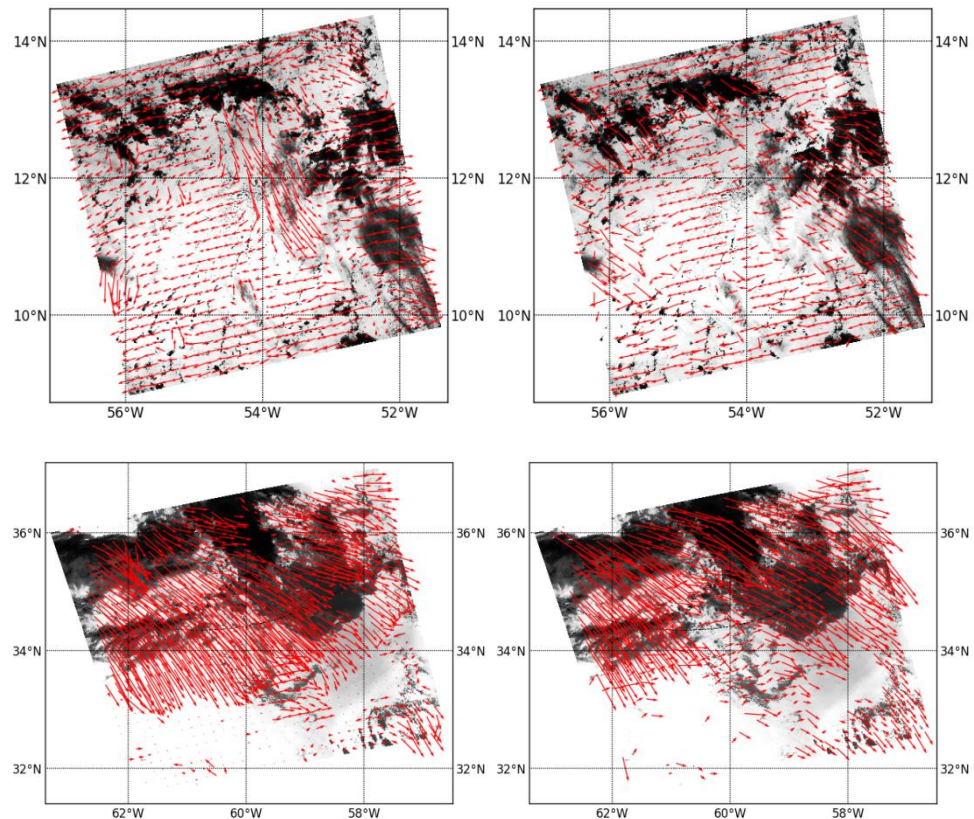
pressure level which minimises the difference between the CTH and geo-potential height is assumed to be collocated, and the reanalysis wind components associated with this pressure level correspond to those observed. Carrying out this process for every pixel provides a MERRA dataset which is collocated with AATSR in both spatial and model level terms.

6.5.2.2 Masking

A mask is generated for the evaluation so that only those locations where the ATSR-2 AATSR tandem analysis observed winds are employed in the statistical analysis. The masking process is straightforward using wind speed and wind elevation constraints. The wind speed threshold is set at 0.5 ms^{-1} , i.e. any observed wind vector which has an absolute value less than this is potentially noise and excluded from the statistical analysis. This threshold was chosen as it defines the quantisation achievable by the ATSR-2 AATSR tandem configuration. The elevation threshold was determined by differencing the AATSR CTH observations with the collocated DEM. Any CTH feature less than 1km above the terrain was excluded, as below this elevation the CTH features may, in fact, be terrain. The derived mask was applied to both the ATSR-2 AATSR and the collocated MERRA wind components.

6.5.3 Results

6.5.3.1 Visual



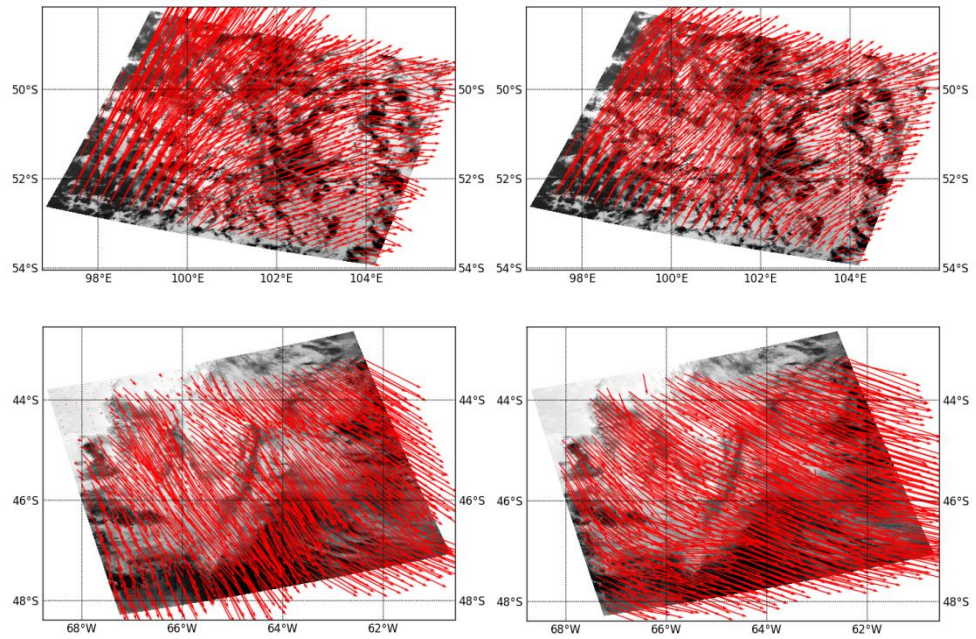


Figure 6-6. The left column presents the AMVs returned by the ATSR tandem instrument in combination with the Farneback algorithm. The right column shows the collocated MERRA reanalysis outputs

6.5.3.2 Statistical

Prior to the statistical analyses presented in this section an outlier removal process was carried out to improve the robustness of the assessment. Outliers were classed as any wind component pairing where the difference between the reanalysis and observational components was more than ± 3 standard deviations from the mean.

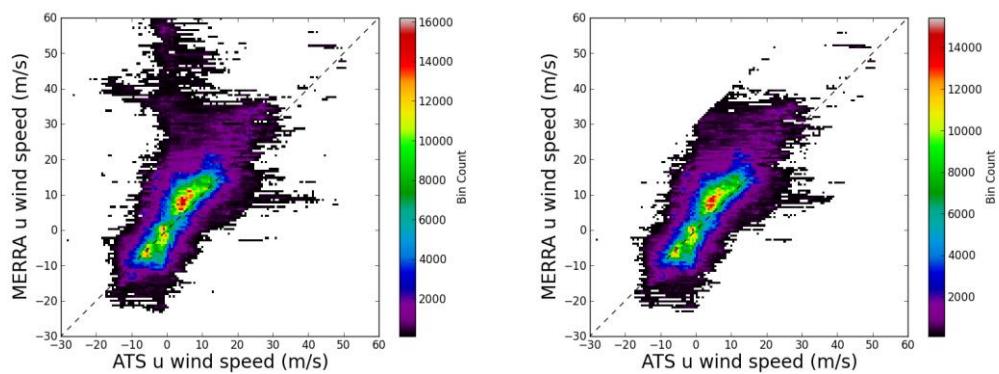


Figure 6-7. The dual histogram of zonal winds speeds as returned by the ATSR tandem observations on the x-axis and the collocated MERRA reanalysis outputs on the y-axis. The left plots is for all collocated observations, the right is with outliers excluded.

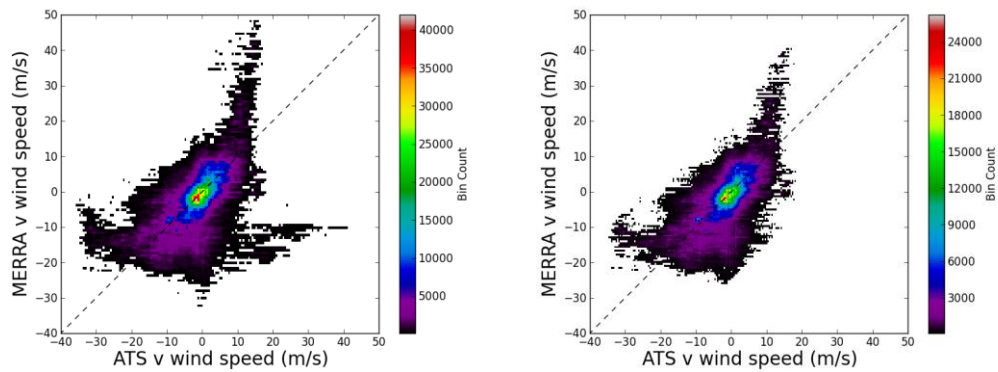


Figure 6-8 The dual histogram of meridional winds speeds as returned by the ATSR tandem observations on the x-axis and the collocated MERRA reanalysis outputs on the y-axis. The left plots is for all collocated observations, the right is with outliers excluded.

6.5.4 Discussion

In the scenes used in the visual assessment, shown in Figure 6-6, the observations and reanalysis show good general agreement in terms of magnitude and direction. The main discrepancies appear to occur in the presence of multi-layer cloud situations with differing wind characteristics. In the statistical analysis derived from the entire orbit the reanalysis and observations again show some good agreement. Following outlier exclusion the zonal wind speed analysis, shown in Figure 6-7, has an RMSE of 7.7 ms^{-1} , this is similar to the error reported in the tracked AMVs (MODIS, GOES: accuracy of $\sim 8 \text{ ms}^{-1}$) assimilated into the MERRA reanalysis, which is encouraging. The bias is only -2.5 ms^{-1} , indicating a slight westward bias in the reanalysis data. The coefficient of determination is 0.58, indicating that 58% of the variance in the observational data is explained by the reanalysis. The meridional wind speed analysis has a reduced RMSE of 6.6 ms^{-1} , again similar to the error reported in the assimilated winds. The bias is reduced at -0.5 ms^{-1} , indicating a bias within the range of the tandem quantisation, which is very encouraging. The coefficient of determination is reduced at 0.42, indicating that 42% of the variance in the observational data is explained by the reanalysis. In summary, considering the inherent differences between the reanalysis and the observations (e.g. operating resolution, operational basis, temporal dislocation, etc.) the results are encouraging.

6.6 Summary and Future Work

6.6.1 Objective Review

As with the previous scientific results chapters, this section assesses the results against scientific objectives outlined in section 6.2 in terms of their successful completion, downsides, and insights gained.

‘Explore the benefits of a tandem satellite imaging system for the determination of tropospheric wind speed, direction and elevation through cloud tracking’.

Two benefits of a tandem ATSR operation over a single satellite system have been identified: the ability to derive both across and along track winds effectively through near identical imaging geometry removing any parallax effects; and, due to the increased time separation between orbits, the ability to observe wind speeds at a significantly improved quantisation (from 8 ms^{-1} to 0.5 ms^{-1}).

‘Locate and apply suitable algorithm to track winds for the ATSR instruments.’

The Farneback optical flow algorithm has been shown to be an excellent technique for the derivation of AMVs from satellite imagery. It is far more robust to the extreme cloud deformations which occur during the 30 minute time lapse between the ATSR image acquisitions, than either the census algorithm or SURF. Furthermore, the algorithm provides a dense field of observations (estimates are generated for every image pixel). This is currently unprecedented, with other observational satellites requiring aggregation in order to generate robust results, therefore providing a wind vector grid at a resolution less than native (e.g. MISR wind components are currently returned at on a $17.6 \text{ by } 17.6 \text{ km}^2$ grid; MODIS wind components are returned on a $26 \text{ by } 26 \text{ km}^2$ grid). This is of importance, as such high resolution AMVs can be used to effectively investigate cloud dynamical processes (e.g. Wu et al., 2010)

‘Asses the accuracy of the AMV retrieval algorithm when applied to the tandem imagery through a ground displacement, zero wind, analysis.’

The AMVs have been assessed using a cloud free scene, where, under the assumption that the scenes are perfectly co-registered, an ideal algorithm would return mean displacements of zero in both axis. The Farneback algorithm, under this evaluation, demonstrates excellent performance, with mean displacements and biases of less than the 0.5 ms^{-1} quantisation achievable by the ATSR tandem operation.

‘Compare the derived ATSR-2 AATSR tandem winds against reanalysis outputs from state-of-the-art reanalyses, in this case MERRA, to provide an initial inter-comparison between observational and reanalysis winds.’

In a visual inter-comparison the observational and reanalysis wind components have been shown to return similar wind vectors. A statistical analysis for an entire tandem orbit was also carried out against the MERRA data. This assessment demonstrated that a good degree of agreement exists between the observational and reanalysis datasets. This is to be expected as the reanalysis outputs are formed from an observational basis, assimilating wind data from both polar orbiting and geostationary imaging systems.

6.6.2 Future Work

From the objective review it is possible to outline a number of avenues to pursue in future studies. Perhaps the most critical is to extend the ground analysis which assesses the performance of the Farneback algorithm. The current evaluation, whilst very promising, is limited in extent. Complete processing of the ATSR-2 AATSR tandem data would facilitate this extension.

Further algorithm developments could also be implemented. Of most importance is an AMV quality control check, something which both the MISR and MODIS products employ. A quality control could be implemented effectively through inter-comparison of the AMVs derived from the nadir image pairing against those from the forward. Any AMVs which differ between the pairings by more than some predefined threshold could then be rejected.

Another potential development is the determination of vertical cloud motion through observation of the change in CTH between AATSR and ATSR-2 acquisitions, providing a complete description (i.e. 3-axis) of cloud motion.

The last obvious path for future work is to assimilate the validated tandem observational data into a current reanalysis to assess whether it leads to improvements in the forecasting ability of the NWP model, along similar lines as that undertaken by (Bormann and Thépaut, 2004). The use of the tandem method is likely improve upon the increased forecasting ability in Polar Regions provided by the assimilation of MODIS data due to the more robust stereo height assignment achievable by stereo in these challenging locations (see Chapter 4).

6.6.3 Chapter Summary

This chapter has demonstrated the potential for deriving tropospheric AMVs from the ATSR-2 AATSR tandem operation. A highly suitable optical flow algorithm, developed by Farneback, has been identified for generating dense cloud tracked displacements across the ATSR thermal channels. The outputs from the algorithm/instrument combination have been shown, though a limited ground displacement assessment, to outperform the MODIS AMVs in terms of accuracy and to be similar to those of MISR. A further, significant, benefit of the Farneback algorithm is that the observed displacements appear to be robust at full ATSR instrument resolution, a feat which neither the MODIS or MISR algorithms are capable of achieving. An inter-comparison against the MERRA reanalysis has also been undertaken and good agreement between the observational and reanalysis dataset has been demonstrated.

7 CONCLUDING REMARKS

7.1 Thesis Review

The overall aim of this thesis has been to apply stereo-photogrammetric techniques to the ATSR instruments in order to study their potential for generating various scientific datasets for application within the field of the atmospheric sciences.

The first development to come out of this study has implications not only for stereo-matching but any other science tasks undertaken with the ATSR instruments that require the forward and nadir views to be accurately co-registered. The poor co-registration between the ATSR images is present across all three instruments, and in terms of stereo matching, leads to significant biases in the returned stereo results (a 1 pixel error in co-registration results in ~1km error in the retrieved CTH). The technique presented in this thesis to improve the image co-registration employs an automated tie-point detection algorithm based on the SURF feature matcher. This approach has been shown to perform very robustly for the detection and matching of features in ATSR imagery. The tie-points, once defined, are employed to derive polynomial warping coefficients which map the forward view onto the nadir. The warping coefficients have been evaluated globally and shown to improve the co-registration between the forward and nadir views to less than a pixel for all ATSR instruments (excepting ATSR-2 post 2000). Currently the warping coefficients are derived for application to 512 by 512 pixel image subsets (due in part to the sub-setting employed in the stereo processing undertaken in the science chapters).

The second development has been the location of a far more suitable stereo matching approach for the determination of CTH from AATSR. The algorithm located, the census transform, is a non-parametric area-based matching algorithm, and this lends it a number of advantages over the more traditional parametric area-based approaches. It is more robust to pixel outliers, that is, pixels with digital values which do not conform to the statistical distribution of the image subset being stereo matched (or normalised). Parametric stereo matching techniques experience a degradation of performance in the presence of such outliers. Effects introduced include smoothing of the disparity fields and erroneous matches at discontinuities (transitions between different disparity groups; often associated with a change in pixel intensity). Due to the robustness of non-parametric transforms to statistical outliers, the quality of the stereo matching outcome is far improved in the presence of discontinuities. This is very important for the stereo matching of cloud features, as they tend to exhibit discontinuities with the underlying land surface. Furthermore non-parametric transforms are very robust to radiometric changes an important

aspect for all stereo algorithms employed in the variable imaging conditions common when employing EO platforms. In nearly all analyses presented in the thesis, the census transform has clearly been the most effective when compared against the other stereo image matching algorithms employed on ATSR.

Three different science tasks have been undertaken using stereo-photogrammetric techniques in this thesis. The principle undertaking has been the development of a new stereo derived CTH and CF climatology over Greenland from the entire AATSR time series. This climatology in the form of CFbA dataset provides stratified estimates of cloud fraction at various elevations above the Earth ellipsoid. This new dataset provides CTH estimates vertically accurate to ~1km, and detects clouds down to ~0.4 optical depths. Furthermore, the CTH retrievals from AATSR stereo have been shown to exhibit much reduced biases (vs. lidar observations) compared to other passive non-geometric CTH assignment approaches. This is particularly the case for the CTH outputs from the 11 μ m channel which have been shown, irrespective of the clouds elevation (see Chapter 4, Figure 4-8), to be practically unbiased compared to lidar retrievals. The stereo climatology has been demonstrated to agree well with other observational climatologies over Greenland, and to provide valuable new insights into cloud characteristics, particularly in the case of high clouds.

The second science task undertaken involved the development and validation of a new SPIH dataset from AATSR for application within the ALANIS Smoke Plumes study. The author's main contribution was the validation of the M6 algorithm and the smoke plume masking routine developed at MSSL (descriptions of both of these processes are given within the thesis). The validation efforts demonstrate that AATSR is effective for retrieving the elevation of large plumes associated with significant forest fire events, and also extensive smoke clouds to ~1km accuracy. Smoke plumes with smaller extents and lower injection heights are typically missed by the masking algorithm and the heights retrieved tend to relate to the terrain rather than the plume.

The final science task undertaken involved the determination of wind speed, direction and elevation through cloud feature tracking across imagery obtained from the ATSR-2 AATSR tandem operation. This study demonstrated the benefits of employing a tandem satellite system, in particular, a substantial improvement in the accuracy of the retrieved winds (for 8 ms⁻¹ to 0.5 ms⁻¹) and also the potential to retrieve both across and along track wind components (vs. across-track only wind components in the case of a single ATSR sensor). Furthermore an optical flow algorithm (Farneback, 2003) which enables determination of wind fields at full ATSR image resolution has been identified. This is unprecedented with the all current feature tracking algorithms employed on other passive observational wind determination systems requiring pixel aggregation (i.e. resolution reduction) in order to obtain reliable results.

7.2 Contributions and Originality

7.2.1 Thesis Contributions

- The first major contribution of this thesis is an automated approach to improve the co-registration between the forward and nadir AATSR views. A set of globally applicable warping coefficients have been derived and evaluated and shown to improve the co-registration between the views of the ATSR instruments (excluding ATSR-2 post 2000) to pixel level accuracy irrespective of location. This development is potentially significant for any science task which requires the ATSR views to be accurately co-registered, such as stereo matching.
- The non-parametric census transform (Zabih and Woodfill, 1994) stereo image matching algorithm has been applied to the AATSR instrument to retrieve disparity estimates. This non-parametric approach has been demonstrated through multiple assessments (against DEM elevations, and collocated CTH retrievals from CALIOP and MISR) to provide very robust stereo measurements and leads significant accuracy gains over the previous stereo algorithms applied to the ATSR instruments.
- A long-term, stereo derived, macrophysical cloud climatology from AATSR over Greenland has been developed. This unique dataset provides the first reliable estimates of CF for an extended time series from an ATSR instrument over any location (AATSR-GRAPE CF observation tend to be unreliable; the previous CF dataset generated over Greenland from ATSR-2 using stereo is only available for 2007 (Griggs and Bamber, 2008)). Furthermore, this dataset represents the first stereo derived CFBA climatology from AATSR for any location. The developed climatology also offers insights into both high and low clouds characteristics, which is a unique feature.
- For the first time an inter-comparison has been made of AATSR stereo derived CTH retrievals against collocated observations from CALIOP. This inter-comparison, using 5 months' worth of finely collocated samples (~5km spatial separation; <10 minute temporal separation) over Greenland demonstrated firstly that the non-parametric census algorithm performs excellently for CTH retrieval when combined with AATSR, and secondly that the biases in the retrieved stereo results are strongly channel dependent. The obtained results demonstrate that the 11 μ m channel exhibits far less bias than either the NIR or visible channels and on average retrieves results that are within 45m of the cloud top as detected by CALIOP. This accuracy is unprecedented, with the majority of other passive remote sensing system tending to retrieve CTHs that are 1km or more below those seen by CALIOP. Further work is required assessing whether this performance is consistent in other regions globally. However, this is potentially a very interesting result in terms of the value of stereo-photogrammetric techniques applied to thermal imagery.
- The inter-comparison of the AATSR stereo CTH retrievals with CALIOP has also enabled a characterisation of the detection limits of the instrument-algorithm pairing. AATSR in

combination with stereo has been shown on average to fail to detect clouds with an optical depth of 0.4. This finding is very similar to the detection limits achievable by other passive CTH determination systems.

- In addition to the CTH and CF studies this thesis has also demonstrated the potential to retrieve smoke plume and smoke cloud heights from AATSR using stereo techniques applied to the 0.55 μ m channel. This undertaking, whilst less successful than the CTH and CF determination studies, has demonstrated that there is some potential for determining smoke feature characteristics through AATSR and stereo photogrammetry, as long as the features are larger than a few tens of pixels and above 1km in elevation.
- The final major contribution from this thesis has been the demonstration of wind speed, elevation and direction retrieval from the tandem ATSR-2 AATSR operation. This study provides the first evaluation of a tandem passive sensing satellite approach for wind retrievals and also provides the first retrievals from a passive sensing system at microscale resolution (≤ 1 km).

7.2.2 Peer Review Journal Papers

Fisher, D., & Muller, J. P. (2013). Global warping coefficients for improving ATSR co-registration. *Remote Sensing Letters*, 4(2), 151-160.

Fisher, D., Muller, J. P., & Yershov, V. N. (2013). Automated Stereo Retrieval of Smoke Plume Injection Heights and Retrieval of Smoke Plume Masks from AATSR and their Assessment With CALIPSO and MISR. *IEEE T. Geosci. Remote Sens.*, in press, 28717,28718

Krol, M., Peters, W., Hooghiemstra, P., George, M., Clerbaux, C., Hurtmans, D., ... & Muller, J. P. (2013). How much CO was emitted by the 2010 fires around Moscow? *Atmospheric Chemistry and Physics*, 13(9), 4737-4747.

7.2.3 Conference Proceedings

Muller, J. P., Walton, D. M., Fisher, D. N., & Cole, R. E. (2010). SMVs (Stereo Motion Vectors) from ATSR2-AATSR And MISRlite (Multi-Angle Infrared Stereo Radiometer) Constellation. In *IWW10-International Winds Workshop* (Vol. 10, p. 8pp).

Fisher, D., J-P. Muller, (2010). Stereo Motion Vectors from ATSR-2 and AATSR. *ESA Living Planet Symposium*. Bergen, Norway 28 June-2 July.

Fisher, D., J-P Muller, (2010). Improved Stereo Retrieval of Cloud-Top Heights and Winds from ATSR2 and AATSR: Techniques and Results. Remote Sensing and Photogrammetry Society Annual Conference. Cork, Ireland 1-3 September.

Fisher D., J-P Muller, V. Yershov., (2011). The Development of Novel Stereo Derived Smoke Plume Products for AATSR and their Application to the 2010 Russian Fire Season. EGU General Assembly. Vienna, Austria 3-8 April.

Muller J-P., V. Yershov, D. Fisher, 2012. Stereo retrievals of cloud and smoke winds and heights from EO platforms: past, present and future prospects. Tenth International National Winds Workshop. Auckland, New Zealand 20-24 February.

Fisher, D., Muller, J. P., & Yershov, V. (2012). The validation and analysis of novel stereo-derived smoke plume products from AATSR and their application to fire events from the 2008 Russian fire season. EGU General Assembly. Vienna, Austria 22-27 April.

Fisher, D., & Muller, J. P. (2012, July). Global analysis of the improvements in AATSR nadir-forward co-registration following the application of an automated registration algorithm. *In Geoscience and Remote Sensing Symposium (IGARSS)*, 2012 IEEE International (pp. 1753-1756). IEEE.

7.3 Future Work

Many of the new contributions presented in section 7.2 can be extended, and in this section the most promising avenues for future work are reviewed. The first obvious direction for future work is to extend and improve upon the stereo derived macrophysical cloud climatology presented in this thesis. The primary improvement required is in the effective screening of problematic, low texture regions of the Earth's surface, which are primarily associated with ice sheets. The current screening method implemented requires all seven AATSR and ATSR-2 channels to effectively screen low texture snow and ice regions, precluding CTH and CF determination during the night-time, and entirely preventing its application to ATSR-1 (which lacks the three visible channels). If an existing or new technique can be implemented to effectively screen problematic low texture regions using only the channel being stereo matched then it would be possible to effectively retrieve diurnal cloud cycles and winter cloud characteristic over ice sheets and other problematic regions which low texture in the ATSR imagery. Once a solution for low texture regions has been located and implemented in the current AATSR-Stereo L2 processing chain, the entire AATSR, ATSR-2 and ATSR-1 global time series could be effectively processed into L3 climatological datasets.

The next proposed avenue for future work is reprocessing of the entire ALANIS AATSR dataset with the census stereo algorithm. When inter-comparing the validation outcomes of census against M6, in every comparable test undertaken in this thesis census offers better performance by a substantial margin. The census algorithm was determined to be a suitable stereo matching algorithm for AATSR after the ALANIS project had been completed; hence it was not evaluated for the task of SPIH determination. However in the future, it should definitely be assessed for this task, as the outcomes are likely to be significantly better than those achieved through the application of M6.

Due to its excellent performance when applied to ATSR-2 AATSR tandem data, the final strongly recommended avenue for future work is an evaluation of the performance of the Farneback optical flow algorithm for cloud tracked wind retrieval with the MISR instrument. The current algorithm employed on MISR is only able to retrieve wind fields at a spatial resolution of 17.6km and an accuracy of 0.5 ms^{-1} . It is proposed that application of the Farneback optical flow algorithm to the MISR instrument would likely

lead to a significant improvement in the spatial resolution of the retrievals, possibly even down to 250m, whilst likely retaining similar or better accuracy levels.

References

- Accica, L. & Goryl, P. (2002). ATSR misregistration problems in mono-gyro mode. Technical note, EMCF-PCSO-EOAD-TN-02-0001, 21 January, ESRIN, ESA.
- Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., & Gumley, L. E. (1998). Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *103*(D24), 32141-32157.
- Ackerman, S., Holz, R., Frey, E. W., Eloranta, B. C., Maddux, and M. McGill, 2008: Cloud detection with MODIS. Part II: Validation. *J. Atmos. Oceanic Technol.*, *25*, 1073–1086.
- Adams, J. M., & Faure, H. (1998). A new estimate of changing carbon storage on land since the last glacial maximum, based on global land ecosystem reconstruction. *Global and Planetary Change*, *16*, 3-24.
- Allan, R. P. (2011). Combining satellite data and models to estimate cloud radiative effect at the surface and in the atmosphere. *Meteorological Applications*, *18*(3), 324-333.
- Ambach, W. (1974). The influence of cloudiness on the net radiation balance of a snow surface with high albedo. *J. Glaciol*, *13*(67), 73-84.
- Andrews, T., Gregory, J. M., Webb, M. J., & Taylor, K. E. (2012). Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters*, *39*(9).
- Ardanuy, P. E., Stowe, L. L., Gruber, A., & Weiss, M. (1991). Shortwave, longwave, and net cloud-radiative forcing as determined from Nimbus 7 observations. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *96*(D10), 18537-18549.
- Arino, O., Simon, M., Piccolini, I., & Rosaz, J. M. (2001, January). The ERS-2 ATSR-2 world fire atlas and the ERS-2 ATSR-2 world burnt surface atlas projects. In *Proc. 8th ISPRS Conference on Physical Measurements and Signatures in Remote Sensing, AUSSOIS* (pp. 8-12).
- Bailey, P. (1994). SADIST products (version 600).
- Baker, S., Scharstein, D., Lewis, J. P., Roth, S., Black, M. J., & Szeliski, R. (2011). A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, *92*(1), 1-31.
- Bamber, J. L., Ekholm, S., & Krabill, W. B. (2001). A new, high-resolution digital elevation model of Greenland fully validated with airborne laser altimeter data. *Journal of Geophysical Research: Solid Earth (1978–2012)*, *106*(B4), 6733-6745.
- Banks, J., Bennamoun, M., & Corke, P. (1997, December). Non-parametric techniques for fast and robust stereo matching. In *TENCON'97. IEEE Region 10 Annual Conference. Speech and Image Technologies for Computing and Telecommunications., Proceedings of IEEE* (Vol. 1, pp. 365-368). IEEE.
- Barnard, S. T., & Fischler, M. A. (1982). Computational stereo. *ACM Computing Surveys (CSUR)*, *14*(4), 553-572.

- Baum, B. A., & Wielicki, B. A. (1994). Cirrus cloud retrieval using infrared sounding data: Multilevel cloud errors. *Journal of Applied Meteorology*, 33(1), 107-117.
- Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110, 346-359
- Bhat, D. N., & Nayar, S. K. (1998). Ordinal measures for image correspondence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(4), 415-423.
- Bolles, R.C., Baker, H.H., and Hannah, M.J. 1993. The JISCT stereo evaluation. In DARPA Image Understanding Workshop, pp. 263–274.
- Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J. L., ... & Webb, M. J. (2006). How well do we understand and evaluate climate change feedback processes?. *Journal of Climate*, 19(15), 3445-3482.
- Bormann, N., & Thépaut, J. N. (2004). Impact of MODIS polar winds in ECMWF's 4DVAR data assimilation system. *Monthly weather review*, 132(4), 929-940.
- Boykov, Y., Veksler, O., & Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(11), 1222-1239.
- Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library*. O'reilly.
- Brockmann, C., Paperin, M., Danne, O., & Ruescas, A. (2013, Sept). Multi-Sensor Cloud Screening and Validation: IdePix and PixBox. In *ESA Living Planet Symposium Proceedings*. Edinburgh 9-13 September.
- Browning, K. A., & Wexler, R. (1968). The determination of kinematic properties of a wind field using Doppler radar. *Journal of Applied Meteorology*, 7(1), 105-113.
- Casadio, S. (2010). On (A)ATSR(-1/2) Co-Location. *ESA Technical Note*.
- Cawkwell, F. G., Bamber, J. L., & Muller, J. P. (2001). Determination of cloud top amount and altitude at high latitudes. *Geophysical research letters*, 28(9), 1675-1678.
- Cawkwell, F. G. L., & Bamber, J. L. (2002). The impact of cloud cover on the net radiation budget of the Greenland ice sheet. *Annals of Glaciology*, 34(1), 141-149.
- Cess, R. D. (1976). Climate change: An appraisal of atmospheric feedback mechanisms employing zonal climatology. *Journal of Atmospheric Sciences*, 33, 1831-1843
- Cess, R. D., Potter, G. L., Blanchet, J. P., Boer, G. J., Ghan, S. J., Kiehl, J. T., ... & Yagai, I. (1989). Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, 245(4917), 513-516.
- Cess, R. D., Potter, G. L., Blanchet, J. P., Boer, G. J., Del Genio, A. D., Deque, M., ... & Zhang, M. H. (1990). Intercomparison and interpretation of climate feedback processes in 19 atmospheric general

- circulation models. *Journal of Geophysical Research: Atmospheres* (1984–2012), 95(D10), 16601-16615.
- Cess, R. D., Zhang, M. H., Ingram, W. J., Potter, G. L., Alekseev, V., Barker, H. W., ... & Wetherald, R. T. (1996). Cloud feedback in atmospheric general circulation models: An update. *Journal of Geophysical Research: Atmospheres* (1984–2012), 101(D8), 12791-12794.
- Chander, G., Killough, B., & Gowda, S. (2010, July). An overview of the web-based Google Earth coincident imaging tool. In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International* (pp. 1679-1682). IEEE.
- Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J. L., Minnis, P., ... & Zeng, S. (2010). The GCM-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP). *Journal of Geophysical Research: Atmospheres* (1984–2012), 115(D4).
- Colarco, P. R., Schoeberl, M. R., Doddridge, B. G., Marufu, L. T., Torres, O., & Welton, E. J. (2004). Transport of smoke from Canadian forest fires to the surface near Washington, DC: Injection height, entrainment, and optical properties. *Journal of Geophysical Research: Atmospheres* (1984–2012), 109(D6).
- Colman, R. (2003). A comparison of climate feedbacks in general circulation models. *Climate Dynamics*, 20(7-8), 865-873.
- Comiso, J. C., Parkinson, C. L., Gersten, R., & Stock, L. (2008). Accelerated decline in the Arctic sea ice cover. *Geophysical Research Letters*, 35(1).
- Conard, S. G., & Davidenko, E. P. (1998, September). Fire in Siberian boreal forests—implications for global climate and air quality. In *Proceedings of the international symposium on air pollution and climate change effects on forest ecosystems. General Technical Report PSW-GTR-166. USDA Forest Service, Albany, California* (pp. 87-94).
- Corlett, G. K., et al. (2009). (A)ATSR Validation Activities: Validation Issues Report 3B. *Space ConneXions Contract: 2004-03-001/CPED10*
- Crutzen, P. J., Heidt, L. E., Krasnec, J. P., Pollock, W. H., & Seiler, W. (1979). Biomass burning as a source of atmospheric gases CO, H₂, N₂O, NO, CH₃Cl and COS. *Nature*, 282, 253-256.
- Curry, J. A., Schramm, J. L., & Ebert, E. E. (1993). Impact of clouds on the surface radiation balance of the Arctic Ocean. *Meteorology and Atmospheric Physics*, 51(3-4), 197-217.
- Danielson, J. J., & Gesch, D. B. (2011). Global multi-resolution terrain elevation data 2010 (GMTED2010). *US Geological Survey Open-File Report, 1073*, 26.
- Davies, R., Horváth, Á., Moroney, C., Zhang, B., & Zhu, Y. (2007). Cloud motion vectors from MISR using sub-pixel enhancements. *Remote sensing of environment*, 107(1), 194-199.
- Day, T., & Muller, J. P. (1989). Digital elevation model production by stereo-matching spot image-pairs: a comparison of algorithms. *Image and Vision Computing*, 7(2), 95-101.
- Denis, M. A., Muller, J. P., & Mannstein, H. (2007). ATSR-2 camera models for the automated stereo

photogrammetric retrieval of cloud-top heights—initial assessments. *International Journal of Remote Sensing*, 28(9), 1939-1955.

Derber, J. C., Parrish, D. F., & Lord, S. J. (1991). The new global operational analysis system at the National Meteorological Center. *Weather and Forecasting*, 6(4), 538-547.

Dessler, A. E., & Yang, P. (2003). The distribution of tropical thin cirrus clouds inferred from Terra MODIS data. *Journal of climate*, 16(8), 1241-1247.

Diekmann, F. J., Mesples, D., Bargellini, P. G., Nilsson, M., Ventimiglia, L., Canela, M., ... & Fehr, T. (2010). Extending the Envisat Mission—Impacts on Ground and Space Segment Operations. In *SpaceOps 2010 Conference, American Institute of Aeronautics and Space Administration* (Vol. 2257).

Di Girolamo, L., & Davies, R. (1994). A Band-Differenced Angular Signature technique for cirrus cloud detection. *Geoscience and Remote Sensing, IEEE Transactions on*, 32(4), 890-896.

Di Girolamo, L., & Davies, R. (1997). Cloud fraction errors caused by finite resolution measurements. *Journal of geophysical research*, 102(D2), 1739-1756.

Di Girolamo, L., & Wilson, M. J. (2003). A first look at band-differenced angular signatures for cloud detection from MISR. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(7), 1730-1734.

Di Girolamo, L., A. Menzies, G. Zhao, K. Mueller, C. Moroney, and D.J. Diner, 2010: MISR Level 3 Cloud Fraction by Altitude Theoretical Basis, JPL D-62358, Jet Propulsion Laboratory, Pasadena, CA, 24 pp.

Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., ... & Verstraete, M. M. (1998). Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. *Geoscience and Remote Sensing, IEEE Transactions on*, 36(4), 1072-1087.

Diner, D., Davies, R., Di Girolamo, L., Horvath, A., Moroney, C., Muller, J., ... & Zong, J. (1999). *MISR level 2 cloud detection and classification*. JPL Technical Report ATBD-MISR-07.

Dong, J., Kaufmann, R. K., Myneni, R. B., Tucker, C. J., Kauppi, P. E., Liski, J., ... & Hughes, M. K. (2003). Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. *Remote Sensing of Environment*, 84(3), 393-410.

Dowling, D. R., & Radke, L. F. (1990). A summary of the physical properties of cirrus clouds. *Journal of Applied Meteorology*, 29(9), 970-978.

Edwards, T., Browning, R., Delderfield, J., Lee, D. J., Lidiard, K. A. (1990). The Along Track Scanning Radiometer measurement of sea-surface temperature from ERS-1. *JBIS J. Bri. Interplanet. Soc.*, 43, 160-180

Farnebäck, G. (2003). Two-frame motion estimation based on polynomial expansion. In *Image Analysis* (pp. 363-370). Springer Berlin Heidelberg.

Ferlay, N., Thieuleux, F., Cornet, C., Davis, A. B., Dubuisson, P., Ducos, F., ... & Vanbauce, C. (2010). Toward new inferences about cloud structures from multidirectional measurements in the oxygen A band:

- Middle-of-cloud pressure and cloud geometrical thickness from POLDER-3/PARASOL. *Journal of Applied Meteorology and Climatology*, 49(12), 2492-2507.
- Fischer, J., & Grassl, H. (1991). Detection of cloud-top height from backscattered radiances within the oxygen A band. Part 1: Theoretical study. *Journal of Applied Meteorology*, 30(9), 1245-1259.
- Fisher, D., & Muller, J. P. (2013). Global warping coefficients for improving ATSR co-registration. *Remote Sensing Letters*, 4(2), 151-160.
- Fisher, D., Muller, J. P., & Yershov, V. N. (2013). Automated Stereo Retrieval of Smoke Plume Injection Heights and Retrieval of Smoke Plume Masks from AATSR and their Assessment With CALIPSO and MISR. *Geoscience and Remote Sensing, IEEE Transactions on*, 99, 1, doi:10.1109/TGRS.2013.2249073.
- Garay, M. J., de Szoeko, S. P., & Moroney, C. M. (2008). Comparison of marine stratocumulus cloud top heights in the southeastern Pacific retrieved from satellites with coincident ship-based observations. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D18).
- Ghent, D., & Remedios, J. (2013, April). Developing first time-series of land surface temperature from AATSR with uncertainty estimates. In *EGU General Assembly Conference Abstracts* (Vol. 15, p. 5016).
- Gillett, N. P., Weaver, A. J., Zwiers, F. W., & Flannigan, M. D. (2004). Detecting the effect of climate change on Canadian forest fires. *Geophysical Research Letters*, 31(18).
- Greenough, G., McGeehin, M., Bernard, S. M., Trtanj, J., Riad, J., & Engelberg, D. (2001). The potential impacts of climate variability and change on health impacts of extreme weather events in the United States. *Environmental Health Perspectives*, 109(Suppl 2), 191.
- Griggs, J. A., & Bamber, J. L. (2008). Assessment of cloud cover characteristics in satellite datasets and reanalysis products for Greenland. *Journal of Climate*, 21(9), 1837-1849.
- Grosse, G., Romanovsky, V., Jorgenson, T., Anthony, K. W., Brown, J., & Overduin, P. P. (2011). Vulnerability and feedbacks of permafrost to climate change. *Eos, Transactions American Geophysical Union*, 92(9), 73-74.
- Gruen, A. (1985). Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*, 14(3), 175-187.
- Handoll, M. (1991). *UK ERS-1 reference manual*.
- Hartmann, D. L. (1994). *Global physical climatology* (Vol. 56). Academic press.
- Hastings, D. A., & Dunbar, P. K. (1998). Development & assessment of the Global Land One-km Base Elevation digital elevation model (GLOBE). *Group*, 4, 6.
- Heidinger, A. K., Evan, A. T., Foster, M. J., & Walther, A. (2012). A naive Bayesian cloud-detection scheme derived from CALIPSO and applied within PATMOS-x. *Journal of Applied Meteorology and Climatology*, 51(6), 1129-1144.
- Heipke, C., Oberst, J., Albertz, J., Attwenger, M., Dorninger, P., Dorrer, E., ... & Neukum, G. (2007).

- Evaluating planetary digital terrain models—The HRSC DTM test. *Planetary and Space Science*, 55(14), 2173-2191.
- Helmer, S., & Lowe, D. (2010, May). Using stereo for object recognition. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on* (pp. 3121-3127). IEEE.
- Hinkelman, L. M., Marchand, R. T., & Ackerman, T. P. (2009). Evaluation of Multiangle Imaging Spectroradiometer cloud motion vectors using NOAA radar wind profiler data. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 114(D21).
- Hirano, A., Welch, R., & Lang, H. (2003). Mapping from ASTER stereo image data: DEM validation and accuracy assessment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(5), 356-370.
- Hirschmüller, H., Innocent, P. R., & Garibaldi, J. (2002). Real-time correlation-based stereo vision with reduced border errors. *International Journal of Computer Vision*, 47(1-3), 229-246.
- Hirschmuller, H. (2005, June). Accurate and efficient stereo processing by semi-global matching and mutual information. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (Vol. 2, pp. 807-814). IEEE.
- Hirschmuller, H. (2008). Stereo processing by semiglobal matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2), 328-341.
- Hirschmuller, H., & Scharstein, D. (2009). Evaluation of stereo matching costs on images with radiometric differences. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(9), 1582-1599.
- Hogan, R. J., C. Jakob, and A. J. Illingworth, 2001: Comparison of ECMWF winter-season cloud fraction with radar-derived values. *J. Appl. Meteor.*, 40, 513–525.
- Holland, M. M., & Bitz, C. M. (2003). Polar amplification of climate change in coupled models. *Climate Dynamics*, 21(3-4), 221-232.
- Holz, R. E., Ackerman, S. A., Nagle, F. W., Frey, R., Dutcher, S., Kuehn, R. E., ... & Baum, B. (2008). Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D8).
- Horstmann, J., Koch, W., Lehner, S., & Tonboe, R. (2000). Wind retrieval over the ocean using synthetic aperture radar with C-band HH polarization. *Geoscience and Remote Sensing, IEEE Transactions on*, 38(5), 2122-2131.
- Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., Van der Linden, P. J., Dai, X., ... & Johnson, C. A. (2001). IPCC, 2001: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge, United Kingdom, New York, USA, Cambridge University Press*, 881, 9.
- Horváth, Á. (2013). Improvements to MISR stereo motion vectors. *Journal of Geophysical Research: Atmospheres*.

- Horváth, Á., & Davies, R. (2001a). Feasibility and error analysis of cloud motion wind extraction from near-simultaneous multiangle MISR measurements. *Journal of Atmospheric and Oceanic Technology*, 18(4), 591-608.
- Horváth, Á., & Davies, R. (2001b). Simultaneous retrieval of cloud motion and height from polar-orbiter multiangle measurements. *Geophysical Research Letters*, 28(15), 2915-2918.
- Hsieh, Y.C., McKeown, D., and Perlant, F.P. 1992. Performance evaluation of scene registration and stereo matching for cartographic feature extraction. *IEEE TPAMI*, 14(2):214–238.
- Istomina, L. G., Hoyningen-Huene, W. V., Kokhanovsky, A. A., & Burrows, J. P. (2010). The detection of cloud-free snow-covered areas using AATSR measurements. *Atmospheric Measurement Techniques*, 3(4), 1005-1017.
- Jakob, C., and G. Tselioudis, 2003: How representative are the cloud regimes at the TWP sites?—An ISCCP perspective. Proc. 13th ARM Science Team Meeting, Broomfield, CO,U.S. Dept. of Energy, 9 pp.
- Jin, Y., & Rossow, W. B. (1997). Detection of cirrus overlapping low-level clouds. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 102(D2), 1727-1737.
- Josset, D., Pelon, J., Garnier, A., Hu, Y., Vaughan, M., Zhai, P. W., ... & Lucker, P. (2012). Cirrus optical depth and lidar ratio retrieval from combined CALIPSO-CloudSat observations using ocean surface echo. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 117(D5).
- Kahn, R. A., Chen, Y., Nelson, D. L., Leung, F. Y., Li, Q., Diner, D. J., & Logan, J. A. (2008). Wildfire smoke injection heights: Two perspectives from space. *Geophysical Research Letters*, 35(4), L04809.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... & Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, 77(3), 437-471.
- Karlsson, K. G., & Dybbroe, A. (2010). Evaluation of Arctic cloud products from the EUMETSAT Climate Monitoring Satellite Application Facility based on CALIPSO-CALIOP observations. *Atmospheric Chemistry and Physics*, 10(4), 1789-1807.
- Kasischke, E. S., & Turetsky, M. R. (2006). Recent changes in the fire regime across the North American boreal region—Spatial and temporal patterns of burning across Canada and Alaska. *Geophysical Research Letters*, 33(9).
- Key, J. R., Santek, D., Velden, C. S., Bormann, N., Thepaut, J. N., Riishojgaard, L. P., ... & Menzel, W. P. (2003). Cloud-drift and water vapor winds in the polar regions from MODISIR. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(2), 482-492.
- Kiehl, J. T., Hack, J. J., & Briegleb, B. P. (1994). The simulated Earth radiation budget of the National Center for Atmospheric Research community climate model CCM2 and comparisons with the Earth Radiation Budget Experiment (ERBE). *Journal of Geophysical Research: Atmospheres (1984–2012)*, 99(D10), 20815-20827.

- Kiehl, J. T., & Trenberth, K. E. (1997). Earth's annual global mean energy budget. *Bulletin of the American Meteorological Society*, 78(2), 197-208.
- Krol, M., Houweling, S., Bregman, B., Broek, M., Segers, A., Velthoven, P. V., ... & Bergamaschi, P. (2005). The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmospheric Chemistry and Physics*, 5(2), 417-432.
- Krol, M., Peters, W., Hooghiemstra, P., George, M., Clerbaux, C., Hurtmans, D., ... & Muller, J. P. (2012). How much CO was emitted by the 2010 fires around Moscow?. *Atmospheric Chemistry and Physics Discussions*, 12(11), 28705-28731.
- Künzli, N., Avol, J. W., Gauderman, W. J., Rappaport, J. M., Bennion, J., McConnell, R., ... & Peters, J. M. (2006). Health effects of the 2003 Southern California wildfires on children. *American journal of respiratory and critical care medicine*, 174(11), 1221.
- Lean, J., & Rind, D. (1998). Climate forcing by changing solar radiation. *Journal of Climate*, 11(12), 3069-3094.
- Li, R. and Liu, C. (2004). Photogrammetry for Remote Sensing. In: Bossler, J. D., Jensen, J. R., McMaster, R. B., & Rizos, C. eds. *Manual of geospatial science and technology*. CRC Press, pp. 285-299
- Liao, X., Rossow, W. B., & Rind, D. (1995). Comparison between SAGE II and ISCCP high-level clouds 2. Locating cloud tops. *Journal of geophysical research*, 100(D1), 1137-1147.
- Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2004). *Remote sensing and image interpretation* (No. Ed. 5). John Wiley & Sons Ltd.
- Llewellyn-Jones, D., Edwards, M., Mutlow, C., Birks, A., Barton, I., & Tait, H. (2001). AATSR: Global-change and surface-temperature measurements from Envisat. *Esa Bulletin-European Space Agency*, 11-21
- Lonitz, K., & Horváth, Á. (2011). Comparison of MISR and Meteosat-9 cloud-motion vectors. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 116(D24).
- Lorenz, D. (1982). Stereoscopic imaging from polar orbit and synthetic stereo imaging. *Advances in Space Research*, 2(6), 133-142.
- Lorenz, D. (1985). On the feasibility of cloud stereoscopy and wind determination with the along-track scanning radiometer. *International Journal of Remote Sensing*, 6(8), 1445-1461.
- Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal Of Computer Vision*, 60, 91-110
- Maddux, B. C., Ackerman, S. A., & Platnick, S. (2010). Viewing geometry dependencies in MODIS cloud products. *Journal of Atmospheric and Oceanic Technology*, 27(9), 1519-1528.
- Marchand, R. T., Ackerman, T. P., & Moroney, C. (2007). An assessment of Multiangle Imaging Spectroradiometer (MISR) stereo-derived cloud top heights and cloud top winds using ground-based

- radar, lidar, and microwave radiometers. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 112(D6).
- Marchand, R., Ackerman, T., Smyth, M., & Rossow, W. B. (2010). A review of cloud top height and optical depth histograms from MISR, ISCCP, and MODIS. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 115(D16).
- Marconcini, M., Fernández-Prieto, D., Simon, P., Hayman, G., Helbert, J., & de Leeuw, G. (2010). ALANIS: a joint ESA–iLEAPS atmosphere–land interaction study over boreal Eurasia. *iLEAPS Newsletter*, 10, 28-33.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*, Henry Holt and Co. Inc., New York, NY.
- Martin, M. V., Logan, J. A., Kahn, R. A., Leung, F. Y., Nelson, D. L., & Diner, D. J. (2010). Smoke injection heights from fires in North America: analysis of 5 years of satellite observations. *Atmos. Chem. Phys*, 10(4), 1491-1510.
- Mazzoni, D., Garay, M. J., Davies, R., & Nelson, D. (2007a). An operational MISR pixel classifier using support vector machines. *Remote Sensing of Environment*, 107(1), 149-158.
- Mazzoni, D., Logan, J. A., Diner, D., Kahn, R., Tong, L., & Li, Q. (2007b). A data-mining approach to associating MISR smoke plume heights with MODIS fire measurements. *Remote Sensing of Environment*, 107(1), 138-148.
- McGill, M. J., Vaughan, M. A., Trepte, C. R., Hart, W. D., Hlavka, D. L., Winker, D. M., & Kuehn, R. (2007). Airborne validation of spatial properties measured by the CALIPSO lidar. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 112(D20).
- Mei, X., Sun, X., Zhou, M., Jiao, S., Wang, H., & Zhang, X. (2011, November). On building an accurate stereo matching system on graphics hardware. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on* (pp. 467-474). IEEE.
- Menzel, W. P., Smith, W. L., & Stewart, T. R. (1983). Improved cloud motion wind vector and altitude assignment using VAS. *Journal of Climate and Applied meteorology*, 22(3), 377-384.
- Menzel, W. "Cloud tracking with satellite imagery: From the pioneering work of Ted Fujita to the present," *Bull. Amer. Meteorol. Soc.*, vol. 82, no. 1, pp. 33–47, 2001.
- Menzel, W. P., Frey, R. A., Zhang, H., Wylie, D. P., Moeller, C. C., Holz, R. E., ... & Gumley, L. E. (2008). MODIS global cloud-top pressure and amount estimation: Algorithm description and results. *Journal of Applied Meteorology and Climatology*, 47(4), 1175-1198.
- Merchant, C. J., Simpson, J. J., & Harris, A. R. (2003). A cross-calibration of GMS-5 thermal channels against ATSR-2. *Remote sensing of environment*, 84(2), 268-282.
- Miller, S. D., G. L. Stephens, and A. C. M. Beljaars, 1999: A validation survey of the ECMWF prognostic cloud scheme using LITE. *Geophys. Res. Lett.*, 26, 1417–1420

- Minnis, P., Harrison, E. F., Stowe, L. L., Gibson, G. G., Denn, F. M., Doelling, D. R., & Smith, W. L. (1993). Radiative climate forcing by the Mount Pinatubo eruption. *Science*, 259(5100), 1411-1415.
- Minnis, P., Sun-Mack, S., Young, D. F., Heck, P. W., Garber, D. P., Chen, Y., ... & Yang, P. (2011). CERES edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data—Part I: Algorithms. *Geoscience and Remote Sensing, IEEE Transactions on*, 49(11), 4374-4400.
- Moon, T., Joughin, I., Smith, B., & Howat, I. (2012). 21st-century evolution of Greenland outlet glacier velocities. *Science*, 336(6081), 576-578.
- Moroney, C., Davies, R., & Muller, J. P. (2002). Operational retrieval of cloud-top heights using MISR data. *Geoscience and Remote Sensing, IEEE Transactions on*, 40(7), 1532-1540.
- Mueller, K., C. Moroney, V. Jovanovic, M. J. Garay, J. P. Muller, L. Di Girolamo, and R. Davies (2013), MISR Level 2 cloud product algorithm theoretical basis, JPL Tech. Doc. D-73327, Jet Propul. Lab., Calif. Inst. of Technol., Pasadena.
- Muller, J. P., Mandanayake, A., Moroney, C., Davies, R., Diner, D. J., & Paradise, S. (2002). MISR stereoscopic image matchers: Techniques and results. *Geoscience and Remote Sensing, IEEE Transactions on*, 40(7), 1547-1559.
- Muller, J. P., Denis, M. A., Dundas, R. D., Mitchell, K. L., Naud, C., & Mannstein, H. (2007). Stereo cloud-top heights and cloud fraction retrieval from ATSR-2. *International Journal of Remote Sensing*, 28(9), 1921-1938.
- Muller, J. P., & Fischer, J. (2007). The EU-CLOUDMAP project: Cirrus and contrail cloud-top maps from satellites for weather forecasting climate change analysis. *International Journal of Remote Sensing*, 28(9), 1915-1919.
- Murray, D., & Little, J. J. (2000). Using real-time stereo vision for mobile robot navigation. *Autonomous Robots*, 8(2), 161-171.
- Mutlow, C. T., M. J. Murray, P Bailey, A Birks. (1999). ATSR-1/2 user guide issue 1, *ESA user Guide*. 2391.
- Naderi, F. M., Freilich, M. H., & Long, D. G. (1991). Spaceborne radar measurement of wind velocity over the ocean-an overview of the NSCAT scatterometer system. *Proceedings of the IEEE*, 79(6), 850-866.
- Naud, C., Muller, J. P., Haeffelin, M., Morille, Y., & Delaval, A. (2004). Assessment of MISR and MODIS cloud top heights through inter-comparison with a back-scattering lidar at SIRTa. *Geophysical Research Letters*, 31(4).
- Naud, C. M., Baum, B. A., Pavolonis, M., Heidinger, A., Frey, R., & Zhang, H. (2007a). Comparison of MISR and MODIS cloud-top heights in the presence of cloud overlap. *Remote sensing of environment*, 107(1), 200-210.
- Naud, C., Mitchell, K. L., Muller, J. P., Clothiaux, E. E., Albert, P., Preusker, R., ... & Hogan, R. J.

- (2007b). Comparison between ATSR-2 stereo, MOS O2-A band and ground-based cloud top heights. *International Journal of Remote Sensing*, 28(9), 1969-1987.
- Nelson, D. L., Chen, Y., Kahn, R. A., Diner, D. J., & Mazzoni, D. (2008, August). Example applications of the MISR INteractive eXplorer (MINX) software tool to wildfire smoke plume analyses. In *Optical Engineering+ Applications* (pp. 708909-708909). International Society for Optics and Photonics.
- Nelson, D. L., Garay, M. J., Kahn, R. A., & Dunst, B. A. (2013). Stereoscopic Height and Wind Retrievals for Aerosol Plumes with the MISR INteractive eXplorer (MINX). *Remote Sensing*, 5(9), 4593-4628.
- NIMA (National Imaging and Mapping Agency), 2000. The Compendium of Controlled Extensions ICE] for the National Imagery Transmission Format [NITF), Version 2.1.
- Norris, J. R., and C. P. Weaver, (2001): Improved techniques for evaluating GCM cloudiness applied to the NCAR CCM3. *J. Climate*, 14, 2540–2550.
- NRC, National Research Council (2003), Understanding climate change feedbacks, The National Academic Press, Washington
- Nyejlukto, M.F., ed. 1994. Russian forests. Federal Forest Service of Russia, All-Russian Research and Information Centre for Forest Resources. Moscow: Company “Kolyev”; 15 p
- Ohta, Y., & Kanade, T. (1985). Stereo by intra-and inter-scanline search using dynamic programming. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2), 139-154.
- Otto, G. P., & Chau, T. K. (1989). ‘Region-growing’ algorithm for matching of terrain images. *Image and vision computing*, 7(2), 83-94.
- Panton, D. J. (1978). A flexible approach to digital stereo mapping. *Photogrammetric Engineering and Remote Sensing*, 44(12), 1499-1512.
- Parkinson, C. L., & Comiso, J. C. (2013). On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm. *Geophysical Research Letters*, 1-6.
- Poulsen, C. A., Siddans, R., Thomas, G. E., Sayer, A. M., Grainger, R. G., Campmany, E., ... & Watts, P. D. (2012). Cloud retrievals from satellite data using optimal estimation: evaluation and application to ATSR. *Atmospheric Measurement Techniques*, 5(8), 1889-1910.
- Prata, A. J., & Turner, P. J. (1997). Cloud-top height determination using ATSR data. *Remote sensing of environment*, 59(1), 1-13.
- Prentice, I. C., Sykes, M. T., Lautenschlager, M., Harrison, S. P., Denissenko, O., & Bartlein, P. J. (1993). Modelling global vegetation patterns and terrestrial carbon storage at the last glacial maximum. *Global Ecology and Biogeography Letters*, 67-76.
- Preusker, R., & Lindstrot, R. (2009). Remote Sensing of Cloud-Top Pressure Using Moderately Resolved Measurements within the Oxygen A Band-A Sensitivity Study. *Journal of Applied Meteorology and Climatology*, 48(8), 1562-1574.

- Ramanathan, V. L. R. D., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., & Hartmann, D. (1989). Cloud-radiative forcing and climate: Results from the Earth Radiation Budget Experiment. *Science*, 243(4887), 57-63.
- Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichet, T., Fyfe, J., ... & Taylor, K. E. (2007). Climate models and their evaluation. *Climate change*, 323.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., ... & Woollen, J. (2011). MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, 24(14), 3624-3648.
- Rossow, W. B., & Garder, L. C. (1993a). Cloud detection using satellite measurements of infrared and visible radiances for ISCCP. *Journal of Climate*, 6(12), 2341-2369.
- Rossow, W. B., Walker, A. W., & Garder, L. C. (1993b). Comparison of ISCCP and other cloud amounts. *Journal of Climate*, 6(12), 2394-2418.
- Rossow, W. B., & Schiffer, R. A. (1999). Advances in understanding clouds from ISCCP. *Bulletin of the American Meteorological Society*, 80(11), 2261-2287.
- Rossow, W. B., Zhang, Y., & Wang, J. (2005). A statistical model of cloud vertical structure based on reconciling cloud layer amounts inferred from satellites and radiosonde humidity profiles. *Journal of climate*, 18(17), 3587-3605.
- Sayer, A. M., Poulsen, C. A., Arnold, C., Campmany, E., Dean, S., Ewen, G. B. L., ... & Watts, P. D. (2011). Global retrieval of ATSR cloud parameters and evaluation (GRAPE): dataset assessment. *Atmospheric Chemistry and Physics*, 11(8), 3913-3936.
- Scharstein, D., & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision*, 47(1-3), 7-42.
- Schaub, M., Schwiegerling, J., Fest, E., Shepard, R. H., & Symmons, A. (2011). *Molded Optics: Design and Manufacture*. CRC Press.
- Schweiger, A.J. and J. Key. 1992. "Comparison of ISCCP-C2 and Nimbus-7 satellite-derived cloud products with a surface-based cloud climatology in the arctic", *J. Climate*, vol. 5, no. 12, pp. 1514-1527.
- Seidel, D. J., Ross, R. J., Angell, J. K., & Reid, G. C. (2001). Climatological characteristics of the tropical tropopause as revealed by radiosondes. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 106(D8), 7857-7878.
- Seiz, G. (2003). *Ground-and satellite-based multi-view determination of 3D cloud geometry* (Doctoral dissertation, PhD thesis, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland. IGP Mitteilungen).
- Shepherd, A., Ivins, E. R., Geruo, A., Barletta, V. R., Bentley, M. J., Bettadpur, S., ... & Zwally, H. J. (2012). A reconciled estimate of ice-sheet mass balance. *Science*, 338(6111), 1183-1189.

- Shin, D., & Pollard, J. K. (1996, February). Cloud height determination from satellite stereo images. In *Image Processing for Remote Sensing, IEE Colloquium on* (pp. 4-1). IET.
- Shin, D., Pollard, J., & Muller, J.-P. (1997). Accurate geometric correction of ATSR images. *IEEE Transactions On Geoscience And Remote Sensing*, 35, 997-1006
- Simmons, A., Uppala, S., Dee, D., & Kobayashi, S. (2007). ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF newsletter*, 110(110), 25-35.
- Smith, D., Delderfield, J., Drummond, D., Edwards, T., Mutlow, C., Read, P., & Toplis, G. (2001). Calibration of the AATSR instrument. In, *Calibration and Characterization of Satellite Sensors and Accuracy of Derived Physical Parameters* (pp. 31-39)
- Smith, D., Mutlow, C., & Rao, C. (2002). Calibration monitoring of the visible and near-infrared channels of the Along-Track Scanning Radiometer-2 by use of stable terrestrial sites. *Applied Optics*, 41, 515-523
- Smith, D., & Poulsen, C. (2008). Calibration status of the AATSR reflectance channels. In, *Proc. of the '2nd MERIS / (A)ATSR User Workshop'*, Frascati, Italy 22–26 September 2008 (ESA SP-666, November 2008).
- Soden, B. J., & Held, I. M. (2006). An assessment of climate feedbacks in coupled ocean-atmosphere models. *Journal of Climate*, 19(14), 3354-3360.
- Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, 21(14), 3504-3520.
- Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., ... & Miller, H. L. (2007). The physical science basis. *Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*, 235-337.
- Stephens, G. L. (2005). Cloud feedbacks in the climate system: A critical review. *Journal of climate*, 18(2), 237-273.
- Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., ... & CloudSat Science Team, T. (2002). The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, 83(12), 1771-1790.
- Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., ... & Midgley, P. M. (2013). IPCC, 2013: Climate Change 2013: The Physical Science Basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.
- Stocks, B. J., Fosberg, M. A., Lynham, T. J., Mearns, L., Wotton, B. M., Yang, Q., ... & McKenney, D. W. (1998). Climate change and forest fire potential in Russian and Canadian boreal forests. *Climatic Change*, 38(1), 1-13.
- Stoffelen, A., & Anderson, D. L. T. (1993). Wind retrieval and ERS-1 scatterometer radar backscatter measurements. *Advances in Space Research*, 13(5), 53-60.

- Stubenrauch, C. J., Chédin, A., Rädcl, G., Scott, N. A., & Serrar, S. (2006). Cloud properties and their seasonal and diurnal variability from TOVS Path-B. *Journal of climate*, 19(21), 5531-5553.
- Stubenrauch, C. J., Cros, S., Guignard, A., & Lamquin, N. (2010). A 6-year global cloud climatology from the Atmospheric InfraRed Sounder AIRS and a statistical analysis in synergy with CALIPSO and CloudSat. *Atmospheric Chemistry and Physics*, 10(15), 7197-7214.
- Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., ... & Zhao, G. (2013). Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX Radiation Panel. *Bulletin of the American Meteorological Society*.
- Sun, J., Zheng, N. N., & Shum, H. Y. (2003). Stereo matching using belief propagation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(7), 787-800.
- Tao, C., & Hu, Y. (2001). A comprehensive study of the rational function model for photogrammetric processing. *Photogrammetric engineering and remote sensing*, 67, 1347-1357.
- Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485-498.
- Thomas, G. E., Poulsen, C. A., Siddans, R., Sayer, A. M., Carboni, E., Marsh, S. H., ... & Lawrence, B. N. (2010). Validation of the GRAPE single view aerosol retrieval for ATSR-2 and insights into the long term global AOD trend over the ocean. *Atmospheric Chemistry and Physics*, 10(10), 4849-4866.
- Trepte, Q., Minnis, P., & Arduini, R. F. (2002, October). Daytime and nighttime polar cloud and snow identification using MODIS data. In *Proc. SPIE 3rd Intl. Asia-Pacific Environ. Remote Sensing Symp* (pp. 23-27).
- Turetsky, M. R., Kane, E. S., Harden, J. W., Ottmar, R. D., Manies, K. L., Hoy, E., & Kasischke, E. S. (2010). Recent acceleration of biomass burning and carbon losses in Alaskan forests and peatlands. *Nature Geoscience*, 4(1), 27-31.
- Turquety, S., Logan, J. A., Jacob, D. J., Hudman, R. C., Leung, F. Y., Heald, C. L., ... & Sachse, G. W. (2007). Inventory of boreal fire emissions for North America in 2004: Importance of peat burning and pyroconvective injection. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 112(D12).
- Vaughan, M. A., Powell, K. A., Winker, D. M., Hostetler, C. A., Kuehn, R. E., Hunt, W. H., ... & McGill, M. J. (2009). Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements. *Journal of Atmospheric and Oceanic Technology*, 26(10), 2034-2050.
- Veksler, O. (2003, June). Fast variable window for stereo correspondence using integral images. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on* (Vol. 1, pp. I-556). IEEE.
- Velden, C. S., Hayden, C. M., Nieman, S. J., Menzel, W. P., Wanzong, S., & Goerss, J. S. (1997). Upper-tropospheric winds derived from geostationary satellite water vapor observations. *Bulletin of the American Meteorological Society*, 78(2), 173-195.

- Velden, C., Daniels, J., Stettner, D., Santek, D., Key, J., Dunion, J., ... & Menzel, P. (2005). Recent innovations in deriving tropospheric winds from meteorological satellites. *Bulletin of the American Meteorological Society*, 86(2), 205-223.
- Wang, J., Rossow, W. B., Uttal, T., & Rozendaal, M. (1999). Variability of cloud vertical structure during ASTEX observed from a combination of rawinsonde, radar, ceilometer, and satellite. *Monthly weather review*, 127(10), 2484-2502.
- Watts, P. D., Mutlow, C. T., Baran, A. J., and Zavody, A. M. (1998): Study on cloud properties derived from Meteosat Second Generation Observations, Eumetsat Report, 2393,2403
- Watts, P. D., Bennartz, R., & Fell, F. (2011). Retrieval of two-layer cloud properties from multispectral observations using optimal estimation. *Journal of Geophysical Research: Atmospheres* (1984–2012), 116(D16).
- Webb, M. J., C. Senior, S. Bony, and J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF, and LMD atmospheric climate models. *Climate Dyn.*, 17, 905–922.
- Webb, M. J., Senior, C. A., Sexton, D. M. H., Ingram, W. J., Williams, K. D., Ringer, M. A., ... & Taylor, K. E. (2006). On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dynamics*, 27(1), 17-38.
- Wetherald, R. T., & Manabe, S. (1988). Cloud feedback processes in a general circulation model. *Journal of the Atmospheric Sciences*, 45(8), 1397-1416.
- Wielicki, B. A., & Parker, L. (1992). On the determination of cloud cover from satellite sensors: The effect of sensor spatial resolution. *Journal of Geophysical Research: Atmospheres* (1984–2012), 97(D12), 12799-12823.
- Wielicki, B. A., Harrison, E. F., Cess, R. D., King, M. D., & Randall, D. A. (1995). Mission to planet Earth: Role of clouds and radiation in climate. *Bulletin of the American Meteorological Society*, 76(11), 2125-2153.
- Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Lee III, R. B., Louis Smith, G., & Cooper, J. E. (1996). Clouds and the Earth's Radiant Energy System (CERES): An earth observing system experiment. *Bulletin of the American Meteorological Society*, 77(5), 853-868.
- Wind, G., Platnick, S., King, M. D., Hubanks, P. A., Pavolonis, M. J., Heidinger, A. K., ... & Baum, B. A. (2010). Multilayer cloud detection with the MODIS near-infrared water vapor absorption band. *Journal of Applied Meteorology and Climatology*, 49(11), 2315-2333.
- Winker, D. M., Pelon, J., & McCormick, M. P. (2003, October). The CALIPSO mission: Spaceborne lidar for observation of aerosols and clouds. In *Proc. Spie* (Vol. 4893, No. 1, p. 11).
- Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., ... & Young, S. A. (2009). Overview of the CALIPSO mission and CALIOP data processing algorithms. *Journal of Atmospheric and Oceanic Technology*, 26(11), 2310-2323.

- Winton, M. (2006). Surface albedo feedback estimates for the AR4 climate models. *Journal of climate*, 19(3), 359-365.
- Wu, D. L., D. J. Diner, M. J. Garay, V. M. Jovanovic, J. N. Lee, C. M. Moroney, K. J. Mueller, and D. L. Nelson (2010), MISR CMVs and multiangular views of tropical cyclone inner-core dynamics, paper presented at Tenth International Winds Workshop, Japan Meteorol. Agency, Tokyo.
- Wylie, D. P., & Menzel, W. P. (1999). Eight years of high cloud statistics using HIRS. *Journal of Climate*, 12(1), 170-184.
- Wylie, D., Jackson, D. L., Menzel, W. P., & Bates, J. J. (2005). Trends in global cloud cover in two decades of HIRS observations. *Journal of climate*, 18(15), 3021-3031.
- Xie, Y., Qu, J. J., Xiong, X., Hao, X., Che, N., & Sommers, W. (2007). Smoke plume detection in the eastern United States using MODIS. *International Journal of Remote Sensing*, 28(10), 2367-2374.
- Yang, Y., Di Girolamo, L., & Mazzoni, D. (2007). Selection of the automated thresholding algorithm for the Multi-angle Imaging SpectroRadiometer Radiometric Camera-by-Camera Cloud Mask over land. *Remote sensing of environment*, 107(1), 159-171.
- Zabih, R., & Woodfill, J. (1994). Non-parametric local transforms for computing visual correspondence. In *Computer Vision—ECCV'94* (pp. 151-158). Springer Berlin Heidelberg.
- Zhang, Z., Deriche, R., Faugeras, O., & Luong, Q. (1995). A Robust Technique for Matching 2 Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry. *Artificial Intelligence*, 78, 87-119
- Zhao, G., & Di Girolamo, L. (2004). A cloud fraction versus view angle technique for automatic in-scene evaluation of the MISR cloud mask. *Journal of Applied Meteorology*, 43(6), 860-869.

Appendix A

Sensor and Year	a_0	a_1	a_2	a_3	b_0	b_1	b_2	b_3
ATSR								
2011	0.0067785	0.0004153	1.0004229	-0.0012567	0.0068850	0.9998457	0.0014864	-0.0018684
2010	0.0068211	0.0000062	0.9999976	-0.0014678	0.0072853	1.0007157	-0.0002441	0.0006383
2009	0.0060237	0.0000678	0.9998340	-0.0011867	0.0073136	1.0010777	0.0009822	-0.0011429
2008	0.0065672	-0.0000050	1.0005713	-0.0012882	0.0081305	1.0010035	0.0013192	-0.0014078
2007	0.0047646	0.0004517	1.0002907	-0.0004731	0.0074255	1.0010364	0.0005494	-0.0006349
2006	0.0043492	-0.0003876	0.9998849	0.0000508	0.0080143	1.0007546	0.0018103	-0.00033572
2005	0.0051383	-0.0000318	0.9996028	-0.0005477	0.0069139	1.0011369	0.0008436	-0.0001474
2004	0.0040982	-0.0001489	1.0001640	-0.0008531	0.0067951	1.0000442	0.0009034	-0.0000299
2003	0.0033437	0.0001441	1.0003393	-0.0004892	0.0079457	1.0001718	0.0017305	-0.0018057
2002	0.0023106	0.0004956	0.9998888	0.0000408	0.0086951	1.0010434	0.0003886	-0.0028782
ATSR2								
1999	0.0005159	0.0001105	1.0005269	-	0.0039706	1.0011284	0.0020219	-
1998	-0.0051764	-0.0000902	1.0002067	-	0.0028822	1.0002171	0.0023228	-
1997	-0.0033432	-0.0002031	1.0001706	-	0.0024121	1.0007225	0.0000206	-
1996	-0.0023599	0.0001611	1.0001775	-	0.0042032	0.9994548	0.0019798	-
1995	-0.0010202	-0.0002010	1.0003297	-	0.0044656	1.0007100	0.0016303	-
ATSR1								
1994	-0.0095473	0.0000062	0.9985013	-	-0.0010717	1.0010025	0.0056120	-

Table A1. The ATSR warping coefficients for application within the warps defined in Equations 2.4 and 2.5. The coefficients are scaled between -1 and 1 and assume a 512 by 512 pixel image.

