Exploiting Weather Forecast Data for Cloud Detection

Shona Mackie

A thesis submitted for the degree of Doctor of Philosophy
The University of Edinburgh

January, 2009

Declaration

This thesis is an account of research undertaken between September 2005 and December 2008 in The School of Geosciences at The University of Edinburgh, Edinburgh, Scotland, United Kingdom.

Except where explicitly stated, this thesis and the material presented in it is my own work and has not been submitted in whole or part for a degree in any university.

Shona Mackie
January, 2009

Acknowledgements

I would like to thank my supervisor Chris Merchant for his advice and assistance, and also Owen Embury and Hugh Pumphrey for their assistance, particularly with computational aspects of the work. I would also like to thank the IT support staff and everyone in the School of Geosciences who helped with the project.

This work was funded by the Natural Environment Research Council, and sponsored through a CASE studentship with the U.K. Met. Office.

Abstract

Accurate, fast detection of clouds in satellite imagery has many applications, for example Numerical Weather Prediction (NWP) and climate studies of both the atmosphere and of the Earth's surface temperature. Most operational techniques for cloud detection rely on the differences between observations of cloud and of clear-sky being more or less constant in space and in time. In reality, this is not the case - different clouds have different spectral properties, and different cloud types are more or less likely in different places and at different times, depending on atmospheric conditions and on the Earth's surface properties. Observations of clear sky also vary in space and time, depending on atmospheric and surface conditions, and on the presence or absence of aerosol particles. The Bayesian approach adopted in this project allows pixel-specific physical information (for example from NWP) to be used to predict pixel-specific observations of clear sky. A physically-based, spatially- and temporally-specific probability that each pixel contains a cloud observation is then calculated. An advantage of this approach is that identification of ambiguously classed pixels from a probabilistic result is straightforward, in contrast to the binary result generally produced by operational techniques. This project has developed and validated the Bayesian approach to cloud detection, and has extended the range of applications for which it is suitable, achieving skills scores that match or exceed those achieved by operational methods in every case.

High temperature gradients can make observations of clear sky around ocean fronts, particularly at thermal wavelengths, appear similar to cloud observations. To address this potential source of ambiguous cloud detection results, a region of imagery acquired by the AATSR sensor which was noted to contain some ocean fronts, was selected. Pixels in the region were clustered according to their spectral properties with the aim of separating pixels that correspond to different thermal regimes of the ocean. The mean spectral properties of pixels in each cluster were then processed using the Bayesian cloud detection technique and the resulting posterior probability of clear then assigned to individual pixels. Several clustering methods were investigated, and the most appropriate, which allowed pixels to be associated with multiple clusters, with a normalized vector of 'membership strengths', was used to conduct a case study. The distribution of final calculated probabilities of clear became markedly more bimodal when clustering was included, indicating fewer ambiguous classifications, but at the cost of some single pixel clouds being missed. While further investigations could provide a solution to this, the computational expense of the clustering method made this impractical to include in the work of this project.

This new Bayesian approach to cloud detection has been successfully developed by this project to a point where it has been released under public license. Initially designed as a tool to aid retrieval of sea surface temperature from night-time imagery, this project has extended

the Bayesian technique to be suitable for imagery acquired over land as well as sea, and for day-time as well as for night-time imagery. This was achieved using the land surface emissivity and surface reflectance parameter products available from the MODIS sensor. This project added a visible Radiative Transfer Model (RTM), developed at University of Edinburgh, and a kernel-based surface reflectance model, adapted here from that used by the MODIS sensor, to the cloud detection algorithm. In addition, the cloud detection algorithm was adapted to be more flexible, making its implementation for data from the SEVIRI sensor straightforward. A database of 'difficult' cloud and clear targets, in which a wide range of both spatial and temporal locations was represented, was provided by Météo-France and used in this work to validate the extensions made to the cloud detection scheme and to compare the skill of the Bayesian approach with that of operational approaches. For night land and sea imagery, the Bayesian technique, with the improvements and extensions developed by this project, achieved skills scores 10% and 13% higher than Météo-France respectively. For daytime sea imagery, the skills scores were within 1% of each other for both approaches, while for land imagery the Bayesian method achieved a 2% higher skills score.

The main strength of the Bayesian technique is the physical basis of the differentiation between clear and cloud observations. Using NWP information to predict pixel-specific observations for clear-sky is relatively straightforward, but making such predictions for cloud observations is more complicated. The technique therefore relies on an empirical distribution rather than a pixel-specific prediction for cloud observations. To try and address this, this project developed a means of predicting cloudy observations through the fast forward-modelling of pixel-specific NWP information. All cloud fields in the pixel-specific NWP data were set to 0, and clouds were added to the profile at discrete intervals through the atmosphere, with cloud water- and ice- path (cwp, cip) also set to values spaced exponentially at discrete intervals up to saturation, and with cloud pixel fraction set to 25%, 50%, 75% and 100%. Only single-level, single-phase clouds were modelled, with the justification that the resulting distribution of predicted observations, once smoothed through considerations of uncertainties, is likely to include observations that would correspond to multi-phase and multi-level clouds. A fast RTM was run on the profile information for each of these individual clouds and cloud altitude-, cloud pixel fraction- and channel-specific relationships between cwp (and similarly cip) and predicted observations were calculated from the results of the RTM. These relationships were used to infer predicted observations for clouds with cwp/cip values other than those explicitly forward modelled. The parameters used to define the relationships were interpolated to define relationships for predicted observations of cloud at 10m vertical intervals through the atmosphere, with pixel coverage ranging from 25% to 100% in increments of 1%. A distribution of predicted cloud observations is then achieved without explicit forward-modelling of an impractical number of atmospheric states. Weights are applied to the representation of individual clouds within the final Probability Density Function (PDF) in order to make the distribution of predicted observations realistic, according to the pixel-specific NWP data, and to distributions seen in a global reference dataset of NWP profiles from the European Centre for Medium Range Weather Forecasting (ECMWF). The distribution is then

convolved with uncertainties in forward-modelling, in the NWP data, and with sensor noise to create the final PDF in observation space, from which the conditional probability that the pixel observation corresponds to a cloud observation can be read. Although the relatively fast computational implementation of the technique was achieved, the results are disappointingly poor for the SEVIRI-acquired dataset, provided by Météo-France, against which validation was carried out. This is thought to be explained by both the uncertainties in the NWP data, and the forward-modelling dependence on those uncertainties, being poorly understood, and treated too optimistically in the algorithm. Including more errors in the convolution introduces the problem of quantifying those errors (a non-trivial task), and would increase the processing time, making implementation impractical. In addition, if the uncertianties considered are too high then a PDF flatter than the empirical distribution currently used would be produced, making the technique less useful. It is hoped that advances in NWP will result in the implementation of this technique in the Bayesian cloud detection algorithm yielding improved results in the future. At present no clear improvement is seen and the computational expense of including the local cloud PDF calcluation in the algorithm is therefore judged unjustified.

The Bayesian method for cloud detection calculates a probability that an observation corresponds to a particular class: clear or cloud. Provided the necessary background information is available, this can be adapted to calculate a probability that an observation corresponds to any number of classes. This was demonstrated here, where the approach was adapted to detect dust, cloud and clear sky simultaneously in a night-time image over sea (generally the most challenging scenario for dust detection). The need for cloud-screening prior to retrieving aerosol observations, which necessarily biases recorded observations of aerosol to those aerosol observations which are spectrally more similar to clear sky than to cloud, is thereby removed for dust. A distribution of simulated Saharan dust observations from another study was used to calculate a PDF, which was made conditional on the pixel NWP Surface Temperature (ST) and Total Column Water Vapour (TCWV). This was combined with the empirical PDF for cloud and the calculated, NWP-conditional, PDF for clear to calculate the normalized posterior probabilities that the pixel observation corresponds to each of the three classes. The latitude- and season-specific prior probabilities required by Bayes Theorem were taken for cloud and clear from International Satellite Cloud Climatology Project (ISCCP) data, and from a dataset of SEVIRI-acquired imagery, for which the Saharan Dust Index (SDI, a measure of the presence of dust) had been calculated, for dust. There being no cloud-clear-dust classified data available for validation, the technique was validated qualitatively through comparison of the three-way classification results against the results of the two-way classification (cloud and clear), and against calculated SDI results (a measure to discriminate between clear and dust). 22 night-time images acquired by the SEVIRI sensor between 2004 and 2006 were used for the validation, and show the technique to produce highly plausible results, although a quantitative assessment is difficult to find.

This thesis presents the work undertaken to carry out these developments and extensions to a Bayesian cloud detection scheme. Through this work, several challenges to the technique,

such as for example ambiguous classification of pixels around ocean fronts and non-latitude specific prior probabilities of cloud and clear, have been investigated and addressed. The project has extended the range of applications for which the cloud detection technique can be useful to include day-time- and land- imagery applications, in addition to the night-time ocean applications for which it was initially designed. In addition, the work undertaken here has resulted in the method has becoming more physically robust, and more thoroughly validated. A further outcome of this work is the application of the cloud detection technique to the successful classification of imagery into cloud, clear and dust observations, providing a potential solution to areas of NWP and climate research.

Acronyms and Nomenclature

The following abbreviations and nomenclature are used throughout this thesis.

Symbol	Meaning
c :	clear sky
$\overline{c}$:	not clear sky
$\mathbf{y^o}$:	observation vector
$\mathbf{x^b}$:	a priori information
$\mathbf{y^b}$:	predicted observation vector
z :	reduced background state vector
B :	error covariance matrix for **z**
R :	matrix of combined channel-dependent sensor noise and RTM error covaraince
H:	tangent linear of the RTM with respect to **z**
$\epsilon^{\mathbf{b}}$:	uncertainties associated with fields of $\mathbf{x^b}$
ϵ^b_{st} :	error in NWP surface temperature field
ϵ_{r_e} :	FM error attributable to assumption of a fixed effective cloud radius
ϵ :	emissivity
Γ :	lapse rate
h :	height above sea level
K :	kernel contributions to land surface reflectance
f :	weighting factors for individual kernel contributions to land surface reflectance
θ :	satellite zenith angle
ν :	solar zenith angle
ϕ :	relative azimuth angle
λ :	waveband for a particular sensor channel

Abbreviation	**Explanation**
AOD:	Aerosol Optical Depth
AATSR:	Advanced Along Track Scanning Radiometer
ATSR:	Along Track Scanning Radiometer
ATSR2:	Along Track Scanning Radiometer 2
AVHRR:	Advanced Very High Resolution Radiometer
BRDF:	Bi-direction Reflectance Distribution Function
BT:	Brightness Temperature
CLARA:	Clustering Large Applications
cwp:	Cloud Water Path
DEM:	Digital Elevation Model
ECMWF:	European Centre for Medium-range Weather Forecasting
ERS:	European Remote Sensing Satellite
ESA:	European Space Agency
FAR:	False Alarm Rate
FM:	Forward Modelling
FPC:	Fixed Point Clustering
GCM:	Global Climate Model
HR:	Hit Rate
ISCCP:	International Satellite Cloud Climatology Project
LSD:	Local Standard Deviation
LUT:	Look-Up Table
MERIS:	Medium Resolution Imaging Spectrometer
MODIS:	MODerate resolution Imaging Spectro-radiometer
MSG:	Meteosat Second Generation
NIR:	Near Infra-Red
NWP:	Numerical Weather Prediction
PDF:	Probability Density Function
PP:	Proportion of Perfect Classifications
RTM:	Radiative Transfer Model
RTTOV :	Radiative Transfer for TOVS
SD:	Standard Deviation
SDI:	Saharan Dust Index
SEVIRI:	Spinning Enhanced Visible and InfraRed Imager
SMC:	Split and Merge Clustering
SR:	Surface Reflectance
SST:	Sea Surface Temperature
ST:	Surface Temperature
TCWV:	Total Column Water Vapour
TIR:	Thermal Infra-Red
TOA:	Top of Atmosphere
TOMS:	Total Ozone Mapping Spectrometer
TSS:	True Skills Score
VisRTM:	Visible Radiative Transfer Model (for 0.6, 0.8 and 1.6μm imagery)
WS:	surface Wind Speed

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Clouds are important to weather and climate, via both their controlling influence on the planet's energy balance and their role in transporting fresh water around the Earth, see for example Liou (1986); Fu (1996); Cronin et al. (2005); Kokhanovsky et al. (2005). Different types of cloud can affect weather and climate systems differently, depending on properties such as height and optical thickness. Such factors have to be taken into account in any model of climate or weather, (Li et al., 2005), for example in that used by the Met. Office to calculate the numerical weather forecast.

There are two problems associated with the treatment of clouds within models. Firstly, the presence and distribution of clouds should agree spatially and temporally with reality, and secondly, sub-grid scale processes within clouds must be represented. Inaccuracies in the representation of clouds within models for weather and climate forecasting are currently recognized to be a significant source of uncertainty in the output from such models, see for example Cess et al. (1990); Derrien et al. (1993) and Li et al. (2005), and there exists a need for a fast, accurate technique to constrain the parameterization. In weather forecasting, these inaccuracies can be addressed to some degree through visual inspection of satellite imagery, which is used, for example, to make small amendments to the Met. Office's Numerical Weather Prediction (NWP) model's representation of clouds, bringing their appearance within the weather forecast into line with reality. If cloud fields can be assimilated more accurately into the model from the imagery, the need for these adjustments would reduce, as clouds within the model would be better constrained. Better representation of clouds in a NWP or climate model at the assimilation stage also means that model fields which depend on the cloud fields are better constrained - in complicated non-linear systems such as weather and climate, manual adjustment of the cloud fields in a model output is unlikely to be accompanied by an appropriate adjustment to other, dependent model fields. The work of this project has contributed to a step towards the improvement of automated assimilation of satellite data into such models.

1.1 Aims of the Project

The aim of this project is to expand a new, generic, cloud detection technique, which calculates a probability of cloud contamination for individual pixels, and is non-sensor-specific (Merchant

et al., 2005). Standard cloud detection for NWP is performed on imagery by threshold testing, which produces a mask of clear and cloudy pixels, (Saunders, 1986; Saunders and Kriebel, 1988). Some inaccuracies in masks produced this way stem from the method's lack of a sound physical foundation, and are both spatially and temporally variable. Algorithms are often sensor-specific, requiring substantial re-writing for new sensors (Merchant et al., 2005). The probability calculated by the new method, which this work aims to expand upon, is based upon the assimilation of background information from climatology and/or NWP forecast fields with satellite imagery, using Bayes Theorem (Bayes and Price, 1763). The technique is therefore physically based and exploits all the available information for each pixel.

This project aims to significantly extend the range of applications for which this Bayesian technique, developed at the University of Edinburgh, is suitable, by developing it to operate over land, see Mackie et al. (submitted 2008a) and section 5.1 on page 43, and to exploit visible and Near Infra-Red (NIR) wavelength imagery, in addition to the thermal infrared wavelengths it was initially designed to exploit, see Mackie et al. (submitted 2008b) and section 5.2 on page 48.

A possible improvement to the results of the Bayesian cloud detection, through clustering of pixels based on their spectral properties prior to running the cloud detection, will also be investigated, with the aim of reducing the occurrence of ambiguously classified pixels around ocean fronts.

Restructuring of the algorithm, which prior to this project was still at a development stage, to be more flexible, so that it can be more easily adapted to applications for new sensors and different NWP models will be a large part of this work, with the goal being to bring the development to a stage where it will be released under a public license.

A further aim of this project was to validate the Bayesian cloud detection technique using data from a sensor other than the Advanced Along Track Scanning Radiometer (ATSR), and an externally provided 'truth' against which to judge the skill of the detection. Previous validation of the technique for imagery acquired by ATSR sensor, see Merchant et al. (2005), was encouraging, but, as ATSR imagery is the basis of the textural PDF (see section 2 on page 12), validation using imagery from another sensor is arguably more appropriate. Furthermore, this work will verify the claim that the technique is not sensor-specific and can be applied to imagery from any sensor for which a Radiative Transfer Model (RTM) exists. The 'truth' used in the previous validation work was constructed by eye by the experts conducting the validation. An externally supplied 'truth', such as the project aims to use, is arguably a more objective measure against which to judge the skill of the cloud detection.

In order for the technique to be fully physically robust, there existed a need for a method to fast forward model cloudy atmospheric states, i.e. to predict observations that would be made by a satellite, given a particular atmosphere, in the presence of cloud. This will be developed as part

of this project. Section 6 on page 74 describes both the motivation for and the validation of this aspect of the work.

Aerosols such as Saharan dust often present a problem to cloud detection because the effect of dust on satellite observations is similar to the effect of cloud. Dust-detection is usually carried out for imagery after cloud detection, that is, after all pixels detected as cloud have been removed. This 2-step detection process means that dust can be erroneously detected as cloud and assimilated as such into a model, raising the uncertainty of the model output, and reducing the accuracy of dust storm forecasting. Through the work of this project, the Bayesian approach will be extended to consider 3 classes - dust, clear and cloud - giving an automated, one-step physically-based probability that a pixel corresponds to a dust observation, a cloud observation, or a clear sky observation, see section 7 on page 124.

1.2 Layout of the Thesis

This introductory chapter describes some of the problems associated with automated cloud detection in satellite imagery. Various methods that have been used to tackle the problem are discussed using examples from the literature. The work for this thesis follows a Bayesian approach and builds on an already existing technique. An understanding of the project work therefore requires an understanding of both the principles behind, and the structure of, this particular technique, and this is given in chapter 2. An investigation into the possible improvement offered by clustering pixels prior to processing for cloud detection is presented in chapter 4. Substantial extensions which were made to the Bayesian method as part of the PhD project are presented with some validation in chapter 5, notably the modifications which have made it possible to apply the technique to land imagery, in section 5.1, and to exploit imagery at visible wavelengths, in section 5.2. A fast forward model for cloudy sky observations at thermal wavelengths, which can be incorporated into the technique and so strengthen its physical robustness, is presented in chapter 6. In chapter 7, the cloud-clear classification algorithm is extended to consider Saharan dust as an additional class, producing a probabilistic 3-class result.

1.3 A Brief Background to Cloud Detection

Almost any estimate of geographical parameters from a satellite image in the reflectance or infra-red parts of the electromagnetic spectrum requires cloudy pixels to be identified. The clouds themselves can be investigated, or that clear pixels can be selected for study of the underlying surface or clear-sky atmosphere (Chen et al., 2003). A better understanding of clouds from detailed cloud studies, which depend on accurate detection, should lead to improvements in NWP and Global Climate Models (GCMs) (Cess et al., 1990) - for example better representation of upper tropospheric ice clouds has been linked to a better representation of monsoon and El Nino

effects (Li et al., 2005). Land (or sea) surface data, particularly temperature and emissivity, which are essential parameters in the energy budget of the Earth, are important to climate and hydrological applications and models, see for example Yu et al. (2005); Jang et al. (2006). Algorithms created to retrieve this information from imagery must either include some process to remove cloud-contaminated pixels, or require imagery to be screened for cloud contamination prior to the application of the algorithm, see for example Wan (1999).

The climate is largely determined by the amount of radiation reaching the earth, which depends on cloud cover (Kokhanovsky and Nauss, 2005), on the chemical composition of cloud (Li et al., 2005; Spang et al., 2005) and on the volume (and droplet size) of condensed water in the atmosphere (Liou, 1986; Cronin et al., 2005). The discrimination of clouds within satellite imagery, and the distribution of their micro-physical properties is essential to an understanding of radiative transfer through the atmosphere (Fu, 1996; Cronin et al., 2005). The applications of such understanding are wide-ranging, for example models of solar energy resources, see for example Wielicki et al. (1996), and studies of biological productivity (Liu and Gautier, 1990). The only realistic way to observe these variables on a global scale, for example for inclusion in a GCM, is through the discrimination of clouds in satellite-acquired imagery (Kokhanovsky et al., 2005). Temperature profiles, measured by satellite sounding infrared wavelengths, are simpler in the absence of cloud, which must otherwise be corrected for. Such profiles are used, amongst other things, for weather forecasting and so it is important to identify which pixels require correction, i.e. to detect cloudy pixels.

The volume of image data typically processed, as well as objectivity considerations, means that cloud detection is generally carried out automatically, i.e. through computational algorithms, a combination of textural and spectral parameters usually being used to identify cloud-contaminated pixels (Ameur et al., 2004). The Met. Office's NWP model, which produces the weather forecast in the U.K., is partially driven by cloud parameters. Algorithms created to retrieve this information from satellite imagery must either themselves include some process to detect cloud-contaminated pixels, or require every image scene to be screened for cloud contamination prior to their implementation Wan (1999).

There are many algorithms for cloud detection in satellite imagery in the literature, some application specific, and some more generic. Section 1.3.1 discusses a threshold-testing approach that is frequently used operationally, while sections 1.3.2, 1.3.3 and 1.3.4 discuss more application-specific methods, and section 1.3.5 introduces the Bayesian method which is focused on in this project. Cloud detection for an image can also be carried out through inspection by an 'expert', but this is a time-consuming process and is generally only used for small datasets for which a 'truth' is required for validation of an automated technique (see, for example Merchant et al. (2005)). The absence of an objective 'truth' against which to validate new cloud detection techniques on a large scale is a problem - 'validation' is often based on comparisons between results from a new technique and from an established one, which does not allow for the new

technique to improve on the old one (see for example Rossow et al. (1985); Ricciardelli et al. (2008)); or on synoptic observations made from the ground, see for example Bréon and Colzy (1998), which relies on clouds being low enough in the atmosphere to be distinct when viewed from the ground, and on the same clouds being seen from below as from above. Notwithstanding these criticisms, such comparisons can indicate the relative strengths and weaknesses of different techniques, and are often used as evidence of validation for a new technique.

1.3.1 Threshold Approaches to Cloud Detection

The discrimination of cloudy from clear pixels can be attempted on the basis of contrasts between the reflectivity, emissivity, temperature or spatial variability of clear areas from cloudy areas. Threshold schemes are based on delineating the boundaries between the signature of these factors in imagery. A 'threshold testing' approach is often used for operational cloud detection, whereby each pixel is individually subjected to a series of tests, failing any one of which leads to its being classified as cloud. Threshold values used for the tests are set by an expert experienced in image processing and in cloud detection. This approach is based largely on the methodology developed by Saunders (1986); Saunders and Kriebel (1988). It should be noted that in Saunders and Kriebel (1988), misclassification of clear pixels as cloud was considered preferable to misclassification of cloud pixels as clear, indicating a bias in that particular set of threshold tests. There are many examples in the literature, each testing pixels on the basis of a different set of threshold values - a cloud mask product is often automatically generated for, and supplied with, imagery in this way. For example, the image of Korea in figure 1.1, acquired by the Advanced ATSR sensor at a wavelength of 1.6μm, is supplied with the cloud mask shown next to it (the land is blanked out in red to make the mask clearer). Some inaccuracy is apparent in the automatically-generated product - for example, the middle section of the image is all classed as cloud, and there are unnatural blocky structures in the detected cloud. Such inaccuracies arise when thresholds defining the cloud masks are inappropriate for the circumstances of the image. Determining a widely applicable set of thresholds is complicated. A threshold value is set to discriminate clouds from the underlying surface using either a particular recorded wavelength, or a function of the recorded wavelengths. An expert, experienced in cloud detection, considers a number of image scenes in order to set these values, but it is unrealistic to expect the same threshold values to be equally appropriate to all variations of climate and of surface, even within a single surface type. The tests therefore suffer from inconsistent (and often unknown) accuracy.

Cloud-detection is generally more straightforward for day-time imagery, when data at both visible and thermal wavelengths are available, and so the demanded accuracy of some detection products is reduced during night-time, for example Hutchison et al. (2005). Some kinds of cloud are more easily detected than others, for example thin cirrus cloud has been noted to be particularly difficult to detect using threshold methods (Saunders, 1986). The accuracy of detection also varies with the properties of the underlying surface, for example detection is more accurate for imagery acquired over surfaces of approximately constant temperature and emissivity (Saunders,

1986). This is because the measurable spectral and textural properties of these surfaces vary slowly, making the contrast between them and those measured for cloud fairly constant. Imagery acquired over coastal areas, ocean fronts, or varying land-type presents problems, as thresholds appropriate for one pixel may not be appropriate for another. This can sometimes be addressed through the use of ancillary data, such as maps of land surface type or surface emissivity (Yu et al., 2005). The validation of a classification algorithm prior to its application to real data is often carried out using simulations which cannot fully allow for the highly complex variability of surface type on the real Earth (Yu et al., 2005), and so the effect of the underlying surface type being non-constant (i.e. of variable surface emissivity or temperature) on the accuracy of the final product often cannot be ascertained prior to the algorithm being applied to real data.

Most threshold methods do not produce an indication of product accuracy, although some algorithms do identify 4 classes rather than 2, separating clear pixels into 'confidently clear' and 'probably clear' classes, and similarly for cloudy pixels, for example Hutchison et al. (2005). Pixels with values close to a threshold value are more likely to be misclassified, and some algorithms do produce an array of confidence values for the classified image scene to allow for this, for example Ackerman et al. (1998), but this is not a measure of the accuracy of the threshold value itself, which is likely to vary spatially across an image scene, but rather of the clarity with which a pixel falls to one side of it. If the threshold value itself is inappropriate, or if its appropriateness varies across the image scene, the accuracy of the product will not necessarily be reflected in the confidence values generated. The variable, largely unknown, accuracy of cloud-detection across a large imaged area, or between repeat images of one place, is a short-coming of the threshold method.

Threshold testing algorithms can be limited by the local uniqueness of the spectral properties of the land (or ocean) surface. The spectral properties of a desert, for example, are different to those of a forest, which are in turn different to those of the ocean, making the algorithms 'scene-specific', rather than generic, and dependent on accurate land/ocean surface classification. Threshold tests are particularly difficult to apply to snow- and ice-covered surfaces, which are often difficult to discriminate from cloud (Ameur et al., 2004).

Threshold techniques have been further criticised as being inappropriate to the modeling of global-scale processes, which often require quantified parameterization of clouds which can be interpreted in terms of cloud processes, rather than the binary (clear-cloudy) mask generally provided by the techniques (Fouilloux and Iaquinta, 1998). Another problem is the unavoidable subjectivity of the threshold values set, which are biased both towards the imaged geographical location which was inspected in order to set them, and towards the expert's own experience, e.g. an expert who is very experienced in detecting cloud over ocean surfaces, may not detect cloud over forested surfaces or over sea ice with the same accuracy.

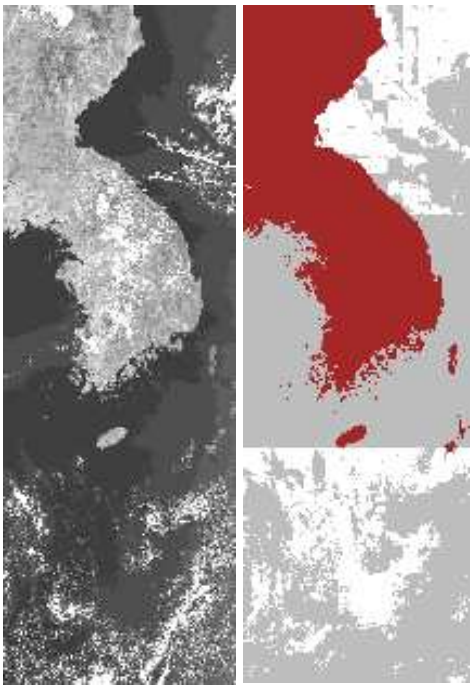


Figure 1.1: Example of an operational cloud mask produced by a threshold testing approach: Left shows the image as recorded at 1.6μm; Right shows the operational cloud mask supplied with the image (with the land blanked out in red). The image was recorded by the AATSR sensor above Korea at 11:11 UTC on May 10th 2005.

1.3.2 A Neural Network Approach to Cloud Detection

An alternative to the threshold approach is to use neural networks for cloud detection, for example Fouilloux and Iaquinta (1998); Simpson and McIntire (2001); McIntire and Simpson (2002). This method, which aims to accurately identify the driving parameters of the classification, has been demonstrated to be both more computationally efficient, and to produce more accurate results than the traditional threshold techniques described in section 1.3.1 (Fouilloux and Iaquinta, 1998).

Surface information such as temperature is particularly significant to atmospheric circulation in regions of sea ice and permanent snow cover (Hall et al., 2004), and so accurate cloud detection is important in these areas to both GCMs and NWP models, but is difficult under the approach of section 1.3.1. Neural networks have been built specifically for these regions - learning the scene-specific differences from a training set of data and 'feeding the information forward' to interpret the subject imagery. An image scene can then be separated into land, cloud and water classes, and the network enables mixed classification for pixels whose spectral and textural properties make their class ambiguous (Simpson and McIntire, 2001; McIntire and Simpson, 2002) (a mixed classification is also a possible interpretation of ambiguous results from the Bayesian technique). Neural networks have the advantage of using the actual data to determine the classification parameters which arguably makes the classification physically robust, as 'natural' boundaries are identified between classes, but also highly scene-specific as the boundaries are likely to vary from site to site. It is unlikely that networks developed for the Arctic, for example, could perform accurately in the Antarctic (McIntire and Simpson, 2002), or that a network optimised for classification in areas of sea ice would be accurate in an area of snow-covered land (Simpson and McIntire, 2001). The subjectivity of the threshold approach is not entirely overcome either, as the initial training data used by a neural network is generally the result of a scene classification carried out by hand by an expert. Any bias in other cloud detection methods, which could potentially be overcome by a neural network, could be inherent in the network if the expertise classifying the initial training data carries the same bias through experience mainly with those other methods.

1.3.3 Use of Oxygen Absorption Bands for Cloud Detection

In the case of sensors that record imagery at wavelengths in which various absorption spectra can been seen, this can be exploited to aid cloud detection. An example of this is land observations made by the MERIS sensor. The quality of atmospheric corrections made to MERIS data products depends on successful cloud detection - being difficult to detect and often of a large horizontal extent, thin cirrus cloud presents a particular problem, and this is addressed by the use of oxygen absorption bands (Borde et al., 2000). Radiance reflected by the surface and measured at the sensor at 0.760μm depends on oxygen absorption within the atmosphere, and so on the surface elevation. Surface pressure is calculated from the ratio of pixel observations made at 0.760μm to observations made at 0.755μm. Thin cirrus cloud corresponds to both high top of atmosphere

radiances in the near infra-red and low pressure, and so can be expected to affect surface pressure that is calculated in this way. Borde et al. (2000) found that the presence of thin cirrus cloud introduced errors of up to 150hPa to the calculated surface pressure. This effect on the surface pressure can be used as an indirect means of detecting thin cirrus cloud. Through study of a number of empirical simulations, Borde et al. (2000) found a threshold that could be set on the difference between surface pressure calculated from the oxygen absorption channel (by the method described), and the surface pressure provided by European Centre for Medium Range Weather Forecasting (ECMWF). This allows atmospheric corrections to imagery recorded by the MERIS sensor over land to be adjusted for the presence of thin cirrus cloud.

Despite being a useful tool to the specific problem of thin cirrus over land, this technique cannot be expected to be useful to the broader problem of detecting all cloud contamination in an image. It is, of course, only useful for sensors which record at the appropriate wavelengths, and therefore is not applicable to any night-time acquired imagery. It relies on a threshold set by an expert who is experienced in cloud detection, making some of the criticisms made in 1.3.1 applicable here also.

1.3.4 Temporal Coherence Techniques for Cloud Detection

A further cloud detection technique uses imagery acquired a short time before the imagery being processed, by the same sensor and for the same location. This technique is usually used to distinguish *candidates* for cloud or clear observations, and the actual distinction is generally made by further tests. The method will be more computationally efficient if it considers fewer images, and so often only two images are used, making it difficult, in the case of a marked difference between the two, to ascertain which image is clear and which is cloud. The work of Massons et al. (1998) on land and sea imagery over the Iberian peninsula observed by the Meteosat sensor demonstrates how the temporal difference method works, and how it can be combined with further cloud detection tests. Each image in that study was 512x512 pixels, and was divided into 8 geographically homogeneous regions for processing. For each region, the visible and infra-red observations were analyzed with the equivalent visible and infra-red observations from 1 hour previously. The temporal coherence between the two image pairs was calculated - pixels corresponding to low temporal variability were considered candidates for clear observations. Such pixels were not assumed at that point to be clear, as static clouds may have high temporal coherence, and similarly not all clear pixels were assumed to correspond to low temporal variability. It *was*, however, assumed that a representative sample of clear pixels had been collected, albeit with some small contamination from cloud observations. Histograms of the spectral properties of high temporal coherence pixels were calculated for each region and used to define thresholds on brightness temperature and reflectance. These thresholds were then used to discriminating cloud from clear observations. In this way, clear pixels missed by the temporal coherence test are collected. Using this technique, Massons et al. (1998) found successful results, and noted that the temporal coherence test performed cloud detection ‘acceptably well’ when used alone, without the further

processing steps described above. The study used only day-time imagery, however, similar results are reported (within 8%) for the same technique when applied to only infra-red observations, indicating that it could be expected to perform as well for night-time imagery (Massons et al., 1998).

As an approach, temporal coherence tests have the advantages of requiring very little a-priori data (only reference imagery from a previous observation), and of discriminating clouds which fill only a fraction of a pixel, and broken clouds, which other methods can have difficulty in detecting. In common with other cloud detection techniques, temporal coherence-based methods generally perform better for sea imagery than for land imagery. Errors in the detection are generally associated with clouds which introduce little variation to the radiance recorded at the satellite - in Massons et al. (1998), detection of thin high level cloud and low level stratus was seen to be less successful than detection of other cloud types. While this method is both successful and fairly computationally efficient (Massons et al., 1998), it can only be applied to cases where the same location is imaged by two identical, or equivalent, sensors within a relatively short period of time.

1.3.5 A Bayesian Approach to Cloud Detection

An approach using physically-based probabilities has been proposed as an alternative to the approaches in sections 1.3.2 and 1.3.1 (Merchant et al., 2005). It aims to avoid the problems associated both with the subjective setting of threshold values, and with the scene- and time-specific appropriateness of the thresholds set. Scene-specific climatology and NWP fields are used as a-priori information, which is combined with the observed image data to calculate a probability of the observations recorded for a pixel having been made in clear-sky conditions, following Bayes' Theorem for Conditional Probabilities (Bayes and Price, 1763). The result is therefore a probabilistic cloud mask - the value for each pixel being the calculated probability of that pixel being clear, given the observation data and the background NWP data.

NWP fields are used to calculate a Probability Density Function (PDF) of observations corresponding to a clear scene using a Radiative Transfer Model (RTM). The value of this PDF for a given observation is the prior probability of that observation assuming a clear (cloud-free) pixel, and an atmosphere described by the NWP fields, with their associated uncertainties. This is combined with both the prior probability of cloudiness at the pixel location, taken from cloud statistics, and the PDF for the observation assuming the presence of cloud, to calculate the posterior probability of clear for the observed pixel using Bayes' Theorem for Conditional Probabilities. Similar work has been done to detect cloud in imagery recorded at microwave wavelengths (English et al., 1999), although, in this case, probabilities were not calculated.

The Bayesian method arguably has several advantages over the threshold methods in section 1.3.1. Not being dependent on subjectively set threshold values, it is more objective than expert-based methods; it produces a probability rather than a binary classification, and so is more flexible, allowing users to vary the 'severity' of the cloud mask to suit the application; it is generic

and so is easier to apply to sensors with different spectral responses than threshold methods; it exploits systematically available prior information about an imaged scene; and it has a sound, transparent, physical basis.

A similar method has been successfully applied in the case of operational cloud detection for the Advanced Very High Resolution Radiometer (AVHRR) sensor (Uddstrom et al., 1999), however in that work PDFs for cloud and clear were constructed in the observation space from empirical training datasets, rather than from scene-specific NWP fields. This shares the advantage of avoiding 'synthetic' threshold values, but does not have the spatial- and temporal- dependence of the method used here.

Bayes' Theorem has already been applied to some areas of image processing, for example to the problem of land cover classification in the absence of up to date training data. Old training data sets were used with the corresponding old image to calculate the most probable class for each pixel in the up to date image. The spectral and textural information from the old training data and image scene provided the a-priori information, which was combined in Bayes' Theorem with the up to date imagery to give the most likely modern class for each pixel (Cossu et al., 2005). The Bayesian approach proposed for cloud detection is a variation on this, with the added difficulty of a more complex data set, i.e. rather than identifying all the different surface types in an image scene, the technique must identify clouds which could appear spectrally and texturally different over each surface type.

The Bayesian approach has the advantage over the other described techniques of being physically valid and, theoretically at least, being easy to apply to data recorded by any satellite sensor, giving it global applicability and meaning it avoids the subjectivity issues of some other methods. It has the potential to improve because it relies on the exploitation of NWP data, the accuracy of which is continually improving, and produces a result in which the certainty for the cloud or clear classification of individual pixels is given, allowing ambiguous, uncertain classifications to be discarded. Meteorological observing satellites often do not image at wavelengths for which the use of absorption spectra is appropriate, and a time series of imagery is not always available for NWP. This work therefore focuses on the Bayesian cloud detection approach as the most promising for real-time, global applications such as Numerical Weather Prediction (NWP). The following chapters present investigations into improving the results from such an approach, and developments that allow it to be adapted to different types of imagery. Some validation of different aspects of the technique is also presented. The skill of the method is judged relative to that of traditional operational methods, which rely mainly on threshold testing, with conclusions included in each chapter, and summarised in chapter 8.

Chapter 2

Structure of the Bayesian Method for Cloud Detection

Much of the work of this project is centred around a Bayesian Cloud Detection technique developed at the University of Edinburgh (Merchant et al., 2005). In order to understand the work presented in this thesis, and some of the motivation for it, it is therefore necessary to understand the general principles and structure of the method as it stood before the PhD project began. This chapter provides an outline of the technique, and contains some of the equations referred to in later parts of the thesis. Except where explicitly stated, for example section 2.3, the work described in this chapter was partly carried out by other people in the research group - the contributions made by this PhD project to this chapter involved testing, modifying and adding to the software already written, bringing it to its present state, in which it has been released under public license.

It was shown in Merchant et al. (2005) that the probability that a pixel in a satellite image records an observation made in clear sky conditions can be found by evaluating equation 2.1, following Bayes' Theorem for conditional probabilities. This technique was initially developed as a means to improving retrievals of sea surface temperature and, prior to this project, was only applicable to Thermal Infra-Red (TIR) imagery acquired over ocean.

$$P\left(c\left|\mathbf{y^o},\mathbf{x^b}\right.\right)=\left[1+\frac{P\left(\bar{c}\right)P\left(\mathbf{y^o}\left|\mathbf{x^b},\bar{c}\right.\right)}{P\left(c\right)P\left(\mathbf{y^o}\left|\mathbf{x^b},c\right.\right)}\right]^{-1} \tag{2.1}$$

The observation vector, $\mathbf{y^o}$, may include brightness temperatures (BTs) and reflectances recorded by the satellite at each of its channel wavelengths, and the local standard deviation of these. The background state vector, $\mathbf{x^b}$, contains information taken from NWP fields, which is used as a priori information to model an expected clear-sky observation for the scene, $\mathbf{y^b}$. The

probabilities on the right-hand-side of equation 2.1 are the season- and latitude-dependent prior probabilities of the pixel location being clear, $P(c)$, and not clear, $P(\overline{c})$. $P\left(\mathbf{y^o}\left|\mathbf{x^b}, c\right.\right)$ is the conditional probability of the observation, $\mathbf{y^o}$, occurring given the prior data, $\mathbf{x^b}$, and clear-sky conditions, c. $P\left(\mathbf{y^o}\left|\mathbf{x^b}, \overline{c}\right.\right)$ is the equivalent probability, given cloudy, instead of clear sky, conditions.

An underlying assumption of this method is that 'clear' and 'cloudy' represent all possible atmospheric states for a pixel, i.e. their probabilities sum to one, even if the true reason for deviation from clear-sky observations is, for example, a heavy loading of atmospheric aerosol. In this chapter, therefore, 'cloudy' is really 'not clear'. The remainder of the chapter discusses in turn how the elements comprising 2.1 are estimated.

2.1 Observation Vector

The basic principle behind the technique is that equation 2.1 finds the probability of clear, given $\mathbf{y^o}$ and $\mathbf{x^b}$, on the basis of contrasts between the likelihood of of the observation vector being recorded under clear and cloudy conditions.

At the beginning of this project, the observation vector $\mathbf{y^o}$ contained the observed Top of Atmosphere (TOA) BTs - chapter 5 describes the extension of the technique to include reflectance data in the observation vector, but here the method as applied to thermal wavelength data is described. Clouds generally correspond to lower-valued observations at thermal wavelengths because they are colder. This difference, however, is not necessarily equally apparent for each imaged thermal wavelength. The typical channels considered are centred around 3.9μm, 11μm and 12μm. For daytime imagery, NIR wavelengths (such as 3.9μm) are not used (the fast RTM which is used does not simulate the solar contribution to TOA BT for this channel).

Cloud-top temperatures are often more spatially variable than underlying STs, especially over ocean, providing additional information that is included in the observation vector for ocean pixels. The local standard deviation (LSD) of the TOA BTs is calculated on a 3x3 grid of pixels centred on the active pixel, $LSD = \left(\frac{1}{(n-1)}\sum_{i=1}^{n}\left(\mathbf{y_i} - \overline{\mathbf{y}}\right)^2\right)^{\frac{1}{2}}$, where $\mathbf{y_i}$ contains the observed channel BTs for the ith pixel, $\overline{\mathbf{y}}$ contains the mean observed BT for each channel over the 3x3 grid of pixels and n is the number of pixels in the grid.

2.1.1 Use of the Observation Vector

The range of possible LSDs and thermal observations that are possible for a clear sky overlaps with the range of those that are possible for a cloudy sky. It is therefore not appropriate to consider observations from a single channel, or LSD, in isolation - while a pixel may correspond to an observation at 12μm that would usually be associated with clear sky, observations made at other

wavelengths and the LSD may make it more probable that the pixel corresponds to a cloud observation. The observations for different channels, and the LSD are therefore considered together as coordinates in a multi-dimensional observation space. A normalized distribution of the possible cloud observations populates the space, and overlaps with an analogous distribution of possible clear sky observations - these are referred to as the cloud and clear Probability Density Functions (PDFs), and are described in sections 2.5 and 2.4. The value of each of these two distributions corresponding to $\mathbf{y^o}$ is read from the distributions and the two values are combined using Bayes Theorem to calculate the probability of clear for the observation (in the clear sky case, the value of the distribution corresponding to $\mathbf{y^o}$ is calculated directly, see section 2.4).

2.2 Background State Vector

The background vector $\mathbf{x^b}$ contains the NWP fields used by the RTM to predict the BTs for clear-sky conditions, typically profiles of atmospheric temperature and water-mixing ratio, and ST. Surface emissivity, calculated from the NWP data, is also included in $\mathbf{x^b}$ and used by the RTM, see section 2.4.2. These data can be sourced from a numerical weather forecast, or from climatology. In the case of imagery at a higher spatial resolution than the NWP information, the modelled BTs are bi-linearly interpolated to each pixel's geographical location from the closest available profiles. At coastal boundaries, only profiles of the same land/sea class as the pixel are used in the interpolation.

2.3 Prior Probabilities of Clear and Cloudy

When this project began, the prior probabilities for clear and cloud, $P(c)$ and $P(\overline{c})$, were fixed at 0.1 and 0.9 respectively. In reality, climatological cloudiness is spatially and seasonally variable, and so prior probabilities were calculated from ISCCP data to address this.

The freely available ISCCP seasonal total cloud amount data (*http://isccp.giss.nasa.gov/products/browsed2.html*) gives the mean seasonal total cloud amount on a 2.5^o equal area grid over the period July 1983-June 2006, as recorded by a number of sensors. The predominant variability in cloudiness is latitudinal and so for computational efficiency the mean total cloud amount was calculated for 2.5^o latitude bands, with land and sea considered separately. Grouping the ISCCP data into four seasons, corresponding to December, January, February (DJF); March, April, May (MAM); June, July, August (JJA); and September, October, November (SON), and taking the mean value for each latitude band in each season, look-up-tables (LUTs) of latitude- and season-specific values for $P(c)$ for land and ocean were calculated, allowing calculation of $P(\overline{c}) = 1 - P(c)$.

Given that the LUTs are averaged over such a large amount of ISCCP data, the estimated error in these LUTs is small, but non-negligible biases may still exist (Rossow et al., 1993). The

ratio between $P\left(\mathbf{y^o}\left|\mathbf{x^b},c\right.\right)$ and $P\left(\mathbf{y^o}\left|\mathbf{x^b},\overline{c}\right.\right)$ typically differs from the ratio between $P\left(c\right)$ and $P\left(\overline{c}\right)$ by several orders of magnitude, and so the former terms usually dominate equation 2.1. The requirement is therefore that $P\left(c\right)$ and $P\left(\overline{c}\right)$ have a correct order of magnitude and biases and uncertainties in $P\left(c\right)$ and $P\left(\overline{c}\right)$, carried from the ISCCP data, are likely to be acceptably small. Modifying the prior probability likewise has an effect on the posterior probability only in borderline cases, where $P\left(\mathbf{y^o}\left|\mathbf{x^b},c\right.\right)$ and $P\left(\mathbf{y^o}\left|\mathbf{x^b},\overline{c}\right.\right)$ are found to be similar.

2.4 Conditional Probability of Observation given Clear Sky

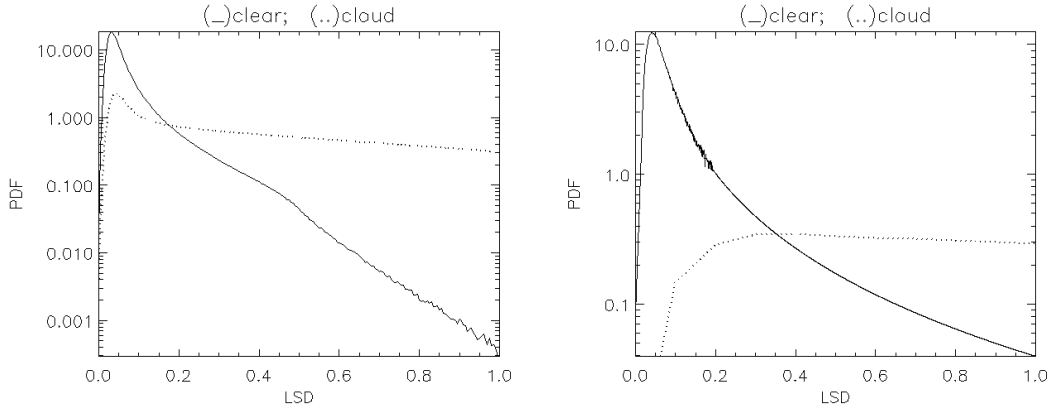
For ocean pixels, the conditional probability of the observation corresponding to clear sky $P\left(\mathbf{y^o}\left|\mathbf{x^b},c\right.\right)$ has a 'textural' component (relating to measures of spatial variability) and a 'spectral' component (relating to radiances or brightness temperatures), which are assumed to be independent. See equation 2.2, where the subscripts s and t refer to the spectral and textural components. $P\left(\mathbf{y^o}\left|\mathbf{x^b},c\right.\right)$ forms part of the denominator in the expression for the posterior probability that the observation corresponds to clear sky in equation 2.1.

$$P\left(\mathbf{y^o}\left|\mathbf{x^b},c\right.\right)=P\left(\mathbf{y_s^o}\left|\mathbf{x^b},c\right.\right)\times P\left(\mathbf{y_t^o}\left|\mathbf{x^b},c\right.\right) \tag{2.2}$$

2.4.1 Textural Probability

The textural conditional probability, given clear sky, $P\left(\mathbf{y_t^o}\left|\mathbf{x^b},c\right.\right)$, is the conditional probability of the observed LSD in BT (the LSD component of $\mathbf{y^o}$), given $\mathbf{x^b}$ and assuming a clear-sky pixel. At present, the technique assumes $P\left(\mathbf{y_t^o}\left|\mathbf{x^b},c\right.\right)=P\left(\mathbf{y_t^o}\left|c\right.\right)$, i.e. the textural probability of the observation corresponding to clear sky is assumed independent of NWP or other pixel-specific background information. Clear-sky atmospheric variability is generally negligible over a 3x3 grid of pixels at the 1 to 5km spatial resolution of meteorological imagers, so the main contributing factors to the LSD are assumed to be sensor radiometric noise and spatial variability in the ST. The LSD in the 11μm channel was taken as the observation, and the effect of radiometric noise was modelled as Gaussian in BT, with standard deviation given by the noise equivalent differential temperature (NEΔT) at 285K. The PDF for the LSD over n realisations of a Gaussian random variable is a chi-squared distribution with number of degrees of freedom n-1 (n-1 rather than n because the mean of the values is used in the calculation of the LSD). Real spatial variability in ST will also increase the LSD, and so must be accounted for to avoid classifying all ocean fronts and eddies as cloud. An empirically-determined PDF for the ST LSD is therefore used in preference

to the chi-squared distribution, since the distribution of ocean fronts is not an analytic function. The empirical PDF was generated from a set of nine full days of daytime-only nadir-viewed imagery acquired at 1km resolution by the sensor ATSR-2. A threshold of 4% was set on the reflectance recorded at 1.6μm to select mainly clear pixels. To minimize contamination of the sample by cloudy pixels, only pixels found by the threshold to be clear, and surrounded by a 5x5 box of pixels also found to be clear, were considered. A one dimensional PDF LUT was formed by binning the LSD values for the 11μm TIR channel (scaled to account for $\frac{}{\partial BT_{11\mu m}}\partial ST$) into bins that clearly resolved the PDF's peak. Nadir-viewed imagery from ATSR-2 was chosen for its high spatial resolution, and the low sensor noise at 11μm (NEΔT $\leq$ 0.05K). To create PDFs for sensors of lower spatial resolution, such as the geostationary imagers, the ATSR-2 clear pixel BTs were aggregated into boxes of the appropriate size to reduce the resolution and the corresponding LSD PDF was re-calculated. To create PDFs for noisier sensors, the PDF for ST LSD at the appropriate resolution is convolved with the chi-squared distribution appropriate to each sensor's noise level. The resulting PDFs for clear for the ATSR sensor, and for the SEVIRI sensor, which has 5km spatial resolution at nadir, are shown in figure 2.1.



(a) (b)

Figure 2.1: Textural Probability Density Functions (PDFs) for cloud(_) and clear(..) calculated for observations made at 11μm: (a) from the ATSR-2 sensor with high spatial resolution; (b) scaled for the lower spatial resolution of the MSG sensor. LSD is in units of Kelvin per pixel. Less variability is expected to be apparent for bigger pixels - differences on a small spatial scale having been 'smoothed' out relative to imagery from a sensor with higher spatial resolution. The textural PDFs for clear and for cloud are described in sections 2.4.1 and 2.5.1 respectively.

2.4.2 Spectral Probability

The spectral probability that the observations correspond to clear, $P\left(\mathbf{y_s^o}\left|\mathbf{x^b}, c\right.\right)$, is calculated using an RTM to model a predicted observation, $\mathbf{y^b}$, based on $\mathbf{x^b}$ and on surface emissivity. ST-dependent refractive indices from Newman et al. (2005) were used for the 11 and 12μm emissivities, and from Pinkley et al. (1977) for 3.9μm. To account for non-direct emissivity, i.e. emissivity reflected from neighbouring wave planes onto the sensor, an additional term was added

from Watts et al. (1996). A fuller description of the emissivity model used to produce the LUT is given in Filipiak and Merchant (submitted 2008).

The spatial resolution of NWP models is generally low compared to climatological features of the ocean surface temperature, and the method allows a high spatial resolution (5km) climatology to be used to estimate observations for clear sky at a higher spatial resolution for the ocean cases. The ST-dependence of the model estimate, $\left(\frac{\partial \mathbf{y^b}}{\partial x^b_{st}}\right)$, is used to adjust the estimate for the spatial variations in ST which are expected to be present, according to the climatology, between the nearby NWP locations. Simulations from the four geographically closest NWP profiles to each pixel location are adjusted in this way and then weighted (by spatial bi-linear interpolation) to give pixel-specific estimates of clear sky BTs.

Although the full state vector, $\mathbf{x^b}$ is used to drive the RTM, the NWP Total Column Water Vapour (TCWV) and ST fields have been shown to dominate any variability in TOA BTs (Merchant et al., 2006c). The full PDF is therefore approximated by considering only the variability in these fields. For the further processing of the BTs modelled by the RTM, a reduced background state vector, **z**, is introduced: $\mathbf{z} = \begin{pmatrix} x^b_{st} \\ x^b_{tcwv} \end{pmatrix}$, as was also used for optimal estimation in Merchant et al. (2008).

Assuming Gaussian distributed errors in **z** and $\mathbf{y^b}$ (Rodgers, 1976), $P\left(\mathbf{y^o_s} \left| \mathbf{x^b}, c \right.\right)$ is given by equation 2.3, where $\mathbf{H'} = \nabla_{\mathbf{z}} \mathbf{y^b}$ is the tangent linear of the RTM, with respect to each of the fields in **z** (see below); **B** is the error covariance of **z**; **R** is the combined error covariance of the RTM and observations, and *n* is the number of spectral elements in the observation vector. Equation 2.3 returns an *n*-dimensional probability density function (PDF) where each dimension represents one of the BT channels in $\mathbf{y^b}$.

$$P\left(\mathbf{y^o_s} \left| \mathbf{x^b}, c \right.\right) = \frac{\exp\left[\frac{-1}{2}\left(\mathbf{y^o} - \mathbf{y^b}\right)^T \left(\mathbf{H'^T B H'} + \mathbf{R}\right)^{-1} \left(\mathbf{y^o} - \mathbf{y^b}\right)\right]}{(2\pi)^{\frac{n}{2}} \left|\mathbf{H'^T B H'} + \mathbf{R}\right|^{\frac{1}{2}}} \tag{2.3}$$

The tangent linear matrix $\mathbf{H'}$ (following the notation of Ide et al. (1997)) is defined as $\mathbf{H'} = \frac{\partial \mathbf{y^b}}{\partial \mathbf{z}}$, and can either be calculated directly from a fast RTM (such as RTTOV (Saunders et al., 2005), which is used for this study), or can be found through perturbation of the elements of **z**. Although a full $\mathbf{H'}$-matrix, considering all the elements of $\mathbf{x^b}$, could be used, it would be computationally expensive and would require a full **B**-matrix to be known (not trivial) and loaded into the algorithm. With the reduced state vector, the covariance matrix is given by equation

2.4, where ϵ^b_{st} and ϵ^b_{tcwv} are the errors in the background state variables ST and TCWV, and are assumed to be independent. The errors in ST and TCWV are specific to the NWP model being used - if they are not provided with the NWP data then they can be found. The ST retrieved for clear sky pixels can be compared to the NWP ST to obtain a value for the error in NWP ST - a typical value is 0.88K. The error in TCWV can then be found by assuming all variance in the (modelled - observed) BTs to stem from errors in TCWV and ST. Equation 2.5, in which VAR refers to the variance, shows how the fractional error, f in TCWV can be found, giving $\epsilon^b_{tcwv} = f \times TCWV$ - a typical value is f = 21%. This method of error estimation was used in the initial validation of the technique (Merchant et al., 2005), and is presented in more detail in Mackie et al. (submitted 2008a).

$$\mathbf{B} = \begin{pmatrix} \left(\epsilon^b_{st}\right)^2 & 0 \\ 0 & \left(\epsilon^b_{tcwv}\right)^2 \end{pmatrix} \tag{2.4}$$

$$f^2 = \frac{VAR\left(\mathbf{y^o} - \mathbf{y^b}\right) - VAR\left(RTM\right) - \left(Ne\Delta T\right)^2 - \overline{\left(\frac{\partial \mathbf{y^b}}{\partial x^b_{st}} \times \epsilon^b_{st}\right)}}{\left(\frac{\partial \mathbf{y^b}}{\partial x^b_{tcwv}} \times TCWV\right)^2} \tag{2.5}$$

The forward modelling error (FM error) from the RTM and the observation uncertainty (sensor noise) are accounted for in the **R**-matrix. The RTM error, ϵ^{FM}, was estimated by comparison with a line-by-line model - only a very weak dependence on atmospheric path length was seen, possibly because a comparison between two model outputs means that systematic errors common to both models will not be found. In the absence of a 'truth', however, this was judged the most practical way to quantify the uncertainty in the RTM predictions. The observation errors, $\epsilon^{NE\Delta T}$ are taken to be the sensor's channel $NE\Delta T$, creating the matrix in equation 2.6, where N is the number of spectral channels used in the calculation (for night-time imagery, N=3, corresponding to 12, 11 and 3.9μm). Off-diagonal elements in the matrix are assumed to be 0.

$$\mathbf{R} = \begin{pmatrix} \left(\epsilon^{FM}_1\right)^2 + \left(\epsilon^{NE\Delta T}_1\right)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \left(\epsilon^{FM}_N\right)^2 + \left(\epsilon^{NE\Delta T}_N\right)^2 \end{pmatrix} \tag{2.6}$$

2.5 Conditional Probability of Observation given Cloud

The conditional probability of the observation corresponding to cloudy sky, $P\left(\mathbf{y^o}\left|\mathbf{x^b},\overline{c}\right.\right)$, is also composed of spectral and textural factors, with the assumption that cloud texture is independent of cloud top temperature, see equation 2.7.

$$P\left(\mathbf{y^o}\left|\mathbf{x^b},\overline{c}\right.\right)=P\left(\mathbf{y_s^o}\left|\mathbf{x^b},\overline{c}\right.\right)\times P\left(\mathbf{y_t^o}\left|\mathbf{x^b},\overline{c}\right.\right) \tag{2.7}$$

Predicting observations for cloud using an RTM, as for the clear sky case, is difficult, as a distribution of predictions is required - representing clouds at different altitudes, with different optical depths and filling different fractions of the pixel. The investigations of this project into a pixel-specific calculation for $P\left(\mathbf{y^o}\left|\mathbf{x^b},\overline{c}\right.\right)$ are presented in section 6. For computational efficiency, however, an RTM is generally not used in the cloudy case. Independence from the background state is assumed, i.e. $P\left(\mathbf{y^o}\left|\mathbf{x^b},\overline{c}\right.\right)=P\left(\mathbf{y^o}\left|\overline{c}\right.\right)$ and normalised PDFs of cloud observations from a global dataset are used as a LUT for the spectral part of $P\left(\mathbf{y^o}\left|\mathbf{x^b},\overline{c}\right.\right)$.

2.5.1 Textural Probability

A LUT for the cloud-sky textural PDF has been developed from an extensive sample of cloud-screened ATSR imagery, by a method analogous to the method for accounting for spatial variability of ocean temperature in clear-sky texture, described in section 2.4.1, see figure 2.1.

2.5.2 Spectral Probability

To generate PDFs of $P\left(\mathbf{y^o}\left|\overline{c}\right.\right)$ as LUTs in BT-space, an RTM and a global cloud climatology were used. The RTM was run for a large number of atmospheric profiles, taken from ECMWF ERA-40 data (available from *http://www.ecmwf.int/products/data/archive/descriptions/e4/*), to calculate clear sky optical depth. Cloud properties, including joint histograms of cloud-top height and cloud optical thickness, were sampled from a cloud climatology (e.g. from ISCCP or MODIS level 3 data, available from *http://isccp.giss.nasa.gov/; http://modis.gsfc.nasa.gov*). The sampled cloud layer was combined with the atmospheric profile, and the TOA BTs were calculated using a delta-Eddington model (Deeter and Evans, 1998) with cloud single scattering albedo and asymmetry properties taken from Hu and Stamnes (1993) for water clouds, and from Baran et al. (2003) for ice clouds. Each atmospheric profile was combined with a large number of the sampled cloud properties, creating a smooth distribution of TOA BTs. This was

normalized and is used as a LUT, each element of which corresponds to a point in BT-space and holds the prior probability of those BTs being observed for a cloudy atmosphere, i.e. $P(\mathbf{y^o}|\overline{c})$. Two-dimensional representations of this (for night-time imagery) 3-dimensional LUT are shown in figure 2.2, for daytime imagery, only 11 and 12μm imagery is considered and a 2-dimensional LUT is used. Although still not scene-specific, this is an advance on the earlier empirical method employed by Merchant et al. (2005), since no existing operational cloud mask is relied upon in generation of this LUT.

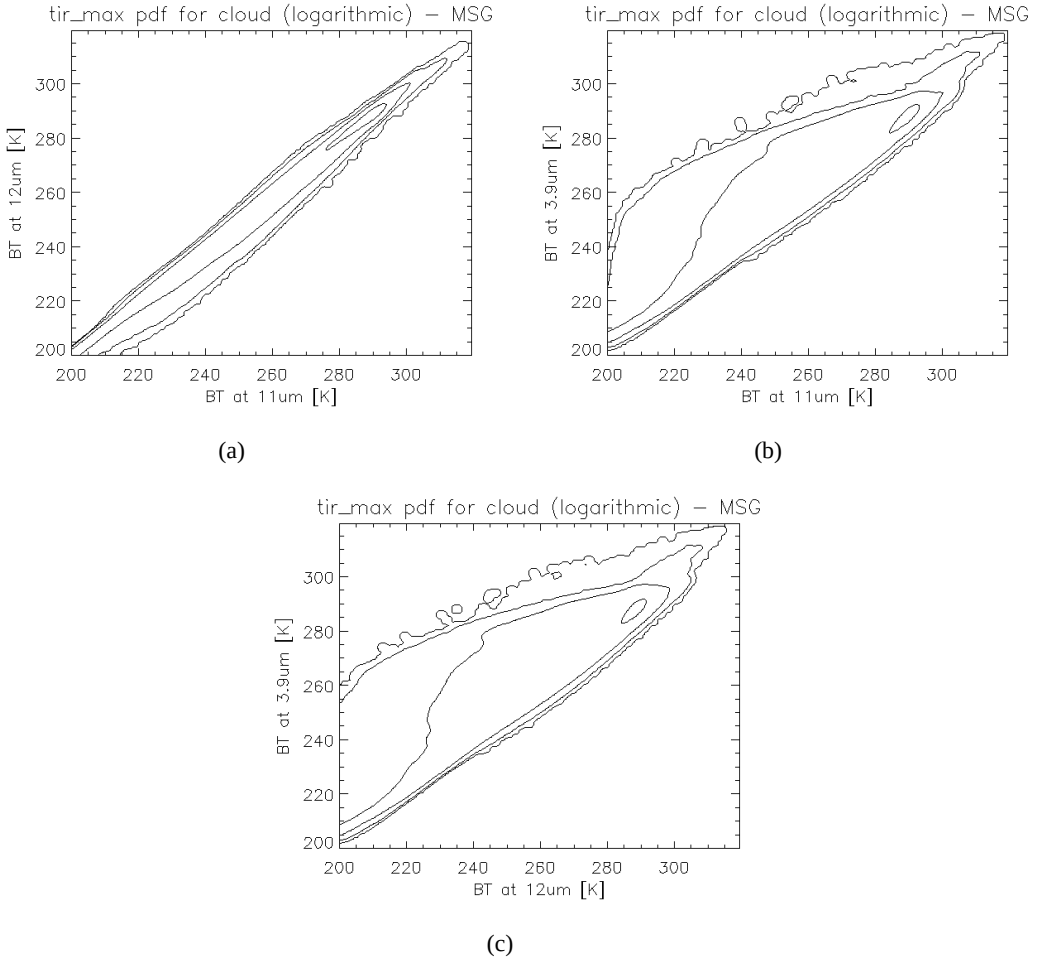


Figure 2.2: Spectral Probability Density Function (PDF) for cloud. The PDF is 3-dimensional, with dimensions corresponding to 3.9, 11 and 12μm; the 2-dimensional plots here show it summed over the dimension corresponding to (a) 12μm, (b) 11μm, and (c) 3.9μm. Contours are a logarithmic intervals, corresponding to $10^{[-8,-6,-4,-2]}$K^{-2}.

2.6 Implementation of the Bayesian Method

The concept of classifying each pixel as clear or cloudy using equation 2.1 is straightforward. For the spectral and textural components of the observation vector, the likelihood of the observations given the background information is found, given that the pixel is clear and given that it is

cloudy. These conditional likelihoods then modify the prior probability of cloudiness to give a final estimate, via equation 2.1. Three of the conditional probabilities (textural for clear and cloudy, spectral for cloudy) are read from pre-calculated LUTs. The fourth requires forward modelling of expected clear-sky observations, their tangent linears, and estimation of the error covariance matrices for NWP and observation fields. An overview of the software framework for implementing the Bayesian technique is given in Figure 2.3.

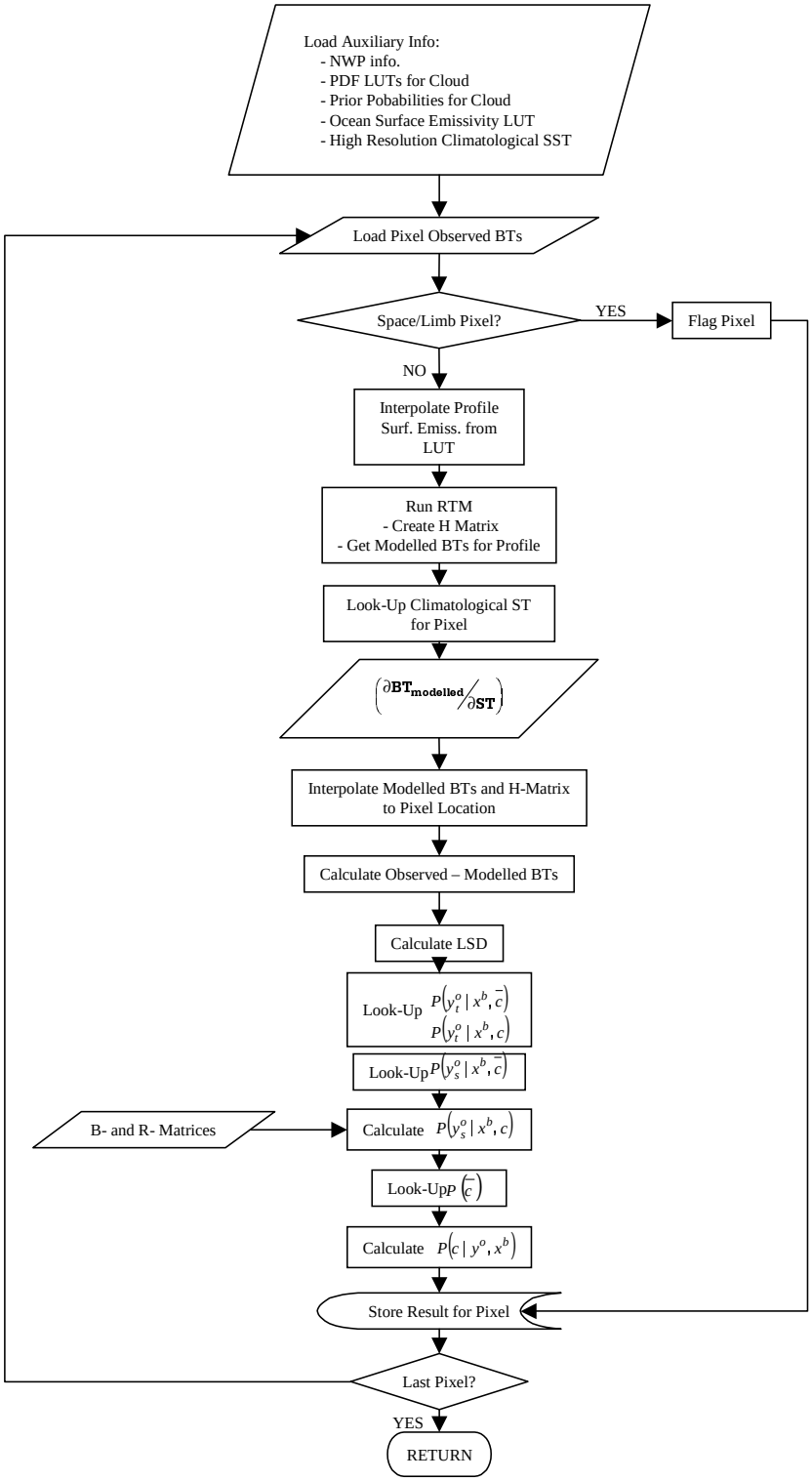


Figure 2.3: Software framework for the Bayesian technique

2.7 Brief Discussion

The work of this project builds on the Bayesian technique described in this chapter, investigating and demonstrating possible improvements to the method. Some of the limitations of this basic form of the technique, which are addressed in later chapters, are discussed briefly here.

When this project began, ambiguous probabilities of clear were being calculated for some pixels in regions of high thermal gradients, for example around ocean fronts. Ideally, the calculation should give a probability of clear very close to either 1 or 0, corresponding to a classification of either clear or cloud. When the probability of clear is ambiguous, an application specific tolerance can be set, depending on whether it is preferable to risk misclassification of clear pixels as cloud or cloud pixels as clear, or ambiguous pixels can be excluded, reducing the volume of data available from the image. The ambiguity seen for pixels around ocean fronts can be attributed to the textural probability of the observation corresponding to clear. Figure 2.1 shows the distribution of LSDs for clear to be narrow, with LSDs that are reasonable for ocean fronts corresponding to a higher value in the textural PDF for cloud than for clear. Chapter 4 investigates a possible method for solving this problem.

Although in principle the technique is not sensor-specific, prior to this project it had only been validated for imagery from one sensor. Further validation, using data from another sensor, and some restructuring of the algorithm to make application to data from new sensors more straightforward, means that the algorithm is now more flexible and useful to a wider range of users. The algorithm was designed as a tool to aid retrieval of sea surface temperatures, and, in the initial form described in this chapter, was not suitable for land imagery applications. Extension of the technique to land imagery makes the algorithm available and useful for a wider range of users. In this basic form, the technique did not exploit imagery at visible wavelengths, limiting its performance for daytime scenes when only observations at 11μm and 12μm channel were used. It has now been extended to exploit both visible and thermal wavelength data.

This basic form of the algorithm had the potential to produce reliable and accurate cloud detection results for any satellite sensor for which an RTM exists. It was, however, only suitable for a limited range of applications (those exploiting thermal wavelength imagery over oceans), and had only been validated in the absence of an objective external ‘truth’ against which to compare the cloud detection results. These shortcomings are described in more detail and addressed in chapter 5.

One of the strengths of the Bayesian approach to cloud detection is that it has a sound physical basis. Whereas traditional threshold-testing approaches generally rely on the expertise of an individual or individuals, this method exploits pixel-specific information and the uncertainties associated with it to make a pixel-specific calculation of the probability that an observation corresponds to clear sky. To be fully physically robust, the probability of the observation corresponding to cloudy sky conditions should also be based on pixel-specific information. The

problems associated with this, and a demonstration of a possible solution, are presented in chapter 6.

The technique relies on the assumption that clear and cloud are the only possible atmospheric states, meaning aerosol must either be identified as cloud or clear sky, when in reality it does not belong in either class. A discussion of some of the problems that stem from this, and a possible solution, is presented in chapter 7.

The basic form of the Bayesian approach to cloud detection which is outlined here is expanded on in the chapters of this thesis, which demonstrate several methods by which it has been, or could be, strengthened.

Chapter 3

Outline of Data Used in the Following Chapters

This chapter describes the data used for validation of the techniques described in chapters 4 to 7, and the imagery used in the investigations described in the the same chapters.

3.1 Satellite Sensors

3.1.1 ATSR Sensors

The Along Track Scanning Radiometer (ATSR) series of instruments have a spatial resolution of 1km at nadir and record observations at one near-infra-red and three infra-red wavelengths: 1.6μm, 3.7μm, 10.8μm and 12.0μ. The instrument is comprised of two sensors, one positioned so as to view Earth at nadir, and the other positioned in order to observe the same area of the Earth a split second later with a viewing angle of 60^o. The first in the series, ATSR-1, was launched in July 1991 on board the European Space Agency (ESA)'s European Remote Sensing Satellite (ERS). ATSR-2 was launched in April 1995 on board ERS-2 and included additional observation channels at 0.55μm, 0.67μm and 0.87μm. In March 2002, the Advanced Along Track Scanning Radiometer (AATSR) was launched. The two modern ATSR instruments benefit from extremely low sensor noise through continuous on-board calibration and infra-red sensors cooled to below 95K. Imagery from the last two of these three instruments, acquired at various dates is used throughout the project - more information on the ATSR instruments and their specifications can be found at *www.atsr.rl.ac.uk/atsr*.

3.1.2 SEVIRI Sensor

The Spinning Enhanced Visible and Infra-red Imager (SEVIRI) sensor is on board the geostationary Meteosat Second Generation (MSG) platform. It records an image of a full disk of the Earth with a spatial resolution of 3km^2 at nadir, using channels centred at the following wavelengths: 0.6μ, 0.8μ, 1.6μ, 3.9μ, 6.2μ, 7.3μ, 8.7μ, 9.7μ, 10.8μ, 12.0μ and 13.4μ. Imagery from this sensor from various dates (given in the text for each specific use of the data) was used to develop, test and demonstrate the Cloud-Clear-Dust Classifier algorithm presented in chapter 7. The database

described in section 3.3, which was used for validation of the application of the Cloud detection technique to both land and daytime imagery, as well as for investigations into an NWP-conditional PDF for cloud observations (in chapters 5 and 6 respectively), contains observations recorded by the SEVIRI sensor (up to 2005 when the database was compiled) for each target. More information on the SEVIRI instrument can be found at *www.esa.int/msg*.

3.1.3 MODIS Sensor

The Moderate Imaging Spectrometer (MODIS) records imagery in 36 wavebands, with spatial resolution varying from 250m (for observations at 0.67μm and 0.87μm), 500m (for observations at 0.46μm, 0.54μm, 1.23μm, 1.6μm and 2.1μm) to 1km (for observations at all other channels). The 16-day averaged land surface emissivity product available from this sensor on a 0.05^o equal-area grid is used as reference data when the cloud detection algorithm is implemented for land imagery in chapter 5. Several land surface reflectance products are also available from this sensor, and the kernel-based model used to derive them forms the basis of the land surface reflectance model implemented in the Bayesian cloud detection, which is described in chapter 5. The parameters required for the model's calculation of surface reflectance are land surface-dependent, and are combined with viewing-geometry-dependent terms in the calculation for surface reflectance. These land-surface dependent parameters are also available as a 16-day average, 0.05^o equal-area grid product from the sensor, and are used with the surface emissivity product, as described in chapter 5. More information on the MODIS sensor, and the products available from it, can be found at *www.gsfc.nasa.gov*

3.2 NWP Data

Other than for the validation work in chapters 5 and 6, the Numerical Weather Prediction (NWP) data used in the project was produced by the European Centre for Medium Range Weather Forecasting (ECMWF)'s European Reanalysis project (ERA-40). This data consists of atmospheric profiles corresponding to specific locations at specific times. The profiles are calculated at 43 model pressure levels and contain fields such as water vapour (which is integrated over the column to find the total column water vapour (TCWV) needed in some parts of the project), temperature, pressure, wind speed (although this field was sometimes missing, see section 5.2.1 on page 48), cloud fraction and cloud liquid and ice water paths (all cloud parameters were fixed to 0 for the forward modelling of observations of clear atmospheric states). Surface fields such as temperature and pressure are also given with the profiles. A subset of profiles from the ERA-40 dataset is used in chapter 6 to define realistic distributions of cloud properties in order to appropriately constrain the simulated distribution of cloud observations.

The NWP model used by Météo-France is the ARPEGE model, and fields from this model are contained in the dataset described in the next section, which was used for validation work in

chapters 5 and 6. This model contains fields analogous to those provided by ECMWF described above, but at 20 model pressure levels, rather than 43. The radiative transfer model used by the project requires NWP data to be provided on 43 fixed pressure levels, corresponding to ECMWF model levels, and so the ARPEGE data fields are linearly interpolated with pressure to the required 43, as described in the text in chapter 5.

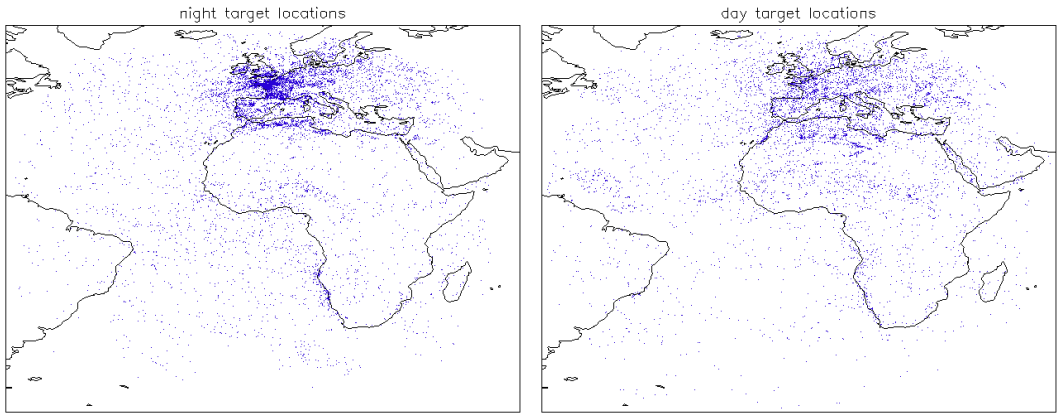
3.3 CMS-Météo-France Target Database

A dataset of 22620 'difficult' cases for cloud detection, compiled by Météo-France (CMS/Météo-France, 2005) from several years of data acquired from the Meteosat-8 sensor, was used to test and validate most of the work described in this project. The targets are described as 'difficult', because they are observations specifically selected as being of interest to nephanalysts and presenting challenges to cloud detection. The dataset is therefore not likely to be representative of commonly occurring atmospheric states, but instead is likely to contain unusual states which are challenging to cloud detection techniques. While this dataset would not generally be chosen for validation work, it was judged appropriate for testing the improvements and extensions made to the Bayesian method. Previous work had already indicated that the Bayesian cloud detection method could produce results with skills scores comparable to operational techniques (Merchant et al., 2005), and it was thought that comparison of the techniques using a representative dataset may be less informative, i.e. both the operational and the Bayesian methods would achieve very high skill scores. A more challenging comparison of the techniques could be carried out with a dataset containing specifically those targets for which all cloud detection techniques could be expected to struggle. For this reason, this dataset was chosen to validate much of the work in this project.

In the database, targets are 5 x 5 pixels. The Brightness Temperatures (BTs) and reflectances measured by the sensor for each pixel are given, while latitude, longitude, solar and satellite zenith angles are given only for the central pixel. In addition to a climatological Sea Surface Temperature (SST), NWP surface pressure, Surface Temperature (ST), water vapour, pressure and temperature fields from the grid-cell of the ARPEGE atmospheric model which is closest to the central pixel location are provided. Each target is assigned to a target class, judged by experts at Météo-France through inspection of the imagery by eye, and each target class is defined as either clear and cloudy as shown in table 3.1. The NWP fields are given for the model profile corresponding to the closest time before and after the image acquisition time - those from before image acquisition were always used in this project. Temperature and water-vapour are provided at 20 altitudes levels, and were interpolated, using fixed pressure levels, to 43 to be forward-modelled by the RTM. For a fuller description of the database, see CMS/Météo-France (2005). The techniques described in this thesis were only applied to the central pixel, as the NWP information corresponded to this pixel and no information on ARPEGE grid cell location was available to interpolate the information to the other pixel locations. Additionally, U.K. Met.

Office operational results were only available for comparison for the central pixel.

The Bayesian technique relies on an RTM being used to model observations for clear sky. The fast RTM, RTTOV, which considers a maximum satellite zenith angle of 75^o, was used for thermal wavelength imagery. Targets with a satellite zenith angle greater than 75^o (591 targets) were not considered. Targets categorised in the database as 'no classification' (69 targets) were removed as being of no use for validation. Some targets were categorized as 'land' or 'open sea' while having a contradicting land-mask value (71 targets) - while the land-mask is provided for every pixel in the target, the whole target is assigned a single category, meaning that for coastal targets the land-mask for the central pixel may disagree with the land/sea categorization which is based on the majority of pixels in the target - such targets were also excluded. For the parts of the project where clear and cloudy are considered the only possible atmospheric states, aerosol targets in the database (5168 targets) were ignored, rather than be judged as cloud or clear. One benefit of the database is that it contains Météo-France's operational results for the targets, giving something against which to compare the performance of the techniques explored in the project. Any targets for which Météo-France's operational result is 'not classified' (67 targets) were therefore removed. A total of 14931 non-aerosol targets were then available for the work. Of these, the 6356 targets with a solar zenith angle $\leq 80^o$ were classed as 'day', and the 8566 with a solar zenith angle $\geq 90^o$ were classed as 'night' - target locations are shown in 3.1. The work of the project focused on day and night conditions, and twilight targets were not used. A description of the target classes, and a breakdown of the number of targets in each, is given in table 3.1.



(a) (b)

Figure 3.1: Location of targets in the CMS-MétéoFrance Database: (a) Day Targets (solar zenith angle less than 80^o); (b) Night Targets (solar zenith angle greater than 90^o).

Code	Description	Cloud/Clear	Day Sea	Day Land	Night Sea	Night Land	Total
101	open sea	0	712	-	1261	-	1973
102	sea with shadow	0	2	-	-	-	2
106	sea with sunglint	0	170	-	-	-	170
151	land	0	1425	-	1587	-	3012
152	land with shadow	0	-	19	-	-	19
181	ice	0	-	5	-	0	5
191	snow	0	-	319	-	54	373
502	stratus	1	508	318	803	2077	3706
503	stratocumulus	1	525	288	574	929	2316
504	shadow over low cloud	1	1	1	-	-	2
601	small cumulus over sea	1	69	-	46	-	115
602	small cumulus over land	1	-	168	-	0	168
606	cumulus congestus over sea	1	34	-	9	-	43
607	cumulus congestus over land	1	-	15	-	1	16
608	cumulonimbus	1	73	160	64	45	342
609	extensive cumulonimbus	1	15	79	37	49	180
701	thin cirrus over sea	1	103	-	66	-	169
702	thin cirrus over land	1	-	233	-	105	338
704	thin cirrus over snow	1	-	5	-	0	5
705	thin cirrus over stratus/stratocumulus	1	79	69	72	112	332
706	thin cirrus over cumulus	1	6	6	2	0	14
707	thin cirrus over altostratus/altocumulus	1	37	48	41	26	152
801	altocumulus/altostratus	1	80	39	57	40	216
802	altocumulus	1	75	241	121	162	599
811	cirrostratus	1	112	173	82	38	405
812	cirrostratus over altocumulus/altostratus	1	106	47	75	31	259
-	-	**Total:**	4132	2233	4897	3669	14931

Table 3.1: Targets that comprise the CMS - MétéoFrance database used in the project. Clear: 0, Cloud: 1.

Chapter 4

Cloud Detection with Pre-Clustering of Satellite Imagery

In the testing of the cloud detection method described in chapter 2, it was noted that a high probability of cloud was being calculated for some clear sky pixels along coasts and at ocean fronts. The beginning of this PhD project sought to address this using the methods described in this chapter. This work has been presented at, and appears in the proceedings of, the 2006 EUMETSAT Meteorological Satellite Conference (Mackie et al., 2006).

The validation of Merchant et al. (2005) showed some pixels around ocean fronts, which were judged clear through inspection of visible imagery, to be classified ambiguously, i.e. to have a reasonable probability of cloud contamination. This is almost certainly the result of textural considerations in the algorithm - areas with steep gradients in brightness temperature or reflectivity, are considered likely to be cloud, despite these properties also being found in clear-sky pixels over ocean fronts. Ocean temperature and reflectivity generally varies gradually relative to cloud top spectral properties and so it is useful to keep these considerations in the algorithm as a useful input to the probability calculation. The problems evident over ocean fronts however, are likely to be seen to greater effect when the algorithm is extended to operate over land surfaces, where spectral properties can be expected to be more variable, and so further ambiguous classifications are to be anticipated.

The concept driving the work of this chapter is that spectral clustering of the image scene and applying the detection algorithm to the mean properties of a cluster, rather than to individual pixels, could significantly reduce the problem of textural considerations. Only the texture of pixels deemed to represent the same 'natural' grouping (e.g. from the same ocean thermal regime) would be considered in determining whether the pixels belonging to that group are cloud-contaminated or not.

Processing clusters of pixels using the mean properties of the cluster risks a loss of spectral information for individual pixels, in effect the resolution of the detection product could change from pixel-scale to cluster-scale if the calculated probability of clear, $P\left(c\left|\mathbf{y^o},\mathbf{x^b}\right.\right)$, is assigned to each cluster rather than to each individual pixel. This can be avoided through the use of an

appropriate clustering procedure to calculate $P\left(c\left|\mathbf{y^o}, \mathbf{x^b}\right.\right)$ values that can be interpolated onto individual pixels through their association to the cluster, allowing pixels to be associated with more than one cluster.

4.1 Choice of Clustering Method

There are many different clustering algorithms available, but the success of all of them is dependent on an appropriate choice of clustering parameters. The spectral threshold tests used in more traditional cloud-detection studies, see section 1.3.1, suggest parameters that can be used to distinguish cloudy- from clear-sky, and various combinations of these were used in assessing the suitability of the particular clustering methods. The aim is to produce a automated technique for cloud detection for day- and night-time imagery, and so the ideal combination of parameters should be derived from only TIR wavelength data.

4.1.1 Split and Merge Clustering

A Split and Merge Clustering algorithm (SMC) was developed by John Marsham and Stephan Mathiesen at the University of Edinburgh in 2002, based on the work of Simpson et al. (1998). A number of spectral parameters are specified and their values are standardized and used to form a vector for each individual pixel. Initially all pixels are grouped as one cluster, and the algorithm iteratively divides this into more clusters, which are subsequently merged or divided again, depending on the specified tolerance, up to the maximum number of clusters, which must be specified. The first cluster (all the pixels) is considered in the same way as all subsequent clusters, that is the two most extreme members are identified, i.e. the two pixels that are furthest apart in the parameter space, and these are used as the basis for two new clusters, to one of which each of the remaining pixels is assigned. The distance (in parameter space) between the new cluster means is tested and, if it is above the specified tolerance, both clusters are split again. If the distance between the two cluster means is below the specified tolerance, they are merged into one cluster for the remainder of the processing time.

This method has the advantage of requiring no initial assumption of the number of clusters present. Clusters found by this method, however, will tend to have the same spherical shape in parameter space, which may not represent the shape of the surface- and cloud-types present in the scene, meaning the clusters can be considered ‘forced’ rather than ‘natural’. For the purposes of cloud detection, it is more useful to find ‘natural’ clusters, so that assigning a probability of cloud contamination to each cluster is equivalent to assigning the same probability to pixels that represent the same surface- or cloud-type.

Various clustering parameters were investigated, and a reasonable separation of imagery into different thermal regimes was achieved, but the SMC method is computationally very expensive,

and considered too slow to be of practical use.

4.1.2 CLARA Clustering

The programming software 'R' has several built-in algorithms which are described in detail in Kaufman and Rousseeuw (1989). One of these, CLARA (Clustering Large Applications), is designed particularly for the handling of large data sets, and so should be computationally more efficient than SMC.

The number of clusters must be specified, which is problematic as an estimate from inspection of the imagery could not be relied upon to be accurate as to the number of natural groupings actually present, and would be impractical for a large image scene/multiple image scenes. If this algorithm was to be implemented in large scale data processing, an automated way of determining the number of clusters (which it is anticipated would be highly variable between image scenes) would have to be found. Once the number of clusters is specified as N, CLARA looks for N 'representative' pixels within the image, and then computes the distance (in parameter space) between all other pixels and those representative pixels. Each pixel is grouped with the representative pixel which is closest to them in the parameter space.

This algorithm creates similarly 'forced' clusters to SMC - there is no flexibility to allow for non-spherical clusters. It does, nevertheless, produce a reasonable separation of an image into different thermal regimes in a much shorter time than the SMC.

4.1.3 Appropriateness of Hard-Edged Clustering

Clustering methods such as SMC and CLARA can be considered 'hard-edged', in that a pixel may only belong to one cluster, and all pixels must be assigned to a cluster. Applying the cloud detection techniques described in chapter 2 to the properties of a cluster (cluster properties being the mean properties of the member pixels) and assigning the resulting probability of clear to all pixels within the cluster could have a negative effect on the results. Pixels with unambiguously clear properties would be assigned mean properties of member pixels in the cluster to which they belong, which necessarily contain some ambiguity from the inclusion of ambiguous pixels in the cluster. It was initially thought that this introduced ambiguity would be small, as the number of ambiguous pixels in a cluster is small relative to the number of unambiguous pixels. The calculation for the probability of clear, however, is non-linear, see section 2.4.2 in chapter 2, and the effect of a very small change in some of the properties read into the algorithm can result in a large change in the calculated probability. The loss of spectral information for individual pixels makes 'hard-edged' clustering methods inappropriate to this cloud detection approach.

To test this, a small region (56x72 = 4032 pixels) of an image acquired by the AATSR

sensor on 10th May 2005 at 11:11 UTC over the Korean East coast was used. The region is free from sunglint, contains some cloud and clear-sky thermal gradients, and some pixels classed ambiguously by the detection algorithm, see figure 4.1. Through inspection of the retrieved surface temperature and of the 1.6μm imagery, the cloud mask in figure 4.1c was constructed manually as a 'truth'. The visible imagery and retrieved surface temperature were inspected to manually create an 'optimum' clustering for the region, see figure 4.2a. The mean properties of each cluster were used in the cloud detection method described in chapter 2, and each pixel was assigned the probability of clear calculated for the cluster to which it belongs, shown in figure 4.2b. The results can be compared to the results of running the detection algorithm on the individual pixels (i.e. without any pre-clustering) in figure 4.1d.

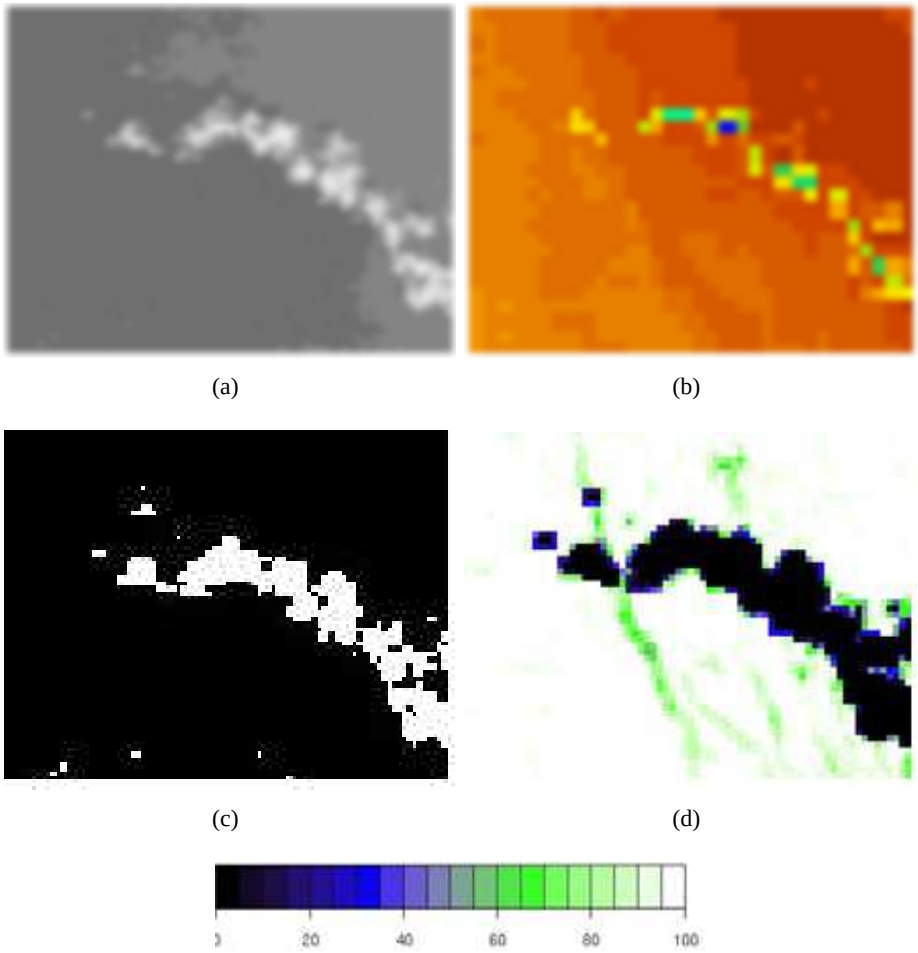


Figure 4.1: Study area for cloud detection with pre-clustering: (a) 1.6μm image, (b) retrieved surface temperature, (c) cloud mask used as 'truth' and (d) calculated probability of clear for each pixel. Study region is off the North East coast of Korea, and is imaged by the AATSR sensor at 11:11 UTC on May 10th 2005.

The results appear to show no benefit to hard-edged clustering of imagery prior to cloud detection. Clear pixels corresponding to strong thermal gradients, see figure 4.1b, have a higher probability of clear, but at the expense of other clear pixels, for which the probability is now

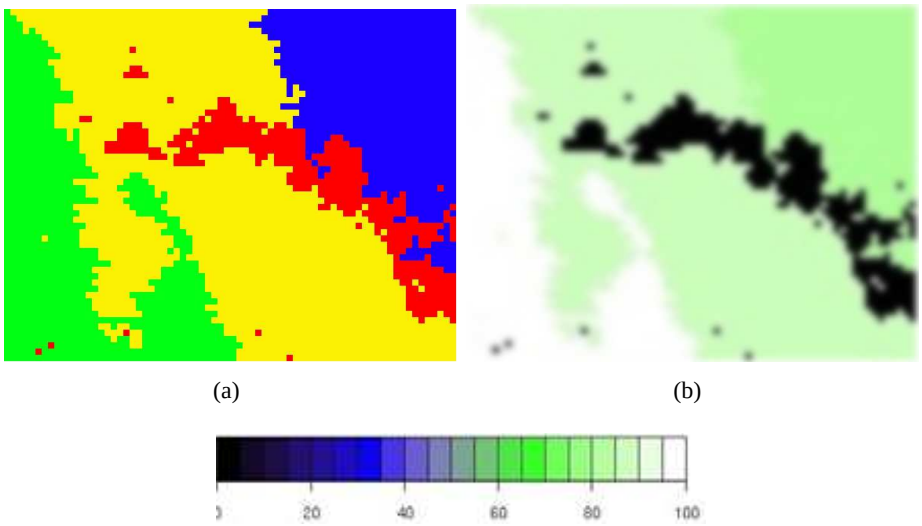


Figure 4.2: (a) Manually selected clusters for study region imaged by the AATSR sensor off the North East coast of Korea on May 10th 2005, (b) the calculated probability of clear for each cluster.

lower. The loss of spectral information for individual pixels makes hard-edged clustering methods inappropriate to the cloud detection techniques of chapter 2. An alternative clustering method, such as Fixed Point Clustering (FPC), which retains some individual pixel information, is more likely to be appropriate.

4.1.4 Fixed Point Clustering

FPC is a clustering algorithm developed by Christian Hennig (Hennig, 2003), with the principle that each 'cluster should be separated from the rest of the data, and the points of the cluster should not split up into further separated sub-classes'. Each cluster is considered independently, and pixels are assigned membership to a particular cluster based on their not being defined as 'outliers' to that cluster. Outliers are defined using Mahalanobis distances[1], and a pixel's membership to cluster A is not considered when membership to cluster B is assessed, meaning that pixels can belong simultaneously to more than one cluster. A pixel may belong partially to a cluster, meaning it may lie on the border of the cluster and be assigned a fractional membership to it, which can be treated as the probability of its membership to that group. Figure 4.3 shows 3 clusters, A, B and C. Pixel a belongs wholly A and partially to B; pixel b belongs wholly to A; pixel c belongs partially to both A and B, but wholly to neither; pixel d belongs wholly to both A and B; pixel e belongs wholly to B and C.

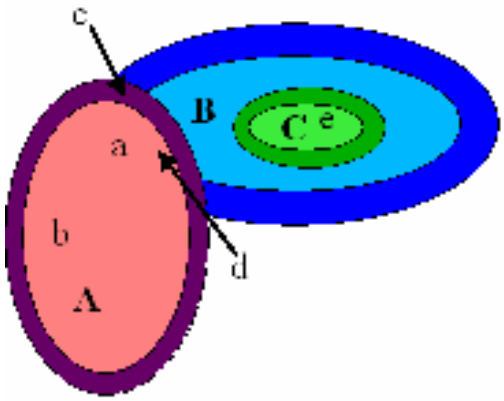


Figure 4.3: Illustration of the Fixed Point Clustering (FPC) Method. Membership of 5 pixels (a-e) to 3 clusters (pink - A, blue - B, and green - C). All pixels in B form one group, while those in C, although belonging to B, also belong to a subgroup within it (which is C). Some of the pixels in A will also belong to B, and vice-versa.

Clustering with FPC has the considerable advantage of minimising the loss of information for individual pixels. Instead of being assigned the mean properties of the cluster to which they belong, pixels can be assigned a combination of the mean properties of all the clusters to which they belong, weighted by the strength of their membership to each one. Rather than being assigned a probability of being clear based on one set of properties, that may not fully represent that specific pixel, a probability can be assigned based on the combination of properties of the classes to which the pixel has been found to be associated. This is in keeping with the general principles of probabilistic cloud detection, whereby probabilities allow decision-making with minimum information loss.

[1] 'Mahalanobis distance' is the distance between 2 points in a distribution, scaled by the covariance. I.e. if x and y belong to a distribution with covariance C, then the Mahalanobis distance between them is $\sqrt{(x-y)^T C^{-1} (x-y)}$.

The method of identifying each cluster individually, by identifying which pixels are *not* associated with a particular pixel, and assigning all other pixels to the cluster, ensures homogeneity within the cluster. By not forcing every pixel to belong to a cluster, or superficially partitioning the data into similar shaped clusters, the algorithm is considered suitable for finding 'natural' groupings in the data. This makes it more likely that clusters will represent real physical types in the imaged surface. The algorithm takes a long time to run, but the time can be significantly reduced if an initial grouping of the data, from which to begin the iterative processing, is specified.

4.1.5 Implementation of FPC for the Bayesian Cloud Detection

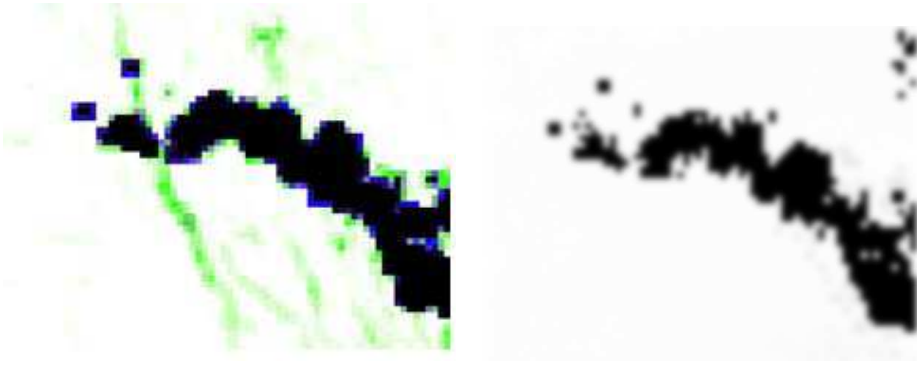
FPC was applied using [11μm 12μm, 11μm 3.7μm, 3.7μm, longitude, latitude] as parameters[2]. Pixels not assigned to any cluster are likely to be cloud, as they do not form any 'natural' spectral grouping with close-by pixels, and were assigned membership to an 'extra' cluster. The mean properties for all clusters were read into the detection algorithm, and the probabilities of clear calculated. The contribution of each pixel to the cluster mean properties was weighted by the strength of its membership to that cluster. Membership strengths range from 0, where a pixel has been defined as an outlier, to 1, where the pixel is a full cluster member. As outlined above, pixels can have membership strengths of 1 to (and so contribute with a weight of 1 to the mean properties of) several clusters. A vector of cluster memberships was calculated and normalized for each pixel. This vector was then used to weight the calculated probability of clear for every cluster, in a pixel-specific sum to calculate the appropriate probability of clear for each pixel.

While being relatively simple to implement, the FPC clustering algorithm is computationally very expensive, making processing time impractically long - an area of several hundred square pixels required more than 24 hours of processing.

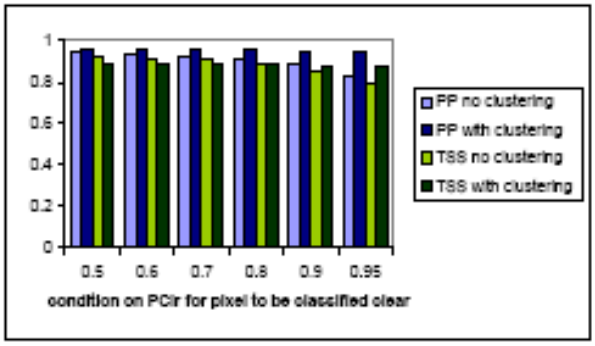
4.2 Case Study Results

The final probabilistic cloud detection result shows improvement when the imagery is pre-clustered, see figure 4.4. A standard verification procedure was used to quantitatively assess the improvement. The Proportion of Perfect Classifications (PP) gives the proportion of pixels correctly identified as clear or cloudy, using the 'truth' mask for reference. The True Skills Score (TSS) is the difference between the proportion of true cloudy pixels detected and the proportion of pixels that are misclassified as cloudy, see table 4.2.

[2]Including latitude and longitude in the clustering parameters reduces the likelihood of pixels that are spectrally similar, but spatially distanced, being associated with each other.



(a) (b)



(c)

Figure 4.4: Results of cloud detection for the study region recorded by the AATSR sensor off the North East coast of Korea at 11:11 UTC on May 10th 2005: (a) without pre-clustering, and (b) with pre-clustering. A more quantitative comparison is given in (c).

-	'truth' clear	'truth' cloud
detected as clear	A	B
detected as cloud	C	D
TSS	$\frac{D}{D+B} - \frac{C}{C+A}$	
PP	$\frac{A+D}{A+B+C+D}$	

4.3 Conclusions from Case Study

The Bayesian technique for cloud detection produces a mask of probabilities for clear, in which pixels with a high probability are judged to be clear, and those with a low probability are judged to be cloud. Pixels for which the calculated probability is at neither extreme are therefore ambiguous and can be classified as cloud or clear depending on the tolerance of the application to cloud contamination (Merchant et al., 2005). The green slanting lines indicating ambiguous classification, in figure 4.4a, do not appear when clustering is carried out prior to cloud detection, see figure 4.4b. These 'slanting line' features in figure 4.4a almost definitely correspond to thermal gradients higher than those generally associated with ocean surfaces, i.e. to ocean fronts, creating an ambiguity in the classification because of the textural part of the cloud detection algorithm, described in section 2.4.1. This case study suggests that clustering imagery prior to applying the detection algorithm results in a more bimodal probability product, particularly in regions of ocean fronts. Low tolerance applications can therefore consider pixels with $P\left(c\left|\mathbf{y^o}, \mathbf{x^b}\right.\right)$ greater than 0.95 to be clear, without losing as significant a volume of information through misclassification of clear pixels. High tolerance applications, for example cloud studies, may not benefit from pre-clustering, as the TSS for the detection when $P\left(c\left|\mathbf{y^o}, \mathbf{x^b}\right.\right)$ is required to be greater than 0.5 (for classification as clear) remains slightly higher when clustering is not applied.

The cloud detection results calculated using clustering as a pre-processing step show some pixels in the top right of the image to be unrealistically classified as cloud, which is not the case when clustering is not carried out, see figure 4.4. This is surprising, as the area does not appear to correspond to a spectrally or thermally distinct region, see figure 4.1. It may be that, being at the edge of the region, these pixels were not assigned with a high membership strength to any cluster - latitude and longitude being used as clustering parameters. If these pixels are spectrally similar to pixels elsewhere in the image, then using latitude and longitude will reduce their association with those other pixels, but they will not necessarily be associated with many other pixels with greater strength, since there are not many pixels nearby. The membership of each pixel to different clusters is normalized to 1, which, for pixels not associated strongly with *any* clusters, could result in an artificial raising of the membership strength to more distant clusters. Figure 4.1 does not show the pixels in the top right to be more reflective or colder than pixels elsewhere, and the cloud classification occurs only for a few pixels, rather than for all pixels at the edge of the image so this explanation may not be appropriate. These pixels may have been assigned to the 'extra' cluster, the properties of which are equally likely to correspond to a cloud or a clear state. Using different clustering parameters may address the problem, but may introduce other clustering artifacts. Weighting the latitude and longitude clustering parameters so as to be less effective than the spectral parameters may also change this result, but may reduce the effectiveness around ocean fronts. If the classification of the pixels in the top right is a result if their being at the edge of the region then it is unlikely that such an artifact would be significant if a larger region were considered. Alternatively, the region could be artificially extended by a border of pixels assigned spectral properties matching those at the edge. This, however, could create artificial clusters, as those 'extra' pixels would effect the clustering of all pixels in the region, not just those at the

edges. It is also possible that these pixels really are cloud which is missed when clustering is not carried out. Clouds which do not appear obvious in either an ST or reflectivity image, as for these pixels - see figure 4.1, are possible, but unlikely. The lack of an absolute 'truth' for this case study means this conclusion must also be considered.

It is interesting that the FPC clustering misses the single-pixel clouds seen at the bottom the region in both the manually clustered image, and in the cloud detection results following the manual clustering, in figure 4.2. Figure 4.1a shows these pixels to be only slightly more reflective than the underlying ocean surface, and it may be that they do not in fact represent clouds. This, however, seems unlikely - a more plausible explanation is that the latitude and longitude parameters used in the clustering mean that single cloud pixels are likely to be associated with their neighbouring pixels. One cloud pixel with strong membership to a cluster containing mainly clear pixels is likely to be classified as clear by the cloud detection. It should also be noted that these clouds are not easily discernible in the ST image in figure 4.1b, and only thermal, not visible, wavelength imagery is used in the clustering and in the cloud detection.

The study is by no means conclusive, as only a small region has been investigated, but clustering has been shown to at least have the potential to improve the Bayesian method of cloud detection. The greatest benefit is seen for ocean fronts, while some weakness in the approach is seen for single pixel clouds. Further work may indicate that pre-clustering of imagery prior to cloud detection is appropriate only for regions where ocean fronts have already been identified as likely. These results, however, suggest that clouds filling only a single pixel, whether over an ocean front or not, are likely to be missed by the cloud detection if imagery is pre-clustered. Without clustering, ocean fronts appear to be classified ambiguously, indicating some uncertainty in the results. Pre-clustering with FPC reduces the ambiguity at ocean fronts, as elsewhere, but this means that pixels which are falsely classified are falsely classified with greater certainty - having a probability of clear closer to 1 or 0 than is the case without pre-clustering. Without a wider case study, it is difficult to judge the effect of this, and so difficult to conclude whether the reduced ambiguity in the classification resulting from pre-clustering is appropriate.

Further work could be done to investigate the effectiveness of pre-clustering using both textural and spectral parameters - essentially all the information used for the cloud detection process itself could be used at the clustering stage. This is likely to be computationally expensive, but may yield improved results. A wider study could investigate the most effective parameters to use for the clustering, which may be different in different regions, for example in regions where ocean fronts occur more or less frequently. It may turn out to be appropriate to use clustering parameters dependent on the NWP information for the pixel.

Since this clustering work was carried out, the Bayesian cloud detection technique has been extended to exploit visible as well as thermal wavelength imagery, see section 5.2 in chapter 5. Including visible wavelength data in the clustering may improve results - further investigations,

ideally using a larger dataset than the case study region, would be needed to find the most appropriate parameters to construct from these extra data. FPC, as it is currently coded, is computationally very expensive, and so any new parameters should replace parameters used here, rather than adding to them. Alternatively, and preferentially, some extensive optimization work could be done, making investigations into different parameters more practical, and allowing a greater number of parameters to be used.

A further development made to the Bayesian cloud detection algorithm since this clustering investigation was carried out, is its extension to land imagery, see section 5.1 in chapter 5. At present, the textural properties of observations made over land surfaces are not used in the calculation of the probability of clear for those observations. This is because land surfaces are often more variable in reflectivity and temperature than ocean surfaces, and so distinguishing clouds on the basis of their higher thermal gradients is less appropriate for land imagery. Clustering may provide a means to exploit some of this textural information and so increase the volume of information on which the calculated probability of clear is based, possibly improving results. An investigation into this would be non-trivial, as a wide variety of land surfaces and clustering parameters would need to be considered, which would not be practical without some optimization of the FPC algorithm. It may be appropriate for the textural information to be weighted so as to contribute less than the spectral information to the calculated probability of clear for land observations, and this would also require careful investigation. Such a study would be interesting, but it is unclear whether any improvement to the results would be found, and the computational expense places it outside the scope of this project.

At present, the computational expense of the FPC algorithm is not judged to be justified by the small improvement it creates in the results of the cloud detection, and clustering is not considered further in this project.

Chapter 5

Extensions Made to the Bayesian Cloud Detection

Prior to this project, the method described in chapter 2 was implemented in an algorithm which had been validated for night-time imagery over ocean from a single sensor - the Along Track Scanning Radiometer (ATSR) (Merchant et al., 2005). In that validation, no external 'truth' was available for assessment of the skill of the technique. Instead, two of the authors, experienced in cloud detection, cloud-screened images by hand to create a 'truth'. Disagreement between the experts (who were also involved in development of the algorithm) may have been biased towards particular atmospheric states for which the technique was then not fully validated, and the 'truth' could have shared biases with the algorithm. The work presented here benefited from an independent 'truth', which was provided with the test data.

This project introduced consideration of a latitude- and season-specific prior probability for the presence of cloud, rather than a fixed value, making it more physically robust, see section 2.3, and extended it to operate over land, see section 5.1, and to consider visible wavelength imagery, making it more appropriate to day-time applications, see section 5.2, making it suitable for a wider range of applications. Validation of the technique, with these extensions, was carried out in this project, see section 5.3, using an externally supplied database of 'difficult' cloud targets from the SEVIRI sensor, usually used by nephanalysts for investigative, rather than validation, work (CMS/Météo-France, 2005) - see section 3.3 for a description of the database. The validation work described in this chapter has been submitted for publication as 2 papers, (Mackie et al., submitted 2008a,s).

One of the benefits claimed for the Bayesian approach is that the principle is generic and the technique can be used on imagery from any thermal sensor for which a RTM exists. In practice, this involved some restructuring of the algorithm to make it more flexible, which has resulted in its application to imagery from other sensors now being more straightforward.

The validation for another sensor - particularly since ATSR-imagery is the basis of the textural PDF for clear, see section 2.4.1, using an external dataset (rather than a subset of data used in the development of the technique), was a useful follow-up to the validation that had already been carried out by Merchant et al. (2005), and the extensions of the technique are anticipated to make

it useful to a wider range of users.

5.1 Extension of the Bayesian Cloud Detection to Land Imagery

The Bayesian cloud detection described in section 2 did not initially consider land imagery, as the original motivation was to provide a tool to aid ocean ST retrieval. The high spatial and temporal variability of land ST make it more challenging to the technique. A part of the work for this project was extending the technique to consider land imagery, while maintaining the computational efficiency required by operational cloud detection applications. An outline of the algorithm, with the extension to land imagery, is shown in figure 5.1.

In order for the algorithm to perform successfully for land imagery, it must be able to forward-model land observations, in the same way it does sea observations. This requires the surface emissivity, and it's temporal variability, to be known. For sea imagery, observations are simulated at NWP profile locations and interpolated to individual pixel locations using ST and spatial distance. The variability of ST over land surfaces makes this inappropriate, but running an RTM for each individual land pixel would require an impractical amount of processing time. Interpolation of predicted observations for land imagery from profile to pixel location is therefore done using elevation and spatial distance, and so requires a Digital Elevation Model (DEM) to be available. Uncertainties in the NWP fields used to simulate clear sky observations are generally more variable and less well known for NWP information at land locations than for sea locations. These uncertainties are used in calculation of the probability of clear sky in section 2.4.2 on page 16, and so an extension of the technique to land imagery requires consideration of appropriate values. Some empirical data which is exploited by the method, for example to calculate the prior probability for cloud, is based on the results of operational cloud detection methods, which are generally more accurate over sea than land. While some clouds may appear differently over land due to topographical effects, making sea cloud data inappropriate for cloud detection in land imagery, this must be weighed against the accuracy of the data itself. The following sections discuss these challenges and how they were addressed in extending the techqniue to operate successfully over land.

5.1.1 Texture

Surface temperature is generally more variable over land than over the sea, and so clouds cannot be discriminated in an image on the basis of thermal texture the way they can for sea, see section 2.4.1 on page 15. For land, therefore, $P\left(\mathbf{y^o}\left|\mathbf{x^b}, c\right.\right) = P\left(\mathbf{y_s^o}\left|\mathbf{x^b}, c\right.\right)$. It is assumed that the spectral properties of clouds over land are the same as those for clouds over sea and so the same prior PDF for cloud is used for both. This is not always valid, for example in the case of sea fog, but in the absence of a reliable dataset of imaged clouds over land, in which all possible cloudy states are represented, it was considered a necessary assumption.

For the technique to be suitable for land imagery, different uncertainties in the a-priori information, i.e. in the NWP fields, must be considered. Brightness temperatures simulated for clear

sky at NWP grid locations should be interpolated to pixel locations with respect to both location and elevation rather than surface temperature. Surface emissivity over land is spatially and temporally variable, and so the constant emissivity used for simulating clear sky radiances over sea is not suitable and new emissivity information is required. The following sections describe solutions to these difficulties, and demonstrate how these adaptations were implemented.

5.1.2 Land Surface Emissivity

A 16-day average land surface emissivity product is available from the Moderate Resolution Imaging Spectro-radiometer (MODIS) on a global 0.05^o equal area grid (freely available from *http://edcdaac.usgs.gov/modis/dataproducts.asp*). This data was used to construct four seasonal maps of land surface emissivity on a 0.1^o equal area grid. The map appropriate to the image acquisition time is loaded into the cloud detection algorithm, and pixel and profile latitudes and longitudes are used to assign surface emissivities to land pixels and to land NWP profiles. The surface emissivity maps used for northern hemisphere spring (March, April and May) are shown in figure 5.2. In the event of surface emissivity data being missing at the profile or pixel location, the RTM uses a fixed surface emissivity. For the fast RTM, RTTOV, which is used to simulate BTs for this project, the default surface emissivity (used when emissivity is not read from the map) is 0.98 for all 3 channels (Saunders, 2002).

Land surface emissivity is dependent on viewing-geometry, and failure to adjust hemispherical emissivities for this can affect retrieved surface temperature by several kelvin (Li and Strahler, 1999). A directional correction to the emissivities, however, would require a land classification map to be used, as the directional-dependence of emissivity is surface-type-dependent, see for example Jupp (1998); Kanani et al. (2007). The accuracy of a surface-classification map is difficult to quantify, and several values for the directional-dependency exist in the literature for any given surface type. The appropriateness of any directional correction would therefore be difficult to judge, and loading in a further reference file would increase processing time. Here, emissivity is assumed independent of viewing geometry, but this is an area that could be developed further.

The surface emissivity becomes an element of $\mathbf{x^b}$ and is used by the RTM to calculate modelled radiances, $\mathbf{y^b}$, and their emissivity-dependence, $\left(\frac{\partial \mathbf{y^b}}{\partial \mathbf{x^b}_{emiss}}\right)$. The emissivity-dependence is used to bi-linearly interpolate the modelled BTs (modelled using the profile emissivity) to the pixel emissivity.

The difference between pixel and profile elevations is also used in interpolation of the modelled BTs to the pixel location. The pixel ST is interpolated from the nearest land NWP profiles, and adjusted for any profile-pixel elevation difference by assuming a fixed lapse rate of Γ=0.0098Km^{-1}, see equation 5.1, in which $\left\langle \mathbf{x^b}_{st_{pixel}} \right\rangle$ is the interpolated pixel ST before the adjustment, and the profile and pixel elevations, *h*, are considered relative to sea level. BTs modelled from the surrounding profiles are bi-linearly interpolated using the profile STs, to this pixel ST, using $\left(\frac{\delta \mathbf{y^b}}{\delta \mathbf{x^b}_{st}}\right)$, which is calculated by the RTM, see equation 5.2.

$$\mathbf{x^b}_{st_{pixel}} = \left\langle \mathbf{x^b}_{st_{pixel}} \right\rangle - (h_{pixel} - h_{profile}) \times \Gamma \tag{5.1}$$

$$\mathbf{y^b}_{pixel} = \mathbf{y^b}_{profile} + \left(\frac{\partial \mathbf{y^b}}{\partial \mathbf{x^b}_{st}} \right) \times \left(\mathbf{x^b}_{st_{pixel}} - \mathbf{x^b}_{st_{profile}} \right) \tag{5.2}$$

5.1.3 Uncertainties in A-priori Information for Land Pixels

While the uncertainty in NWP TCWV is anticipated to be comparable for land and sea profiles, temporal and spatial variability makes the accuracy of NWP ST for land generally poorer than for ocean. The uncertainty is likely to be spatially variable and is also likely to be greater in the daytime, as the variable absence or presence of direct sunlight controls the rate and duration of warming (and/or cooling) periods. The uncertainties chosen, ϵ^b_{st}, are 2.4K and 1.6K for day and night respectively. The accuracy of the NWP ST field is not expected to vary with atmospheric path length, but these uncertainties are nevertheless multiplied by the pixel's atmospheric path length to allow for some dependence of the FM error on atmospheric path length, which is otherwise not considered. These are similar to figures found for the accuracy of the 2m air temperature field for the ARPEGE NWP model (*H. Roquet, Météo-France, personal communication, 25 July 2008*). More work into the accuracy of NWP fields for specific NWP models would give a more appropriate expression for these uncertainties, but at present this information is not available in the literature.

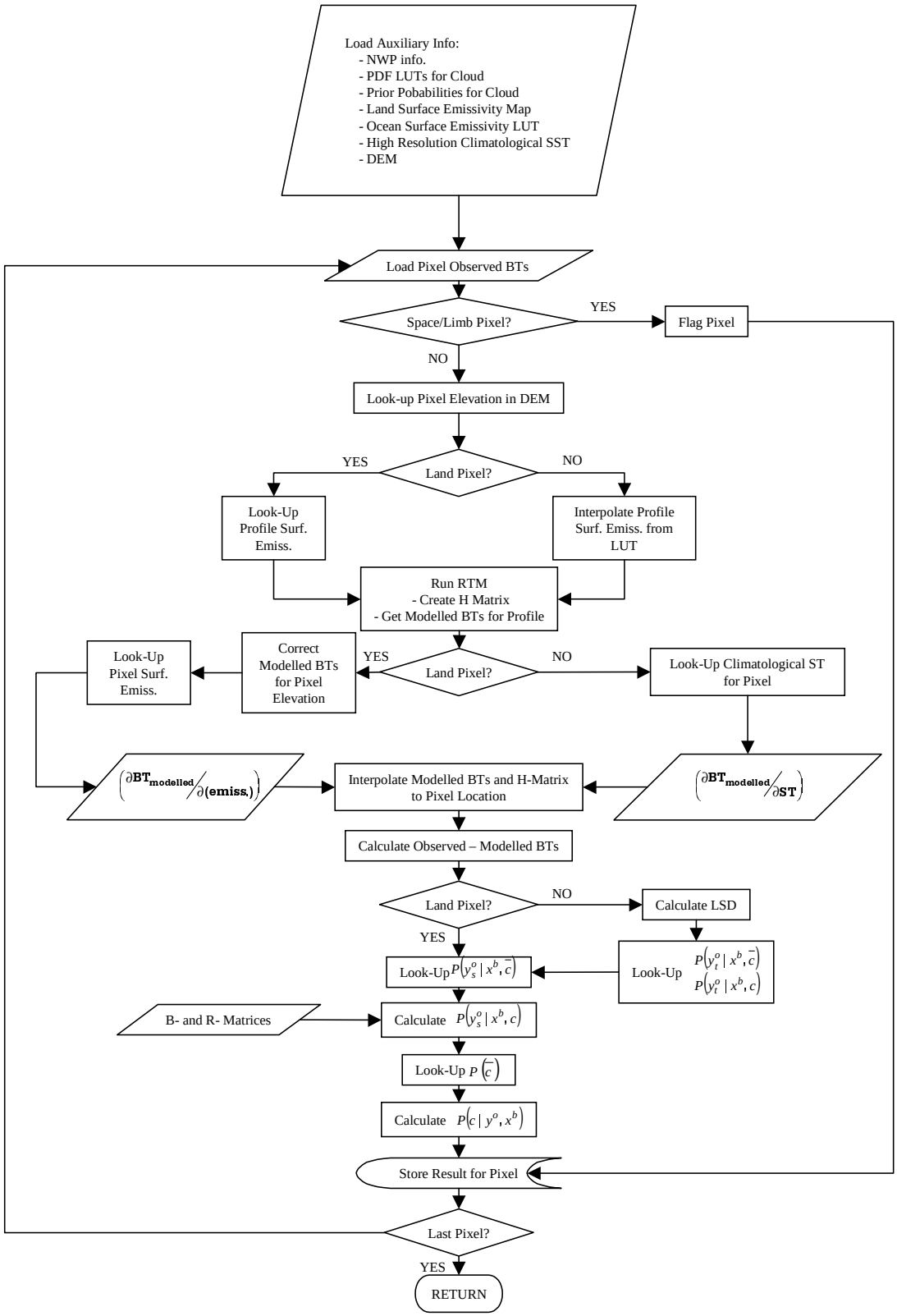
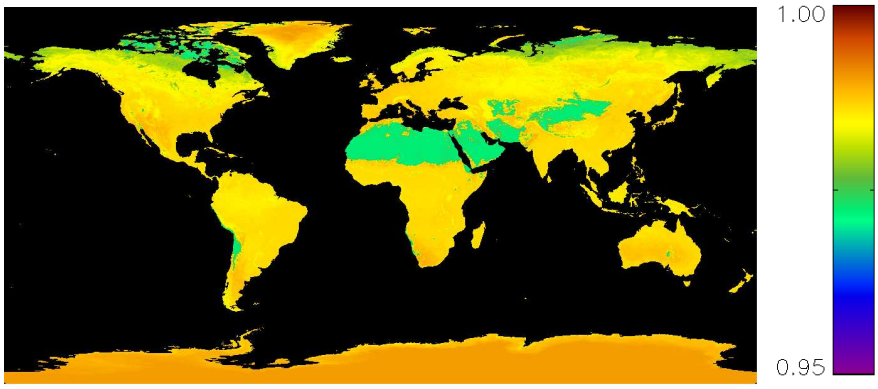
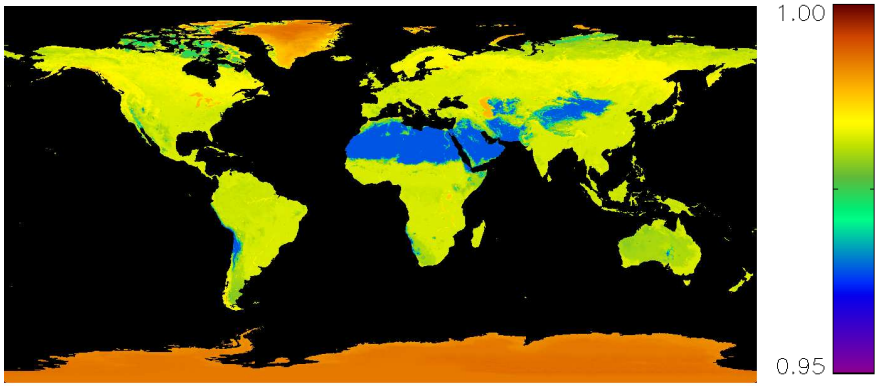


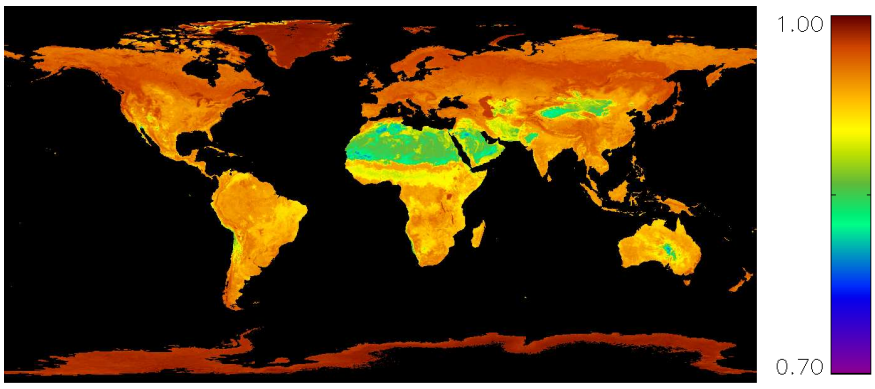
Figure 5.1: Flowchart outlining the Bayesian cloud detection algorithm, with the extension to land imagery.



(a)



(b)



(c)

Figure 5.2: An example of land surface emissivity maps calculated from data provided by the MODIS sensor. These maps are used for images acquired in March, April and May: (a) 12μm; (b) 11μm; and (c) 3.9μm. The colourscale corresponds to emissivities of 0.95-1.0 for 12μm and 11μm, and to 0.7-1.0 for 3.9μm.

5.2 Extension of the Bayesian Cloud Detection to Visible Wavelength Imagery

The Bayesian cloud detection algorithm was not initially designed to exploit visible wavelength imagery, as the motivation was to aid retrieval of sea surface temperatures, and, while a fast RTM, RTTOV, was freely available to simulate observations at thermal wavelengths, there was not one available for visible wavelengths. During the time of this project, a fast RTM for visible and near infra-red (NIR) wavelengths (VisRTM) was developed at the University of Edinburgh by Owen Embury, and this was incorporated into the cloud detection algorithm to facilitate exploitation of visible and NIR wavelength imagery, at 0.6μm, 0.8μm and 1.6μm, in addition to the wavelengths already exploited. Only those aspects of implementation of VisRTM into the Bayesian cloud detection for which this PhD project is responsible are described here, Mackie et al. (submitted 2008b) includes a fuller description of VisRTM. An outline of the algorithm, with the extension to visible and NIR wavelength imagery, is shown in figure 5.3.

5.2.1 Implementing the Visible and Near-Infrared RTM

For TIR wavelengths, the RTM is run for each NWP profile, and the resulting BTs interpolated to each pixel, using surface emissivity, ST, and spatial location. For visible and NIR wavelengths, spatial variability in surface reflectance makes an analogous linear interpolation of simulated reflectances from profile to pixel inappropriate, at least for land imagery. VisRTM must therefore be run for every pixel, and so it is required to be very fast, even if somewhat less accurate than proprietary software as a result.

The VisRTM is run for each individual pixel, driven by the background state vector, $\mathbf{x^b}$, which contains NWP information (interpolated from the closest four NWP profile locations onto the pixel location), surface emissivity, and, for land pixels, the parameters used for calculation of surface reflectance, see section 5.2.2, resulting in a predicted observation at visible and NIR wavelengths: $\mathbf{y^b}_{vis,NIR} = \mathbf{F}\left(\mathbf{x^b}\right)$, where **F** represents the VisRTM calculations, and $\mathbf{y^b}$ is the vector of predicted observations for clear sky at all wavelengths. Whereas the profile-pixel elevation difference is used for interpolation of modelled BTs, there is no such proxy for the spatial variability of land surface reflectance, and so results of the VisRTM cannot, at least for land pixels, be linearly interpolated from profile to pixel locations and the VisRTM must be run for every pixel.

Additional A-Priori Data Needed for Forward-Modelling

VisRTM requires Aerosol Optical Depth (AOD) and surface Wind-Speed (WS) data to be provided in the background state vector, $\mathbf{x^b}$. NWP profiles do not always include these fields, in which case a default value of AOD=0.1 is used. For WS, an average wind speed is used, based on the average of all the ocean profiles in a global dataset of atmospheric profiles from the European

Centre for Medium-range Weather Forecasting (ECMWF), see figure 5.4 (Chevalier, 2001). The average WS value, set as the default WS value, was $\mathbf{x^b}_{WS} = 8.63ms^{-1}$. WS only affects the reflectance simulated by VisRTM for ocean surfaces, where it is used to identify areas of sunglint, and so a value appropriate to ocean can be used as a default WS over both land and ocean.

5.2.2 Modelling Surface Reflectance

A surface reflectance field is also required by VisRTM. For ocean pixels, a WS- and viewing geometry- dependent surface reflectivity LUT is provided, pre-calculated from a model developed by Filipiak and Merchant (submitted 2008). The calculations for the LUT reflectivities follow the methods of Watts et al. (1996), which follow from Masuda et al. (1988), using Cox and Munk's wind-speed dependent distribution of wave slopes (Cox and Munk, 1954), which is similar to the work of Wu and Smith (1997). For land pixels, a surface reflectance model was formulated following the kernal-driven method used for calculation of the MODIS Bi-directional Reflectance Distribution Function (BRDF) (Strahler and Muller, 1999). The BRDF describes surface reflectance for a given viewing geometry - this can be integrated to give total hemispherical albedo. The MODIS albedo, product calculated in this way, has been successfully validated through comparison with in-situ measurements for a variety of land surfaces, for example over a range of cropland, soils and man-made structures in America, (Liang et al., 2002), over Africa, (Privette et al., 2004) and over Greenland (Stroeve et al., 2004). A theoretical study of the accuracy of this albedo product, using model-model comparisons, indicates typical uncertainties are below 10%, with a large part of the error being attributable to uncertainties in the cloud mask (Lucht, 1998), although it should be noted that a quality assessment study of the MODIS product found significantly higher root mean square values at higher latitudes, (Shuai et al., 2008). The land surface reflectance model used in cloud detection algorithm is described in the following section.

Land Surface Reflectance Model Predicting the amount of visible and near infra-red radiation reaching a satellite sensor requires consideration of reflection at the surface, which depends on the viewing geometry. This dependence on satellite and solar zenith angles, and on relative azimuth angle, can be described by a Bi-directional Reflectance Distribution Function (BRDF). Different surfaces reflect differently, for example while specular reflection may be dominant for a mirror-like surface, volumetric scattering must be considered for a leafy forest canopy (to account for reflections between leaf surfaces) and in sparse forest, or in mountainous terrain, shadows and the obscuring of objects from view (geometric scattering) must be accounted for. This dependence on the spatial distribution and density of vegetation means the BRDF can be used to characterize vegetated land surfaces and a considerable amount of work has been done in this area, for example Kriebel (1979); Barnsley et al. (1997); Privette et al. (1998). More frequently, a BRDF allows a series of reflectances, observed with different viewing geometries, to be adjusted to a common viewing geometry, which is a prerequisite for surface albedo calculations.

A global surface albedo product, calculated using a BRDF kernel, is available from the Moderate Resolution Imaging Spectroradiometer (MODIS), and the same kernel method is followed in this work to predict surface reflectance. The kernel follows the work of Roujean et al. (1992), and is a sum of three terms, corresponding to isotropic, volumetric and geometric surface scattering, described here in principle. For a fuller description of the algorithm, see Strahler and Muller (1999).

The MODIS BRDF calculation is described in equation 5.3, where the weighting parameters, *f*, depend on the waveband of the sensor channel, λ, and the kernel contributions, $\mathbf{K}$, depend on the satellite zenith angle, θ, solar zenith angle, ν, and the relative azimuth angle, ϕ. The value of the BRDF for a given viewing geometry and given weighting parameters is the modelled directional surface reflectance.

$$BRDF\left(\theta,\nu,\phi,\lambda\right) = f_{iso}\left(\lambda\right) + f_{vol}\left(\lambda\right) \times \mathbf{K}_{vol}\left(\theta,\nu,\phi\right) + f_{geo}\left(\lambda\right) \times \mathbf{K}_{geo}\left(\theta,\nu,\phi\right) \qquad (5.3)$$

The gradient of the difference between this modelled directional surface reflectance and that which is actually observed is minimised with respect to f in order to determine each f value. If a negative value is found for *f*, it is set to 0 and the other f values recalculated, maintaining semi-orthogonality.

The kernel contribution, $\mathbf{K}_{vol}$, is based on the radiative transfer work of Ross (1981) for a dense leaf canopy with homogeneous leaf angle distribution, and a Lambertian background, known as the RossThick kernel. It is described in equation 5.4, using the same notation for angles as for equation 5.3, and $\cos\left(\epsilon\right) = \cos\left(\theta\right) \times \cos\left(\nu\right) + \sin\left(\theta\right) \times \sin\left(\nu\right) \times \cos\left(\phi\right)$.

$$\mathbf{K}_{vol} = \frac{\left(\frac{\pi}{2} - \epsilon\right) \times \cos\left(\epsilon\right) + \sin\left(\epsilon\right)}{\cos\left(\theta\right) + \cos\left(\nu\right)} - \frac{\pi}{4} \qquad (5.4)$$

A sparse arrangement of three-dimensional objects on a Lambertian background is

assumed for $\mathbf{K}_{geo}$, taken from Wanner et al. (1995), based on the work of Li and Strahler (1992). These studies found the shape of tree crowns to be more important to the BRDF than their height from the ground. For the surface reflectance calculations in the Bayesian cloud detection, spherical tree crowns at a height from the ground equal to half their diameter are assumed. Equation 5.5 describes $\mathbf{K}_{geo}$, where the angles are defined as for equations 5.3 and 5.4 and Ω describes overlapping shadows: $\Omega = \frac{1}{\pi} \times (t - \sin(t)\cos(t))(\sec(\theta) + \sec(\nu))$, using $\cos(t) = 2\left(\frac{\sqrt{D^2 + (\tan(\theta)\tan(\nu)\sin(\phi))^2}}{\sec(\theta)+\sec(\nu)}\right)$ and $D = \sqrt{\tan^2(\theta) + \tan_2(\nu) - 2\tan(\theta)\tan(\nu)\cos(\phi)}$.

$$\mathbf{K}_{geo} = \Omega - \sec(\theta) - \sec(\nu) + \frac{1}{2} \times (1 + \cos(\epsilon)) \times \sec(\theta) \times \sec(\nu) \tag{5.5}$$

Atmospherically corrected observations for an image pixel can be used to determine f for each sensor channel (which will change with surface type) and the kernel contributions,$\mathbf{K_{vol}}, \mathbf{K_{geo}}$, can be calculated from the viewing geometry. Temporal variations in sensor noise make it appropriate to use an average of calculated values of f for a given location, and seasonal variations in land surface cover make it appropriate to use a different f value for each season. A map of f values can be combined with the viewing geometry of another sensor to predict the surface reflectance at a given image pixel location.

The MODIS sensor retrieves the weighting parameters, *f*, for land pixels, and makes 16-day averages available on a 0.05^o equal area grid. This data, acquired between 2006 and the present, was grouped into 4 seasonal maps, corresponding to December, January and February; March, April and May; June, July and August; and September, October and November. While issues of computational efficiency make it prudent to limit the volume of reference data required by the Bayesian cloud detection, seasonal changes in land surface cover mean it would be inappropriate to group all the data together. The average value for each parameter for each of the 3 required wavelengths was calculated for each season for each pixel, creating 4 reference maps. For computational efficiency when loaded into the cloud detection code, these maps were upscaled to a 0.1^o grid (after first screening out ocean grid cells to ensure that the 0.1^o grid cell contains the average of only the land 0.05^o grid cells it contains). The map for the appropriate season is loaded into the cloud detection code and the pixel latitude and longitude are used to retrieve the f values for each pixel, which are combined with the contributions $\mathbf{K}$, calculated from the pixel's viewing geometry as described above, to predict surface reflectance, SR, for the pixel, which is then passed to VisRTM.

5.2.3 New Error Considerations Required

The reduced background state vector **z** introduced in section 2.4.2, is now extended to include the fields needed by VisRTM, as shown in equation 5.6. The dependency of the VisRTM calculations on these additional fields is required for equation 2.3, as $\mathbf{H}' = \frac{\partial \mathbf{y^b}}{\partial \mathbf{z}}$, and these dependencies are calculated for each VisRTM run. The errors associated with these fields are added into the **B**-matrix, see equation 5.7. For all pixels, $\frac{\partial \mathbf{y^b}_{TIR}}{\partial \mathbf{z}_{AOD,WS,SR}} = 0$.

$$\mathbf{z} = \begin{pmatrix} \mathbf{x^b}_{st} \\ \mathbf{x^b}_{tcwv} \\ \mathbf{x^b}_{AOD} \\ \mathbf{x^b}_{WS} \\ \mathbf{x^b}_{SR} \end{pmatrix} \tag{5.6}$$

$$\mathbf{B} = \begin{pmatrix} \left(\epsilon^b_{st}\right)^2 & 0 & 0 & 0 & 0 \\ 0 & \left(\epsilon^b_{tcwv}\right)^2 & 0 & 0 & 0 \\ 0 & 0 & \left(\epsilon^b_{AOD}\right)^2 & 0 & 0 \\ 0 & 0 & 0 & \left(\epsilon^b_{WS}\right)^2 & 0 \\ 0 & 0 & 0 & 0 & \left(\epsilon^b_{SR}\right)^2 \end{pmatrix} \tag{5.7}$$

Wind-Speed Error. For land pixels, $\frac{\partial \mathbf{y^b}}{\partial \mathbf{x^b}_{WS}}$ is 0, and so it is acceptable to use a value for ϵ^b_{WS} that is appropriate for only ocean pixels (since the tangent linear and corresponding error are multiplied together in equation 2.3). The standard deviation of the WS values which are used to set the default value, see section 5.2.1, can be used as the uncertainty required for the **B**-matrix. The distribution of WS values in the dataset, however, is not Gaussian, and only 66% of the profiles fall within one standard deviation of the mean. Uncertainties in the Bayesian cloud detection technique are all treated as Gaussian, see section 2.4.2, and so the algorithm was run using both $\epsilon^b_{WS} = 4.19ms^{-1}$= 1 standard deviation, and 1.5 standard deviations (which includes 87.7% of the profiles). The results were identical, except in regions of sunglint, where differences in the calculated probability of clear were still generally too small to affect a binary mask made

using a threshold of 50% on the calculated probability of clear. It was therfore decided that, not-withstanding the non-Gaussian nature of the error, an error of one standard deviation could be assigned, making the **B**-matrix term $\left(\epsilon^b_{WS}\right)^2 = 17.56m^2s^{-2}$. In cases where a WS field is provided with the NWP data, this error is probably conservative, however, in the absence of an alternative value, it is always used. It is likely that this error, and all errors in the **B**-matrix, are spatially- and temporally variable, but to account for this would increase the computational cost of the algorithm and is considered unlikley to have a significant effect on the result.

Land Surface Reflection Error. The SR error needed for the **B**-matrix depends on the errors in the parameters, *f*. An estimated error appropriate to a seasonal average over a 0.1^o grid is required. Theoretical accuracies for the MODIS retrieved parameters product have been found through inter-model comparisons (Lucht, 1998). However model to model comparisons can only produce a relative error, which may differ from the absolute error. The errors found were less than 10% for all 3 parameters, for all channels, and it was noted that averaging the retrieved data over 16 day periods did not significantly affect the product accuracy. The figure of 10% is in agreement with more recent studies that have compared in-situ measurements against the MODIS albedo product, which is calculated form the retrieved f parameters using the same surface reflectance model implemented here, for example Privette et al. (2004); Stroeve et al. (2004). On this basis we use, a single error term of 10% for the calculated surface reflectance, ϵ^b_{SR}, in the **B**-matrix of equation 5.7, intended to reflect the uncertainty in all 3 *f* parameters. A full treatment of the error, accounting for spatial, seasonal and viewing geometry- variability would extend the **B**-matrix to computationally expensive proportions, and, provided the error is small, this approximation for the uncertainty is considered appropriate. The dependence of the total estimated visible and NIR radiances on the surface reflectance is calculated by the VisRTM and added to the $\mathbf{H}'$-matrix.

Forward-Modelling Uncertainty for VisRTM. The **R**-matrix in equation 2.3 accounts for the FM-uncertainty in each channel, and so it is now extended from equation 2.6.

The database described in section 3.3 is used as a 'truth', against which to compare VisRTM results. The difference between the two gives an estimate of the error in the VisRTM calculations, see figure 5.5 and table 5.1. Sunglint targets were excluded from this error assessment, as they resulted in unacceptably high FM errors. A higher uncertainty is already allowed for in sunglint areas through consideration of the WS error and this is judged sufficient compensation for excluding them from the FM error calculation. Targets classed as 'sea with shadow' were also not considered. These targets have a lower reflectance than open sea without shadow, meaning both that they could increase the spread of reflectances considerably, increasing the standard deviation used for the FM error, and that they already appear as a low probability of cloud in the PDF for cloud.

Agreement between modelled and observed reflectances for snow and ice targets is lower, probably as a result of inaccurate surface reflectivity: the surface reflectance obtained from seasonally-averaged data cannot effectively account for short-term snow cover. Increasing the surface reflectance error term in the **B**-matrix would address this, but would reduce the success of the cloud detection for all classes of snow-free land surfaces. NWP fields could in principle be used to judge the likelihood of snow/ice cover could be used to set an NWP-dependent surface reflectance error: however, for this work a single value including the effect of snow and ice targets was included in calculation of the VisRTM errors and biases.

A solar-zenith-dependent error and bias could be considered, but the technique performs well assuming independent errors and biases, and the distribution of solar zenith angles represented in the database for clear sky targets does not allow for meaningful solar-zenith-dependent errors to be calculated.

A quadratic fit to the output from a line-by-line model was used in development of VisRTM to define atmospheric transmission in terms of TCWV, see figure 5.6. Some uncertainty in simulations calculated by VisRTM stems from this fit, which is not equally accurate for all three channels, or for all atmospheric path lengths. Further uncertainty could stem from the fact that the VisRTM does considers only first-order scattering, an approximation referred to as the *single scattering approximation*.

Channel Wavelength (μm)	Land (1425 targets)	Land with Shadow (19 targets)	Ice (5 targets)	Snow (319 targets)
Mean: modelled - observed reflectance				
0.6	+0.0008	+0.0567	+0.0759	-0.0979
0.8	+0.0144	+0.0751	-0.0054	-0.0899
1.6	-0.0232	+0.1007	-0.0368	+0.0291
Standard Deviation: modelled - observed reflectance				
0.6	0.0276	0.0419	0.1489	0.0764
0.8	0.0291	0.0515	0.1441	0.0679
1.6	0.0298	0.0915	0.0317	0.0287
Cross-Channel Covariance: modelled - observed reflectance				
0.6, 0.8	0.0005	0.0021	0.0214	0.0047
1.6, 0.6	0.0001	0.0037	0.0041	0.0012
1.6, 0.8	0.0004	0.0045	0.0039	0.0007

Table 5.1: Table showing the errors and biases in the VisRTM calculations for different land surface types - reflectances are fractions, not percentages.

This method of estimating errors is not ideal in two ways. Firstly, the database used is the same as that used for validation of the technique, see section 5.3, which may mean that the errors are somewhat 'tuned' to the validation data. Secondly, the differences in figure 5.5 result from both uncertainty in the VisRTM calculation and uncertainty in the NWP information used for

the calculation. The latter source should already be accounted for in the **B**- and $\mathbf{H}'$- matrices, and so taking the whole difference to be the FM-error results in a conservative error estimation. Separating the differences into NWP uncertainty, and uncertainty in the VisRTM, however, is challenging, and no straight-forward method could be devised. The results are found not to depend strongly on the precise value of these errors. Therefore, the uncertainty in VisRTM calculations for each channel for sea and land pixels was set at the standard deviation of the difference between modelled and observed visible and NIR radiances for that channel for clear sea and land pixels respectively. The squares of these errors were then added to the TIR FM errors and sensor noise along the diagonal of the **R**-matrix.

Some correlation was noted between the VisRTM error for different channels, see figure 5.7. For land targets, the correlation between modelled-observed difference at NIR and at visible wavelengths can mostly be attributed to snow and ice targets, and so including these covariances as off-diagonal elements in the **R**-matrix improves the performance of the algorithm over snow and ice land surfaces. Off-diagonal terms in the **R**-matrix were also set for the covariance between the errors at 0.6μm and 0.8μm for both sea and land targets.

A bias correction was set according to the peak of the histograms in figure 5.8, which include only data points within 1.5 standard deviations of the mean, outliers being considered to correspond to high uncertainty in NWP information and therefore excluded. VisRTM does not consider higher order scattering, which is more significant at shorter wavelengths and this effect can be seen in figure 5.8a, where the model's under-prediction of reflectance increases as the wavelength is decreased. The value of the peaks in this histogram were used to apply a bias correction to the VisRTM-estimated reflectances at visible and NIR wavelengths over sea. For land targets, variability in the uncertainty of NWP information and of surface reflectance parameters masks this effect and results in the histogram peaks being less well-defined, and not positioned as expected, see figure 5.8b. Unknown biases in the NWP information used in the VisRTM for land targets may be creating a positive bias in the modelled reflectances, countering the intrinsic negative bias resulting from the single scattering approximation. Scattering at 0.6μm is more significant, causing the negative bias to increase and the peak to move to the left relative to the other channel peaks. The increased dominance of the positive bias for 0.8μm and 1.6μm is indicated by the position and ill-definition of their peaks, suggesting a variable uncertainty, which is expected in the NWP information for land targets. A further explanation for the difference between the sea and land histograms is the intrinsic variability of the spectral properties of land surfaces, relative to those of sea surfaces. No bias correction is made to the estimates of the VisRTM at 0.8μm and 1.6μm for land targets, and a 0.6μm bias correction is made using the peak of the histogram in figure 5.8b.

5.2.4 Visible and NIR PDFs

Spectral PDFs for Cloud. To calculate LUTs for $P\left(\mathbf{y}^{\mathbf{b}}{}_{s}\,\middle|\mathbf{x}^{\mathbf{b}},\bar{c}\right)$, cloud pixels were selected from imagery acquired by Meteosat-8 sensor between 2004 and 2006. Cloud pixels were detected

using the Bayesian detection with TIR channels only, and only ocean pixels were selected (the Bayesian detection being more accurate over ocean when only TIR imagery is used). A threshold of 30% was set on the resulting probability of clear. This means that the final PDF cannot represent all clouds - those which were not detected with relatively high certainty by the technique using TIR imagery are excluded and so will not necessarily be considered within the technique when VIS and NIR are considered. In the absence of readily available cloud observations, however, this was judged the most convenient method to construct a PDF. PDFs are constructed separately for the four seasons, with observations from each image weighted by the time of day of the image's acquisition. The weighting is based on the frequency with which that acquisition time occurs in the dataset for that season, and so ensures for example that GMT-early-morning observations are not over-represented in Spring. Separate PDFs are made for solar zenith angles 0^o-10^o, 10^o-20^o,70^o-80^o. One-, two- and three-dimensional PDFs were constructed in this way for each VIS-NIR channel combination, giving the user the flexibility to use whichever channels are available (useful if data from a particular channel on the sensor malfunctions) whilst defaulting to the optimum channel combination for efficiency. The PDFs are shown in figure 5.9 and 5.10. As in section 5.1, it is assumed that the spectral properties of clouds over land are the same as those for clouds over sea and so the same PDF for cloud, constructed from cloud imagery over sea, which was considered more reliable than cloud imagery over land, is used for both. This is not always valid, for example in the case of sea fog, but in the absence of a reliable dataset of imaged clouds over land, in which all possible cloudy states are represented, it was considered a necessary assumption.

Textural PDFs for Clear and Cloud. A textural PDF at 1.6μm, to be used only for ocean pixels, was constructed from the same imagery as for the spectral PDFs described in the preceding paragraph. The data was grouped into 10^o solar zenith angle bins and a separate PDF constructed for each group. An array of local standard deviations (LSDs) was calculated using a 'sliding window' to assign to each pixel position the mean 1.6μm reflectance observed for pixels with the same cloud/clear classification in a 3x3 box centred on it. Cloud screening was carried out using TIR channels only. Ocean pixels with a calculated probability of clear lower than 30% and higher than 90% were deemed cloud or clear respectively for inclusion in the appropriate textural PDF. While the high threshold for clear pixels makes cloud contamination of the clear PDF unlikely, a corresponding threshold of 10% for cloud pixels would be likely to exclude many cloud edge pixels, which should be represented in the PDF for cloud, and so the higher threshold of 30% was chosen. The textural PDF for clear and for cloud is shown in figure 5.11, summed over all solar zenith angles.

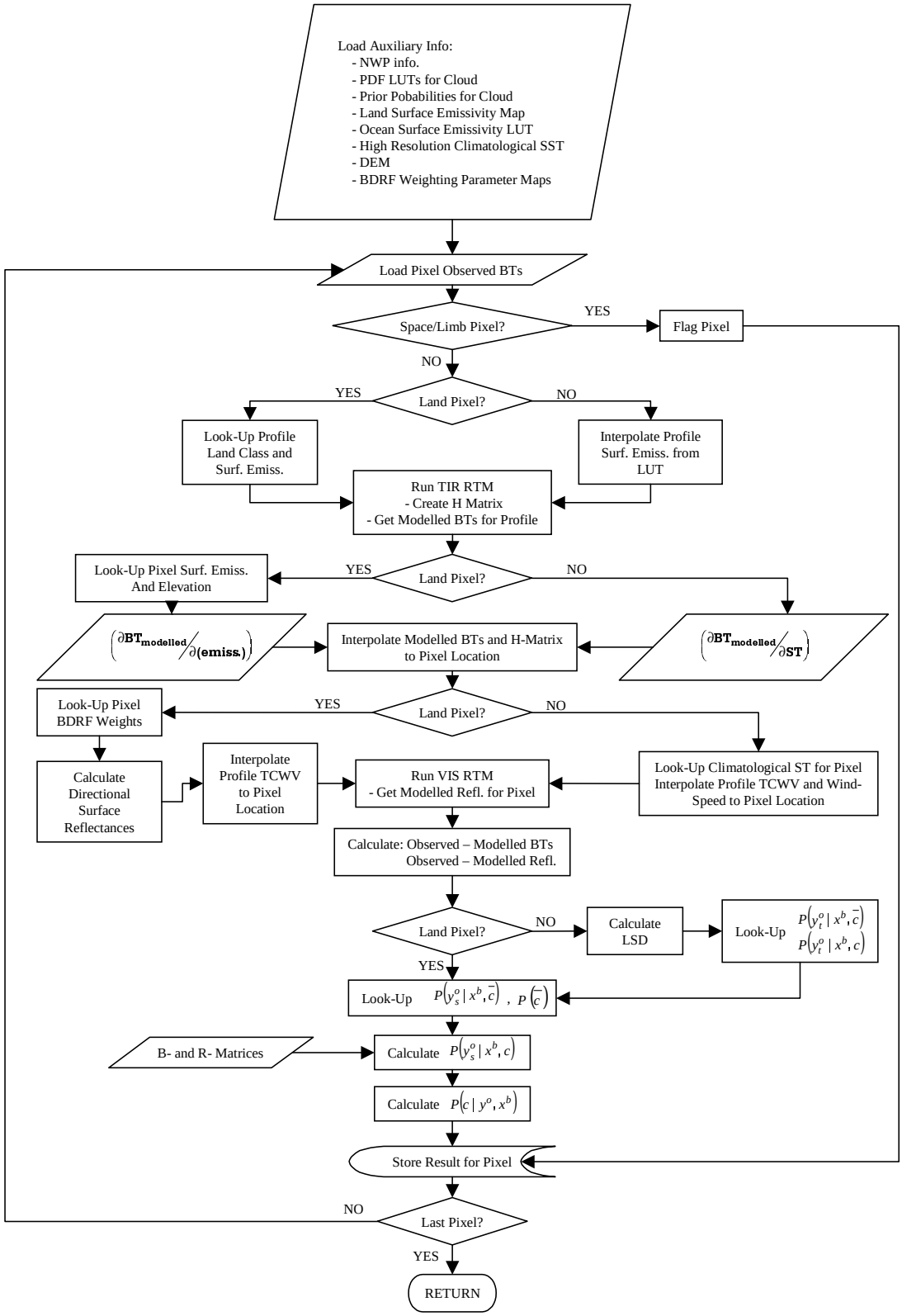


Figure 5.3: Flowchart outlining the Bayesian cloud detection algorithm, with the extension to visible and NIR wavelength imagery.

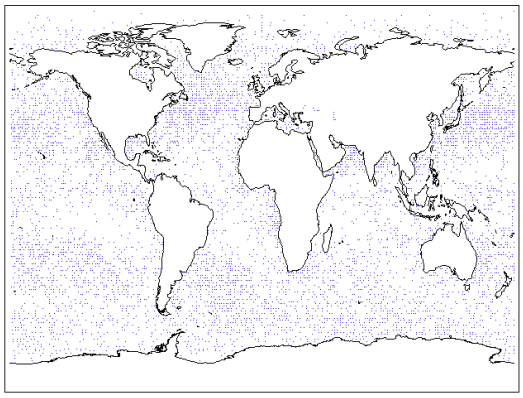


Figure 5.4: Location of ocean profiles in the dataset of ECMWF profiles used to calculate a default surface wind-speed value.

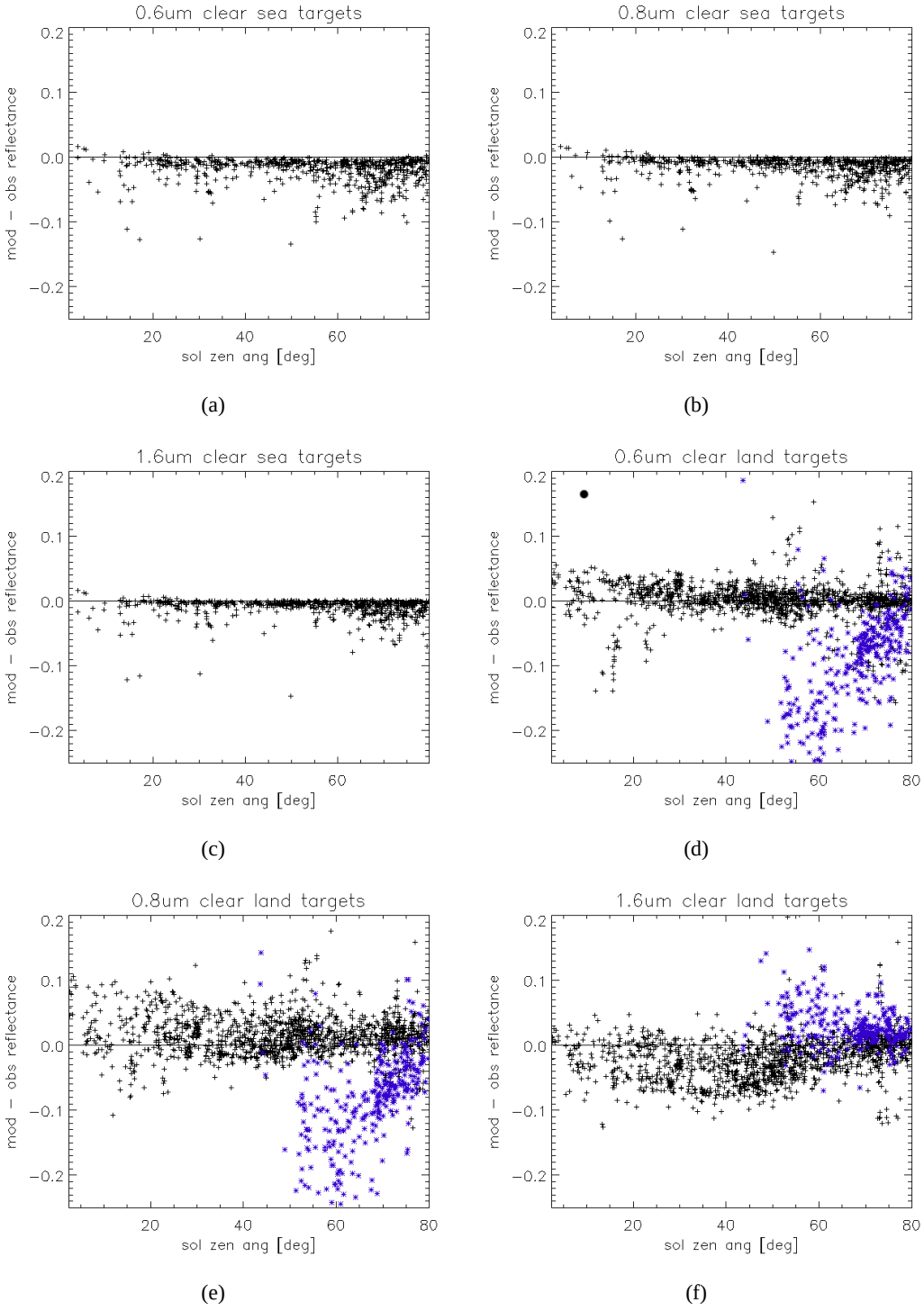


Figure 5.5: Plots show differences between observed and modelled visible and NIR reflectances for clear sky targets from the database described in 3.3 for observations recorded over sea at (a) 0.6μm, (b) 0.8μm and (c) 1.6μm; and over land at (d) 0.6μm, (e) 0.8μm and (f) 1.6μm. Snow and ice targets are shown as blue stars, and other land/sea targets as black crosses. The standard deviation of the difference was considered the FM-error in VisRTM.

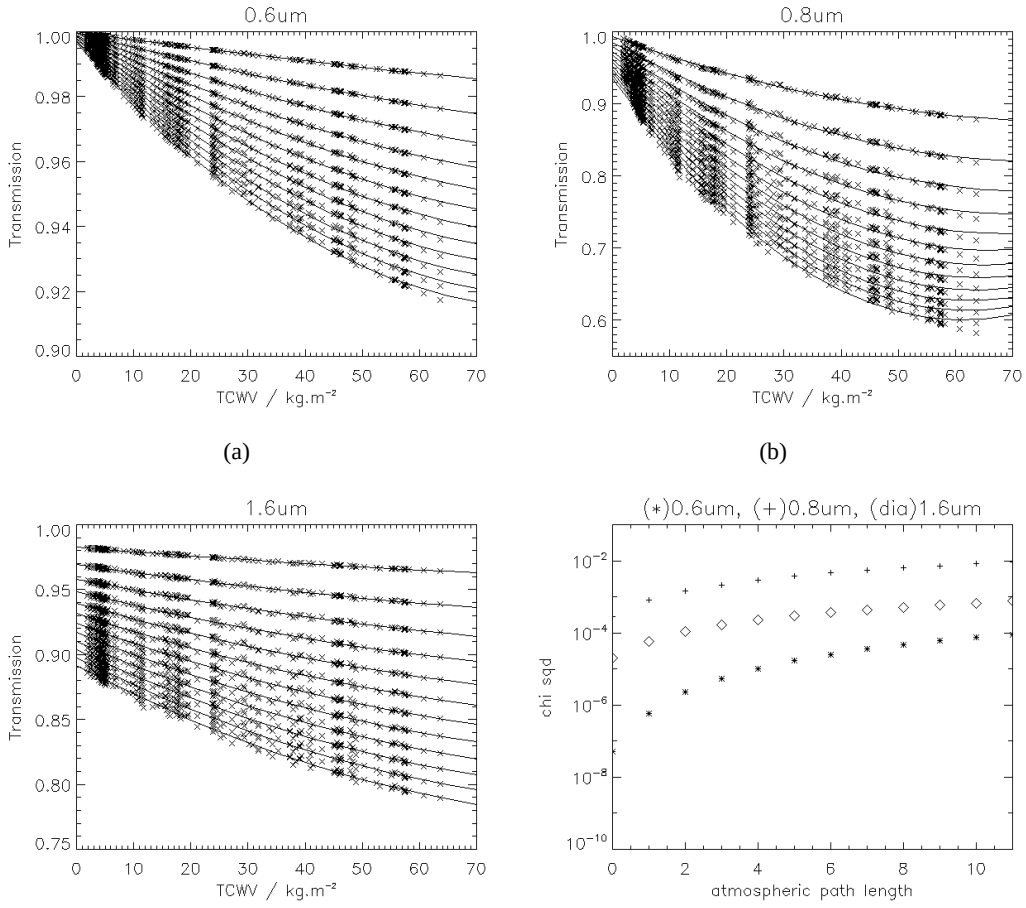


Figure 5.6: Crosses mark the results of line-by-line model calculations. The fitted curves show the relationship assumed by VisRTM. The 12 different curves on each plot correspond to atmospheric path lengths of 1-12: (a) 0.6μm, (b) 0.8μm and (c) 1.6μm. The chi squared values calculated for each curve are shown in (d).

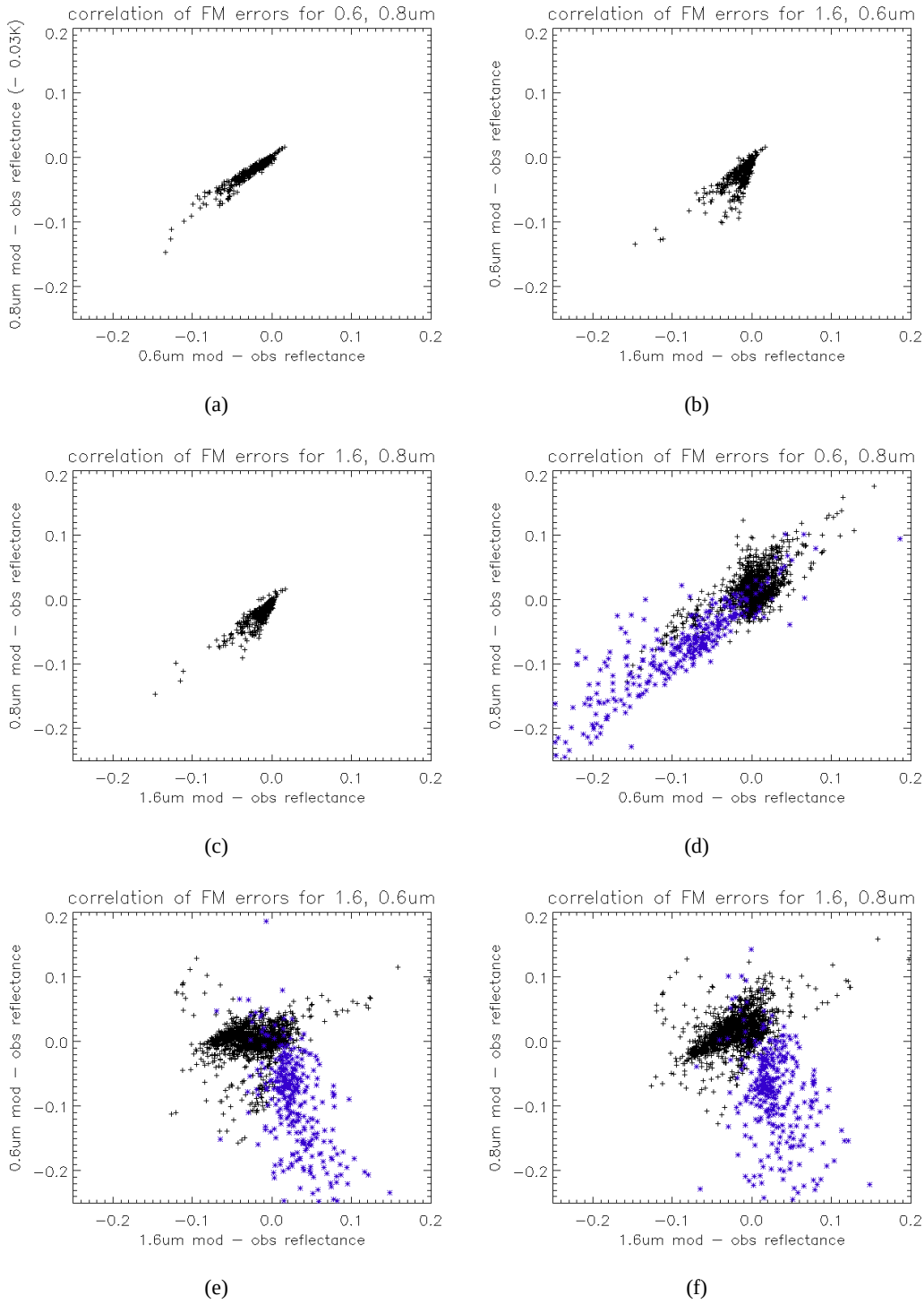
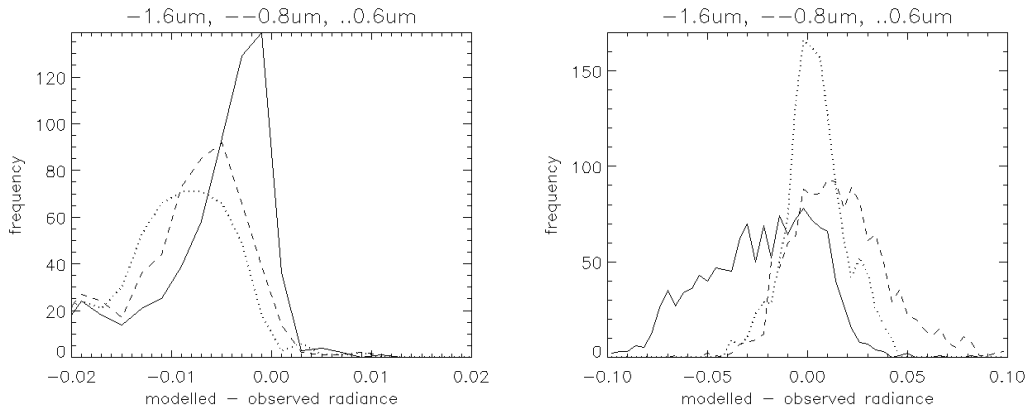
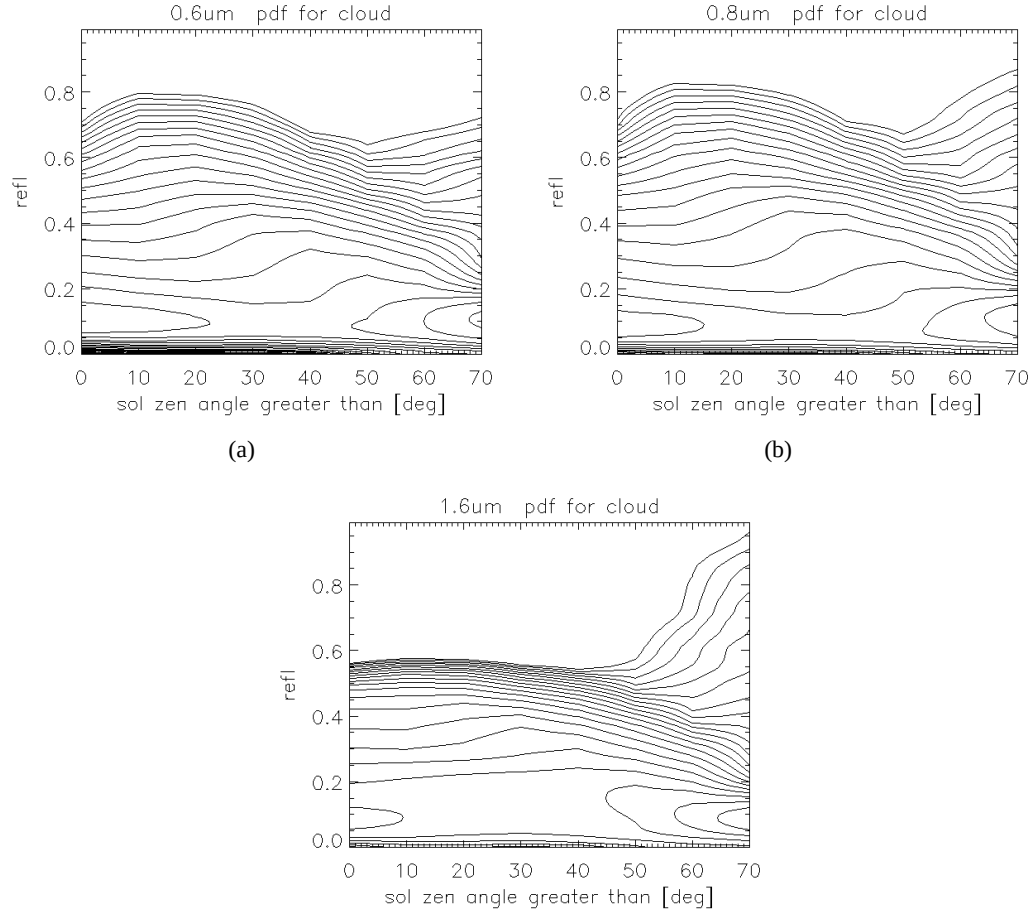


Figure 5.7: Correlation between modelled-observed reflectance differences for sea targets in the database described in 3.3: (a) 0.6 and 0.8μm, correlation coefficient=0.97; (b) 1.6 and 0.6μm, correlation coefficient=0.82; (c) 1.6 and 0.8μm, correlation coefficient=0.90; and for land targets from the same database, with snow and ice targets shown as blue stars, and other land targets as black crosses: (d) 0.6 and 0.8μm, correlation coefficient=0.88; (e) 1.6 and 0.6μm, correlation coefficient=-0.35; (f) 1.6 and 0.8μm, correlation coefficient=-0.18.



(a) (b)

Figure 5.8: Histograms of the difference between modelled and observed visible and NIR reflectances for clear sky targets from the database described in 3.3, for (a) sea targets and (b) land targets. The value at the peak was applied as a bias correction to the visible and NIR reflectances simulated by VisRTM.



(a) (b)

(c)

Figure 5.9: Plots of the single-channel spectral PDFs for cloud used for visible and NIR imagery: (a) 0.6μm, (b) 0.8μm and (c) 1.6μm, with contours marking a logarithmic scale of $10^{-0.7,-0.8,...,+1.3}$ in units of probability per % reflectance. These plots show the single channel, solar-zenith angle dependent PDFs.

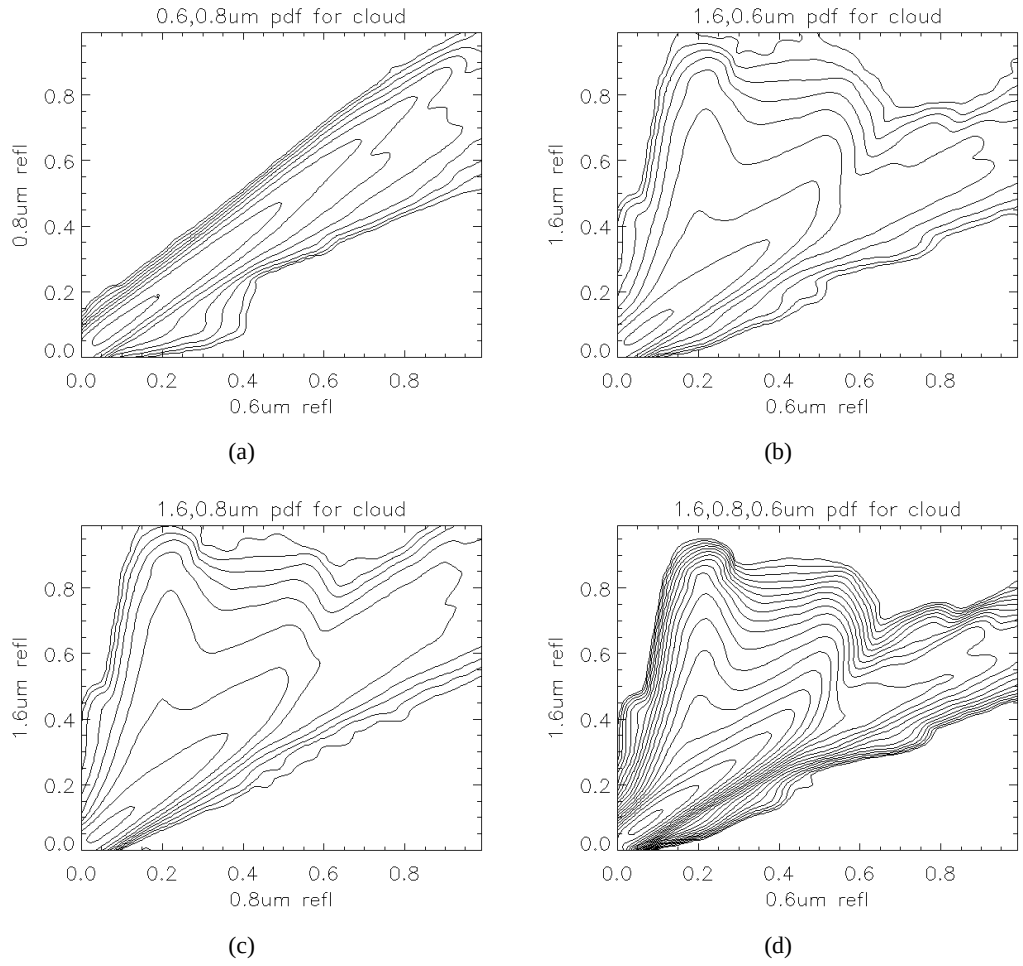


Figure 5.10: Plots of the 2- and 3-channel spectral PDFs for cloud used for visible and NIR imagery: 2-channel PDFs are shown summed over the solar zenith dimension for: (a) 0.6μm, (b) 0.8μm and (c) 1.6μm, with contours marking a logarithmic scale $10^{-3.15,-2.45,...,+2.15}$ in units of probability per % reflectance (the lowest contour value is represented by the outermost contour). The 3-channel PDF is shown in (d), summed over both the solar zenith dimension and the 0.8μm dimension, with contours spaced logarithmically at $10^{-0.25,-0.5,...,+4.75}$ in units probability per of % reflectance.

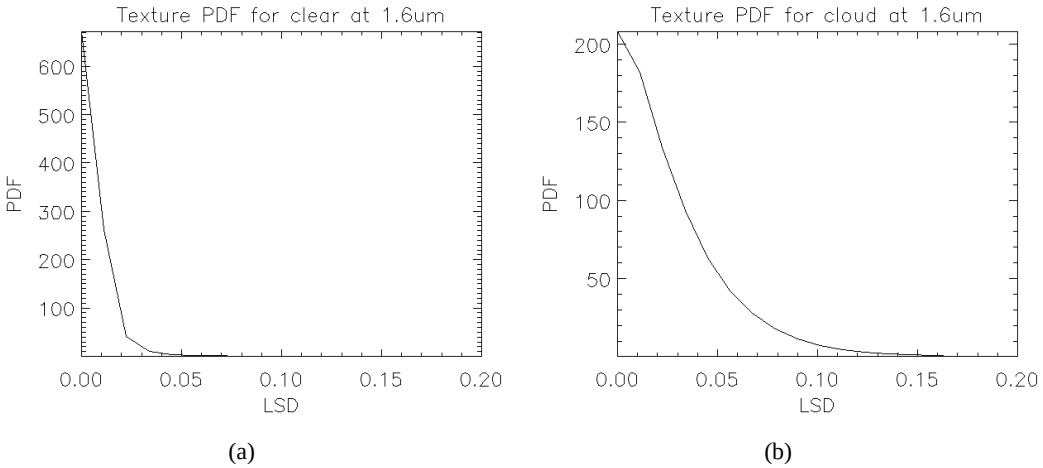


Figure 5.11: Plot of the textural PDF for 1.6μm for (a) clear, and (b) cloud.

5.3 Validation of Bayesian Cloud Detection with Extensions

This section describes the validation work carried out for the extensions to the Bayesian cloud detection method described in sections 5.1 and 5.2. The database described in appendix 3.3 was used for the validation, providing results from the U.K. Met. Office and MétéoFrance for comparison with the results of the Bayesian technique. It should be emphasised that these targets are considered 'difficult' cases, and operational results from both Météo-France and the U.K. Met. Office are therefore not representative of their performance overall.

An ocean ST climatology averaged from several years of data, with very high spatial resolution, was used (Faugere et al., 2001). This climatological ST is altered such that its average value over a 11x11 grid cell area, centred on the cell to which it is assigned, is the same as the weekly Reynolds ST (Reynolds et al., 2002) for that area (the Reynolds ST is at 1^o resolution), thereby correcting for year-to-year variation. This corrected climatological ST is then used as the NWP ST ($\mathbf{x^b}$) for the RTM.

Estimates are required for the uncertainties in **z**, to enter into the **B**-matrix in equation 2.4. Work was done by Dr. Old at the University of Edinburgh to estimate these as $\epsilon^b_{st} = 0.4K$ and $\epsilon^b_{tcwv} = 0.18$, see Mackie et al. (submitted 2008a) for details.

Sensor noise is set to +/- 0.17, 0.11 and 0.15K for channels at 3.9, 11 and 12μm respectively, and +/- 0.0005, 0.0006 and 0.0009 (fractions, not percentages) for channels at 0.6, 0.8 and 1.6μm respectively, taken from Schmetz et al. (2002).

5.3.1 Definition of Skills Scores

The probabilistic results of the Bayesian cloud detection are converted to a binary result using a threshold of 50% on the calculated target probability of clear. The Proportional of Perfect classifications (*PP*), Hit Rate (*HR*), False Alarm Rate (*FAR*) and True Skill Score (*TSS*) are then used to give a measure of the skill of the technique. These are calculated from the number of correctly classified clear targets (*z*); the number of falsely classified clear targets (*f*); the number of correctly classified cloud targets (*h*); and the number of falsely classified cloud targets (*u*), see equations 5.8, with *TSS* being arguably the best measure of overall skill, since it considers both the skill in detecting cloud and the skill in avoiding false detection.

$$\begin{aligned} PP &= \frac{(h+z)}{(h_z+f+u)} \times 100 \\ HR &= \frac{h}{(h+z)} \times 100 \\ FAR &= \frac{f}{(f+z)} \times 100 \\ TSS &= (frach(h+u) - fracf(f+z)) \times 100 \end{aligned} \tag{5.8}$$

5.3.2 Channel Combination Selection

Exploiting as many channels as possible, and so making the distinction between cloud and clear based on the maximum amount of sensor-recorded information, should give the best possible estimated classification, provided the forward models have known, zero-mean Gaussian-distributed errors. In practice, deviation from these ideal conditions make it worth considering channel selection. For example, the maps of BDRF weighting-parameters may not be equally accurate for all three visible and NIR channels, but the errors used would not reflect this. VisRTM relies on a fit between TCWV and transmission, the accuracy of which may vary between channels, see Mackie et al. (submitted 2008b). Scattering is generally less significant at longer wavelengths, and so the accuracy of the scattering contribution to VisRTM should be higher for 1.6μm than for 0.8μm and 0.6μm, see section 5.2.3, and sensor noise also varies between channels. The treatment of TIR imagery relies heavily on the NWP ST field, which is considered less accurate for land than for ocean, see section 5.1.3. The high value ascribed to ST uncertainty in the **B**-matrix for land pixels may make the clear PDF less distinct from the cloud PDF, resulting in a more ambiguous interpretation of TIR observations. For daytime imagery, the Bayesian technique can consider two visible and one near NIR channel (at 0.6, 0.8 and 1.6μm), referred to collectively as VIS, and two TIR channels at 11 and 12μm. The Bayesian cloud detection was run using different combinations of these channels on the validation dataset, and the results converted to a binary cloud mask using a threshold of 50% on the probability of clear calculated for each target. The True Skills Score (*TSS*) was then used to indicate the skill of the cloud detection for each combination. Only daytime targets were considered, the best results at night-time having already been seen when all 3 TIR channels are exploited (Merchant et al., 2005). The results are shown alongside results of the operational techniques of Météo-France and, for land targets, of the U.K. Met. Office in table 5.2.

It is encouraging that the highest skills score for land imagery is achieved when all sensor channels are considered, indicating that the surface reflectance model is working well. Over sea, consideration of imagery at 1.6μm does not appear to help the detection, although it is interesting that consideration of 1.6μm alone with the TIR channels results in a higher skills

Channel Combination	***TSS* [%]**	
	Day, Sea	**Day, Land**
all channels	87.50	**72.84**
all TIR	72.52	44.55
all TIR, 1.6, 0.8μm	86.60	65.56
all TIR, 1.6, 0.6μm	86.04	72.24
all TIR, 0.8, 0.6μm	**89.44**	55.89
all TIR, 1.6μm	87.19	59.96
all TIR, 0.8μm	88.89	49.43
all TIR, 0.6μm	88.23	52.67
1.6, 0.8, 0.6μm	82.03	67.34
1.6, 0.8μm	78.83	60.13
1.6, 0.6μm	76.70	65.60
0.8, 0.6μm	85.19	51.89
U.K. Met. Office	**86.30**	-
MétéoFrance	**90.37**	**70.23**

Table 5.2: Skill of the Bayesian cloud detection using different channel combinations, and of the operational techniques.

score than considering 1.6μm with the TIR channels and either 0.8 or 0.6μm. This is surprising, as atmospheric transmission is thought to be represented reasonably accurately in the VisRTM for 1.6μm, see figure 5.6. There is no heightened sensitivity to the surface reflectance model at 1.6μm, and the single-order approximation for Rayleigh scattering in the VisRTM is anticipated to be less valid at shorter wavelengths. The sensor calibration at 1.6μm may be less accurate for the other channels, or it may be a peculiarity of the targets in this dataset - for example, it could be that some targets in the database correspond to atmospheric states where this is not the case, for example some aerosol with unusual scattering properties may be present. In view of these results, the channel combination 11, 12, 0.8 and 0.6μm is used for sea targets, while for land targets data from all observation channels is exploited.

5.3.3 Results

The results of the validation are separated into land and sea results for night-time targets in tables 5.3 to 5.6, and the same separation is made for day-time targets in tables 5.7 to 5.10, using the skills scores defined in section 5.3.1. The target codes are described in appendix 3.3.

Night Targets

	Bayesian			
-	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**
Sea	93.69	93.80	6.50	**87.30**
Land	78.06	84.51	36.14	**48.37**

Table 5.3: Skill of the Bayesian cloud detection for night-time targets

Night Targets

-	**U. K. Met. Office**				**MétéoFrance**			
-	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**
Sea	88.43	84.58	5.31	**79.26**	85.56	80.92	6.90	**74.02**
Land	-	-	-	-	62.86	51.92	13.04	**38.88**

Table 5.4: Skill of the operational methods of the U.K. Met. office and MétéoFrance, for night-time targets for comparison with table 5.3

Night Targets Over Sea

Code	**Bayesian**				**U.K. Met. Office**				**MétéoFrance**			
	z	f	h	u	z	f	h	u	z	f	h	u
101	1179	82	-	-	1194	67	-	-	1174	87	-	-
502	-	-	715	88	-	-	537	266	-	-	502	301
503	-	-	559	15	-	-	545	29	-	-	514	60
601	-	-	23	23	-	-	26	20	-	-	17	29
606	-	-	8	1	-	-	9	0	-	-	9	0
608	-	-	64	0	-	-	64	0	-	-	64	0
609	-	-	37	0	-	-	37	0	-	-	37	0
701	-	-	66	0	-	-	65	1	-	-	65	1
705	-	-	72	0	-	-	72	0	-	-	72	0
706	-	-	2	0	-	-	2	0	-	-	2	0
707	-	-	41	0	-	-	41	0	-	-	41	0
801	-	-	57	0	-	-	57	0	-	-	57	0
802	-	-	121	0	-	-	121	0	-	-	121	0
811	-	-	82	0	-	-	82	0	-	-	82	0
812	-	-	75	0	-	-	75	0	-	-	75	0

Table 5.5: Number of correctly classified clear targets (z), incorrectly classified clear targets (f), correctly classified cloud targets(h) and incorrectly classified cloud targets (u) for all night targets over sea.

Night Targets Over Land

Code	**Bayesian**				**MétéoFrance**			
	z	f	h	u	z	f	h	u
151	1039	548	-	-	1418	169	-	-
191	9	45	-	-	9	45	-	-
502	-	-	1627	450	-	-	958	1119
503	-	-	825	104	-	-	404	525
607	-	-	1	0	-	-	1	0
608	-	-	45	0	-	-	45	0
609	-	-	49	0	-	-	49	0
702	-	-	103	2	-	-	84	21
705	-	-	108	4	-	-	82	30
707	-	-	26	0	-	-	26	0
801	-	-	40	0	-	-	40	0
802	-	-	161	1	-	-	123	39
811	-	-	38	0	-	-	34	4
812	-	-	31	0	-	-	31	0

Table 5.6: Number of correctly classified clear targets (z), incorrectly classified clear targets (f), correctly classified cloud targets(h) and incorrectly classified cloud targets (u) for all night targets over land.

Day Targets

Bayesian				
-	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**
Sea	89.44	95.67	6.22	**89.44**
Land	86.60	91.96	19.12	**72.84**

Table 5.7: Skill of the Bayesian cloud detection for day-time targets

Day Targets

-	**U. K. Met. Office**				**MétéoFrance**			
-	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**
Sea	93.13	93.09	6.79	**86.30**	95.12	95.01	4.64	**90.37**
Land	-	-	-	-	85.27	89.63	19.40	**70.23**

Table 5.8: Skill of the operational methods of the U.K. Met. office and MétéoFrance, for day-time targets for comparison with table 5.7

Day Targets Over Sea

Code	Bayesian				U.K. Met. Office				MétéoFrance			
	z	f	h	u	z	f	h	u	z	f	h	u
101	672	40	-	-	665	47	-	-	673	39	-	-
102	1	1	-	-	1	1	-	-	1	1	-	-
106	156	14	-	-	158	12	-	-	169	1	-	-
502	-	-	482	26	-	-	437	71	-	-	470	38
503	-	-	505	20	-	-	504	21	-	-	505	20
504	-	-	1	0	-	-	1	0	-	-	1	0
601	-	-	37	32	-	-	35	34	-	-	38	31
606	-	-	34	0	-	-	34	0	-	-	33	1
608	-	-	73	0	-	-	73	0	-	-	73	0
609	-	-	15	0	-	-	15	0	-	-	15	0
701	-	-	102	1	-	-	103	0	-	-	102	1
705	-	-	79	0	-	-	79	0	-	-	79	0
706	-	-	6	0	-	-	6	0	-	-	6	0
707	-	-	37	0	-	-	37	0	-	-	37	0
801	-	-	80	0	-	-	80	0	-	-	80	0
802	-	-	75	0	-	-	75	0	-	-	75	0
811	-	-	112	0	-	-	112	0	-	-	112	0
812	-	-	106	0	-	-	106	0	-	-	106	0

Table 5.9: Number of correctly classified clear targets (z), incorrectly classified clear targets (f), correctly classified cloud targets(h) and incorrectly classified cloud targets (u) for all day targets over sea.

Day Targets Over Land

Code	Bayesian				MétéoFrance			
	z	f	h	u	z	f	h	u
151	1171	254	-	-	1389	36	-	-
152	6	13	-	-	17	2	-	-
181	1	4	-	-	5	0	-	-
191	252	67	-	-	19	300	-	-
502	-	-	283	35	-	-	261	57
503	-	-	268	20	-	-	252	36
504	-	-	0	1	-	-	0	1
602	-	-	114	54	-	-	112	56
607	-	-	15	0	-	-	15	0
608	-	-	160	0	-	-	160	0
609	-	-	79	0	-	-	79	0
702	-	-	193	40	-	-	188	45
704	-	-	3	2	-	-	4	1
705	-	-	69	0	-	-	69	0
706	-	-	6	0	-	-	6	0
707	-	-	48	0	-	-	48	0
801	-	-	39	0	-	-	39	0
802	-	-	241	0	-	-	241	0
811	-	-	173	0	-	-	173	0
812	-	-	47	0	-	-	47	0

Table 5.10: Number of correctly classified clear targets (z), incorrectly classified clear targets (f), correctly classified cloud targets(h) and incorrectly classified cloud targets (u) for all day targets over land.

5.3.4 Discussion

Night Results

The three methods show differing balances between missing cloud and false detection. The Bayesian approach has a relatively high *FAR* over land, but also a relatively high *HR*, so that the *TSS* is higher than for the operational approach. Over sea, the *FAR* of the Bayesian method falls between those of the operational methods, but the *HR* is higher than both, resulting in a higher *TSS*.

The Bayesian *TSS* is 9.5% higher than the operational method's *TSS* for land targets, and more than 8% higher than both operational approaches' *TSS* for sea targets. The *FAR* is higher for land targets in both the operational and Bayesian techniques, and is particularly large in the Bayesian case, at 36.1%. This could be due to the high uncertainty assigned to NWP ST for land targets. Increasing uncertainty in the NWP ST field lowers the probability of the observation given clear conditions, therefore making it less likely for a target to be classified as clear. Improvements in both NWP itself, and in quantification of NWP accuracy, both of which will be aided by improved cloud detection, will address this.

Clear targets over land are classified less successfully for snow surfaces by both techniques (code 191). Snow is generally problematic for cloud detection, being cold and white, as clouds are, and appearing and disappearing on the surface at variable time-scales which are difficult to predict. In the Bayesian case, the false alarms are probably at least partially due to the seasonally-averaged surface emissivity maps, which, for non-persistent snow-cover, is likely to mean an inaccurate surface emissivity is passed to the RTM. All cloud targets over land were detected more successfully by the Bayesian technique, particularly stratus (code 502) and stratocumulus (code 503), for which the hit rates are 32.2% and 45.3% higher, respectively. The hit rates for thin cirrus over stratus/stratocumulus (code 705) and altocumulus (code 802) are 23.2% and 23.5% higher for the Bayesian technique, making the overall skill of the detection over land encouragingly high, despite the high *FAR*.

Open sea targets containing ocean fronts are sometimes mistaken for cloud under the Bayesian approach, due to their relatively high 11μm LSD, and this could explain some of the miss-classification in this category (code 101). Overall, however, the technique performed well for the open sea category, with 93.5% of targets correctly classified. Over sea, the Bayesian technique detected 89.0% of stratus clouds (code 502), which compares favourably to the 66.9% and 62.5% hit rates of the operational methods. Similarly, the Bayesian hit rate for stratocumulus over sea (code 503) is 2.5% and 7.9% higher than the hit rates of the U.K. Met. Office and Météo-France respectively. The U.K. Met. Office's method hit rate for small cumulus over sea (code 601) is 6.5% higher than the Bayesian, and 13% higher than that of Météo-France. Small cumulus can be a difficult category for detection in infra-red imagery when the cloud fraction is sufficiently low and the cloud top temperature sufficiently warm that they are spectrally indistinguishable at thermal wavelengths, causing misses in the Bayesian detection. No mean-

ingful conclusion can be drawn from the difference between the hit rates for cumulus congestus and thin cirrus clouds over sea (codes 606 and 701 respectively), being just one target in each case.

For both land and sea, the results are very good when compared against operational results for the same set of 'difficult' targets.

The Bayesian method could be improved by using a dynamic, rather than a fixed, surface wind-speed value for ocean targets where the field is not provided in the NWP information. This is not currently done, as the technique is designed to be suitable for real-time applications and it is anticipated that reading in more data would slow it down without yielding significantly improved results. Similarly, consideration of more NWP fields in the **B**-matrix could improve results, but at the cost of processing time. The NWP-dependency of the technique means that as NWP fields continue to improve, in spatial resolution and in clear-sky temperature and humidity accuracy, the Bayesian cloud detection statistics will improve also.

Day Results

Over sea, the *HR* scores for all three techniques are similarly high, although the Bayesian approach and Météo-France are higher than the U.K. Met. Office, by 2.6% and 1.92% respectively. This difference in *HR*, accompanied with almost equal *FAR* scores, results in the Bayesian approach achieving a *TSS* 3.1% higher than the U.K. Met. Office. The *TSS* achieved by MétéoFrance, which is 0.9% higher than the Bayesian result, is due to the low *FAR* achieved by this operational method, 1.6% lower than the Bayesian *FAR*. Over land, the *FAR* scored by the Bayesian approach is similar to the operational score, which is just 0.3% higher The *HR* over land, however, is 2.3% higher. This combination of a slightly lower *FAR* and a higher *HR* result in a higher *TSS* for the Bayesian technique over land.

Both *FAR* scores for land are higher than for sea, reflecting the fact that detecting cloud over land is intrinsically more difficult. Land surfaces are generally more variable in reflectivity and emissivity than sea surfaces. The Bayesian method relies on these properties being estimated with known accuracy and the results indicate that the seasonal reference maps used for emissivity and for reflectivity calculations generally produce sufficiently accurate results to be suitable. The operational method successfully classified 11 more targets than the Bayesian approach in the 'land with shadow' target category (code 152), which may suggest a sub-optimal geometric term in the BRDF surface reflectivity model, see section 5.2.2. This is unavoidable so long as the reference data used to consider the relative importance of each term in the BRDF is read from a grid with cells large enough to contain a range of land surfaces. Reading in reference data with a higher spatial resolution would be computationally expensive (and potentially those data would be more error prone). Another likely explanation for 'land with shadow' targets being classed as cloud is that they are in fact cloud shadows on land, a class which is not currently modelled in the Bayesian technique. The high *HR* of 82.2% achieved for land using the current reference data,

indicates that in most cases a higher spatial resolution is not required.

Snow and ice present challenges to cloud detection (codes 191 and 181 respectively), as both surfaces often appear very similar to clouds. The Bayesian approach does not match the operational results over ice, although it is difficult to draw a meaningful conclusion from just 5 targets. With a hit rate of 79%, the Bayesian technique improves significantly on the 60.0% hit rate of the operational method over snow. This is an especially encouraging result as the reference emissivity and reflectivity data (which have been seasonally-averaged) were shown above to have large errors for non-permanent snow-cover, which was anticipated to result in poor performance for such targets.

The Bayesian detection of clouds over land is generally more successful than the operational technique, the *HR* being higher by 6.9%, 2.5%, 1.2%, and 2.1% for stratus (code 502), stratocumulus (code 503), small cumulus (code 602) and thin cirrus (code 702) respectively. The 82.8% *HR* for thin cirrus, which are often difficult to distinguish, is particularly encouraging. Although only three of five of thin cirrus targets over snow (code 704) are detected, the small number of targets in this category make any conclusion impossible, and overall the Bayesian detection of clouds over land improves on the operational detection.

All three techniques achieved a high proportion of correctly classified targets in the open sea category (code 101), Météo-France scoring 94.5%, the Bayesian approach 94.4% and the U.K. Met. Office 93.4%. While Météo-France were more successful in cases of sea with sunglint (code 106), correctly classifying 99.4% of targets, wile the Bayesian and U.K. Met. Office approaches achieved 91.8% and 92.9% respectively, all three approaches performed well for this target category. Stratus over sea (code 502) was detected most successfully by the Bayesian method, with a *HR* of 94.9%, compared to the *HR*s achieved by the U.K. Met. Office and Météo-France of 86.0% and 92.5% respectively. The *HR* of the Bayesian method for small cumulus clouds over sea (code 601), 53.6%, falls between the two operational results of 50.7% and 55.1% for the U.K. Met. Office and Météo-France respectively. For all other cloud categories over sea, the Bayesian results match those of the operational methods, demonstrating its suitability for daytime ocean applications as well as daytime land applications, where the results exceed the operational results.

5.4 Conclusions and Further Work

In the validation described in section 5.3, the Bayesian approach to cloud detection was demonstrated to perform for especially selected ‘difficult’ cloud targets, achieving a true skills score which exceeds current operational results for for both land and sea at night, and for land during daytime. For sea during day-time, the result of one operational technique is exceeded and the Bayesian result is within 1% of the other. Prior to this work, the algorithm to implement the Bayesian approach had only been validated for a night-time sea imagery from a single sensor - it

has now been shown to be suitable for day and night imagery over both land and sea, and has been restructured to be more easily adapted to more sensors - as this validation work, using imagery from the SEVIRI sensor demonstrates. The range of clear and cloudy in the database covers a large geographical area and a wide range of cloud types are represented.

The success of the Bayesian approach depends on accurate NWP data being available, and on the performance of a reliable RTM. As advances in NWP are made, which improved cloud detection will feed into, it is anticipated that the Bayesian approach should produce further improved results. In addition to being arguably more physically justified, through its exploitation of scene-specific information, this approach to cloud detection has been demonstrated to yield results with skills scores that either exceed or match to within 1% the skills scores of current operational techniques. It is therefore judged to be a useful and suitable tool for cloud detection applications over land and ocean for day and night-time imagery.

Future work to validate the algorithm for a dataset other than that which was used to create the visible and NIR PDFs for cloud would be useful, and could also serve to further develop the flexibility of the technique, making it more easily adaptable to imagery from other sensors. Better estimates for the uncertainties used in the technique could improve the results, for example sensor noise could be made observation-specific, depending on the radiances and reflectances reaching the sensor. Errors in the NWP data are likely to be spatially variable, especially over land, and so some location-dependence of the values used could be introduced. A directional adjustment to values read from the land surface emissivity LUT could benefit the performance of the algorithm, but it is not clear how this could be done practically. The FM-errors should probably have some dependence on atmospheric path length, which dependency could then be removed from **B**-matrix, and entered into the **R**-matrix. As the climate changes, and observational techniques improve, it may be worthwhile to re-calculate some of the reference data used by the algorithm, such as land surface emissivity and reflectance parameters and the prior probability of cloud, in order to use more appropriate data as they become available. In its current form, the algorithm is suitable for operational applications of cloud detection.

Chapter 6

Calculating a Local PDF for Cloud

The Bayesian cloud detection method described in chapters 2 and 5 is temporally- and spatially-specific because it uses local NWP information to predict clear sky radiance and reflectance reaching a satellite sensor, as described in equation 2.1 on page 12. To be fully physically robust the distribution of predicted radiances and reflectances reaching a sensor for a cloudy sky should also be locally- and spatially-specific. Currently the probability density function (PDF) used for cloud is a normalized distribution of cloud observations from a global dataset. The technique should be strengthened if an NWP-conditional PDF for cloud observations were used. An NWP-dependent PDF for cloud should represent pixel-specific predicted cloud observations, and so is expected to have a sharp peak and steep sides, in contrast to the global PDF used previously, which must necessarily contain a range of predicted cloud observations wide enough to cover those appropriate to any pixel, and is therefore relatively broad and flat. The increased steepness and less extensive range of the NWP-dependent PDF means it can be expected to lead to a more certain classification, i.e. pixel observations are more likely to be associated with a very high or very low conditional prior probability of cloud, making the final discrimination between cloud and clear less ambiguous. This is not implemented in the technique presented in previous chapters because of the difficulty in forward-modelling all the possible cloudy atmospheric states for a given imaged scene. Rather than forward-modelling one set of atmospheric conditions, as for the clear-sky case, the cloudy PDF must represent observations for a range of atmospheric states, with clouds at different altitudes, with different optical depths and filling different fractions of the pixel. To be useful operationally, the PDF must be generated quickly, making it impractical to cover the full range of possible cloudy conditions with individual RTM runs. This chapter describes a method of calculating a NWP-conditional spectral PDF for cloud at thermal infrared (TIR) wavelengths and presents some results to show the effect of using a local PDF for cloud instead of a global distribution.

6.1 NWP-Based Predictions for Observations of Cloud

The flow-chart in figure 6.1 outlines the method described in the following sections. The NWP-profile is forward-modelled with the RTM, RTTOVcld (Merchant et al., 2006b), with 60 altitude levels. Single phase clouds are added to 14 of the modelled altitudes in separate model-runs, with

cloud pixel-coverage varying from 10% to 100% in 30% increments, and cloud ice- or liquid-water path (cwp) varying from 0 to 0.1kgm^{-2} in 11 exponentially spaced increments. The cloud altitudes correspond to the altitude of the model level 2nd away from the Earth's surface (i.e. the closest level to the surface that does not actually touch it), up to the 40th model level from the surface, with cloud only being modelled at every second model altitude within this range. Treating ice- and liquid-phase clouds separately, the forward-modelled brightness temperatures (BTs) are plotted against cwp for each modelled altitude and cloud fraction and an exponential curve fitted, see figure 6.2. The equation for the curve has the form $BT = a + b \times \left(1 - \exp^{\frac{cwp}{c}}\right)$ where a, b and c are fitting parameters found by iteration for each forward-modelled cloud altitude-fraction combination. With the curve defined, BTs can be read for clouds with cwp values other than those modelled, so reducing the number of necessary RTM runs. The steepness and minimum BT value of the curve changes with altitude and with cloud pixel coverage. This is because a lower pixel coverage requires a thicker cloud for an optically saturated observation, and higher altitude clouds will saturate at lower BTs. The parameters are interpolated to give the fitting parameters for BT-cwp curves at altitudes and with pixel coverages other than those that are forward-modelled. In this way, BTs for clouds with tops at 10m intervals through the atmosphere, and with pixel coverages varying from 10-100% in 1% increments, and with cwp varying from 0-0.1kgm^{-2} in 99 exponentially spaced increments, are predicted without being explicitly modelled. Section 6.2 explians how these BTs are represented within the PDF with weights corresponding to the relative liklihood of each cloud being present.

It is notable that the relationship between cwp and BT for some clouds modelled at low altitudes appears to be the inverse of that for clouds modelled at higher altitudes. Clouds at lower altitudes can emit thermal radiation at radiometric temperatures greater than those at which the Earth's surface emits. This can occur because the surface emissivity of the Earth is less than that of the cloud, or because the temperature of the cloud is greater than that of the Earth's surface (temperature inversion). The presence of some cloud can therefore have the effect of supplementing thermal emission from the surface and increasing the recorded BTs. At sufficiently high altitudes, clouds always emit at lower radiometric temperatures than the Earth's surface, and generally lower BTs are recorded.

The curves describing the cwp-BT relationship are not assumed to be the same for ice- and liquid-phase clouds, and so the two cases are processed separately, creating two separate distributions of predicted BTs. Each distribution is convolved with a combination of sensor noise, forward-modelling uncertainties and an assumed Gaussian uncertainty in the NWP TCWV and ST fields (analogous to equation 2.3 on page 17 for the clear sky case) to create not-normalized phase-specific PDFs. The two PDFs are combined and normalized to create a NWP-conditional PDF for cloud.

The remainder of this chapter discusses the many factors which need to be considered in creating a NWP-dependent PDF for cloud in sections 6.2 to 6.4. A case study is used to demonstrate

the method in section 6.5. The technique is applied to a larger dataset in section 6.6 and in section 6.7 the results are used to discuss and assess the effect of the local, NWP-dependent PDF on the overall cloud detection skill. Some conclusions on the method and on its effect on the performance of the Bayesian cloud detection technique are given in section 6.8.

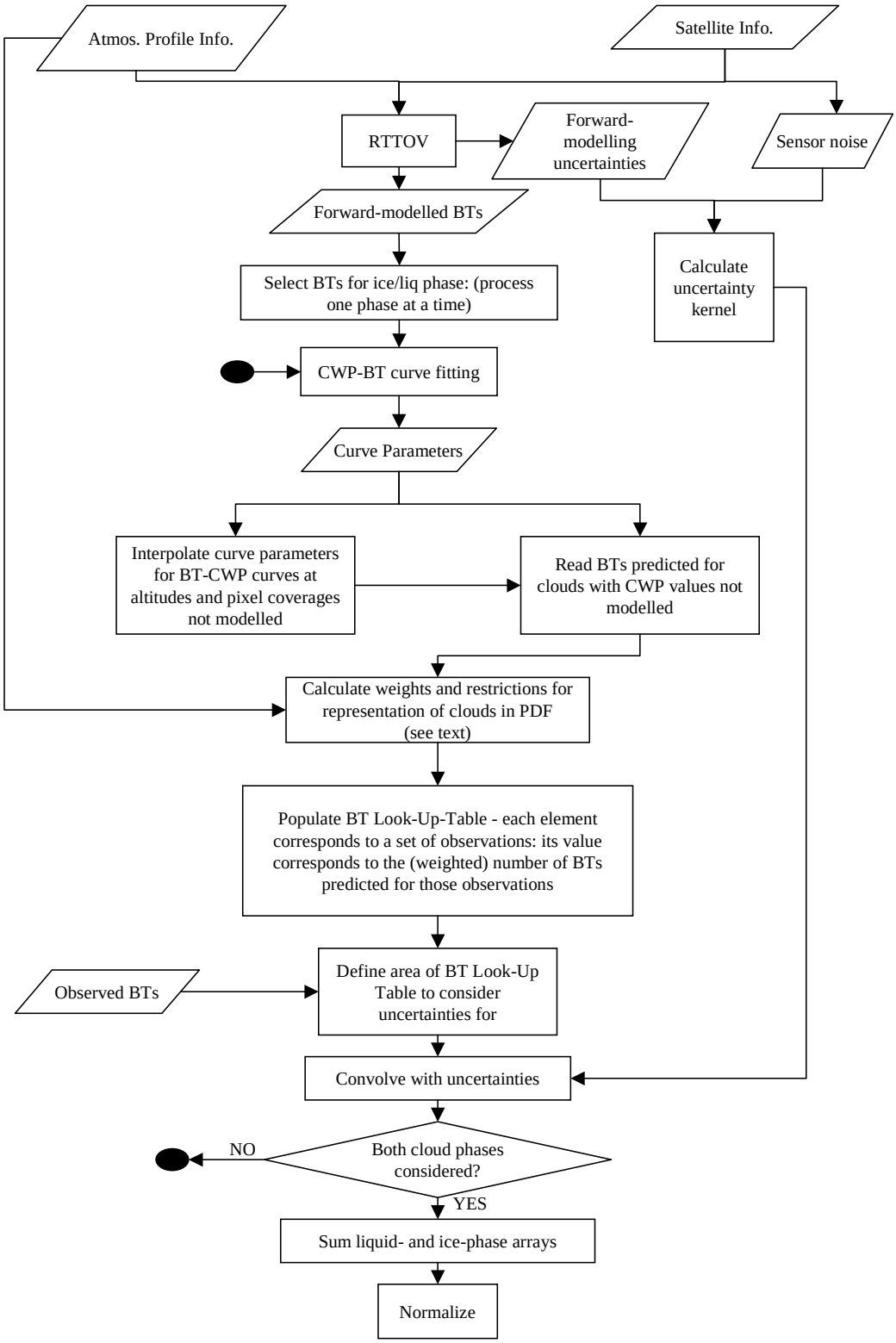
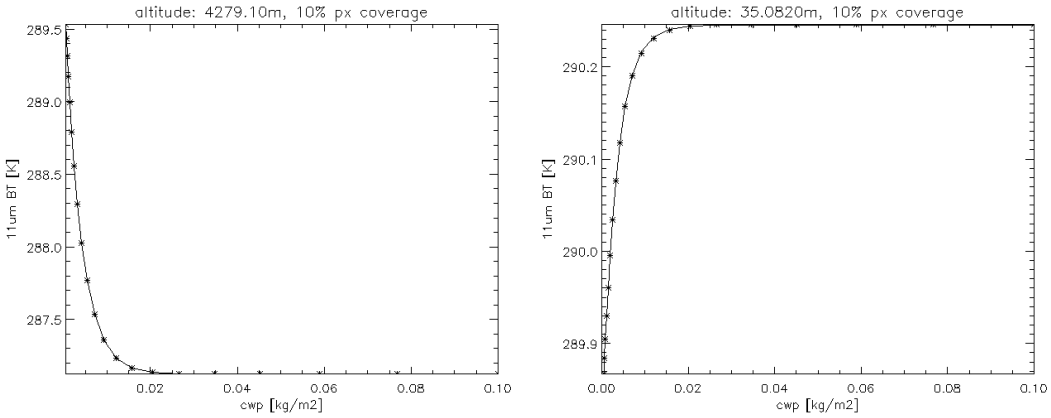


Figure 6.1: Flowchart outlining the steps taken to generate a NWP-conditional PDF for cloudy observations.



(a) (b)

Figure 6.2: Forward-modelled BTs plotted against cwp for 2 liquid phase clouds, with curves fitted. The asterix represent the modelled BTs, and the lines show the fitted curves: (a) high altitude clouds - in fact BTs forward-modelled for clouds at all altitudes except the very lowest (35m above the Earth's surface) show this exponentially decaying relationship to cwp, and (b) BT-cwp relationship from forward-modelling for a low-altitude cloud, which has a higher radiometric temperature than the cloud.

6.2 Restricting which Clouds are Represented in the PDF

The simulated clouds whose BTs contribute to the PDF must be realistic, given the NWP-profile. For example, it would be physically unreasonable for an ice phase cloud to occur close to the Earth's surface in a region where the Earth's surface temperature is very warm, or for a liquid phase cloud to occur at an altitude where the ambient air temperature is below 230K. This section describes the conditions that are placed on modelled clouds for representation within the PDF to address this.

6.2.1 Temperature

It is judged appropriate for liquid phase clouds to be represented in the PDF only at altitudes where the ambient air temperature makes their presence likely, and similarly for ice phase clouds. Cloud observations are included in the PDF for liquid- and ice-phase clouds at 10m intervals through the atmosphere. The NWP temperature field is linearly interpolated between profile altitudes to calculate the ambient air temperature for these cloud top positions, see figure 6.3. Predicted observations for ice phase clouds are only included in the PDF if the interpolated temperature is less than 273.15K, and for liquid phase clouds if the temperature is greater than 233.15K.

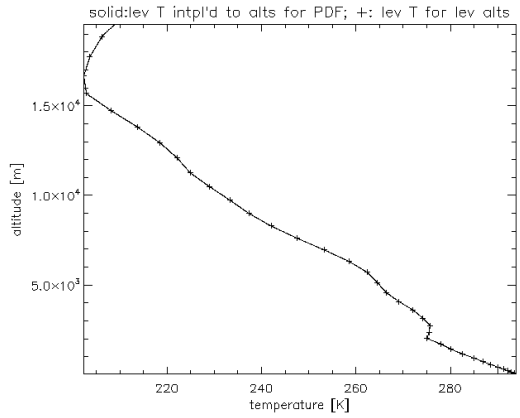


Figure 6.3: The line shows the curve fitted to the temperature at the 60 RTM altitudes (+), from which temperatures at 10m intervals through the atmosphere are read.

6.2.2 Altitude-specific Maximum Cloud Water Path

This subsection describes how cloud water path (cwp) used to restrict which clouds are represented within the PDF to clouds considered to be realistic at each altitude.

The dataset used in section 5.2.1 contains 6813 NWP profiles over ocean, each modelled at 60 altitudes. The maximum cwp for each model altitude in the dataset was set as the altitude-specific maximum cwp for clouds in the PDF. In other words, clouds with a cwp greater than the maximum cwp seen in the dataset (for the model altitude corresponding most closely to the cloud top altitude) are not represented in the PDF, see figure 6.4. In the PDF, all clouds are represented as single-layer clouds, i.e. with all water- or liquid content compressed into one model layer. The cwp taken from each profile in the dataset for each altitude is therefore the total integrated cwp between the surface and that altitude in the profile. The maximum of these (from all the profiles) for each altitude is set as the maximum cwp for clouds represented in the PDF.

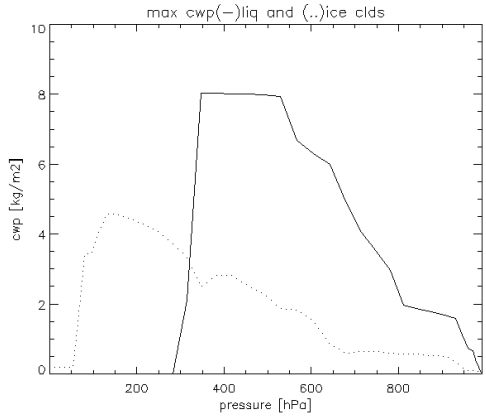


Figure 6.4: Altitude- and phase- specific maximum cwp for clouds represented in the PDF for cloud. Solid line: liquid phase clouds; dotted line: ice phase clouds. Pressure is used on the x-axishere as a proxy for altitude.

6.2.3 Bias to Retrieved Surface Temperature

Radiances recorded for pixels that are partially covered by cloud, and/or contain thin cloud may be so little affected by the cloud that retrievals of surface parameters are not significantly affected. The initial aim of the Bayesian cloud detection algorithm was to aid retrievals of sea surface temperature (SST), and so it was considered appropriate to limit the clouds represented in the PDF to those which effect a bias in the retrieved SST.

A pixel can contain a small fraction of thin cloud without the Earth-surface parameters it records at thermal wavelengths being significantly affected. The Bayesian cloud detection was initially designed as a tool to aid retrieval of sea surface temperature from satellite imagery, and so it was judged appropriate to place a condition based on ST on clouds represented in the PDF. A surface temperature is calculated from the BTs predicted for each cloud, and required to be at least 0.2K different to that calculated from the BTs predicted for clear sky conditions in order for that cloud to contribute to the PDF.

6.3 Weighting Clouds' Representation in the PDF

The PDF is an estimate of the probability of a set of observations, given that there is cloud present, and given the prior NWP profile information. The contribution of each represented cloud to the PDF should therefore be weighted by its likelihood of being present, relative to the likelihood of other clouds being present.

6.3.1 Optically Saturated Clouds

Optically saturated clouds may correspond to a range of cwp values, corresponding to 'optically saturated' and 'optically super-saturated' clouds, but will have the same effect on recorded radiances. It is therefore unnecessary to simulate observations for more than one optically saturated cloud at each cloud altitude represented in the PDF (as optically saturated clouds at the same altitude for the same pixel will correspond to the same predicted observation). It is, however, important to weight the representation of the saturated prediction in the PDF to account for the fact that it may represent many clouds. The following paragraph describes how this saturation weight is defined.

Cloud observations are forward-modelled for construction of the PDF with cwp set to the same maximum value for each forward-modelled cloud altitude. This value is beyond the optical saturation point for clouds at every forward-modelled cloud altitude. A weight is found for the representation of clouds with this maximum cwp value using the same dataset as used in section 6.2.2. Using the atmospheric profiles in the dataset, the ratio of the number of clouds with cwp $\geq$ this maximum forward-modelled value to the number of clouds with cwp $<$ this maximum forward-modelled value is calculated for each model altitude. The representation of cloud observations within the PDF should match this ratio, i.e. if all predicted cloud observations are represented in the PDF with a value of 1, and there are n predicted cloud observations, then the observation corresponding to the cloud with the maximum forward-modelled cwp should be represented with a weight of n times the ratio found from the dataset for that cloud altitude. This means that the proportion of cloud observations with cwp $\geq$ the maximum forward-modelled cwp represented in the PDF matches the equivalent proportion of clouds seen in the dataset. The ratios for liquid- and for ice-phase clouds are found separately, and are shown in figure 6.5.

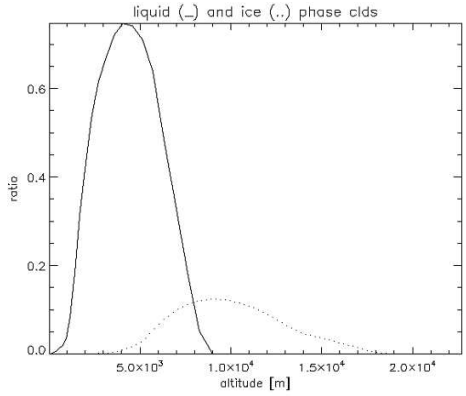


Figure 6.5: Ratio of the number of clouds with cwp $\geq$ the maximum forward-modelled cwp, to the number of clouds with cwp $<$ the maximum forward-modelled cwp for each profile altitude in the dataset. The text explains how this ratio is used to set a weight on the representation of cloud observations corresponding to the highest simulated cloud cwp value in the PDF.

6.3.2 Cloud-Filled Pixels

The model accounts for partially- and fully-filled pixels, but it is appropriate to consider the relative frequency of these states in reality, and to account for how this changes with pixel size. An empirical dataset of ATSR2 nadir-acquired imagery, consisting of measurements from January, April, July and October, was used to calculate a latitude- and season-specific ratio of cloud-filled to cloud-edge pixels, as follows. The ATSR2 cloud mask was used to define a pixel as cloud-edge if it was classed as cloud but had at least 1 neighbour classed as clear, or if it was clear but had at least one neighbour classed as cloud. The 8 immediate neighbours of a pixel were used for this, and cloud pixels for which all 8 neighbours were also cloud, were defined as cloud-filled . This ratio is used to weight clouds with 100% pixel coverage more heavily in the PDF. The accuracy of the ATSR2 cloud mask is lower at very high latitudes, and so the ratio calculated for 75^o latitude is used for profiles corresponding to locations at latitudes greater than 75^o (similarly, for latitudes less than -75^o, the ratio for -75^o is used). Figure 6.6 shows the calcluated ratios of cloud filled to cloud edge pixels. The ratio found from the ATSR2 data is scaled to be appropriate to the nadir-pixel size for the sensor being used, and is then further scaled for each pixel using atmospheric path length. This makes the weight appropriate for nadir and non-nadir imaged pixels.

This method of calculating a weight for cloud-filled pixels relies on the accuracy of the ATSR2 cloud mask, which has been noted to over-detect clouds in some instances, see the example in Merchant et al. (2005). It is also subjective, as a pixel is judged as 'cloud edge' if just one neighbouring pixel is clear, which may lead to an over-estimation of the ratio of cloud-filled to cloud edge pixels. There is, however, a good argument for using these subjective criteria: clouds do not generally have edges that line up with pixel boundaries. Despite these shortcomings, it is considered the most practical available method of calculating an appropriate weight.

6.3.3 NWP Cloud Fraction

The NWP data from which the predictions of cloud observations are forward-modelled contains a cloud fraction field at 60 model level altitudes. In theory, this should indicate the overall likelihood of cloud being present at that altitude in the area covered by the NWP grid cell. This was initially considered for use as a weight for the likelihood of any cloud at the altitudes represented in the PDF (using the closest profile altitude). Several studies indicate high uncertainty in the cloud fraction predicted by NWP models, see for example Jakob (1999); Hogen et al. (2001); Dybbroe et al. (2005), who all found problems with ECMWF's cloud fraction field. It was therefore decided not to use the profile cloud fraction field for the generation of the cloudy PDF. Instead, a method of calculating a 'maximum' cloud fraction from the NWP humidity data is implemented and used to calculate an alternative weight. This is described in section 6.3.4.

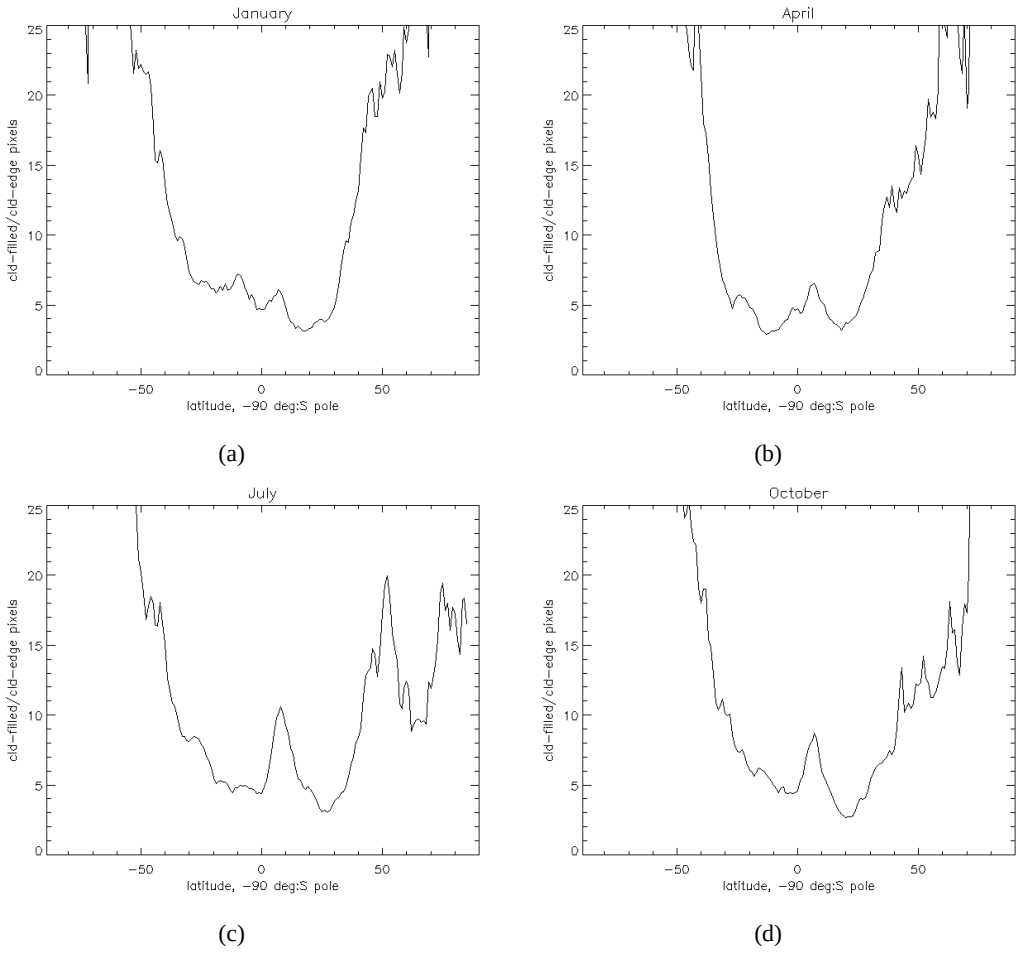


Figure 6.6: Ratio of cloud-filled to cloud-edge pixels from imagery acquired by the ATSR-2 sensor in 1997, in: (a) January, (b) April, (c) July and (d) October.

6.3.4 Relative Humidity

Clouds are only expected to be present in saturated, or near-saturated if sea salt is present, air masses. The amount of water vapour at a given altitude in a NWP grid cell can therefore be used to calculate a measure of the likelihood of cloud being present at each NWP model altitude.

The NWP water vapour field is interpolated to the altitudes at which clouds are represented in the PDF. For imagery from most sensors, a NWP grid cell is larger than a pixel and so a maximum relative humidity is calculated for each altitude by assuming all the water vapour at that altitude to be concentrated in an area the size of the pixel. In the presence of sea salt (rarely observed in significant quantities more than 2km above the surface, see for example Woodcock (1953); Junge et al. (1969)), clouds are able to form in air with at least 70% relative humidity (Moran and Morgan, 1997). Below 2km, the maximum NWP grid-cell fraction that could be at a relative humidity of 70% (the rest of the grid cell containing no water vapour), given the water vapour field, is considered the maximum cloud fraction for the grid cell. This maximum grid cell fraction is used directly as an altitude-specific weight for the representation of clouds at altitudes

up to 2km above the surface[1]. Above 2km, the relative humidity used to calculate the maximum grid cell fraction is 100%. There are cases where, below 2km, the NWP water vapour field can contain enough to cover more than the area of the grid cell at 70% relative humidity, leading to a maximum grid cell cloud fraction greater than 1. At these altitudes, clouds are even more likely than at altitudes where the amount of water vapour allows a maximum grid-cell cloud fraction of 1, so it is not appropriate to set the weight to 1 in both places. Clouds at these altitudes are assigned a weight greater than 1 - an example is shown in figure 6.7.

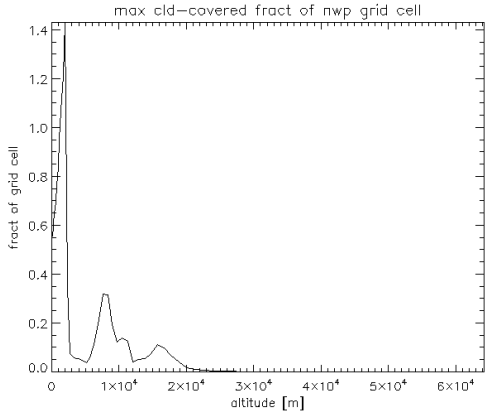


Figure 6.7: The relative humidity weights calculated for an ECMWF NWP profile located off the SW coast of Korea. NB. All 60 NWP model altitudes are shown here, but clouds are only represented in the PDF at altitudes up to the 40th model altitude from the surface.

[1]The altitude to which sea salt appears in significant quantity in the atmosphere depends on the speed of winds at the surface, which break the crests of waves, releasing salt into the lower atmosphere (O'Dowd et al., 1996). Using the NWP surface wind-speed, when provided, to decide between using a relative humidity of 70% or 100% was therefore considered. The NWP data, however, represent a 'snap-shot' in time and do not contain information on the history of the surface winds - a strong surface wind can cause the release of a lot of sea salt, which is dispersed in the atmosphere and does not disappear as soon as the wind weakens. The surface wind-fields were therefore not used here and the decision was made dependent only on altitude.

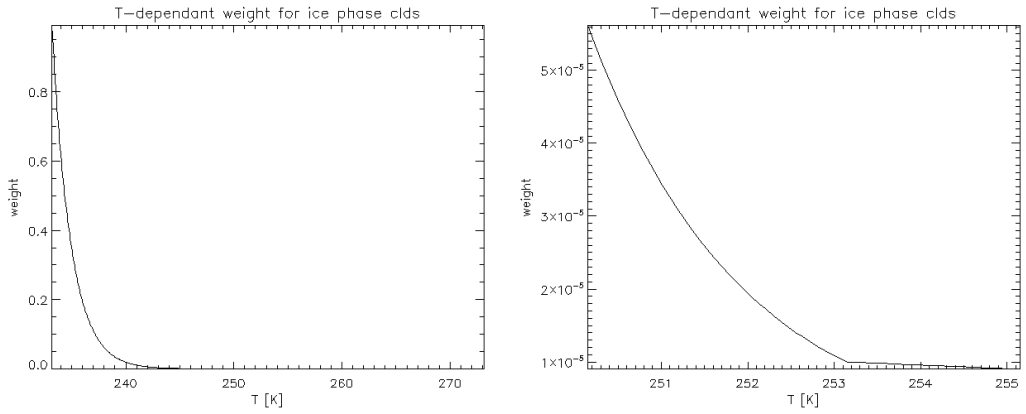
6.3.5 Ice- and Liquid- Phase Clouds

Ice clouds are not represented in the PDF for altitudes at which the ambient temperature exceeds 273.15K. Their representation at ambient temperatures below this is weighted, the argument supporting the weighting scheme being as follows:

For ice clouds to form at ambient temperatures above 233.15K (below which homogeneous freezing of water droplets occurs), the presence of ice nuclei is required. These are usually less abundant than condensation nuclei. Freezing nuclei generally become active when the ambient air temperature drops below approximately 264.15K , and deposition nuclei become active at around 253.15K (Rogers and Yau, 1989; Moran and Morgan, 1997). In the presence of ice nuclei, ice clouds are therefore likely to be present at ambient air temperatures of around 264K, but become more likely at around 253K, when both freezing and deposition nuclei are available. The exact temperature depends on the relative abundance of freezing and deposition nuclei. The presence of ice nuclei is highly spatially- and temporally variable, and where ice nuclei are available, the readiness with which heterogeneous nucleation (either freezing or deposition) takes place depends on the properties of the foreign particles - the concentration of nuclei has been observed to be up to 104 times less than the measured concentration of ice crystals (Rogers and Yau, 1989). The problem is therefore not easy to parameterize, see for example Knight (1979); Mossop (1985). Fletcher (1962), summarizing the findings of studies at the time, found typical concentrations of ice nuclei to increase by a factor of 10 for each 4K of cooling below 253.15K and this exponential relationship is still generally accepted, although deviations of up to a factor of 10 have been observed (Rogers and Yau, 1989; Moran and Morgan, 1997).

A relationship between the presence of ice nuclei and temperature is used to weight the representation of ice clouds, relative to that of liquid phase clouds at the same ambient air temperature. Although rare, observations of ice clouds at temperatures above 268.15K have been recorded. As a synthesis of this information, a linear weighting was assigned to ice clouds from 273.15K to 253.15K. To clouds at ambient temperatures below 253.15K, an exponential weighting was applied following Fletcher (1962), reaching a weighting of 1 at 233.15K, see figure 6.8.

Super-cooled liquid phase clouds can exist in ambient air temperatures down to 233.15K , but are more likely at warmer temperatures (Rogers and Yau, 1989). The relative likelihood of super-cooled liquid clouds depends on the number of ice nuclei present, and so is highly spatially and temporally variable. In the PDF, liquid phase clouds are assigned a weight of 1 at ambient temperatures of 273.15K and above, and a weight of 0 at 233.15K and below, with a weighting for temperatures between these that it 1 - the weighting for an ice cloud at the same temperature, as described above.



(a) (b)

Figure 6.8: Temperature-dependent weighting assigned to the representation of ice clouds in the PDF: (a) curve showing all the weights between 1 and 0, (b) the intersection between the 2 different regimes for the temperature-weight relationship.

6.3.6 Atmospheric Stability

The most common process through which air becomes saturated and clouds form is the expansional cooling of air rising through the atmosphere (Moran and Morgan, 1997), and so a weight based on atmospheric stability was considered, but was rejected, following the reasoning outlined below.

In a stable atmosphere vertical air motion is suppressed as rising air becomes denser (cooler) than its surroundings and falls back to its original position, and so cloud formation by expansional cooling is suppressed. In an unstable atmosphere, rising air can remain less dense than its surroundings, and so vertical air motion (and expansional cooling) is enhanced, making cloud formation more likely.

Suppression of cloud formation through suppression of vertical air motion does not affect clouds that have already formed - a saturated air parcel can persist without rising or falling in the atmosphere. It should also be noted that air parcels descending in an unstable atmosphere undergo compressional warming, decreasing their relative humidity and so tending to dissipate any cloud that has already formed. It therefore does not follow simply that the presence of clouds is more or less likely in a stable or unstable atmosphere. Cloud formation is not equivalent to cloud persistence, and without knowledge of the vertical air movement (upwards or downwards) and of the history of the column of atmosphere, cloud persistence cannot be reliably estimated. No weight is therefore assigned to clouds in the PDF on the basis of atmospheric stability.

6.4 Uncertainties

The clear-sky PDF is a Gaussian distribution centred on a simulated point in BT space (the result of forward-modelling for a clear sky). Within the distribution, each element corresponds to the probability of those BT values having been recorded in clear sky conditions. Section 2.4.2, in chapter 2, explains that the Gaussian distribution is based on the error defined by the uncertainty in the NWP ST and TCWV, plus noise in the sensor channels (assumed to be independent).

Whereas in the clear sky case there is only one central point in BT-space, in the cloudy sky case there are many, corresponding to the range of BTs predicted for a range of cloudy atmospheric states, each of which should in principle also have a distribution estimated for it to account for NWP uncertainties and noise. Proper convolution of these distributions would be prohibitively intensive computationally. The observation predicted for each cloud represented in the cloud PDF contributes initially to a single point in BT-space. The value of the contribution depends on the weights described in section 6.3, as applied to that particular cloud. An array in observation space, corresponding to the observations predicted for all represented clouds, is then created before uncertainties are accounted for. A normalized kernel of weights is then calculated and moved as a 'sliding window' over the array to create a new array, in which each element of the original array is represented as a Gaussian distribution. The weighting function used for this kernel is described in the next paragraph. To save processing time, the expansion is calculated only for a finite area of the array, and each element in BT-space is expanded over a distance of four standard deviations in every direction, where standard deviation is calculated as $\sqrt{\left(\frac{\delta BT}{\delta ST}\epsilon^b_{st}\right)^2 + \left(\frac{\delta BT}{\delta TCWV}\epsilon^b_{tcwv}\right)^2}$, with $\epsilon^b_{st,tcwv}$ referring to the error in the NWP ST and TCWV fields, as previously. For computational efficiency, this distance in BT-space was restricted to +/-8K.

The sensitivity of the forward modelling to changes in ST and TCWV is combined with the errors in ST in TCWV, and added to the sensor noise, **R**, to create a weighting function, **WF**, see equations 6.1 and 6.2, where **H'** and **B** are the values used for the clear sky calculation discussed in section 2.4.2 in chapter 2. This weighting function is used to construct a kernel, which is convolved with the array in BT-space using the 'sliding window' mentioned above, to expand the value of each element into a distribution over an area of 4 standard deviations in each direction. The proportion of each original value assigned to elements in the distribution is a function of distance from the distribution centre (the BT-coordinates of the original value being expanded), dBT. Forward-modelling sensitivities to ST and TCWV are unlikely to be the same for the range of cloudy atmospheric states which are simulated to create each cloud PDF, however the method requires that the weighting function be applied *after* the BT-space has been populated with predicted cloud observations, meaning that individual clouds cannot be associated with separate weighting functions. For computational efficiency, therefore, the **H'**-matrix calculated for the simulation of clear sky radiances, and used to calculate the PDF for clear sky, is also used in the cloudy case.

$$\mathbf{HBH} = \mathbf{H'^{T}BH'} + \mathbf{R} \tag{6.1}$$

$$\mathbf{WF}\left(dBT\right) = \frac{dBT \times \mathbf{HBH}^{-1} \times dBT}{2\pi^{\frac{3}{2}}\sqrt{|\mathbf{HBH}|}} \tag{6.2}$$

For the cloudy PDF, the noise term, **R**, in equation 6.1, includes an estimation of FM error in addition to sensor noise: $\mathbf{R} = \sqrt{(FM\ error)^2 + (sensor\ noise)^2}$. The values for sensor noise are the same as for the clear PDF, but the FM error includes uncertainties stemming specifically from modelling cloud parameters, and the following section contains a discussion of these.

6.4.1 Estimating Forward Modelling Uncertainty

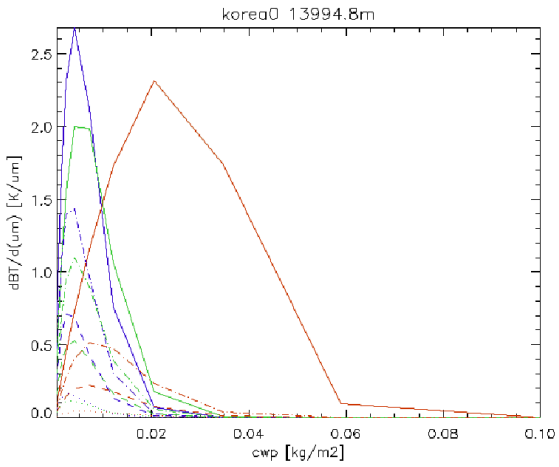
Several assumptions are made in RTTOV's calculations of radiances for cloudy states, and these can be considered major sources of FM uncertainty. The following section describes how appropriate values were assigned to FM uncertainties.

It should be noted that, in order for the cloudy PDF to be generated quickly, FM uncertainties are not considered until the BT-space has been populated with entries for the clouds it represents. More than one cloud may be represented in the same element of BT-space, and so the uncertainty-based expansion cannot be cloud-specific.

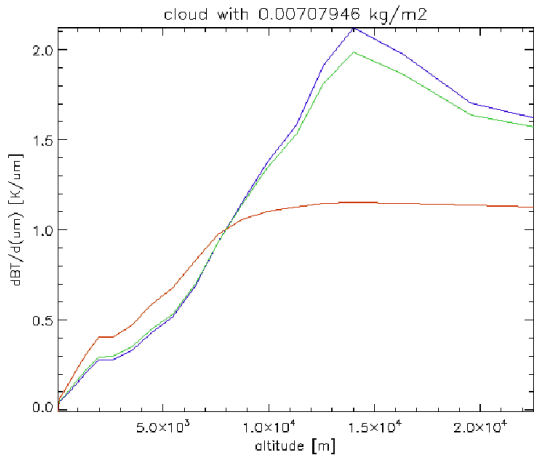
Assumption of Optically Thin Clouds and Negligence of Scattering. RTTOV does not allow the attenuation of cloud radiance within a model layer, in other words, clouds are assumed to be optically thin. This can lead to an error of up to 0.5K for radiances predicted in the presence of clouds (Merchant et al., 2006b). Scattering effects are neglected in RTTOV in order to allow radiances to be calculated from a linear combination of black-body cloud radiance and clear-sky radiance. The error resulting from this has been found (through comparison with another RTM which does consider scattering effects) to be 1K for 11m and 12m, and 2-5K for 3.7m for water clouds; and 1-2K for 3.7m and 12m and 1K for 11m for ice clouds (Merchant et al., 2006b).

Fixed Effective Radius for Water Clouds. A constant effective radius, r_e, is used in RTTOV to determine the absorption coefficients for forward-modelling BTs for water clouds. Empirical studies have observed a range of values for r_e, even at the same geographical location, see for example Han et al. (1994); Martin et al. (1994); Minnis et al. (1998), and so this assumption is considered a source of FM uncertainty. Different models assign different values to r_e, for example in the 1990s the ECMWF Operational Model used 15μm (Morcrette, 1990), while the U.K. Met. Office Unified Model used 7μm (later changed to 10μm) (Martin et al., 1994) and the ISCCP cloud optical depth retrieval algorithm assumed 10μm (Rossow and Schiffer, 1991), while Bower and Choularton (1992) found 14μm to be the optimum parameterisation, all for water clouds. Model parameterisations of r_e can strongly influence cloud optical properties - a fixed constant value can result in an under-estimation of both outgoing long-wave- and down-welling short-wave radiation (Somerville et al., 1999). In RTTOV, the r_e for water clouds is set to 10μm over land, and 13μm over ocean. To test FM sensitivity to this assumption, BTs were forward-modelled for clouds added to an atmospheric profile over ocean, with the land-flag set to 'ocean', and then set to 'land'. The emissivity was fixed at the value for ocean in both cases, meaning that the only difference between the 2 sets of forward-modelling was r_e for water clouds. It is accepted that this method of sensitivity-testing assumes that FM sensitivity to r_e does not change for r_e outside the 10-13μm interval tested here. This assumption is not necessarily valid, but it was considered the only practical way to investigate the uncertainty, as RTTOV has only two r_e settings.

The sensitivity of the forward model to the fixed effective radius assumption, in Kelvin per μm, is seen to be dependent on the liquid water content of the cloud, a relationship that is acknowledged in more sophisticated radiative transfer models, as noted for example by McFarlane et al. (1992); Daum and Liu (2005). It is also dependent on the temperature at which the cloud emits, and so on the altitude of the cloud, see figure 6.9. As noted at the start of this section, however, FM-uncertainties must be held constant for all clouds.



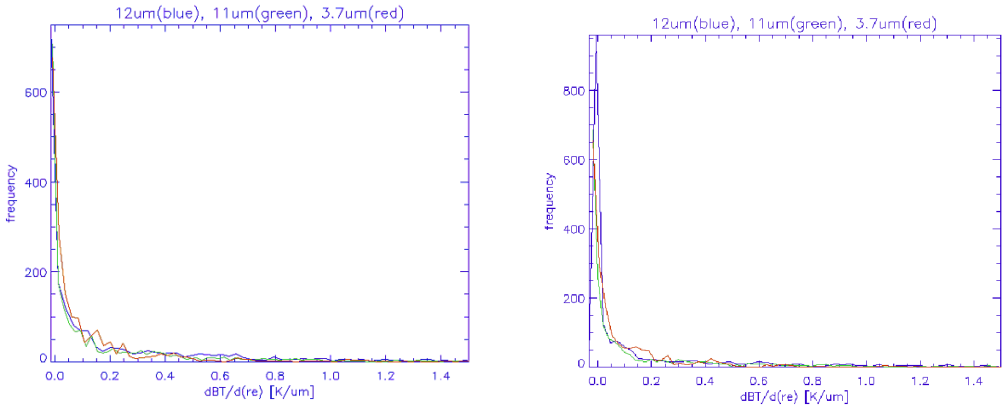
(a)



(b)

Figure 6.9: Plots showing (a) the cwp-dependence (for a fixed cloud altitude) and (b) the altitude-dependence (for fixed cwp) of the FM sensitivity to $\mathbf{r_e}$. Sensitivities for 12μm, 10.8μm and 3.9μm are shown in blue, green and red respectively. The sensitivity for clouds filling 100%, 70%, 40% and 10% of the pixel are shown as solid, dash-dot, dash and dot respectively.

The sensitivity test was repeated for the same clouds, added to 4 different atmospheric profiles - 2 corresponding to nadir-viewed imagery and 2 to to forward-viewed imagery from the same sensor AATSR - and the same peak location found (within 0.01K). The nadir sensitivity distributions were very similar to each other, as were the forward distributions, which were narrower. The sensitivity was therefore calculated for nadir-view- and forward-view- forward-modelling by combining the results from the 2 corresponding profiles, see figure 6.10.



(a) (b)

Figure 6.10: Histograms showing the sensitivities of FM BTs to $\mathbf{r_e}$ for (a) nadir viewing geometry, and (b) for forward-viewing viewing geometry.

Using only the difference between the results for the two fixed r_e values means the distribution can be expected to represent one side of the 'true' distribution of sensitivities. These FM runs represent the distribution of clouds forward-modelled for the PDF, but this is not the same as the distribution of clouds which are represented in the final PDF (the FM uncertainties are carried on to all represented clouds - those not explicitly forward-modelled are represented through interpolation of FM results, as described in section 6.1). Thick clouds, which figure 6.9 shows to correspond to low FM sensitivity to r_e, and which account for the low-FM-sensitivity peaks in figure 6.10, are weighted so as to be represented more strongly in the PDF. Sensitivity distributions corresponding to the distribution of clouds represented in the PDF, rather than those that are explicitly forward-modeled, would therefore have higher peaks than those in figure 6.10, making the long tails in in those figures less significant.

Assuming figure 6.10 shows one half of a Gaussian distribution, the standard deviation (SD) was found through scaling the 34th percentile of a normalized Gaussian with unit SD, centred on 0, to the 34th percentile of the found distribution (the distance from the peak in which 34% of data points are contained). One SD of a Gaussian contains 68% of the data, and so this method finds the robust SD from one half of a Gaussian distribution. Data lying more than 4 SDs from the peak were removed and the process repeated until less than 5% of the data lay further than 4 SDs from the peak. This was then taken as the SD. For the nadir and forward-view sensitivities, over 85% and 90% of the data were within 2 SDs for all channels respectively. In both cases, the absolute value of the peak plus 2SD was taken as the uncertainty, which the method requires to be symmetric.

To calculate the uncertainty from the r_e parameter in RTTOV, the sensitivity must be combined with a consideration of the range of differences, ϵ_{r_e}, between RTTOV's 13μm r_e and r_e values measured for real clouds. A U.K. Met. Office image of r_e retrieved for water

clouds over the ocean around the U.K. (image from 06 March 2007 at 1000UTC, available at: *http://www.metoffice.gov.uk/research/nwp/satellite/imagery/cloudtop_files/EREA11_200703061000.png*) shows a range of 5-25μm. The more extreme values in this plot appear at phase boundaries between water and ice clouds, and so are probably artifacts of the retrieval, disregarding these leaves a range of approximately 5-18μm. A near-global study found r_e for cloud droplets over ocean to fall between 10-15μm, although some latitude-dependence was observed (Han et al., 1994). This agrees with the earlier-mentioned range of values found in models with a fixed r_e parameter. In accordance with this, ϵ_{r_e} was set to 3μm for the validation. This gives the contribution from the r_e parameter to the FM uncertainty, shown in table 6.1.

This estimate of the uncertainty follows from simulations for the AATSR sensors, but the values for sensors with the same channel windows are expected to be similar. It is therefore considered appropriate to use these values in calculating uncertainties for producing a PDF for cloud detection in SEVIRI imagery, which has channel windows at 12μm, 10.8μm and 3.9μm, although it is accepted that this assumption may be less valid for the 3.9μm channel. The uncertainty appropriate to each target is interpolated linearly, using the target atmospheric path length, from these values, approximating the satellite zenith angle for forward-view AATSR to 60^o.

Ice Particle Size Distribution and Crystal Shape. Several parameterisations are available for the definition of the particle size distribution (PSD) for ice clouds in RTTOV - the default is the temperature-dependent Ou-Liou parameterisation (Ou and Liou, 1995), which calculates the 'generalised effective diameter' described in Fu (1996), using only temperature. This was used in Merchant et al. (2006b) and is also used here. There is also a choice of particle shape - aggregates or hexagonal columns. The crystal geometry used in this work is aggregates, as stated in Baran (2005) to be most often observed [2]. The absorption coefficients for ice clouds are calculated for the specified particle shape, using the generalised effective diameter, which is calculated from the chosen PSD (Saunders et al., 2005).

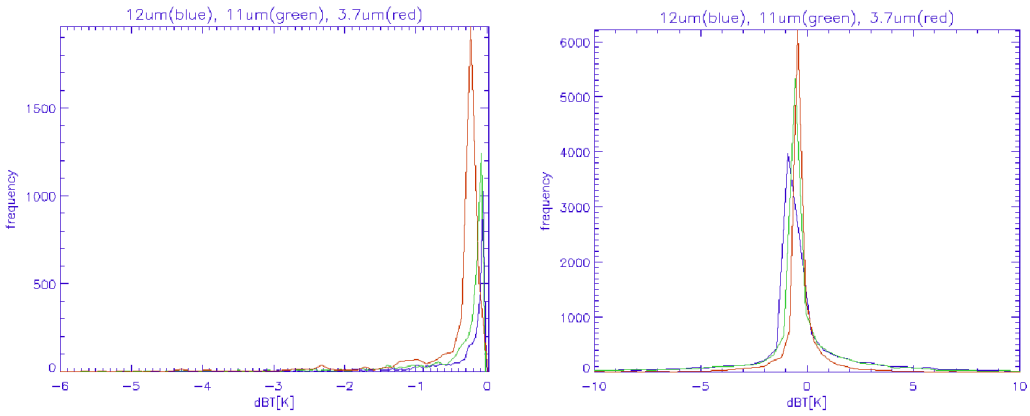
Investigations into FM sensitivity to PSD, and to ice crystal shape, have found overall differences of up to 12K, see for example Baran (2005), although the highest sensitivities in that study corresponded to BTs at 25μm and the uncertainties seen for lower wavelengths were smaller. The PSDs that were compared in Baran (2005) do not include the Ou Liou parameterisation, but the large uncertainties found indicate that the sensitivity of modelled cloud radiances to this parameter are high. The crystal shapes that were compared are ice aggregate and hexagonal column, and the maximum difference between BTs calculated using these shapes was 5K, which occurred at a wavelength of 25μm. Smaller differences were found for wavelengths of 20μm and 8.2μm.

Radiances were forward-modelled for the same ice clouds as those forward-modelled for the PDF, for the same sensor and profile information as for the investigation into the effective radius parameter for water clouds. The PSD was set to the Ou Liou distribution and the ice crystal shape set first to hexagonal columns, and then to aggregates. The difference in the resulting radiances gives an indication of the uncertainty introduced by the choice of particle shape, although the criticisms made for the fixed effective radius study also apply here. Distributions for nadir-view and forward-view forward modelling were very similar and so data from all 4 tests were considered together. The same procedure was followed to discard the tail and find the robust SD, and therefore the uncertainty (using 3 SDs, which contained over 79% data for all 3 channels) as previously. This gives +/- 0.127K, +/- 0.215K, +/- 0.653K for 12μm, 10.8μm and 3.7μm respectively.

To investigate the FM sensitivity to the choice of the Ou Liou PSD, the ice crystal shape was set to aggregates and BTs were forward-modelled using RTTOV with the same input data as previously, with each of the available PSD choices: Ou Liou PSD, the Wyser PSD (Wyser, 1998), the Boudala PSD (Boudala et al., 2002) and the McFarquar PSD (McFarquhar et al., 2003).

Distributions for nadir-view and forward-view forward modelling were again very similar and so data from all 4 tests were considered together, shown in figure 6.11. The robust SD of the distributions was found as before, taking the 34th percentile from the left hand side of

[2]The study in Baran (2005) considered only cirrus clouds, while this project considers all clouds, however there have been no studies, to the author's knowledge, of the relative prevalence of the 2 crystal shapes in real clouds, or of the relative cost to RTTOV's accuracy of using one in preference to the other



(a) (b)

Figure 6.11: Histograms showing FM sensitivity to a change in (a) the ice crystal shape (results(hexagonal columns) - results(aggregates)), and (b) the ice cloud particle size distribution (results(Ou Liou PSD) - results(other PSDs))

the distribution. Using 2 SDs, which contained over 93% of the data for all 3 channels, the uncertainties were calculated as +/- 1.809K, +/- 1.130K or +/- 0.913K for 12μm, 10.8μm and 3.7μm respectively.

Combining the FM uncertainties from the choice of ice crystal shape and PSD in quadrature gives the FM uncertainties in table 6.1. It should be noted that these errors may be pessimistic, as the comparisons are between parameterisations, not a single parameterisation and reality. The possibility of all the parameterisations introducing a similar bias is not accounted for, but in the absence of full line-by-line calculations of ice cloud radiances, these uncertainties are considered the most appropriate available to this study.

Total Forward Modelling Uncertainty. The FM errors discussed in the preceding 2 sections, and those which are also relevant to the clear sky calculations, see section 2.4.2 on page 16, are summarised in table 6.1. Combined in quadrature, they give the total FM errors that were used for simulations of cloudy observations from the SEVIRI sensor. PDFs were also created with a maximum and minimum FM uncertainty, using +/-6μm (instead of +/-3μm) for ϵ_{r_e} and using the peak plus 1SD from the sensitivity test results (instead of the peak plus 2SDs) to find the FM uncertainty due to PSD and r_e. The results of the cloud detection for the SEVIRI images considered were unchanged when either of these PDFs were used in place of the PDF created with the values in table 6.1, which were judged the most appropriate FM uncertainties.

Channel	**Cloud Phase**	**Viewing Geometery**	**Error Source**	**Error [K]**
All	All	All	RTTOV assumption of optically thin clouds	+/-0.5
12μm, 10.8μm 3.9μm 10.8μm 12μm, 3.9μm	Water Ice	All	RTTOV negligence of scattering effects	+/- 1.0 +/- 2.5 +/- 1.0 +/- 1.5
12μm 10.8μm 3.9μm 12μm 10.8μm 3.9μm	Water	0^o 60^o	RTTOV's assumption of a fixed cloud effective radius	+/- 0.089 +/- 0.051 +/- 0.058 +/- 0.041 +/- 0.121 +/- 0.111
12μm 10.8μm 3.9μm -	Ice	All	RTTOV's PSD parameterisation and fixed ice crystal shape	+/- 1.813 +/- 1.151 +/- 0.915 -
12μm 10.8μm 3.9μm 12, 10.8μm 3.9μm	All	0^o 60^o	Clear-sky comparison with line-by-line calculations (Embury, 2006)	+/-0.04 +/- 0.03 +/- 0.05 +/- 0.10 +/- 0.07

Table 6.1: Errors assigned to particular aspects of forward modelling for clouds using RTTOV.

6.5 Example of Use of Local PDF for Cloud

An NWP profile centred on an area off the coast of Korea, shown in yellow in figure 6.13, on page 100, was used to calculate a local PDF for cloud for pixels imaged in that region by the Advanced ATSR sensor at 11:11UTC on 10th May 2005. The local PDF is shown alongside the global distribution of cloudy observations (which is currently used for the Bayesian cloud detection) in figure 6.12. $P\left(\mathbf{y^b}\left|\mathbf{x^b},\overline{c}\right.\right)$ can be read from figures 6.12a and b and used in equation 2.1. Currently, $P\left(\mathbf{y^b}\left|\overline{c}\right.\right)$ is read from figures 6.12c and d and entered in equation 2.1 as $P\left(\mathbf{y^b}\left|\mathbf{x^b},\overline{c}\right.\right)$. The case study compares the skill resulting from using the local PDF and using the global distribution for this region, and also for the region marked in cyan in figure 6.13, for which an NWP profile centred on that area was exploited.

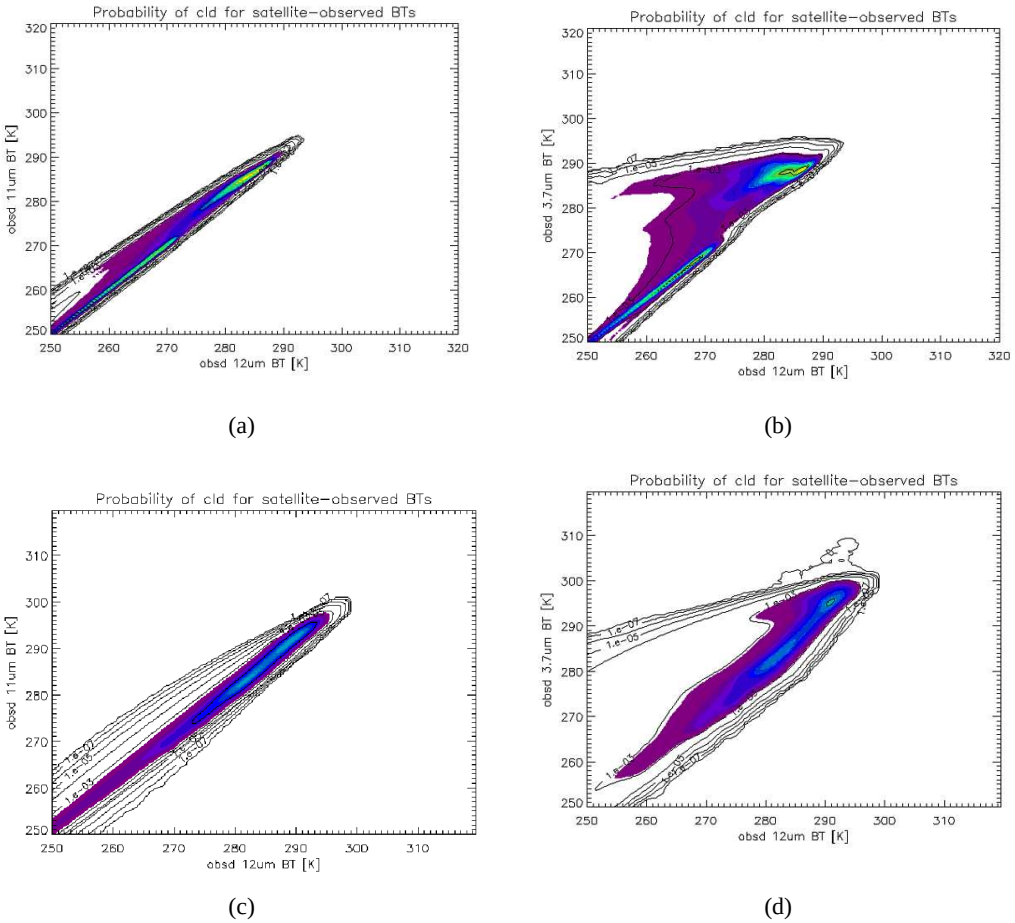


Figure 6.12: (a) and (b) Local PDF for cloud, and (c) and (d) global LUT currently used in place of a local PDF. Both are plotted in 2 dimensions (summed over the third dimension) with filled contours at 20 equally spaced intervals, spanning range of distribution (0-0.0127). Black contours plotted on a logarithmic scale.

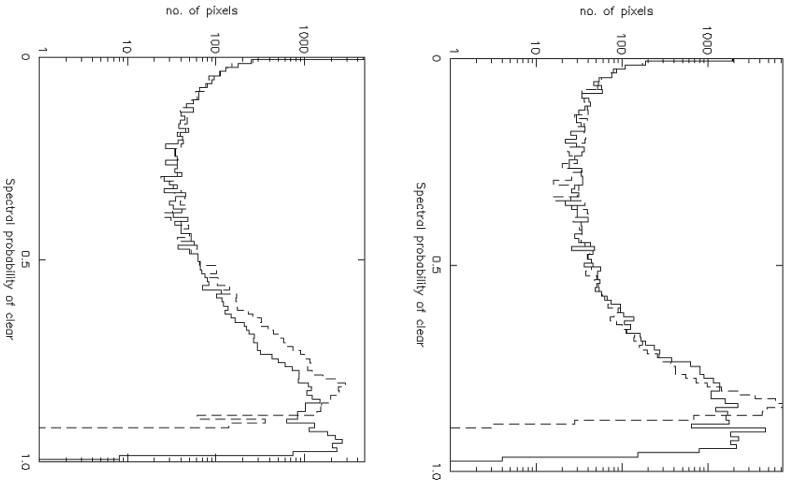
It is anticipated that using a local PDF for cloud, calculated as described earlier in this chapter, should improve the results of the Bayesian cloud detection by making the distribution of calculated posterior probabilities more bimodal - ideally a clear pixel should have a 100% probability of clear, while a cloud pixel should have a 0% probability of clear. A NWP-dependent

PDF for cloud has the advantage of being narrower than the global distribution, see figure 6.12, as only clouds realistic to the profile are represented. Values read from a narrower PDF are likely to be higher or lower than those read from a broader distribution, i.e. observations can be said with more certainty to correspond to a cloud observation or not. When entered into equation 2.1, this results in a more polarised posterior probability of clear being calculated. A region of 100x100 pixels centred on the 2 profile locations marked in figure 6.13 was selected and the probability of clear calculated using the appropriate NWP-dependent PDF and the global distribution. The distributions of posterior probability of clear values found using the two different PDFs, see figure 6.14, show that using the local PDF for cloud increases the certainty with which clear pixels are identified, i.e. higher posterior probabilties of clear are calculated. Although this indicates an increased certainty in the result, it does not compare the accuracy of the results - it is generally more desirable to achieve a less certain result which is the correct side of a 50% probability of clear, than a highly certain result, which is the wrong side. An assessment of the accuracy of the two detections is needed for comparison.

It should be noted that in both cases in figure 6.14, the distributions result from using a local PDF for cloud calculated from a single NWP profile, on which the region is centred, and not from NWP fields interpolated to each individual pixel from the nearest surrounding NWP profile locations (which occurs for calculation of the PDF for clear sky). The resulting posterior probability of clear can therefore be expected to have an accuracy which decreases with distance from the regions' centres. Although not anticipated to produce optimum results, using a single NWP profile allows the calculations to be carried out quickly, which is especially useful at this investigative stage. Ignoring the textural component of the PDFs, since it is the same whether a global or local PDF is used for spectral component, a 'spectral-only' probability of clear[3] can be calculated for pixels in the study regions, as shown in figures 6.15 and 6.16. Both figures show the local PDF for cloud to result in a more definite classification but without any obvious change to which pixels are identified as cloud. This suggests that the local PDF for cloud should only affect more ambiguous pixels, i.e. those with a probability of clear mid-way between 0 and 1. To determine if and how the accuracy of the cloud detection is affected in these regions, a 'truth' is needed for comparison with these results.

[3]This is equivalent to setting the textural PDF for both clear and cloud to 1, $P\left(\mathbf{y_t^o}\left|\mathbf{x^b}, c\right.\right) = P\left(\mathbf{y_t^o}\left|\mathbf{x^b}, \bar{c}\right.\right) = 1$

Figure 6.13: Image acquired at 1.6μm with marked regions centred on the location of the NWP profiles used to generate PDFs. North is at the top of the image.



(a) (b)

Figure 6.14: Histogram of calculated posterior probability of clear for pixels, with dashed and solid lines showing the result of using the global distribution and local PDF respectively. Result (a) for region in yellow box in figure 6.13, and (b) for region in cyan box.

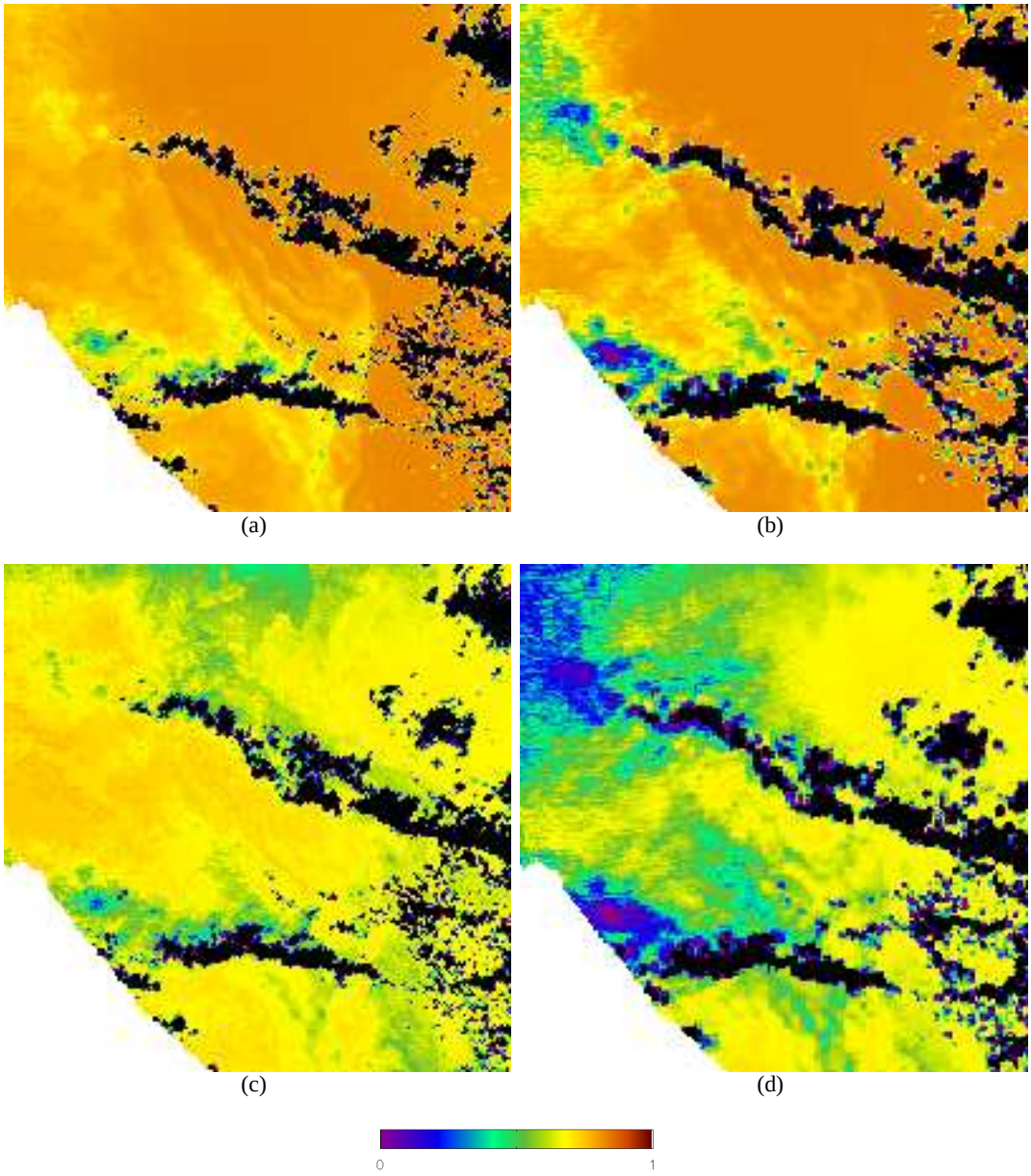


Figure 6.15: The posterior probability for clear calculated for pixels in the Northern study region in figure 6.13 (the lower left corner corresponds to land and so is masked out): (a) using a local PDF for cloud and a nadir-view; (b) a local PDF for cloud and a forward-view; (c) using the global PDF for cloud and a nadir-view; (d) a global PDF for cloud and a forward-view.

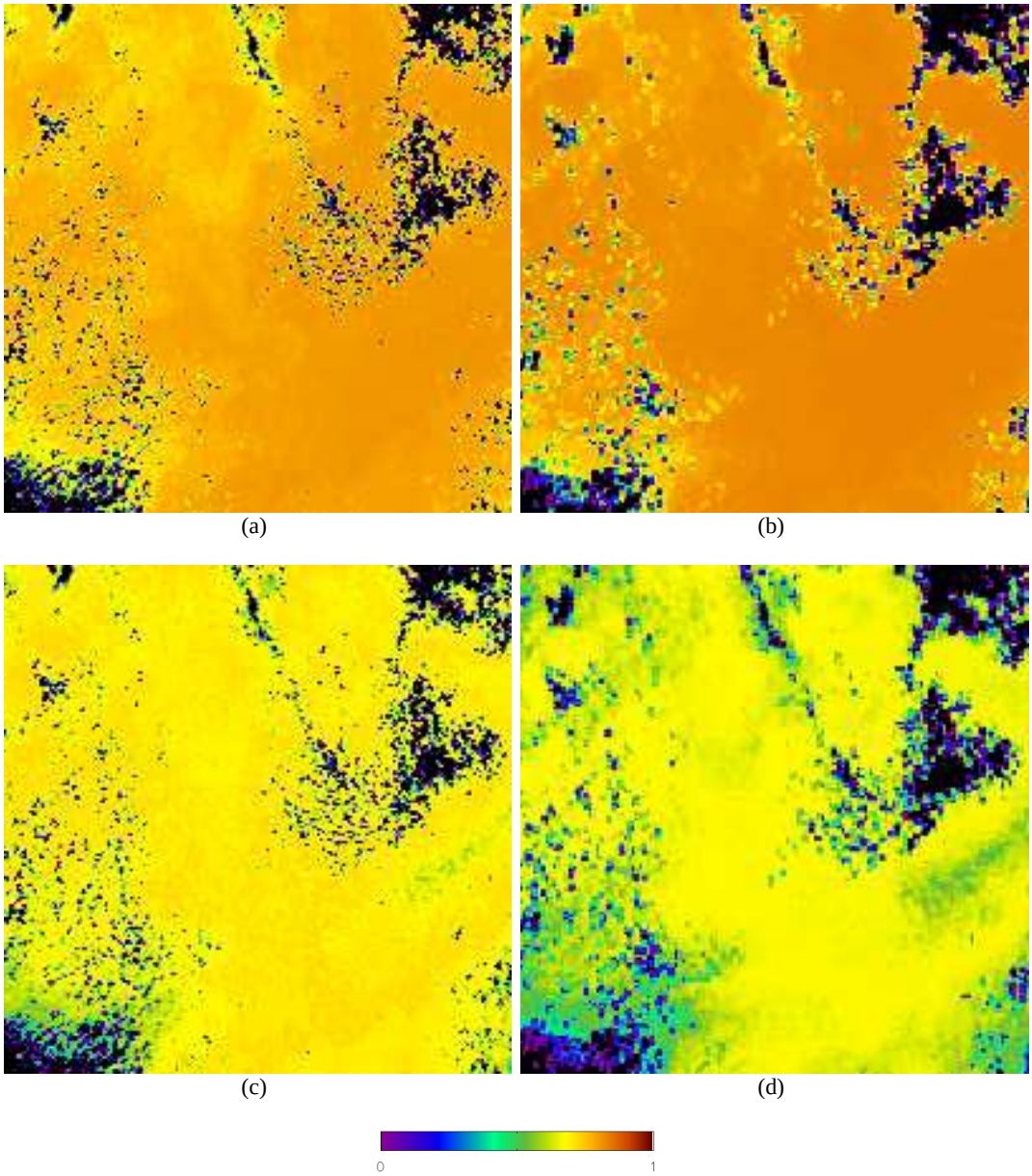


Figure 6.16: The posterior probability for clear calculated for pixels in the Southern study region in figure 6.13: (a) using a local PDF for cloud and a nadir-view; (b) a local PDF for cloud and a forward-view; (c) using the global PDF for cloud and a nadir-view; (d) a global PDF for cloud and a forward-view.

6.5.1 Constructing a 'True' Cloud Mask

To further study the distribution of calculated posterior probability of clear values, a 'truth' image was constructed for each of the regions in figure 6.13. The regions were classified into 'clear' and 'cloud' by hand by two experts[4] experienced in cloud detection for satellite imagery. The hand-classification was done using a widget (a graphical user interface) to set dynamic thresholds on pixel values at visible and thermal wavelengths of the imagery. The widget is described in Appendix A. Pixels within 3 pixels of land (as defined by the Advanced-ATSR land-mask product) were excluded from the 'truth'. Land pixels were excluded because at this time, the Bayesian cloud detection had only been validated for imagery acquired over ocean surfaces. Coastal pixels were excluded partly because of the lack of confidence expressed by both experts in their screening next to the land-mask, and partly to ensure that all land pixels are excluded (the land-mask is temporally-static, while the area of land covered by water changes with tides). This resulted in 1.18% of pixels outside the land-mask in the Northern (yellow box) region being excluded, and none in the Southern (cyan box) region.

Intuitively, it seems that all available information, i.e. both visible and thermal imagery, should be considered in order to construct the most accurate 'truth' for the scene, however there exist clouds which only affect observations at visible wavelengths. For example, there exist very thin clouds, and clouds which only fill a very small fraction of the pixel, which may be seen in visible wavelength imagery, while having an effect at infra-red radiances that is smaller than can be identified by expert inspection. A 'truth' constructed through consideration of both visible and thermal imagery will include such clouds, which the automated detection was not, at this time, designed to identify [5]. To demonstrate the difference between a 'truth' constructed from thermal- and from visible-imagery, figure 6.17 shows both 'truths' constructed using the judgment of the same 2 experts for a nadir-viewed image of the Northern study area in figure 6.13. The agreement on 'true' pixel classification between the 2 experts was higher for both thermally- and visibly-constructed 'truths' than for either expert between their own thermally- and visibly- constructed 'truths', see table 6.2. Problems in pixel co-location between the forward- and nadir-viewed imagery were overcome by constructing a separate 'truth' mask for each of the views (this also overcomes the problem that forward-viewed imagery necessarily views a 'thicker' slice of atmosphere).

The two experts constructed their masks at different times, and neither was made aware of the other's results until both were completed. The two experts agreed on the classification of over 96% of pixels in each of the areas being investigated, and only these were used as a 'truth', against which to judge the results of the automated cloud detections. Although this results in less than 4% of pixels being excluded, it is accepted that disagreement between the two masks

[4] C. Merchant and O. Embury, IAES, University of Edinburgh - it was decided inappropriate for the author, having seen the results of the automated detection before the construction of a 'truth' mask, to hand-classify the image and contribute to the 'truth' used for comparison.

[5] At this point in the project the visible wavelength RTM used for the work described in chapter 5 had not been developed and the cloud detection algorithm exploited only imagery recorded at thermal wavelengths.

'truth' masks compared	% of pixels with the same classification
Northern region - nadir view	
OE (V) and CM (TIR)	97.77
CM (V) and CM (TIR)	93.33
OE (V) and OE (TIR)	94.09
OE (TIR) and CM(TIR)	97.68
Northern region - forward view	
OE (TIR) and CM (TIR)	96.28
Southern region - nadir view	
OE (TIR) and CM (TIR)	96.26
Southern region - forward view	
OE (TIR) and CM (TIR)	97.50

Table 6.2: Table showing the agreement between 'truths' constructed by 2 experts using visible (V) and (TIR) wavelength imagery. The individual experts' masks are referred to by the experts' initials, OE and CM.

appears biased towards cloud edges, and so the calcluated skill may not reflect the accuracy of the automated detection of cloud edges. It is also possible that those pixels for which the two experts disagree are likely to be those which present the greatest challenge to the automated detection and so the results of testing may be slightly biased towards 'easier' classifications. It should be further noted that, while clouds represented in the PDF are required to introduce a bias of at least 0.2K to the retrieved surface temperature, there was no such restriction placed on clouds detected by eye in construction of the 'truth' cloud mask. It is therefore possible that the 'truth' mask contains clouds that the automated detection is not designed to detect. Despite these faults, this method of constructing a 'truth' was considered superior to the alternative of using another automated detection algorithm, and is similar to that used in other validation studies, for example in Merchant et al. (2005).

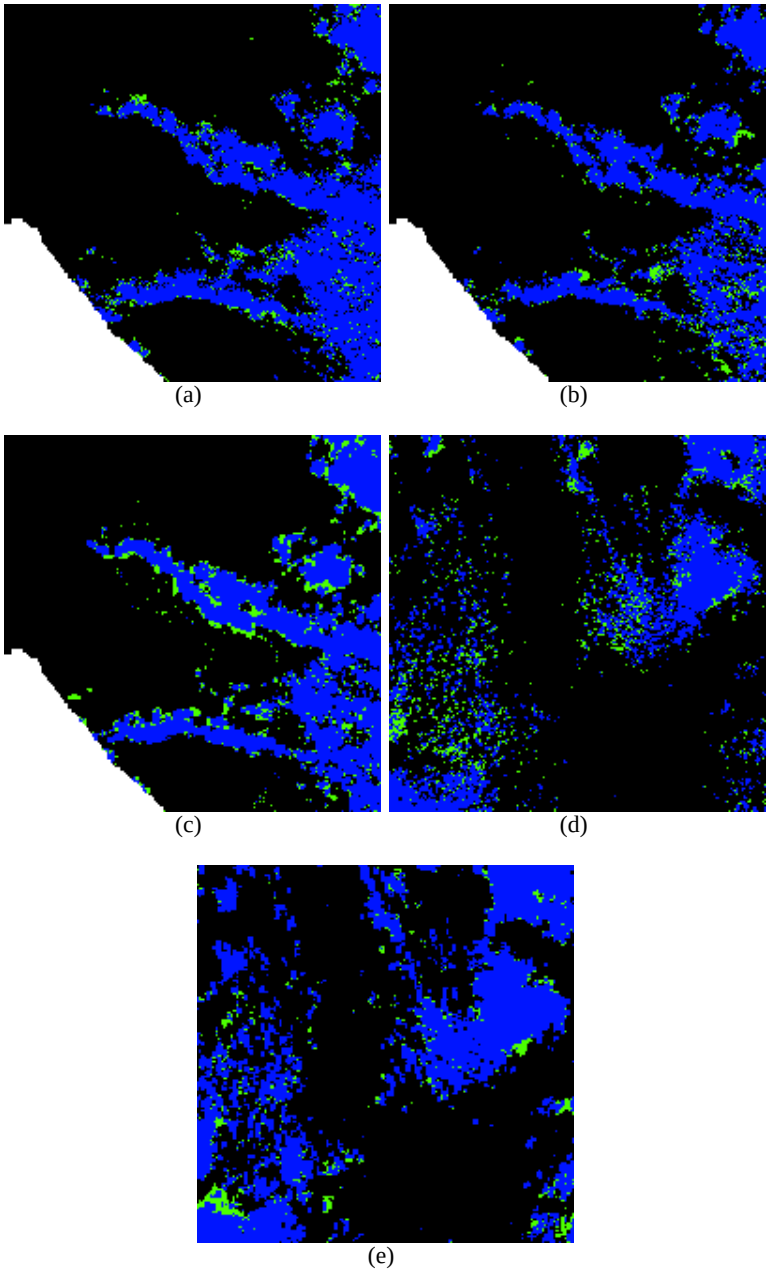


Figure 6.17: Agreement between the hand-classification of the two experts for the Northern study region in figure 6.13, using: (a) visible wavelength-, nadir-viewed- imagery; (b) TIR wavelength-, nadir-viewed- imagery; (c) TIR wavelength-, forward-viewed- imagery; agreement between both experts for the Southern study area in figure refkorea, using (d) TIR wavelength-, nadir-viewed- imagery, (e) TIR wavelength-, forward-viewed- imagery. Blue: both experts classified as cloud, Black: both experts classified clear, Green: experts disagreed, White: land-mask.

6.5.2 Case Study Results

The 'truth' mask is used to assess and compare the performance of the cloud detection when the global and local PDFs for cloud are used. As mentioned in section 6.5, the accuracy of results from the local PDF should decrease with distance from the regions' centres, and so bands of imagery, 10 pixels thick, were initially identified for skills testing, see figure 6.18. The distance between the centre of each band and the pixel with latitude and longitude closest to the profile latitude and longitude was incremented by 10 pixels for successive bands, and the skill assessed using the methods defined in section 5.3.1 on page 64 to define HR, FAR and TSS. The effect of the local cloud PDF on the skill fluctuated between positive and negative values, and it was concluded that there were too few pixels in each band to facilitate a meaningful analysis. Skills scores were therefore calculated for each region as a whole, setting different thresholds on the probability of clear to define clear pixels, to give some quantitative indication of the skill, see figure 6.19.

The skill of the two detection methods appears to be very similar until a high threshold is applied to the probability of clear (as could occur for applications requiring strict cloud screening), which supports the qualitative findings in figures 6.15 and 6.16, and is shown in more detail in table 6.3. The small sample size, the lack of an objective 'truth', and the failure to make the local PDF for cloud truly pixel-specific means that this case study can only be used to demonstrate the computationally-successful implementation of the technique, and the resulting enhancement to the polarization of the calculated probabilities of clear. A larger dataset, with an externally supplied 'truth', and for which pixel-specific calculations of the local PDF for cloud are made, is needed to validate the technique.

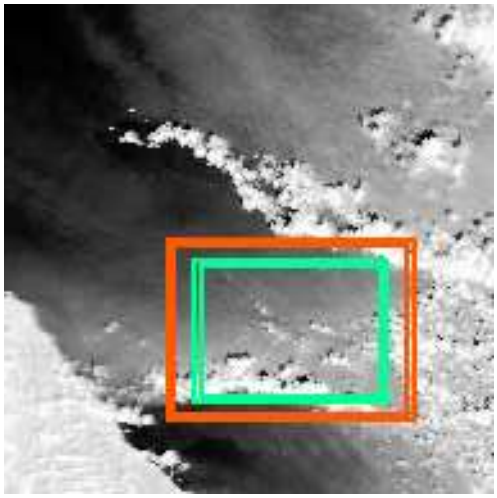


Figure 6.18: Region surrounding location of Northern NWP profile in figure 6.13, with 2 areas (bands 10 pixels deep between lines of the same colour) in which the skill if the automated detection, using the global and local PDFs for cloud, was assessed.

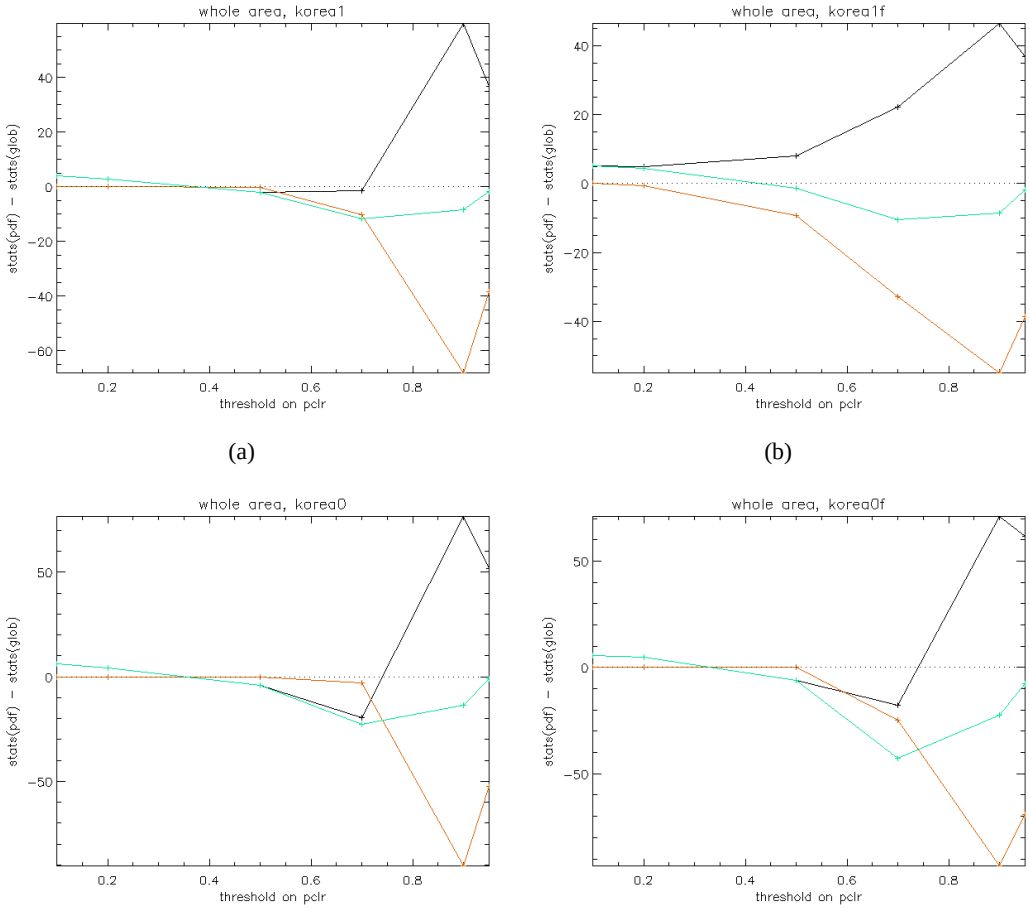


Figure 6.19: Skills scores calculated using the results from the local PDF - skills scores calculated from results from the global PDF for cloud for the study regions shown in figure 6.13: (a) Northern region, nadir-view; (b) Northern region, forward-view; (c) Southern region, nadir-view; (d) Southern region, forward-view. Black shows TSS, Red shows FAR and Cyan shows HR - see text for definitions.

Northern region

-	**Glob-PDF**	**Loc-PDF**
% of clear pixels with pclr $\geq$ 85%	26.9	56.3
% of clear pixels with pclr $\geq$ 90%	0	37.5
% of cloud pixels with pclr $\leq$ 15%	70.5	71.9
% of cloud pixels with pclr $\leq$ 10%	66.6	68.0

Southern region

-	**Glob-PDF**	**Loc-PDF**
% of clear pixels with pclr $\geq$ 85%	20.7	60.4
% of clear pixels with pclr $\geq$ 90%	1.5	44.8
% of cloud pixels with pclr $\leq$ 15%	12.9	12.1
% of cloud pixels with pclr $\leq$ 10%	11.9	11.1

Table 6.3: Comparison of the calculated posterior probability of clear (pclr) for 'truly clear' and 'truly cloud' pixels (as defined by the truth mask described in section 6.5.1) in the regions marked in figure 6.13, using the global PDF (Glob-PDF) and the local, NWP-dependent PDF calculation (Loc-PDF).

6.6 Use of a Local Cloud PDF for a Larger Dataset

The results of the case study above are encouraging, but are not accompanied by an objective 'truth' for assessment of the skill, and are based on too small a data sample to be conclusive. Furthermore, the local PDFs which were used are not pixel specific, having been calculated from two single NWP profiles and applied without any interpolation to every pixel in the surrounding region. A more appropriate assessment of the effect on the cloud detection skill of replacing the global PDF for cloud with a NWP-dependent PDF for cloud requires a larger dataset, ideally with a 'truth' for comparison. Also, an NWP profile is either required for every pixel, or the NWP-dependency of the forward-modelling must be used to interpolate PDFs calculated from the closest NWP profile locations to the pixel location, using appropriate pixel properties, as occurs in calculation of the PDF for clear sky.

The dataset used for the validation work in chapter 5, which contains both NWP data and observation data for the central pixel in targets of 5x5 pixels, as described in Appendix 3.3, was used to compare the skill of the Bayesian cloud detection technique with a global, and with a local, NWP-dependent, PDF for cloud. Some targets, such as those belonging to aerosol classes (and so corresponding to neither clear nor cloud) were removed as described in Appendix 3.3, leaving 14931 targets for this study. The targets cover a range of cloud types, imaged over sea and land from day- and night-time scenes, at the locations shown in figure 3.1. The number of targets in each individual class for day-sea, day-land, night-sea and night-land is shown in table 3.1, also in the Appendix. The database has the advantage of containing an independent 'truth' for comparison with the algorithm results, and all the required NWP information specific to every target, i.e. with no interpolation required.

The methods described in this chapter were used to calculate a local spectral PDF for each target from its NWP information. For daytime targets, a 2-dimensional NWP-dependent spectral PDF was calculated, with dimensions corresponding to 11μm and 12μm, while for night-time targets a third dimension, corresponding to 3.9μm was included. These PDFs were substituted for the spectral global PDF for cloud used in chapters 2 and 5. The textural considerations of the algorithm were not held constant, as they were for the case study, but were used as described in chapter 2, when both PDFs were used, i.e. there was no difference between the texture considerations used when the local spectral PDF was implemented, and when the global spectral PDF was implemented. For daytime imagery, the spectral PDF was multiplied by the empirical visible and near-infrared-wavelength spectral PDF for cloud described in section 5.2 in chapter 5. The algorithm was run using the global PDF for cloud described in preceding chapters, and then for the same targets using the local PDF for cloud which is described in this chapter. The results compared in order to assess the relative skill of the method using the different PDFs. The algorithm has been shown to give successful cloud detection when textural and visible-wavelength information is exploited, and so it was decided to include these. This means that some measure of the skill of the technique using both PDFs can be given, in addition to the relative assessment that can be made through comparison of the two sets of results.

For day-time imagery observations at five wavelengths were exploited: 0.6, 0.8, 1.6, 10.8 and 12.0μm; for night-time imagery observations at 3.9, 10.8 and 12.0μm were used. This is slightly different to the work in section 5.3, chapter 5, where not all the observation channels were exploited for daytime imagery. The results for night- as well as for day-time imagery may also be slightly different to those in chapter 5 due to the ongoing development of the Bayesian detection algorithm - the work presented in this chapter was undertaken at a later date, and the results may be affected by some minor changes that were made to the algorithm. The skill is quantified using the same measures as for the case study in section 6.5.2, which are defined on page 64, and the results are summarised in table 6.4. The database is comprised of targets specially selected as 'difficult' cases for cloud detection, and so the results are not necessarily representative of the general skill of the method. For comparison, therefore, results from the operational techniques of MétéoFrance and the U.K. Met. Office for the same targets are shown in table 6.5. The results are broken down into individual target classes in tables 6.6 to 6.9. Distributions of the calculated probabilities of clear, resulting from using the global- and using the local-PDF are shown in figurelocPCLRplot1 for day sea targets, for both 'truly' clear and 'truly' cloudy targets. The equivalent distributions for day land-, night sea- and night land- targets are shown in figures 6.21 to 6.23.

-	**Global PDF**				**Local PDF**			
-	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**
Overall	86.5	90.6	20.5	**70.1**	76.8	67.6	7.6	**60.0**
Day Sea	94.3	96.2	9.5	**86.7**	94.7	93.1	1.9	**91.2**
Day Land	86.1	92.3	20.6	**71.7**	77.9	66.7	10.1	**56.5**
Night Sea	93.6	93.9	6.9	**87.0**	87.1	80.2	1.7	**78.6**
Night Land	78.2	84.9	36.6	**48.3**	60.4	48.0	12.6	**35.5**

Table 6.4: Skill of the Bayesian cloud detection technique when a global PDF for cloud is employed, and when it is substituted with a local (i.e. pixel-specific), NWP-dependent PDF for cloud. For the definition of PP, HR, FAR and TSS see section 5.3.1 on page 64.

-	**MétéoFrance**				**U.K. Met. Office**			
-	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**	***PP*[%]**	***HR*[%]**	***FAR*[%]**	***TSS*[%]**
Overall	79.2	74.2	12.3	**61.9**	77.8	84.9	34.2	**50.8**
Day Sea	95.1	95.0	4.6	**90.4**	93.1	93.1	6.8	**86.3**
Day Land	85.3	89.6	19.4	**70.2**	-	-	-	**-**
Night Sea	85.6	80.9	6.9	**74.0**	88.4	84.6	5.3	**79.3**
Night Land	62.9	51.9	13.0	**38.9**	-	-	-	**-**

Table 6.5: Skill of the operational techniques for the same targets as those for which the Bayesian results are given in table 6.4. Only results for sea targets were available from the U.K. Met. Office. For the definition of PP, HR, FAR and TSS see section 5.3.1 on page 64.

Day Sea Targets

-	**Global PDF**				**Local PDF**			
Class	***z***	***f***	***h***	***u***	***z***	***f***	***h***	***u***
101	667	45	-	-	703	9	-	-
102	1	1	-	-	1	1	-	-
106	132	38	-	-	163	7	-	-
502	-	-	492	16	-	-	456	52
503	-	-	505	20	-	-	503	22
504	-	-	1	0	-	-	1	0
601	-	-	37	32	-	-	31	38
606	-	-	34	0	-	-	33	1
608	-	-	73	0	-	-	73	0
609	-	-	15	0	-	-	15	0
701	-	-	102	1	-	-	90	13
705	-	-	79	0	-	-	79	0
706	-	-	6	0	-	-	6	0
707	-	-	37	0	-	-	37	0
801	-	-	80	0	-	-	80	0
802	-	-	75	0	-	-	75	0
811	-	-	112	0	-	-	112	0
812	-	-	106	0	-	-	106	0
TOTAL:	**800**	**84**	**1754**	**69**	**867**	**17**	**1697**	**126**

Table 6.6: Skill of the Bayesian cloud detection technique when a global PDF for cloud is employed, and when it is substituted with a NWP-dependent PDF for cloud for day-time sea targets. For the definition of z, f, h and u, see section 5.3.1 on page 64.

Day Land Targets

-	**Global PDF**				**Local PDF**			
Class	***z***	***f***	***h***	***u***	***z***	***f***	***h***	***u***
151	1146	279	-	-	1258	167	-	-
152	6	13	-	-	18	1	-	-
181	0	5	-	-	4	1	-	-
191	252	67	-	-	309	10	-	-
502	-	-	284	34	-	-	84	234
503	-	-	268	20	-	-	129	159
504	-	-	0	1	-	-	0	1
602	-	-	114	54	-	-	48	120
607	-	-	15	0	-	-	13	2
608	-	-	160	0	-	-	160	0
609	-	-	79	0	-	-	79	0
702	-	-	198	35	-	-	151	82
704	-	-	3	2	-	-	0	5
705	-	-	69	0	-	-	59	10
706	-	-	6	0	-	-	6	0
707	-	-	48	0	-	-	45	3
801	-	-	39	0	-	-	36	3
802	-	-	241	0	-	-	231	10
811	-	-	173	0	-	-	172	1
812	-	-	47	0	-	-	47	0
TOTAL:	**1404**	**364**	**1744**	**146**	**1589**	**179**	**1260**	**630**

Table 6.7: Skill of the Bayesian cloud detection technique when a global PDF for cloud is employed, and when it is substituted with a NWP-dependent PDF for cloud for day-time land targets. For the definition of z, f, h and u, see section 5.3.1 on page 64.

Night Sea Targets

-	**Global PDF**				**Local PDF**			
Class	***z***	***f***	***h***	***u***	***z***	***f***	***h***	***u***
101	1174	87	-	-	1240	21	-	-
502	-	-	717	86	-	-	502	301
503	-	-	559	15	-	-	509	65
601	-	-	23	23	-	-	11	35
606	-	-	8	1	-	-	8	1
608	-	-	64	0	-	-	64	0
609	-	-	37	0	-	-	37	0
701	-	-	66	0	-	-	63	3
705	-	-	72	0	-	-	72	0
706	-	-	2	0	-	-	2	0
707	-	-	41	0	-	-	41	0
801	-	-	57	0	-	-	57	0
802	-	-	121	0	-	-	121	0
811	-	-	82	0	-	-	82	0
812	-	-	75	0	-	-	75	0
TOTAL:	**1174**	**87**	**1924**	**125**	**1240**	**21**	**1644**	**405**

Table 6.8: Skill of the Bayesian cloud detection technique when a global PDF for cloud is employed, and when it is substituted with a NWP-dependent PDF for cloud for night-time sea targets. For the definition of z, f, h and u, see section 5.3.1 on page 64.

Night Land Targets

-	**Global PDF**				**Local PDF**			
Class	***z***	***f***	***h***	***u***	***z***	***f***	***h***	***u***
151	1031	556	-	-	1411	176	-	-
191	9	45	-	-	24	30	-	-
502	-	-	1637	440	-	-	732	1345
503	-	-	829	100	-	-	447	482
607	-	-	1	0	-	-	1	0
608	-	-	45	0	-	-	45	0
609	-	-	49	0	-	-	49	0
702	-	-	104	1	-	-	95	10
705	-	-	108	4	-	-	85	27
707	-	-	26	0	-	-	26	0
801	-	-	40	0	-	-	39	1
802	-	-	161	1	-	-	149	13
811	-	-	38	0	-	-	38	0
812	-	-	31	0	-	-	31	0
TOTAL:	**1040**	**601**	**3069**	**546**	**1435**	**206**	**1737**	**1878**

Table 6.9: Skill of the Bayesian cloud detection technique when a global PDF for cloud is employed, and when it is substituted with a NWP-dependent PDF for cloud for night-time land targets. For the definition of z, f, h and u, see section 5.3.1 on page 64.

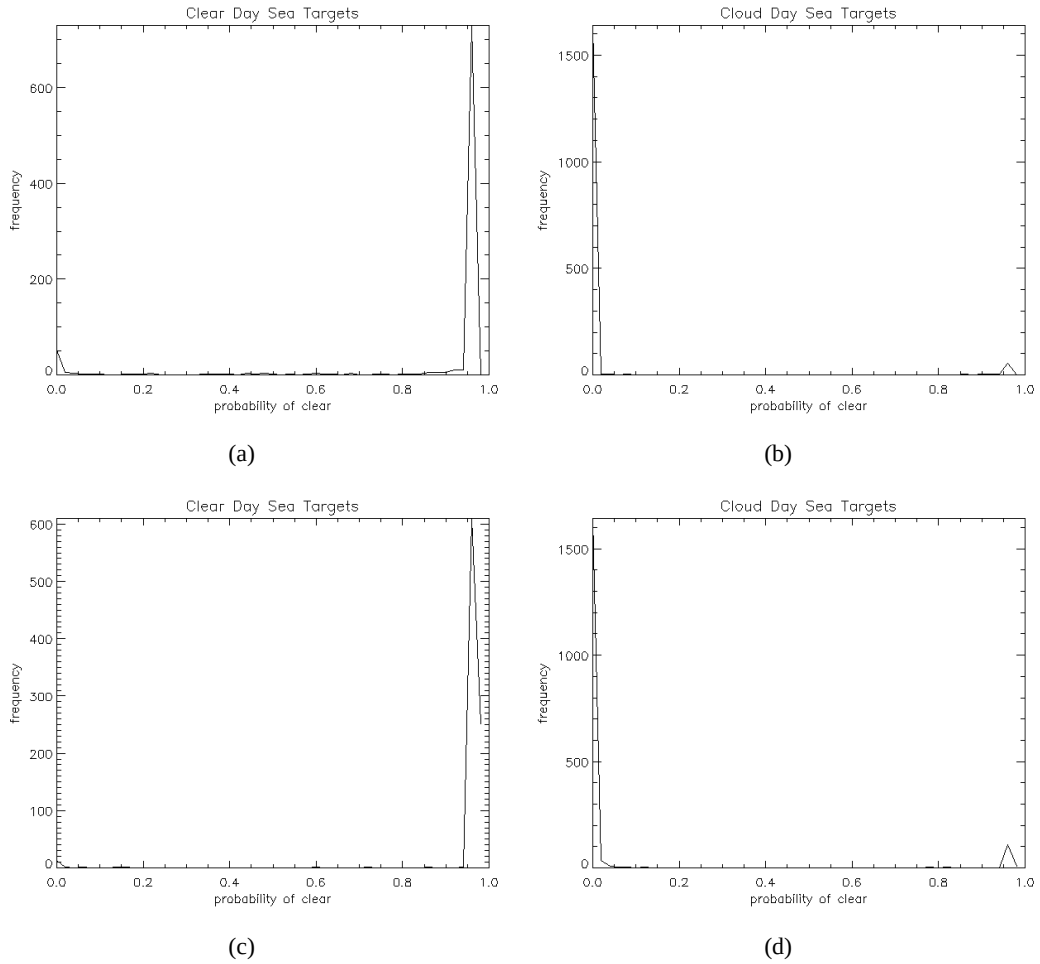


Figure 6.20: Distribution of calculated probabilities of clear for day sea targets: (a) results for ‘truly’ clear targets, using global PDF for cloud; (b) results for ‘truly’ cloud targets, using global PDF for cloud; (c) results for ‘truly’ clear targets, using local PDF for cloud; (d) results for ‘truly’ cloud targets, using local PDF for cloud.

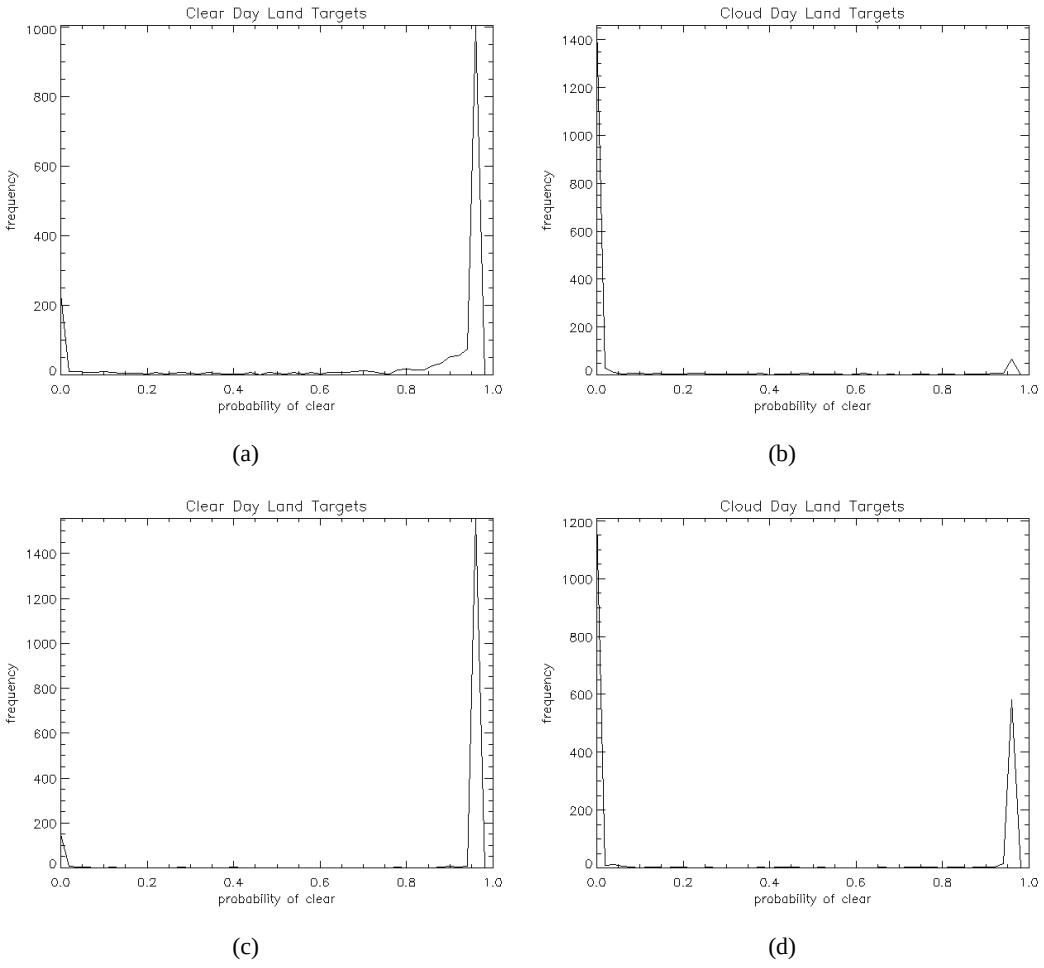


Figure 6.21: Distribution of calculated probabilities of clear for day land targets: (a) results for 'truly' clear targets, using global PDF for cloud; (b) results for 'truly' cloud targets, using global PDF for cloud; (c) results for 'truly' clear targets, using local PDF for cloud; (d) results for 'truly' cloud targets, using local PDF for cloud.

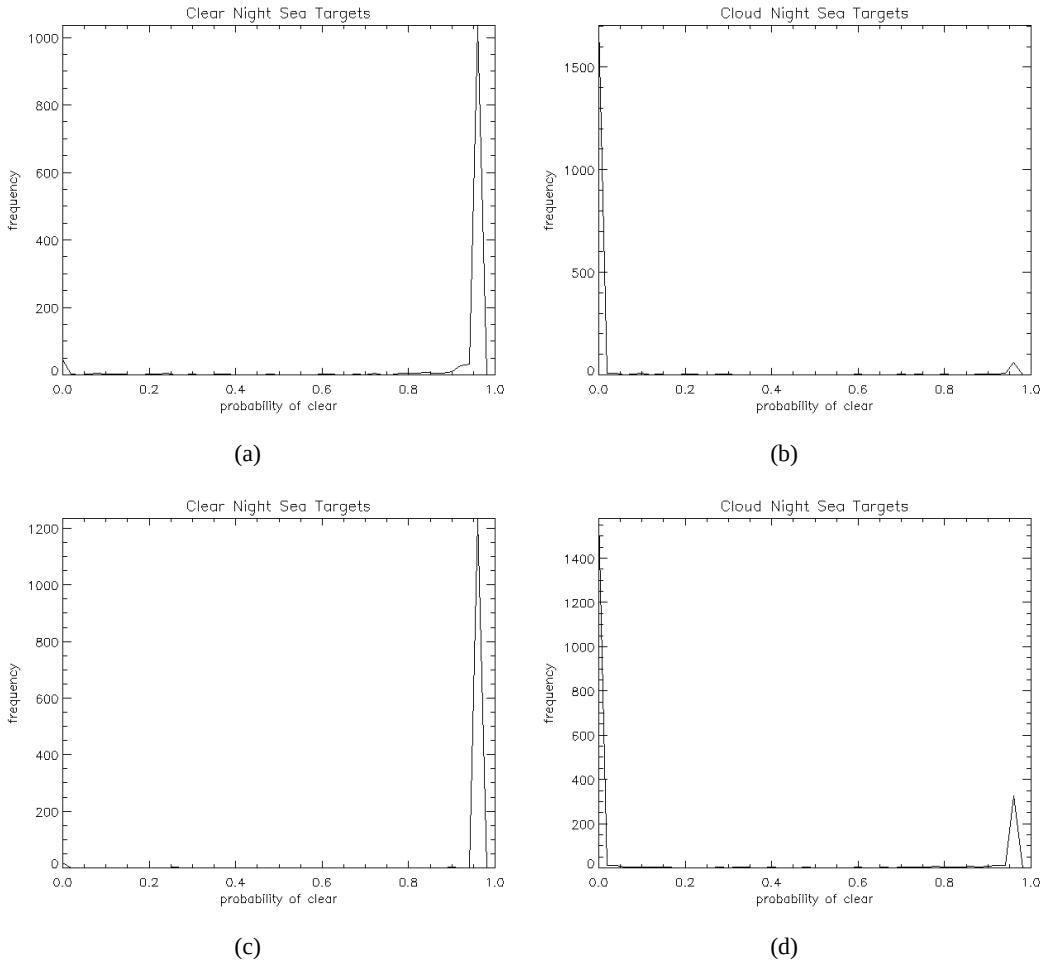


Figure 6.22: Distribution of calculated probabilities of clear for night sea targets: (a) results for 'truly' clear targets, using global PDF for cloud; (b) results for 'truly' cloud targets, using global PDF for cloud; (c) results for 'truly' clear targets, using local PDF for cloud; (d) results for 'truly' cloud targets, using local PDF for cloud.

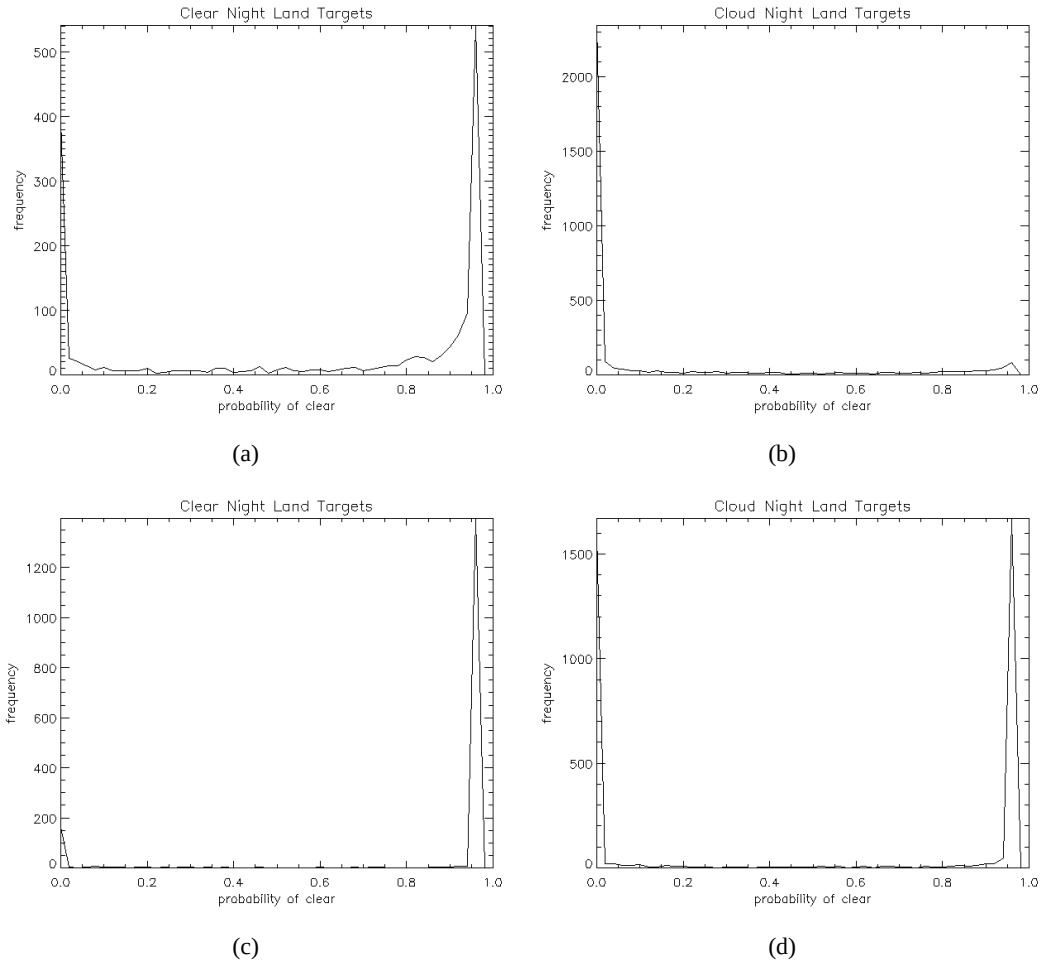


Figure 6.23: Distribution of calculated probabilities of clear for night land targets: (a) results for 'truly' clear targets, using global PDF for cloud; (b) results for 'truly' cloud targets, using global PDF for cloud; (c) results for 'truly' clear targets, using local PDF for cloud; (d) results for 'truly' cloud targets, using local PDF for cloud.

6.7 Discussion

Overall, the results show that using the local PDF for cloud on this dataset does not improve the results and in most cases, actually reduces the skill of the cloud detection. This is mainly due to a decrease in *HR*, indicating that pixels are less likely to be classified as cloud when the local PDF is used. A 4.6% improvement is seen in *TSS* for day sea targets, but this is countered by the 15.2%, 8.4% and 12.8% reductions in *TSS* seen for day land, night sea and night land respectively. For day sea targets, the skill of the operational tecniques is exceeded whichever PDF is used, and for night sea targets the skill achieved using the local PDF is greater than that demonstrated by one operational method, and similar to that achieved by the other operational techqniue, but lower than that achieved by the gloabl PDF. For night land targets, the local PDF results in a similar *TSS* to the operational method, however this is again lower than that achieved if the global PDF is implemented. For day land targets, a significant reduction in the *TSS* for the Bayesian technique follows from the implementation of the local PDF, making it lower than both the operational result and that from the Bayesian approach using the global PDF for cloud. The distributions of posterior probability of clear values calculated for clear and cloudy targets in each day-, night-, land-, sea- category show fewer ambiguously classed pixels, indicating a more certain classification. This is the result of using the narrower, steeper local PDF for cloud, which means that observations are more likely to be assigned a very high or very low probability of cloud, and less likely to be assigned an ambiguous value. A more certain classification of pixels into clear and cloud classes would usually be considered a strength, but the skills scores indicate that pixels are also *mis-classified* with greater certainty when the local PDF for cloud is used, which should be considered a weakness. In the following discussion some reasons for these disappointing results will be presented. The discussion should be read with reference to table 3.1 in Appendix 3.3, where a description of the target code for each class is given. Specific target classes are referred to in the discussion by their code, which is given in parenthesis.

The local PDF for cloud is generally narrower and steeper than the global PDF, see for example figure 6.16. It is therefore more crucially important that it's position be centred on a realistic prediction of cloudy sky radiance for the target. The predictions are based on forward modelling of the NWP information and it may be that, given this a dataset comprised of 'difficult' targets for cloud detection, the spectral properties of the cloud targets cannot be accurately described by the approximations made in the fast forward model. The modelling of cloudy radiances is likely to be sensitive to more fields in the NWP data than the simulation of clear sky radiances, for example slightly different temperature profiles through a clear sky atmosphere would result in slightly different radiances being simulated, but could significantly affect the radiance of a cloud at a given altitude, and so the radiance received at the sensor. This makes it arguably more appropriate for more terms to be included in the **H'**- and **B**-matrices than at present, where only those considered necessary for the clear sky calculation are included. Calculation of the local PDF for cloud is also more sensitive to the NWP data than the clear sky PDF because the data is used in the forward modelling, and in weighting and restricting the representation of specific clouds in the final PDF. The position of the cloudy PDF's peak therefore depends on the

accuracy of the NWP information to a greater degree than for the clear sky case. Including terms to account for uncertainties in more NWP fields in both the **H'**- and **B**-matrices would broaden the PDF for cloud and make it shallower, making the algorithm less sensitive to the exact position of the peak. The extra computational expense, as well as the difficulty of accurately quantifying uncertainties in the NWP data (which the algorithm treats as constant, although in reality they are likely to vary in space and time), means that only uncertainties in TCWV, ST, AOD, wind-speed and surface reflectance are considered, see chapters 2 and 5.

The use of the same **H'**-matrix for both the clear and the cloudy PDFs is also arguably inappropriate. The sensitivity of the RTM to the NWP data can be expected to be different when clouds are modelled and when simulations of clear sky radiances are carried out. The sensitivity of simulated cloud radiances to the NWP fields is also likely to differ between the clouds which are modelled - for example radiance simulated for an atmosphere containing an optically opaque cloud is likely to be less affected by the NWP surface temperature field than a radiance modelled for a thin cloud, or for no cloud. For computational efficiency, the RTM sensitivity to NWP data is only calculated once for every NWP profile. Computing a separate **H'**-matrix for each modelled cloud would slow the algorithm considerably, not only because of the additional **H'**-matrix calculations, but also because the method by which uncertainties are considered for creation of the cloud PDF does not allow the uncertainties to be cloud-specific. There is no straightforward alternative to the structuring of this part of the algorithm which does not involve a prohibitively high computational cost.

It is also possible that uncertainty in the RTM, used to predict cloud observations from the NWP data, causes the centre of the PDF to be positioned inappropriately. The forward modelling uncertainty attributed to the modelling of cloud radiances, detailed in table 6.1, was found mostly through model-model comparisons, rather than from a comparison between RTM predictions and a 'truth'. It is therefore possible that some bias in the predicted cloud observations may exist that is not accounted for in the tabulated uncertainties. Increasing the values accounting for forward-modelling in the **R**-matrix would broaden the PDF in a similar way to increasing the values in the **B**-matrix, but the problem of finding appropriate values, other than through model-model comparisons remains.

The local PDF for cloud improves the results for day sea targets, achieving a *TSS* that is greater than both operational techniques, and greater than when it is substituted for the global PDF for cloud. As the one target category for which the skill of the Bayesian technique was not seen to exceed operational methods in section of 5.3 chapter 5, this is encouraging. Comparing the two sets of day sea Bayesian results, the 4.5% rise in *TSS* is explained by the 7.6% fall in *FAR* - a drop which counters the 3.6% drop in *HR*, giving a positive effect on the overall skill. Multilayer clouds (705, 706, 707 and 812) and thicker clouds, such as cumulonimbus (target classes 608, 609), altocumulus (801, 802) are detected equally successfully using both PDFs. This is not surprising, as the effect of these clouds on observed radiances is likely to be greater than

for thinner clouds, making them easier to distinguish from clear sky. Thin cirrus (701) and small cumulus clouds (601) are not detected as successfully when the local PDF is used. As everything other than the spectral PDFs for cloud is extactly the same in the two versions of the Bayesian algorithm, this shows that these cloud observations correspond to a lower value in the local PDF than in the global PDF for cloud. In the fast forward model used to predict cloud radiances, the liquid- or ice- water path of the cloud is spread over the area covered by a pixel, meaning that the effect of small thick clouds on observed radiances is treated as equivalent to that of extensive thin clouds. The weight given to the representation of full-pixel cloud observations, see figure 6.6, may result in increasing the value of the PDF peak so that thin or small clouds are represented at the edge of the PDF, making their successful detection less likely. For the same reason, the detection of thin or small clouds is likely to depend more heavily on the position of the PDF peak in observation space, than the detection of thicker clouds. Missed stratus targets (502) could be explained by the textural considerations in the algorithm, see 2.4.1 in chapter 2, as stratus cloud is sometimes observed with a cloud top temperature variability more often associated with clear sky observations of sea surfaces than clouds. This effect, however, does not explain why the two different PDFs result in different detection rates for stratus. The fact that these targets are from a database of 'difficult' cloud targets makes it likely that at least some clouds have unusual spectral properties, which may not be predicted by the fast forward modelling on which calculation of the local PDF for cloud is based. This could also explain the difference in the the detection skill for stratocumulus targets (503). Clear sky targets (101, 102, 106) are identified more successfully when the local PDF for cloud is implemented, almost definitely because the steeper, narrower cloud PDF means than a clear sky observation is less likely to correspond to a high value in the PDF for cloud. This is the expected effect of using the local PDF for cloud.

The cloud detection for day land targets using the local PDF for cloud is far less successful than when the global PDF for cloud is used. Clear sky targets (151, 152, 181 and 191) are correctly identified more often when the local PDF is used, making the *FAR* 10.5% lower, but this is countered by the increased mis-classification of cloud targets, which reduces the *HR* by 25.6%, giving a 15.2% lower *TSS* when the local PDF for cloud is used. The reduction in *FAR* can be explained in the same was as for day sea targets above, while the decreased *HR* is probably attributable to the local PDF's dependency on NWP data. NWP information for land targets is likely to be less accurate than for sea targets because of temporal and spatial variability in the parameters represented in the NWP fields. The local PDF for cloud is therefore more likely to be positioned inappropriately in observation space for land targets, i.e. it could represent inappropriate predictions of cloudy-sky observations. Although this is also true for the clear sky PDF, the greater dependency of the cloud PDF on NWP fields make its position and shape more sensitive to NWP uncertaintes. This could explain the failed detection of some thin cirrus and small cumulus targets (602, 702 and 704), observations of which are likely to represented at the edge of the PDF. It is disappointing that more stratus (502), stratocumulus (503) and altocumulus (802) targets are missed when the local PDF for cloud is used. This could be partly due to the fact that the targets are 'difficult' cloud targets, and so unlikely to be spectrally representative of all

clouds, as mentioned in the preceding paragraph. It is, however, more likely that the NWP data over land is not accurate enough to justify the limited number of uncertainties used in expansion of the PDF around the predicted cloud observations. If the issues of accurately quantifying NWP and RTM uncertainties, and of computational expense were overcome, and the **B**-matrix were extended to include all uncertainties in the NWP data which affect the local cloud PDF, it is likely that the PDF for some land targets would become too broad and shallow to be useful. A PDF based on local NWP data is physically more justifiable than a global distribution of observations, and should in theory be capable of distinguishing cloud and clear observations based on spatially- and temporally-specific criteria, and so give a more accurate result. If the NWP data is associated with too high an uncertainty, however, the cloud PDF will be a flat distribution in which a wide range of observations will correspond to very similar probabilities for cloud, which are likely to be low as the PDF is expanded over a greater part of the observation space.

The skill of the cloud detection using the local cloud PDF for night sea targets is similar to that of the operational techniques, but corresponds to a *TSS* 8.4% lower than that achieved when the global PDF for cloud is used. Predicted cloud observations at 3.9μm are likely to be affected more by the scattering approximations made in the forward model than predictions at 11μm and 12μm, which could mean that the PDF for some night targets is centred inappropriately even if the NWP data were reliable. This could result in cloud observations falling outside the area of observation space covered by the cloud PDF, and so being assigned a low probability of cloud even if the uncertainties in the **B**-matrix are appropriate. Small cumulus (601) and thin cirrus (701) observations, which are likely to be represented at the edge of the PDF, are missed more often when the local PDF is used, which may be explained by the position of the PDF peak in the same way as above for day sea and day land targets.

Night land targets present the greatest challenge to all three of the compared techniques. The Bayesian method significantly outperforms the operational approach when a global PDF for cloud is used, but achieves a slightly lower *TSS* than the operational method when the local PDF for cloud is used. Despite the relatively successful *TSS* achieved using the global PDF, the *FAR* for that technique, at 36.6%, is very high. A better *FAR*, similar to the operational value, is achieved when the local PDF is used, however this is attained at the cost of reducing the *HR*, resulting in a *TSS* 12.8% lower than if the gloabl PDF is used. It is not surprising that the least skillful results are seen for night land targets. Cloud detection for land targets is intrinsically more difficult, as clouds are more difficult to distinguish from a variable-temperature land surface than from a relatively uniform sea surface. Less information is available for cloud detection at night, as only thermal infrared observations can be exploited, and less information can be expected to lead to a less skillful result. Both Bayesian techniques rely on NWP data to predict clear sky observations and calculate a clear sky PDF, and NWP data is likely to be less accurate over land. Implementation of the local cloud PDF in the Bayesian method increases the sensitivity to NWP data, as both the clear- and cloudy- PDFs rely on NWP fields, and the dependency of the cloudy PDF is considerably greater, as discussed above. The forward-modelling of cloud

observations may be less accurate when the 3.9μm channel is included, which could make results based on the local cloud PDF less skillful than those based on the gloabl PDF. The increased uncertainty that can be associated with both the forward-modelling and the NWP data, combined with the lack of visible wavelength observations, make night land cloud detection difficult. Implementation of the local cloud PDF in the Bayesian cloud detection scheme increases its sensitivity to the increased uncertainties in NWP and in the RTM, and so gives a less skillful result.

6.8 Conclusions

The dependency of the local PDF for cloud on the forward-modelling of NWP data is both its strength and its weakness. The sound physical basis on which cloudy and clear observations are discriminated between should give a cloud detection scheme that is robust, providing spatially- and temporally-specific classifications. The pixel-specific PDF should be narrower and steeper than the global PDF used in previous chapters, meaning observations are more likely to be either very strongly (near the PDF peak), or very weakly, associated with cloud - leading to fewer ambiguous classifications. The way in which uncertainties are considered should allow the probabilistic result to be interpreted in terms of the confidence the user should have in the classification. In practice, the narrowness and steepness of the local cloud PDF makes the classification of cloud observations that correspond to the edge of the PDF highly sensitive to the position of the peak and to the extent of the PDF, i.e. to the value of the uncertainties used to expand the predicted observations into a PDF and to which clouds are represented. This steepness and narrowness may be exaggerated by the limited number of terms considered in the **B**-matrix - it may be more appropriate to add further terms to the b-matrix to account for uncertainties in more NWP fields. Any extra term, however, adds an extra dimension to the calculations, and this would make the algorithm, as it currently stands, too computationally expensive to be practical.

For observations of small or thin clouds, an inappropriately centred PDF can result in a very low PDF value being assigned to the observation, making it less likely to be classed as cloud. The position of the PDF peak depends on the NWP data, which is judged less reliable for land targets and may explain why lower skills scores are seen for these than for sea targets. It also depends on uncertainties and biases within the RTM, which are considered greater for the forward-modelling of cloud observations than for clear sky observations, particularly at night when the scattering approximations within the RTM may result in inaccurate simulations of cloud observations at 3.9μm. Without a 'truth' against which to compare the performance of the RTM for cloud observations, the uncertainties used here are judged to be the most appropriate available values. Quantifying uncertainties in the NWP data, particularly over land, is challenging, as it is likely to vary both spatially and temporally. If the uncertainties are assigned values that are high, however, the technique will fail to yield useful results, because a broad, flat PDF would result in all observations corresponding to similar conditional prior probabilities of cloud: $P\left(\mathbf{y^o}\left|\mathbf{x^b},\bar{c}\right.\right)$. The method can therefore only accurately detect cloud provided the NWP data uncertainties are

quantified reasonably accurately, and are not unreasonably high.

The dataset used here is comprised of ‘difficult’ targets for cloud detection, which may mean that the spectral properties of some of the cloud targets are specifically those for which the approximations made in the forward-modelling of cloud observations are not appropriate. Further work on a more representative dataset may give more reliable insights into the effect of using a local, NWP-dependent PDF for cloud.

The effect of implementing the local PDF for cloud in this study is to reduce the number of clear targets misclassified as cloud, but also to simultaneously raise the number or cloud targets misclassified as clear, and so to reduce the overall skill of the detection. This can probably be explained by the increased dependency of the technique on both NWP data and forward-modelling assumptions when the local PDF for cloud is implemented. As advances are made in NWP and the data becomes more reliable, and the uncertainties more accurately quantified, particularly over land, it is hoped that the local PDF for cloud will give a more skillful detection. Similarly, as the capability of RTMs to accurately and quickly simulate cloudy atmospheric states increases, and uncertainties become better quantified, the skill of the cloud detection with the local PDF for cloud should improve.

Chapter 7

Adding a Dust Class to the Bayesian Classifier

In previous chapters, a cloud detection scheme has been developed using a classification into clear and cloudy pixels. In the real atmosphere, however, non-cloudy skies can contain varying amounts of aerosol. If aerosol is not accurately discriminated in an image, it can be falsely classified as cloud, leading to inaccurate cloud data being recorded, or can result in erroneous retrievals of surface parameters, as the retrievals are made assuming clear sky conditions. It would therefore be useful to extend the Bayesian cloud-clear classifier to a cloud-clear-aerosol classifier. In this chapter, the cloud detection scheme is extended to treat desert dust aerosol as a third class.

Desert dust is important to models of climate and weather - studies have shown that it could account for up to 50% of naturally-occurring tropospheric aerosol (Gobbi et al., 2000), and that it could have a similar effect on the Earth's radiation budget to a greenhouse gas, with a radiative forcing effect of the order of tens of Watts per m^2 at both the surface and at the top of the atmosphere (the sign of the forcing in the two instances being opposite) (Highwood et al., 2003). The high spatial and temporal variability of dust loading in the atmosphere makes accurate representation within models difficult, a problem which is heightened by the actual distribution of dust at any given time being largely unknown (Tegen and Fung, 1994). Many studies of dust have been done using ground-based observations from sun-photometers (Holben et al., 2001), giving highly localized records of data for dust events (many studies focus specifically on dust events, such as dust storms, when most, if not all, of the observed aerosol is known to be dust (Ogunjobi et al., 2004)). For example Pinker et al. (2001) provides a detailed analysis of a dust event in the sub-Sahel region of Africa in 2000, in which ground-based observations were compared to model data and space-borne observations for the same site. The models were noted to under-estimate aerosol optical depth (AOD), while some of the space-borne observations appeared to over-predict AOD. Without a reliable estimate of the global distribution of dust at a specific time, it is difficult for the dust-scheme of any GCM to be fully tested. While some of the most reliable aerosol observations are from ground-based measurements, a global distribution can only realistically be observed from space, although such in-situ measurements are often used as ground truths in the development and validation of aerosol observation retrievals from satellite sensors, for example (North, 2002). One problem associated with the identification of dust in satellite imagery is cloud contamination, particularly from thin cirrus clouds (Pinker et al., 2001). For this reason, most dust-detection is carried out after imagery has been cloud-screened, meaning that satellite observations of dust are

often biased towards dust which appears spectrally more similar to clear sky than to cloud (Tegen and Fung, 1994). This bias could be avoided if cloud and dust were simultaneously discriminated from clear sky in an image. This chapter presents, and demonstrates, an extension of the technique outlined in chapter 2 which discriminates between clear sky, cloud and desert dust to probabilistically classify satellite-acquired imagery into observations of clear sky, cloud and desert dust.

7.1 A Bayesian Classifier for Dust, Cloud and Clear

The Bayesian cloud detection algorithm was written to calculate the NWP-dependent posterior probabilities of cloud and clear for individual pixels in a satellite image, see chapter 2. An assumption implicit in the method used to do this is that 'clear' and 'cloud' are the only possible atmospheric states. The technique calculates the probability of the observations for the pixel having been made under clear and under cloudy conditions, and combines these, weighted by the latitude- and season-dependent prior probability for each of these 2 states. The prior probabilities sum to 1, excluding the possibility of the observation corresponding to a non-clear but cloud-free state of the atmosphere, such as a dust or aerosol-contaminated atmosphere.

The calculation of the final probabilistic result follows from Bayes Theorem for Combining Conditional Probabilities, which theorem can be applied to multi-state problems, as well as to the 2-state problem described in chapter 2. Equation 7.1 is the probability of a system being in state *c*. This is the basis from which equation 2.1 in chapter 2 was derived in Merchant et al. (2005). For any system, equation 7.1 describes the probability of it being in a state *c*, given observations of the state, **y**, and a-priori information **x**. In the cloud-clear classification algorithm outlined in chapter 2, *c* is clear sky, **x** is the background data (i.e. NWP and any reference data, such as surface emissivity, albedo) and **y** is a vector of satellite observations. The term on the denominator, $P(\mathbf{y}\,|\mathbf{x})$ includes both the state c, and the state 'not c', denoted $\overline{c}$, i.e. $P(\mathbf{y}\,|\mathbf{x}) = P(\mathbf{y}\,|\mathbf{x}, c)\,P(c) + P(\mathbf{y}\,|\mathbf{x}, \overline{c})\,P(\overline{c})$. For a system with more states than c and $\overline{c}$, the probability that the true state of the system is c_i can also be found from equation 7.1, now written as equation 7.2. In the multi-state case, $P(\mathbf{y}\,|\mathbf{x})$ must include c_i and all the states that make up 'not c_i', i.e. all other states which are possible for the system, see equation 7.3.

$$P(c\,|\mathbf{y}, \mathbf{x}) = \frac{P(\mathbf{y}\,|\mathbf{x}, c)\,P(c)\,P(\mathbf{x}\,|c)}{P(\mathbf{y}\,|\mathbf{x})\,P(\mathbf{x})} \tag{7.1}$$

$$P\left(c_i\left|\mathbf{y},\mathbf{x}\right.\right)=\frac{P\left(\mathbf{y}\left|\mathbf{x},c_i\right.\right)P\left(c_i\right)P\left(\mathbf{x}\left|c_i\right.\right)}{P\left(\mathbf{y}\left|\mathbf{x}\right.\right)P\left(\mathbf{x}\right)} \tag{7.2}$$

$$P\left(\mathbf{y}\left|\mathbf{x}\right.\right)=\sum_j P\left(\mathbf{y}\left|\mathbf{x},c_j\right.\right)P\left(c_j\right) \tag{7.3}$$

If a-priori information is assumed independent of the class of the system, i.e. $P\left(\mathbf{x}\left|c_j\right.\right)=P\left(\mathbf{x}\right)$ - an assumption made in formulation of the original 2-state cloud detection calculation in Merchant et al. (2005) - then equation 7.1 becomes equation 7.4.

$$P\left(c_i\left|\mathbf{y},\mathbf{x}\right.\right)=\frac{P\left(\mathbf{y}\left|\mathbf{x},c_i\right.\right)P\left(c_i\right)}{\sum_j\left(P\left(\mathbf{y}\left|\mathbf{x},c_j\right.\right)P\left(c_j\right)\right)} \tag{7.4}$$

Equation 7.4 is an expression for calculating the probability that the system is in state c_i. The denominator is the sum of the conditional probabilities for a given observation corresponding to each possible state (including c_i), multiplied by the prior probability of that state being present. An appropriate prior probability for each possible state and a method of calculating the probabilities of an observation corresponding to each individual state are required, i.e. a prior probability and a PDF is required for each state.

The more ‘not clear’ atmospheric states that are considered by the cloud detection, the more realistic the results are likely to be - in chapters 2-5, some of the information, such as the prior probability, for the ‘not clear’ state actually corresponds to cloud and so is not appropriate to all ‘not clear’ states . This means that observations of other ‘not clear’ states may not be classified into the ‘not clear’ class, and may instead end up in the ‘clear’ class. Other ‘not clear’ states are likely to include aerosol-contaminated atmospheric states, which can often appear similar to cloud in observations. This means that such states are actually more likely to end up being classified as ‘not clear’ than clear. If ‘not clear’ is to be interpreted as ‘cloud’, which is the aim of the algorithm, then these other states need to be accounted for in a third class.

The information required to implement a third class in the algorithm is a prior probability and a PDF for aerosol. Neither of these are readily available. An aerosol of particular interest to NWP groups is Saharan dust. Work done on detecting it includes the Saharan Dust Index (SDI), which value is a function of the BTs recorded for a pixel at 3.7, 8.7, 11 and 12μm (Merchant et al., 2006a). The SDI is founded on a Principal Components Analysis (PCA) of simulated dust and clear-sky observations in BT-difference space. The second principal component was judged to correspond to the presence of dust, and so it's value, when calculated for real observations, can be treated as an indicator for dust. Cloud observations were not simulated for comparison and so the SDI can only be considered a tool for differentiation between dusty and clear sky observations, and not between dusty, clear and cloudy observations, or observations contaminated by the presence of any other aerosol. A full description of the technique is given in (Merchant et al., 2006a).

A limitation of the SDI method is that it requires observations at four wavelengths, while some sensors used in NWP only record imagery at 3.7, 11 and 12μm, and so do not have enough information to detect dust using SDI. This chapter demonstrates how this lack of spectral information can be compensated for by exploiting NWP fields, in addition to three observation channels, to detect dust.

Most methods, SDI included, require imagery to be cloud screened before any test for dust can be applied. If a cloud screening algorithm inadvertently screens dust as well as cloud, then retrievals of dust will be biased towards those cases where the dust appears spectrally more similar to clear sky than to cloud. It is also possible that clouds missed by the cloud screening process could be picked up as dust - the SDI theory was based upon comparisons of simulated cloud-free atmospheres with and without dust contamination. For some clouds, the apparent SDI is within the range associated with dust. A probabilistic one-step cloud-clear-dust classification could therefore be a useful tool, and was developed as part of this PhD project for night-time imagery over ocean.

Discrimination of dust in daytime imagery, when observations at visible wavelengths provide extra information to aid separation of dust, cloud and clear pixels, is more straightforward than for night-time scenes. This investigation therefore focused on night-time imagery, as the more challenging case, where improvement on current methods is anticipated to be most beneficial.

7.2 Prior Probability for Dust

A latitude- and season-specific prior probability LUT for dust was calculated from a dataset of 358 midnight images from the SEVIRI sensor for which SDI had been calculated - these covered nearly a full year of data, all from 2005, with only 7 days missing. The SDI data was calculated

after cloud screening had been carried out using standard operational algorithms, and it is likely that some dust was removed with the cloud screening - nonetheless, it is the best available measure on which a prior probability for dust can be based. A threshold of -0.2 was used to decide if a pixel contained dust or not. This is the threshold at which retrieved SSTs are corrected for dust contamination (MétéoFrance, 2006). The correction is very small for pixels with this lowest SDI value, and a higher threshold could be argued to be more suitable. A higher threshold, however, would reduce the number of pixels considered dust-contaminated to those likely to contain a lot of dust - thereby reducing the algorithm's ability to identify pixels containing only a small amount of dust. SDI data are only available for pixels which have been classed by an operational cloud detection algorithm as cloud-free, making it likely that some dust pixels have already been excluded, meaning that the number of dusty pixels found using the SDI data is likely to be lower than the number of truly dust-contaminated pixels. A lower threshold on the SDI would add to this synthetic reduction, and so -0.2 was judged an appropriate threshold for identifying and counting dust pixels.

The data were supplied on an equal area grid, and the number of instances in which an individual pixel held an SDI value above the threshold in a given season, divided by the number of data for that pixel for that season is shown in figure 7.1. Seasons were chosen December, January and February (DJF), March, April and May (MAM), June, July and August (JJA) and September, October and November (SON). The SDI indicates the presence of dust in the Atlantic Ocean to the South of the bulge of North-West Africa, where dust is in fact unlikely to be present. This is a common artifact in SDI imagery, for which no explanation has been found (*personal communication, C. Merchant, Nov. 2008*). It may be that dust from an unknown source is actually present in this region, or it may be that there is a spectral anomaly in the ocean surface in this region.

Loading data at this spatial resolution into the algorithm would be computationally expensive, and the prior probability should have only a very small effect on the end result for any given pixel. The difference between PDF values for the different classes should be orders of magnitude greater than the difference between the prior probabilities - only in cases where the background information, $\mathbf{x}$, is associated with particularly high uncertainty, resulting a flat distribution from which $P(\mathbf{y}\,|\mathbf{x}, c_i)$ is read, is it anticipated that the prior will have a significant effect. The data in figure 7.1 were therefore grouped into boxes of 10^o latitude x 10^o longitude, considered a large enough scale to capture an appropriate level of detail - figure 7.2 shows this grid superimposed on the JJA SDI data.

The mean from each of the higher resolution grid cells was used to fill the new grid cells. These values can be interpreted as the average number of dusty days observed in a non-cloudy sky in each season in each 10^ox10^o cell. After this upscaling, the anomalous dust suggested by the SDI over the Atlantic is no longer significant. This is preferable to removing it from the calculation for the prior, as it is possible, although unlikely, that it does actually correspond to

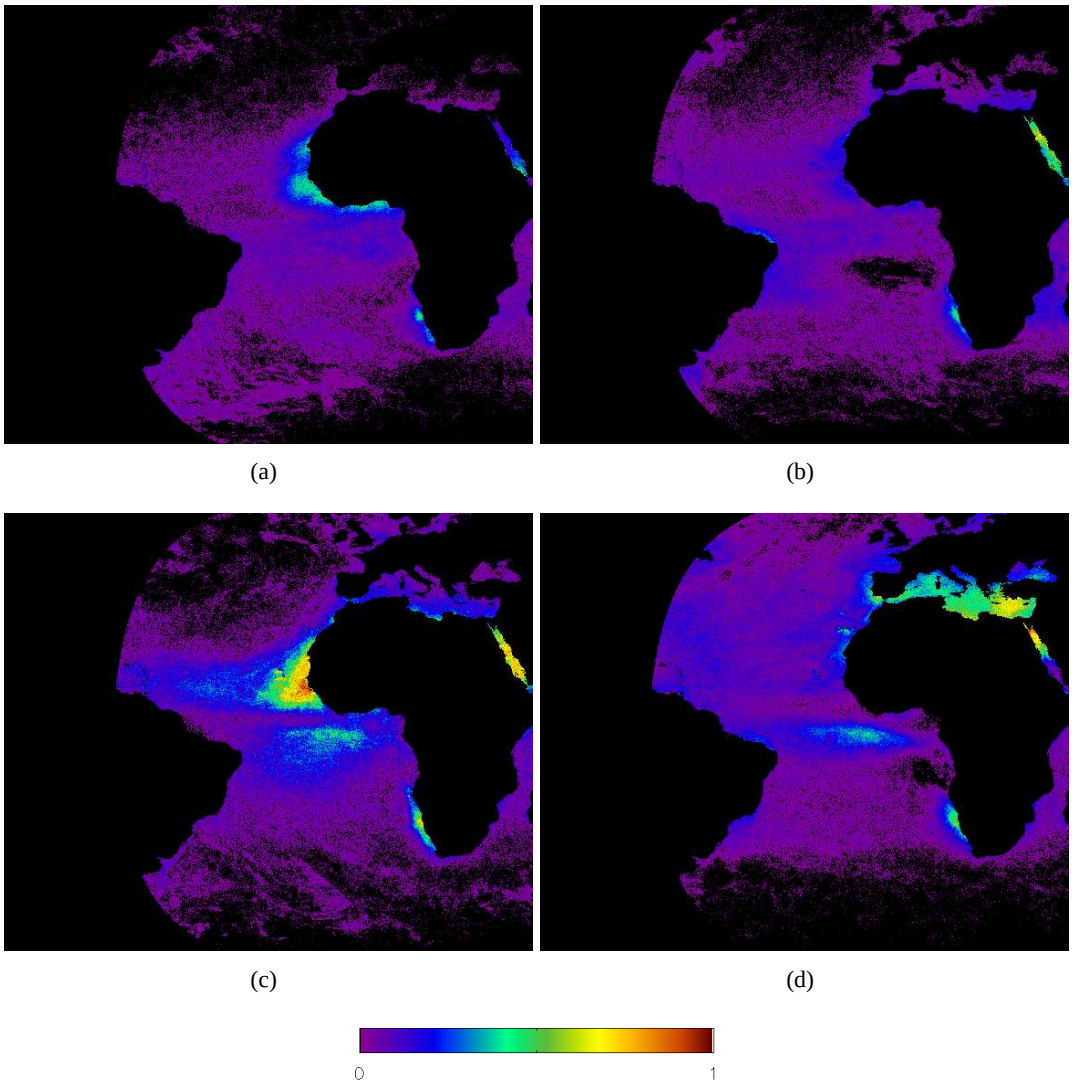


Figure 7.1: Fraction of images from a dataset of 358 midnight images acquired by the SEVIRI sensor in 2005 containing dust for each season, calculated by setting a threshold of -0.2 on the SDI. The colourscale ranges from 0-1, and seasons correspond to (a) DJF, (b) SON, (c) MAM, (d) JJA

dust. The upscaled maps are used as prior probability for dust, see figure 7.3.

The dataset covers the region -60º to +60º latitude, -100º to +45º longitude. Outside this region, the prior probability of dust is set to the minimum value for inside the region for the appropriate season. The values found in this way for the prior probability for dust are within the range of figures found by other studies which have counted the number of 'dusty days' in the region, for example Jankowiak and Tanré (1992). The boundaries of the region and the differences seen between the seasons also agree roughly with other studies (except for the anomalous region in the Atlantic mentioned earlier) for example Jankowiak and Tanré (1992); Chiapello and Moulin (2002).

The prior probability for cloud,$P\left(c_{cloud}\right)$, is taken from the ISCCP total cloud amount

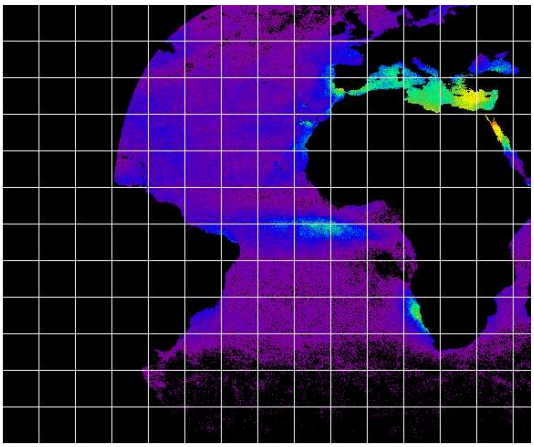


Figure 7.2: Grid used for prior probability of dust.

product, and the prior probability for clear is currently set to be $1 - P\left(c_{cloud}\right)$. This is because the only states allowed by the calculation are clear and cloud - that assumption can now be altered to build a technique assuming clear, cloud and desert dust to be the only possible atmospheric states. The SDI data on which the prior probability for dust is based was calculated after cloud pixels had been removed, and so the prior probability for dust from the LUT is multiplied by the prior probability for clear, $1 - P\left(c_{cloud}\right)$, to give the prior probability for dust used in the algorithm. The prior probability for clear is $1 - \sum_{j \neq clear} P\left(c_j\right)$.

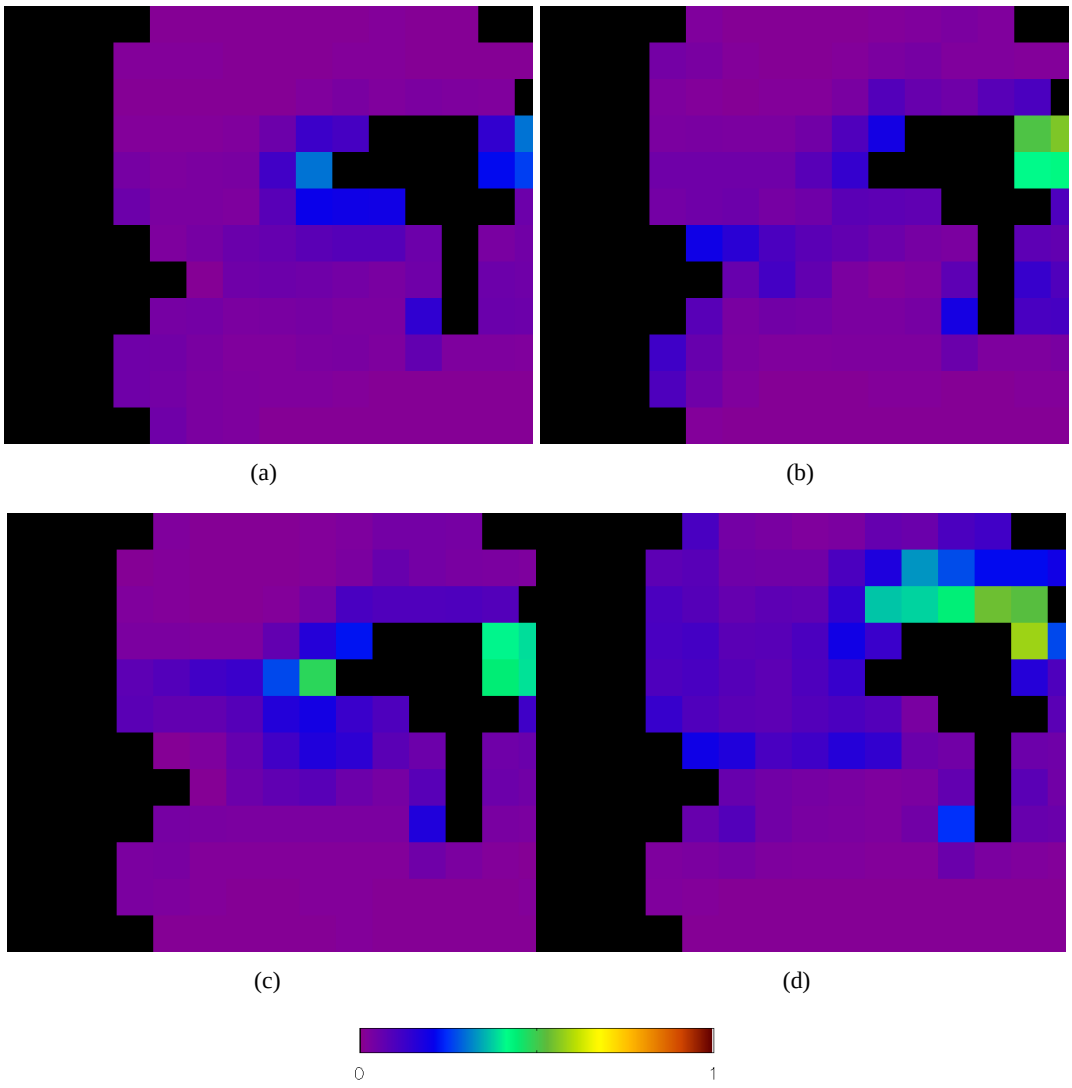
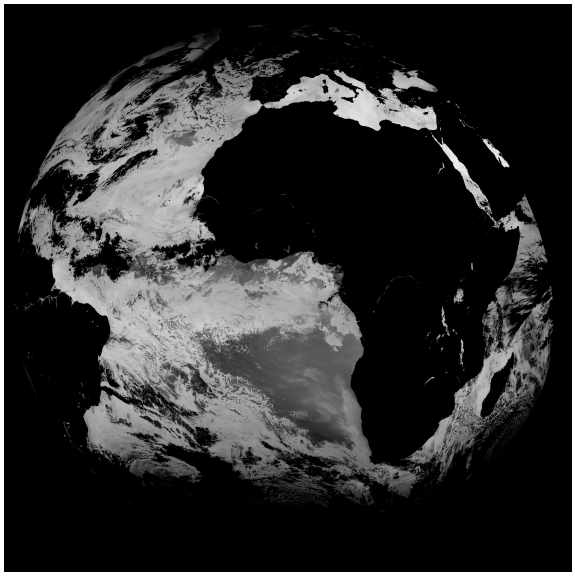


Figure 7.3: Seasonal maps of the prior probability for dust. Seasons correspond to (a) DJF, (b) SON, (c) MAM, (d) JJA

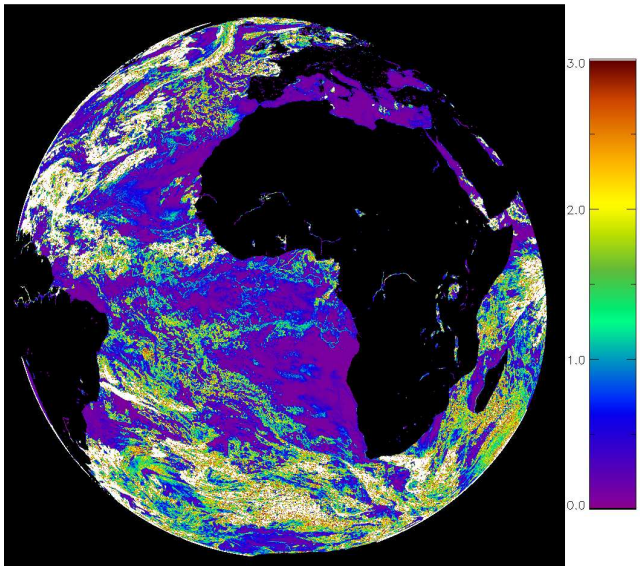
7.3 Calculating a Textural PDF for Dust

Although dust often appears spectrally similar to cloud, it can generally be associated with smaller horizontal temperature gradients than cloud. This suggests that, just as texture is considered helpful for discrimination between observations of cloud and of open ocean (see section 2.4.1 in chapter 2), texture should aid discrimination between observations of dust and of cloud. The local standard deviation of the observations at 11μm (over 3x3 pixels) was calculated for an image acquired by the SEVIRI sensor on 30$^{\text{th}}$ July 2005 at 2am UTC, see figure 7.4. The highest texture for dust was judged through visual inspection to be 1.15, and this was used as a threshold to mask out cloud pixels, see figure 7.5. The PDF for dust is assumed to have the general form $PDF = ax\exp^{-bx}$, where a and b are constants, and must satisfy the conditions that it's definite integral is 1, i.e $\int_0^\infty PDFdx = 1$, and that a texture of 1.15 corresponds to the 95$^{\text{th}}$ percentile, i.e. $\int_0^{1.15} PDF\delta x = 0.95$. Solving this problem for a and b results in a=17.016, and b=4.125, giving the expression for the textural PDF in equation 7.5, which is plotted as a function in figure 7.6. The shape of this PDF between the LSD limits of 0 and 1.15 is somewhat arbitrary, there not being enough information to justify any particular form. It is, however, supported by the histogram in figure 7.7, which shows the distribution of LSDs for pixels with an SDI greater than 0.2, and a LSD at 11μm of less than 1.15. The SDI threshold of 0.2 was set through consultation with those experienced in SDI processing *(personal communication, C. Merchant, Nov. 2008)*, and could be argued to be somewhat arbitrary. For this reason, and the fact that only one image contributes to the distribution, figure 7.7 cannot be taken as conclusive evidence for the shape of the textural PDF for dust, but it does support the shape shown in figure 7.6. The texture at 11μm is calculated for each pixel in an image and used to calculate the textural PDF for dust. The textural component of the probability that observations correspond to dust, for the image in figure 7.4, is shown in figure 7.8, alongside the textural components of the probability for cloud and clear. The range of textures associated with clear and dusty states is smaller than that associated with cloud, which includes very high textures. The textural PDF for cloud therefore has a broader peak than that for dust or clear and, since all three must integrate to 1, the broader peak must necessarily be lower than the peaks in the other PDFs. This means higher textural PDF values can be expected for dust and clear than for cloud, as illustrated in figure 7.8. Dust observations are anticipated to be spectrally distinct from clear observations, and so should be distinguished by the spectral component of the PDF. The main function of the textural component of the PDF is to help distinguish between cloud and dust observations. Although these can appear spectrally similar, figure 7.8 shows them to be separable on the grounds of texture - only the small fraction of clouds that are both spectrally and texturally similar to dust will appear as dust. The textural PDF for dust calculated in this way is multiplied by the spectral PDF described in section 7.4 to give $P\left(\mathbf{y^o}\left|\mathbf{x^b}, c_{dust}\right.\right) = P\left(\mathbf{y_t^o}\left|c_{dust}\right.\right) \times P\left(\mathbf{y_s^o}\left|\mathbf{x^b}, c_{dust}\right.\right)$.

$$P\left(\mathbf{y_t^o}\,|c_{dust}\right) = 17.016 \times \exp^{-4.125 \times texture} \tag{7.5}$$

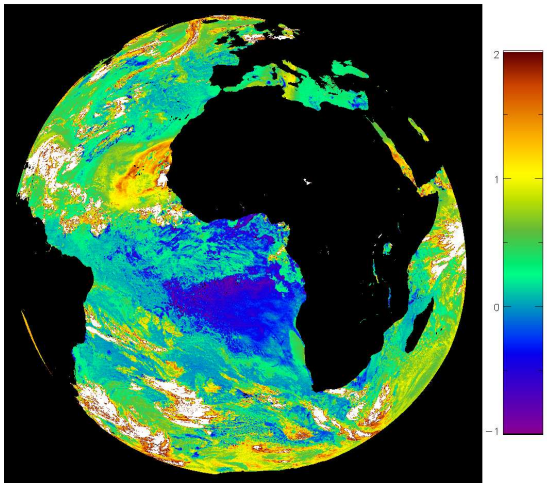


(a)

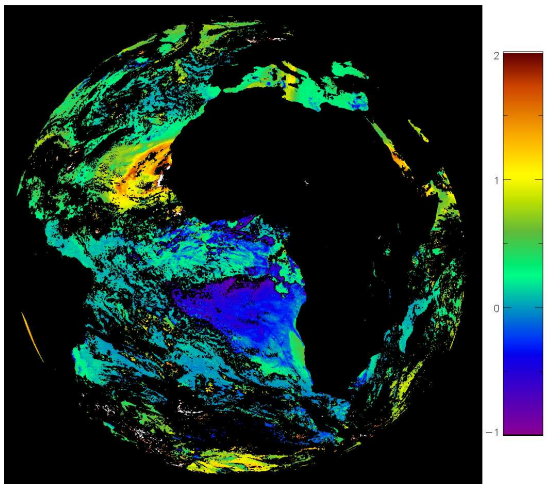


(b)

Figure 7.4: Image used for calculating textural PDF for dust, recorded by the SEVIRI sensor at 2am UTC on July 30th 2005. (a) Image recorded at 11μm, (b) texture calculated over 3x3 pixels, the colourscale corresponds to a range of 0-3K.



(a)



(b)

Figure 7.5: (a) SDI calculated for the case study image; (b) the SDI with a texture mask (constructed from 11μm observations) applied. The case study image was recorded by the SEVIRI sensor on July 30th 2005 at 2am UTC.

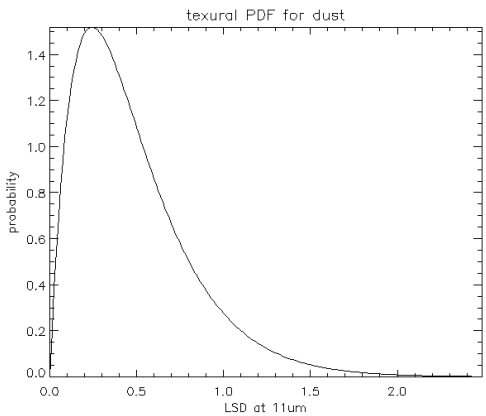


Figure 7.6: Textural PDF for dust.

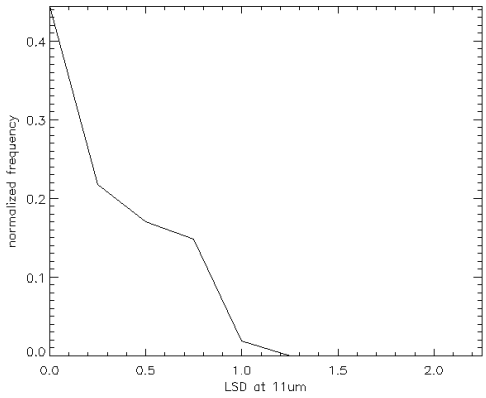
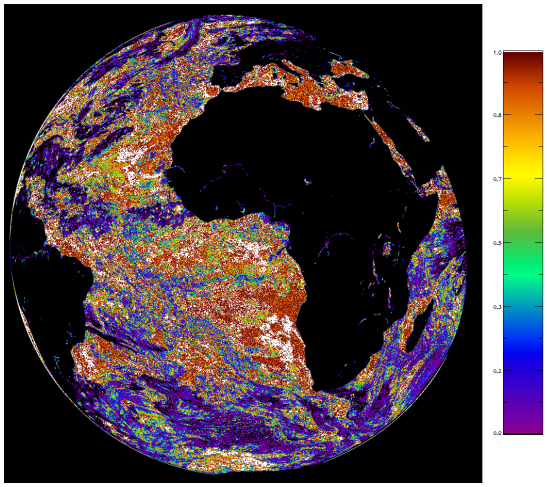
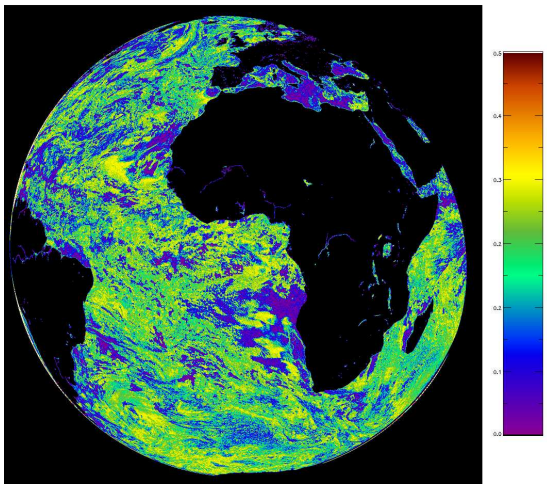


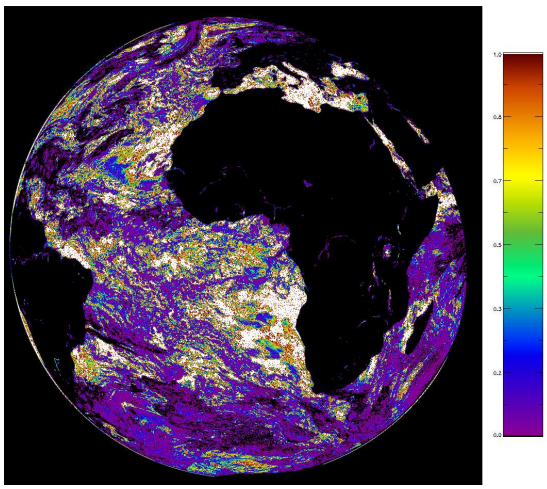
Figure 7.7: Support for the form of the textural PDF for dust. The histogram shows the 11μmLSDs calculated for pixels in the case study image with an SDI greater than 0.2, and a 11μm LSD of less than 1.15.



(a)



(b)



(c)

Figure 7.8: Textural component of the probability that the observations correspond to (a) dust, (b) cloud and (c) clear. The colourscale corresponds to a range of 0.0-1.0 for dust and clear, and a range of 0.0-0.5 for cloud. The case study image for which these probabilities were calculated was recorded by the SEVIRI sensor at 2am UTC on July 30th 2005.

7.4 Calculation of a Spectral PDF for Dust

This section describes a method to achieve a probability density function for spectral observations of dust. It is based on all plausible simulated dust simulations in a dataset described in section 7.4.1, with some dependency on the surface temperature and total column water vapour expected from NWP for the pixel. Section 7.4.2 explains how uncertainties in the NWP ST and TCWV are factored into the calculation of the PDF value corresponding to the pixel observation.

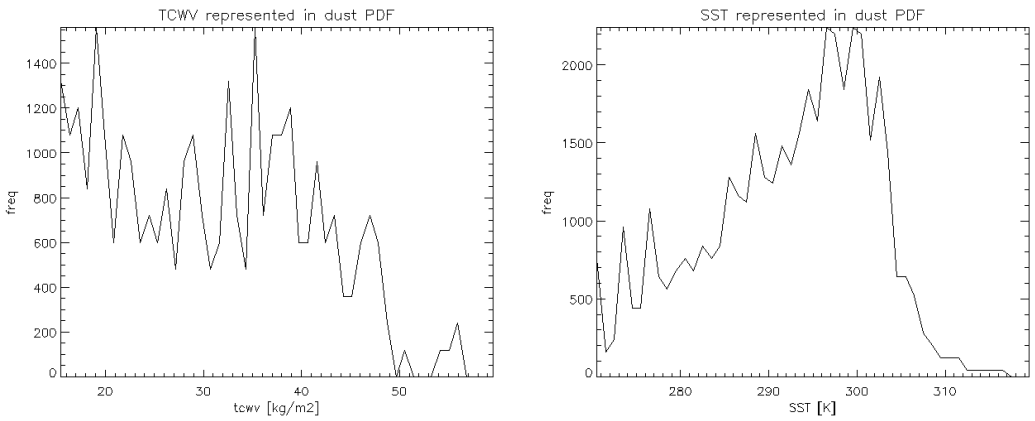
7.4.1 Data for Calculating a Spectral PDF for Dust

The SDI technique for discriminating dust-contaminated from clear pixels in an image was based upon simulations of observations at thermal wavelengths (the objective was to aid SST retrieval, and the development of the technique involved modifying a radiative transfer model (RTM) which, at the time of that work, was not readily available for visible wavelengths). A global set of atmospheric profiles over ocean, slightly biased towards mid-latitudes (Francoise et al., 2002), was used for the work. Distributions of Brightness Temperature (BT) observations were simulated from the profiles for the SEVIRI sensor with satellite zenith angles between 0^o and 75^o with layers of desert dust 1km thick, positioned at 0, 2, 3 and 4km above the Earth's surface, with aerosol optical depths (AODs) between 0.1 and 1.0 at 10μm. The same atmospheric profiles were also forward-modelled without the dust contamination, and it was through comparison of the 2 simulated datasets that the SDI measure for dust was formulated (Merchant et al., 2006a). The SDI measurement is based on a principal components analysis of ***only*** the dust-free simulations. The simulated dust BTs are therefore independent of the information behind the methodology used to calculate the SDI data (which is used for the prior probability of dust) and therefore independence has been preserved. Dust observations simulated from the Haywood set of dust profiles (Highwood et al., 2003) [1], used by Merchant et al. (2006a), form the basis of the PDF used for dust observations, from which $P\left(\mathbf{y}\,|\mathbf{x}, c_{dust}\right)$ is read.

Rather than constructing a density function in BT-space from all the simulations, which would create a PDF which was not NWP-dependent, analogous the global PDF of cloud observations in chapter 2, some dependency on the profile ST and TCWV is assumed. The simulations cover an appropriately wide range of water vapour amounts, and surface temperatures, see figure 7.9.

For each pixel, the pixel ST and TCWV (interpolated from the nearest 4 ocean NWP profiles) are used to select simulated dust observations from atmospheric profiles with TCWV and ST within twice the error ascribed to ST and TCWV (ST and TCWV errors are read from the **B**-matrix - the same NWP errors are assumed throughout the cloud detection method, see chapter 2). In cases where the pixel ST is below the minimum ST represented in the distribution of simulated dust BTs, an ST of 270K is used to select simulated observations rather than the pixel's

[1] Haywood and OPAC datasets were used to define the properties of the dust added to the profiles - the Haywood dataset appeared to give more reliable results, and so was used in derivation of the SDI, and so was also used from this work.



(a) (b)

Figure 7.9: Histograms of atmospheric variables simulated to model a distribution of dusty observations: (a) atmospheric total column water vapour (TCWV), and (b) sea surface temperature (SST).

own ST, similarly, if the pixel TCWV is below 1.6gm^{-2}, then 1.6gm^{-2} is used for selection. The simulated observations form the basis of the spectral PDF for dust.
The simulated BTs are seen to vary almost linearly with profile ST - a relationship that is more pronounced at lower AODS as expected, see figure 7.10. Plots of simulated BT against TCWV show more scatter, but a relationship can still be seen. The extent of the scatter appears to be independent of the AOD, see figure 7.11.

Simulated BTs were seen to depend linearly on atmospheric path length, with a gradient which is dependent on AOD. At each simulated AOD, the mean gradient over all the profiles was taken as the AOD-specific path length dependency for the simulated observations, see figure 7.12. This was initially used as a scaling factor, meaning only the BTs simulated for a satellite zenith angle of 0^o were needed to calculate the PDF for dust. Limiting the volume of information read into the algorithm ensures that it runs quickly enough to be practical at this development stage, if not fast enough for operational applications. Adjusting the simulated BTs in this way to predict observations for dust appropriate to the path length for each observed pixel makes the PDF more pixel-specific, and so more appropriate. The BTs resulting from this adjustment, however, are generally colder than pixel observations of dust, which results in calculation of a lower probability of these observations corresponding to dust, which is likely to reflect a bias, either in the modelling behind the dust simulations, or in the finite selection of atmospheric profiles to which dust was added to perform the simulations. To make a decision on the appropriateness of the path-length adjustment, an image from the SEVIRI sensor was selected, in which dust could be seen through calculation of the SDI. The classification for the case study in section 7.5 was carried out using a path-length dependent PDF for dust, and a PDF independent of path length.

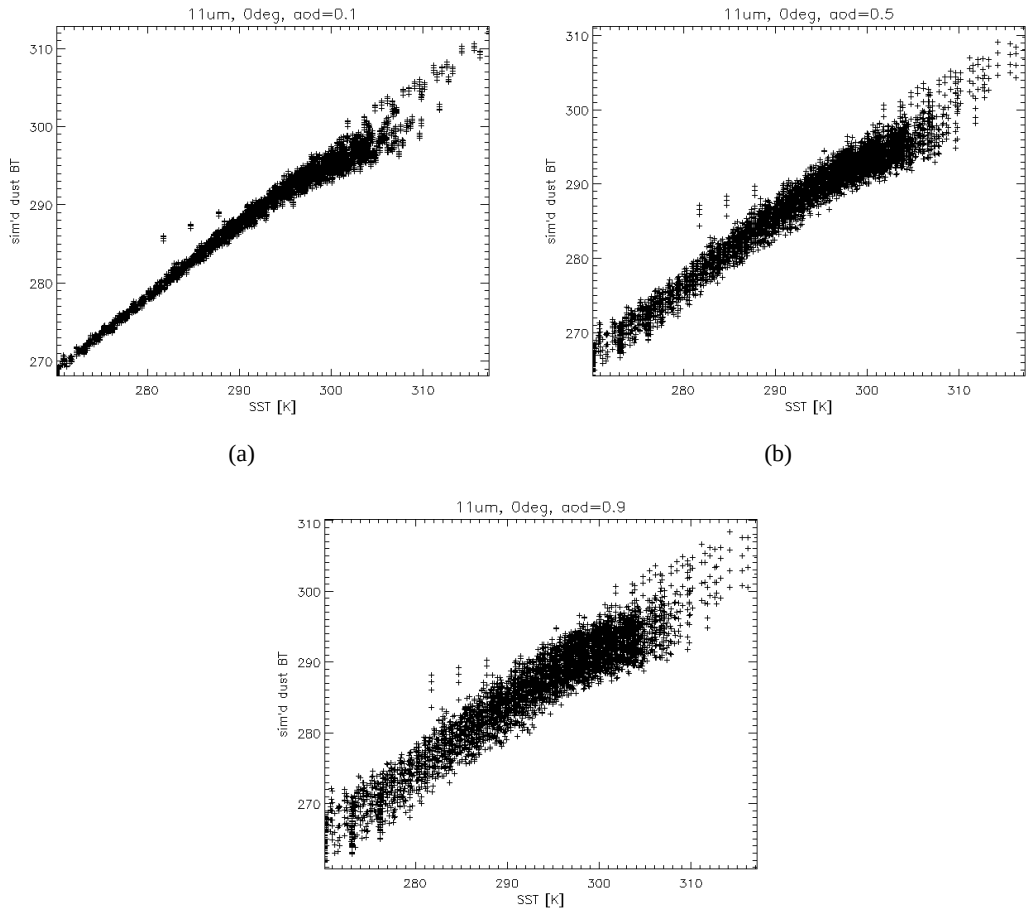


Figure 7.10: Relationship between simulated BTs and profile ST (labeled SST for sea surface temperature) for a satellite zenith angle of 0°, for (a) AOD=0.1, (b) AOD=0.5, and (c) AOD=0.9. Plots for other wavelengths and satellite zenith angles look very similar.

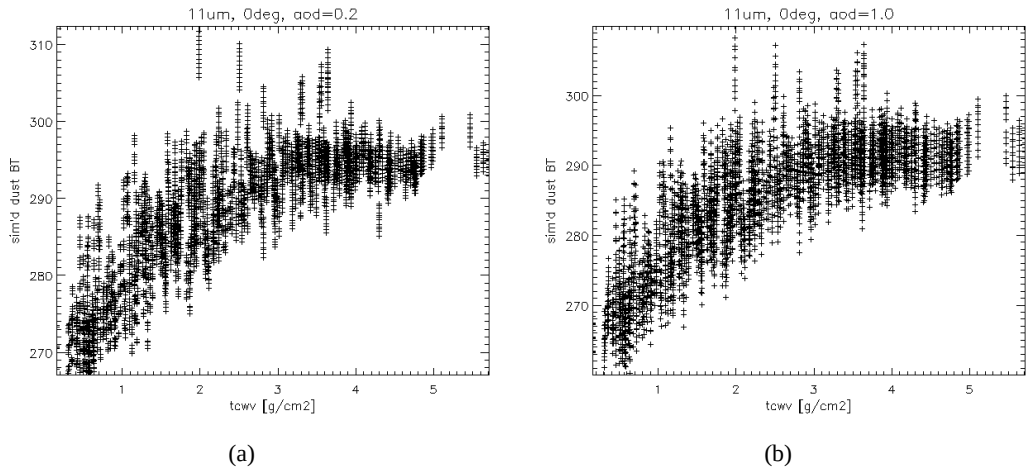


Figure 7.11: Relationship between simulated BTs and profile TCWV for a satellite zenith angle of 0°, for (a) AOD=0.2, (b) AOD=1.0. Plots for other wavelengths and satellite zenith angles look very similar.

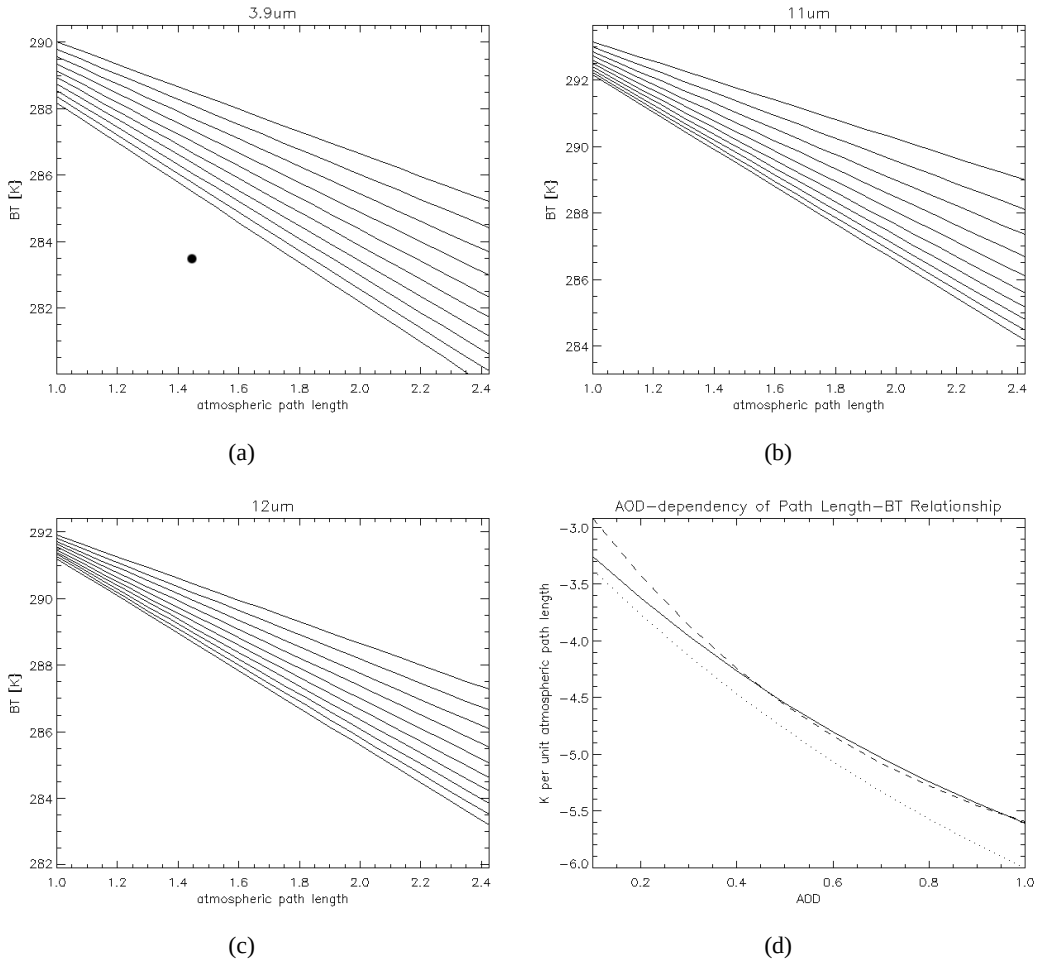


Figure 7.12: The dependency of simulated observations for dust on atmospheric path length. The lines show simulations made for one atmospheric profile with dust added at different different AODs for (a) 3.9μm, (b) 11μm and (c) 12μm. The AOD-specific mean dependency of the simulated BTs on atmospheric path length is shown in (d), where the solid line represents 12μm, the dashed line 11mum, and the dotted line 3.9μm.

7.4.2 The Spectral PDF for Dust Calculation

All the simulations of dust BTs for a satellite zenith angle of 0^o are read into the Bayesian cloud detection scheme once each time it is run. If none of these selected dust BTs are within +/-5K of the observation for the pixel, the probability of the observation corresponding to dust is set to 0. Otherwise, the PDF value for the observation ($P(\mathbf{y}\,|\mathbf{x}, c_{dust})$) is the sum of equation 7.6 calculated for each of the selected simulations. This is equivalent to equation 2.3 in section 2.4.2, in chapter 2, page 17 and the $\mathbf{H}'$, $\mathbf{B}$ and $\mathbf{R}$ matrices are defined there. The pixel ST and TCWV constitute the elements of $\mathbf{x}$; the simulated observation is $\mathbf{y^b}$ and the actual observation is $\mathbf{y^o}$).

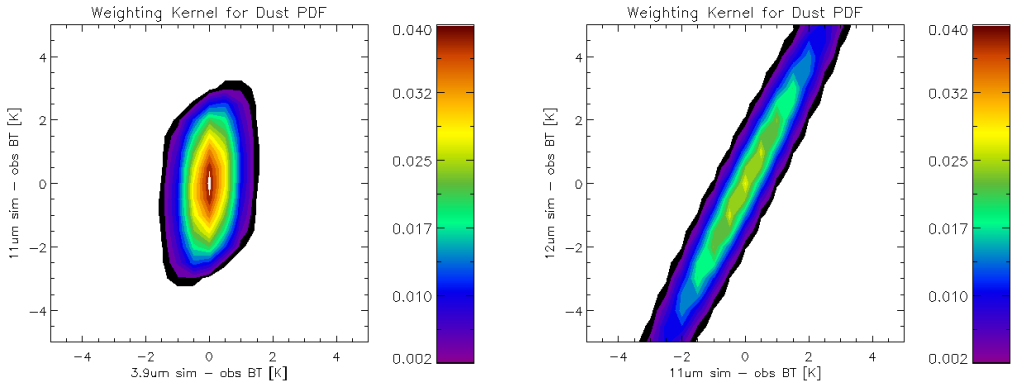
$$P(\mathbf{y_s^o}\,|\mathbf{x}, c_{dust}) = \frac{\exp\left[\frac{-1}{2}\left(\mathbf{y^o}-\mathbf{y^b}\right)^T\left(\mathbf{H'^T B H'}+\mathbf{R}\right)^{-1}\left(\mathbf{y^o}-\mathbf{y^b}\right)\right]}{(2\pi)^{\frac{n}{2}}\left|\mathbf{H'^T B H'}+\mathbf{R}\right|^{\frac{1}{2}}} \tag{7.6}$$

The same uncertainties are assumed for the dust simulations, as for the clear sky forward modelling, because the dust BTs were simulated using an adaptation of the same fast radiative transfer model as used to simulate clear sky radiances, see section 2.4.2 in chapter 2 and 7.4. The PDF is normalized by dividing this number by the total number of simulated observations within ST- and TCWV- range of the pixel. For computational efficiency, equation 7.6 is only calculated for simulations within +/-5K of the actual observation, the contributions for other simulations being assumed to be so close to 0 as to be negligible (such simulations still contribute to the normalization). This means that all simulations within +/-5K of the observation can add to the probability that the observation corresponds to dust.

If the number of selected simulations within +/-5K of the observation is greater than 21^3, then it becomes more efficient to pre-calculate a normalized weighting array of 21x21x21 elements in BT-space. The central element holds the value of equation 7.6 for $\mathbf{y^o}-\mathbf{y^b}=0$. Each dimension corresponds to possible values of $\mathbf{y^o}-\mathbf{y^b}$ for a particular channel, in increments of 0.5K. The elements hold the corresponding values of equation 7.6. The elements of the array corresponding to $\mathbf{y^o}-\mathbf{y^b}$ for each simulated dust observation $\mathbf{y^b}$ are added to form the PDF. The PDF is then normalized as before, through division by the number of selected simulations. An example of the weighting function is given in figure 7.13.

Although the former method is almost always carried out (it being extremely rare that more than 21^3 simulations are within the required ST-, TCWV- and BT- range of the pixel), and is arguably more appropriate (as it avoids forcing the observation-simulation difference into 0.5K bins), it is necessary to have both methods available in the algorithm, as the biggest challenge to

including a third class in this way is to do it in a manner that is computationally efficient enough for development work to be practical.



(a) (b)

Figure 7.13: Example of a 3-dimensional weighting function used to create the PDF for dust. The plots are summed over the dimension corresponding to: (a) 12μm, and (b) 3.9μm.

Weighting the Representation of Dust Observations for Higher AODs

In order for the PDF for dust to realistically represent all possible observations of dust, it may be appropriate to weight the representation of different dust AODs with their relative likelihood. This adjustment should be small, as spatial and temporal variability for the likelihood of the presence of dust are already accounted for in the prior probability for dust, $P\left(c_{dust}\right)$. Only very high AODs, not observed with high frequency *anywhere* (even in regions of high dust loading), should therefore be represented with less weight than other AODs.

Ground-observation data from a site in Banizoumbou in Niger, in the Sahel region of Africa (Holben et al., 2001) are looked at. These data correlate well with AOD retrieved from the TOMS satellite (Torres et al., 2002), and agree broadly with AOD measurements from other areas where high dust loading is expected, e.g. Ogunjobi et al. (2004), and with AVHRR-retrieved data (Tegen and Fung, 1994). Adjusting the representation of observations for dust with high AOD according to a measured AOD-frequency distribution for a dusty area such as this would only affect the representation of very high AODs in the PDF for dust. The frequency of observations of AOD at 0.5μm falls to below 5% for AOD $\geq$0.65, but does not reach 0 until AOD $\geq$ 2.85. To represent this, an exponentialy decreasing weight can be applied to the representation of dust with AOD(0.5μm) $\geq$0.65, reaching approximately 0 at AOD(0.5μm)=2.85.

The simulations used for the PDF were carried out using AOD values ranging from 0.1 to 1.0 at 10μm, which corresponds to an AOD range of 0.19 to 2.0 at 0.5μm (Merchant et al., 2006a). The simulated observations vary linearly with AOD, see figure 7.14, and so could be interpolated to include representation of higher AOD observations, weighted to be decreasingly likely as AOD

increases.

AOD values greater than 2.0 at 0.5μm, however, are rarely observed and some models, for example Tegen and Fung (1994), consider a maximum AOD much lower than this. It could therefore be argued to be appropriate *not* to explicitly represent observations for AOD $\geq$2.0 at 0.5μm, but to slightly *increase* the weight with which observations of dust at AOD(0.5μm)=2.0 are represented. This argument counters the argument for decreasing the weighting of higher AOD observations, and so in conclusion, no weight was applied to the representation of observations of dust with different AODs.

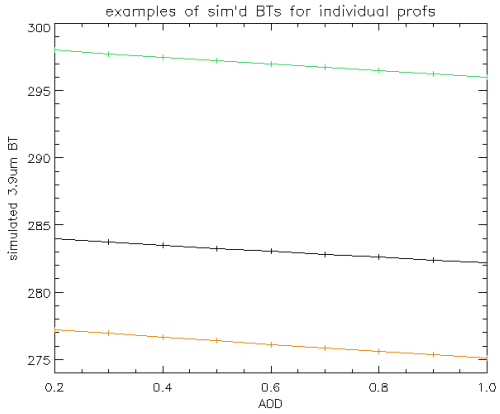
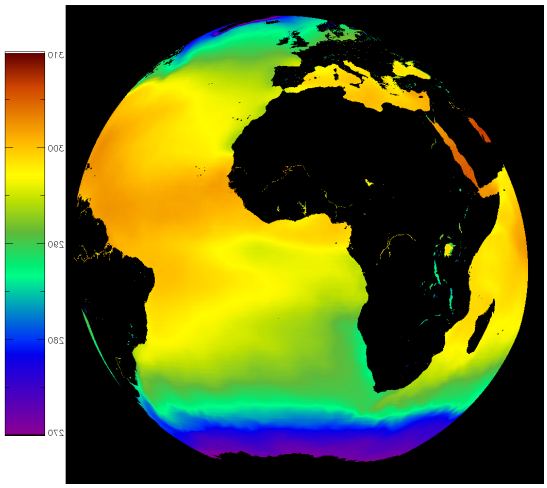


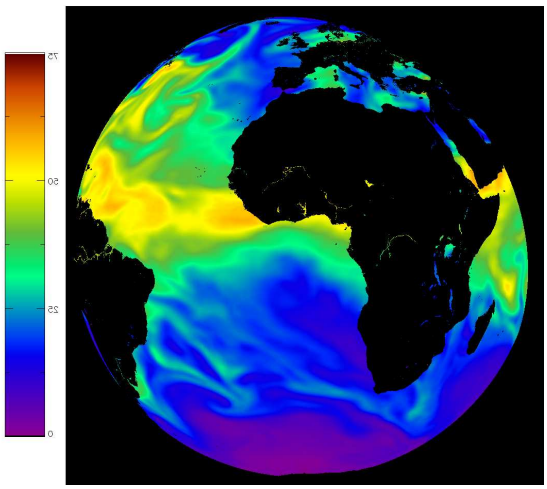
Figure 7.14: The dependence of simulated observations on dust AOD at 10 μm for 3 atmospheric profiles, shown in different colours.

7.5 Case Study

The image used in formulation of the textural PDF for dust, see figure 7.4 is used as a case study to investigate the performance of the clear-cloud-dust classifier, with and without a path-length dependency for the spectral PDF for dust. The NWP ST and TCWV fields, interpolated to every pixel in this image and used in selection of simulated dust observations for the spectral PDF for dust, are shown in figure 7.15. The spectral PDF for dust calculated with and without the path length dependency is shown in figure 7.16. The resulting probability-of-clear, -of-cloud and of-dust images are shown in figures 7.17 and 7.18. Pixels are assigned to the class for which their calculated posterior probability is highest, creating the images in figure 7.19.



(a)



(b)

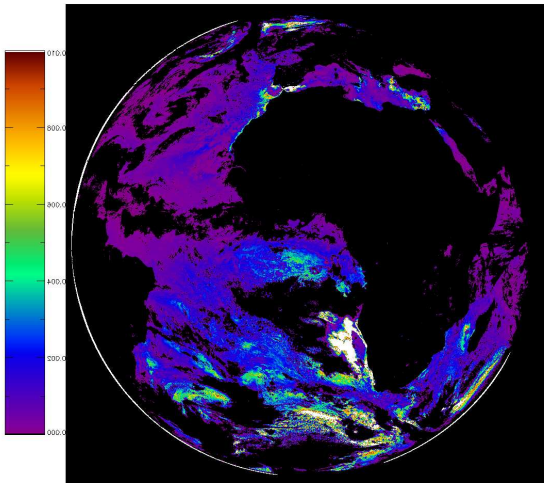
Figure 7.15: NWP ST and TCWV fields interpolated to pixels for the case study: (a) ST, colourscale corresponds to a range of 270-310K, and (b) TCWV, colourscale corresponds to a range of 0-75kgm^2.

The SDI images in figure 7.5 should be used as a reference to assess the skill with which dust is identified in the classification algorithm. The algorithm was also run as a 2-way classifier, as described in chapter 2, and the calculated posterior probability of cloud is shown

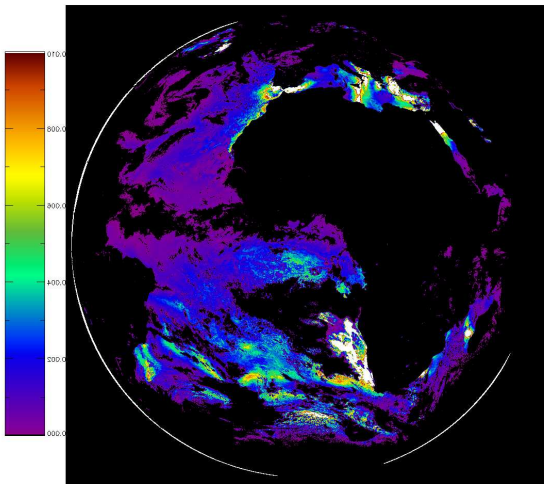
in figure 7.20 (the probability for clear in the 2-class case is 1- the probability calculated for cloud).

Assessing the accuracy of the 3-class results is problematic: the SDI data can be used comparatively to indicate the 'true' presence of dust, but SDI is only designed to identify dust in imagery after cloud screening. Similarly, the results of the 2-class probability of clear are calculated assuming clear and cloud to be the only possible atmospheric states, and so must necessarily assign every dust pixel to one of these classes. In other words, the SDI can discriminate between clear and dust, assuming the absence of cloud, while the 2-class results can discriminate between cloud and clear, assuming the absence of dust. A qualitative assessment can, however, be made through examination of the classification results with reference to both these quantities.

Histograms of both the SDI and the 2-class probability of clear for all 3 classes are shown in figure 7.21 for the classification using the path length-dependent spectral PDF for dust, and in figure 7.22 using a spectral PDF independent of path length. An indication of the certainty with which each pixel is classified by the algorithm can be taken from figure 7.23, which, for pixels classed as belonging to each separate class, shows a distribution of the posterior probability of the pixels' membership to that class.

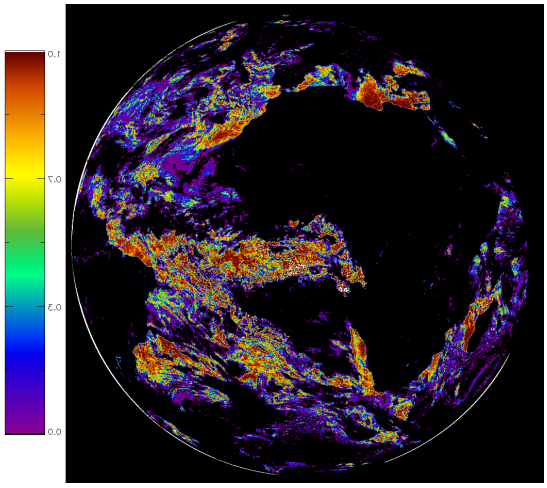


(a)

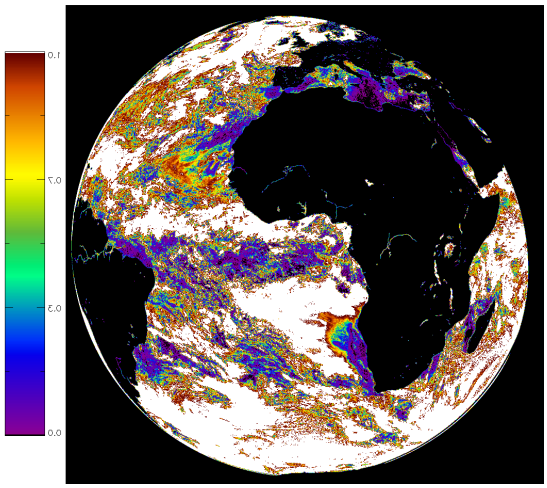


(b)

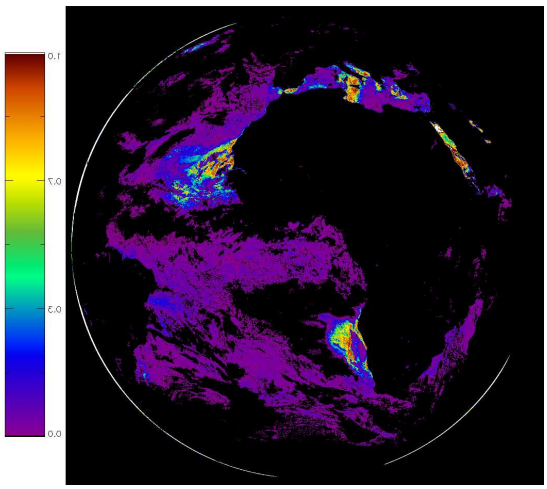
Figure 7.16: Spectral PDF for dust calculated for the case study image: (a) with a path length-dependency, and (b) without any path length-dependency. The colourscale corresponds to a range of 0-0.01.The case study image was recorded by the SEVIRI sensor at 2am UTC on July 30th 2005.



(a)

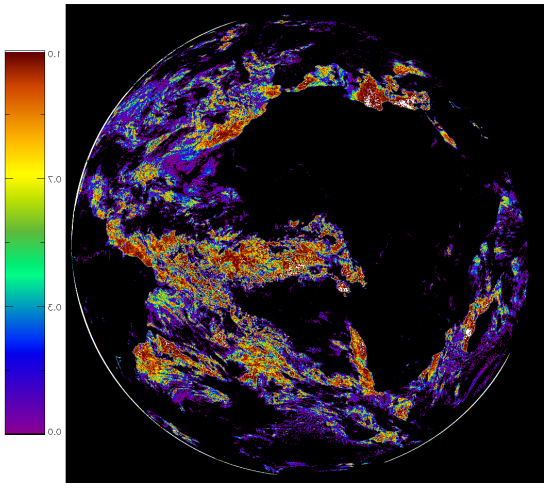


(b)

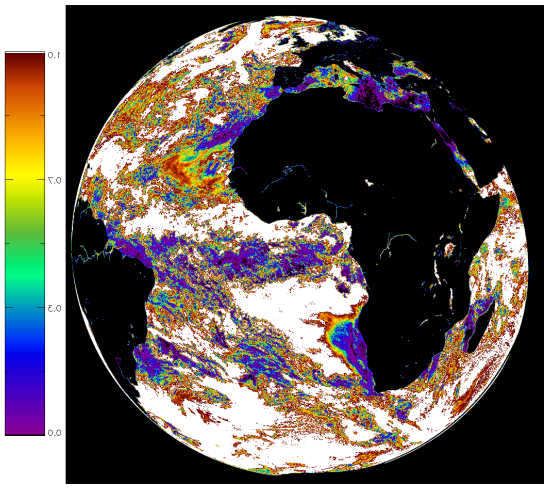


(c)

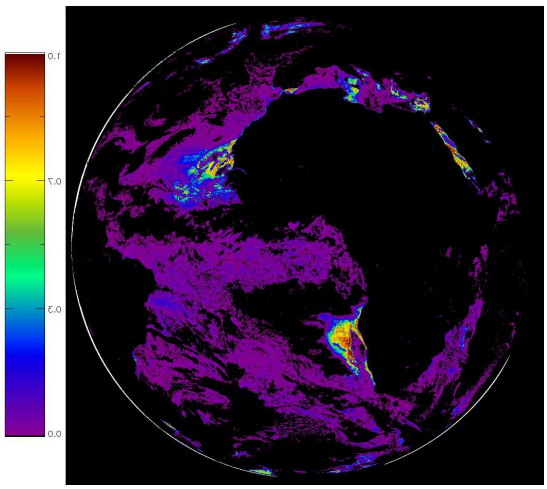
Figure 7.17: The calculated posterior probability of (a) clear, (b) cloud and (c) dust with no dependence on atmospheric path length for the spectral dust PDF. The colourscale corresponds to a range of 0-1.



(a)

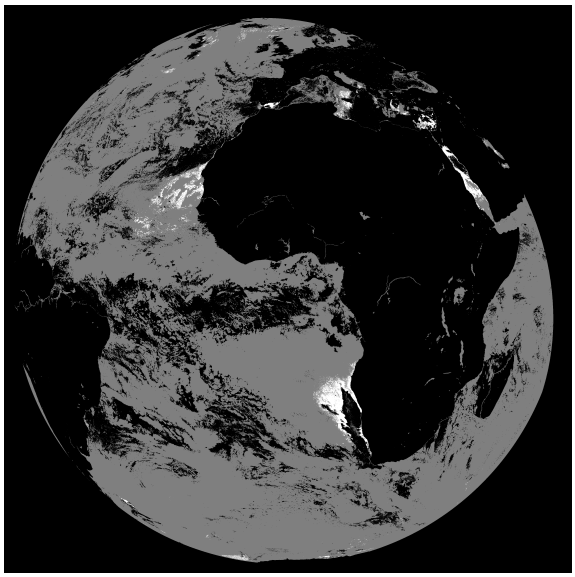


(b)

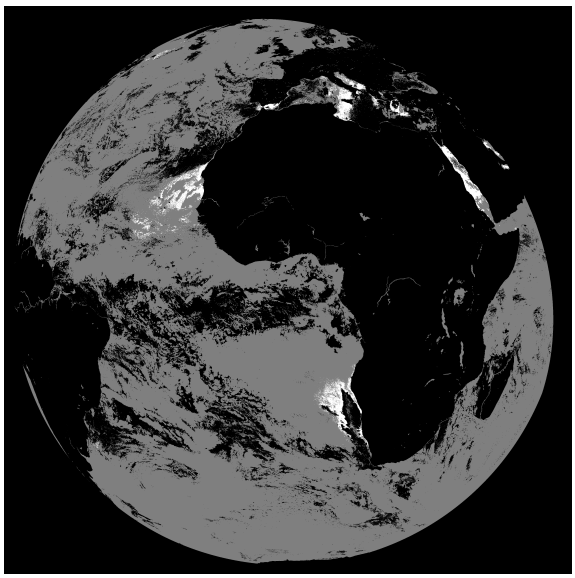


(c)

Figure 7.18: The calculated posterior probability of (a) clear, (b) cloud and (c) dust with a path length dependent spectral PDF for dust. The colourscale corresponds to a range of 0-1.

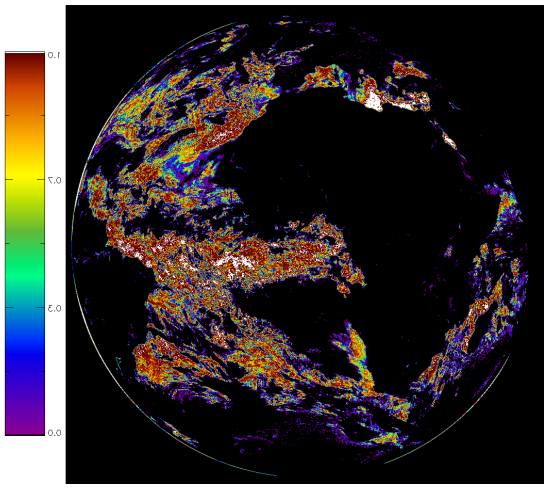


(a)

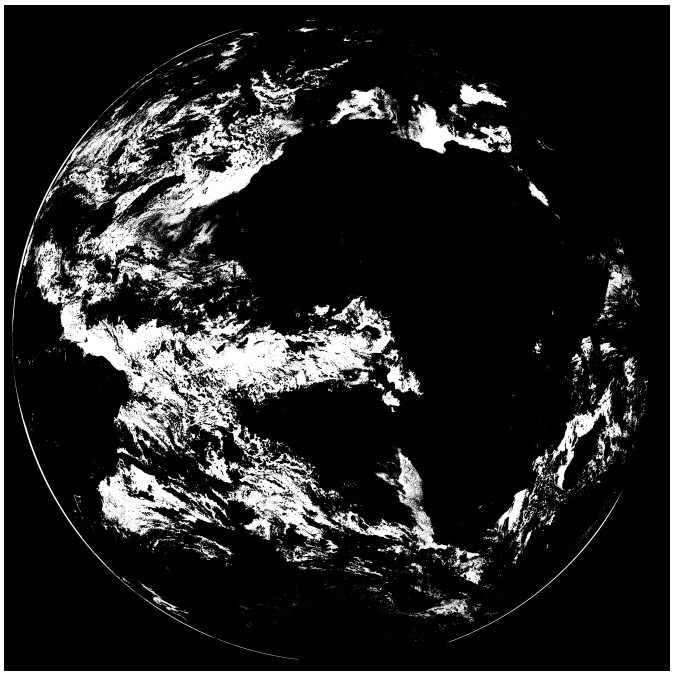


(b)

Figure 7.19: Classification results for case study image recorded by the SEVIRI sensor at 2am UTC on July 30th 2005: (a) using a path length-dependent spectral PDF for dust, and (b) using a spectral PDF for dust with no path length-dependency. Clear is black, cloud is grey and dust is white.

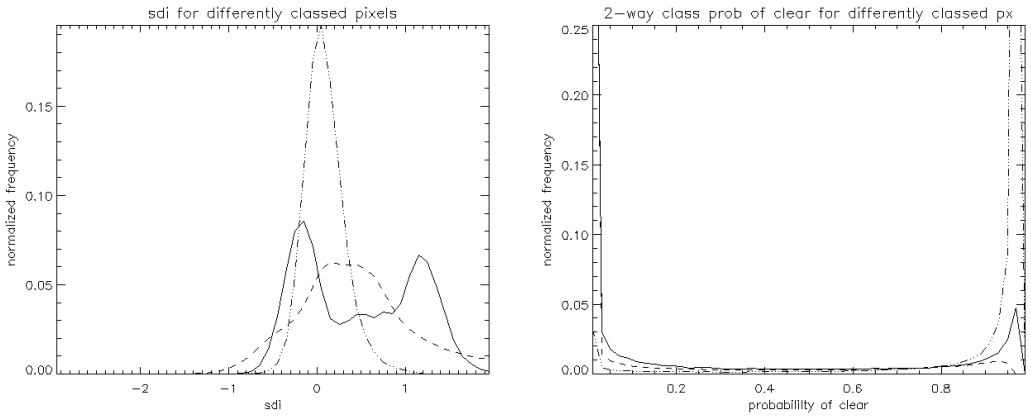


(a)



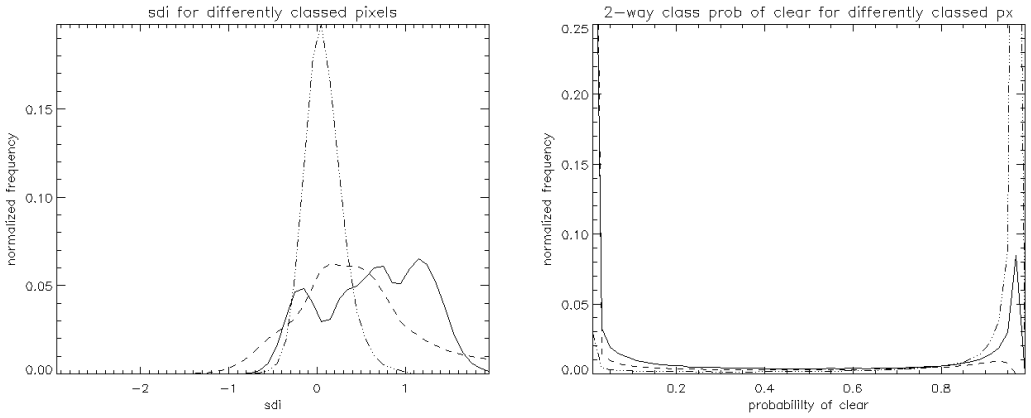
(b)

Figure 7.20: Result of the 2-class algorithm for the case study: (a) the colourscale corresponds to a range of 0-1 for the posterior probability of clear, and (b) a binary mask formed using a threshold of 0.5 on the probability of clear - black is cloud and white is clear.



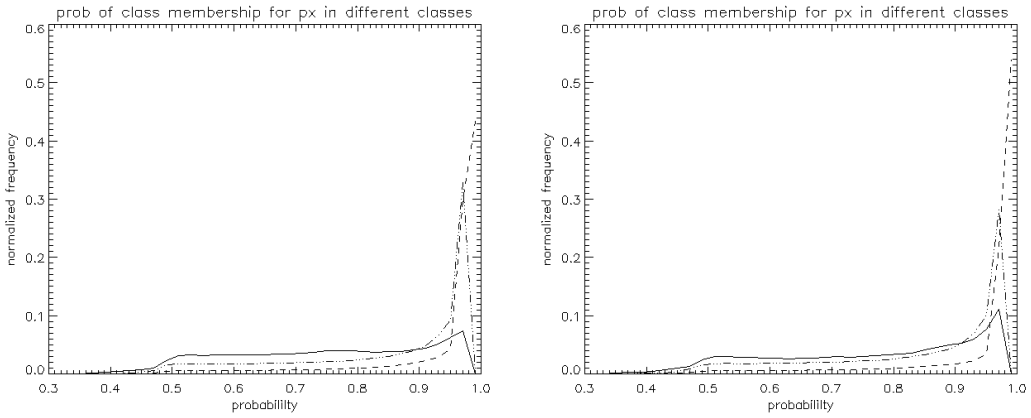
(a) (b)

Figure 7.21: Histograms of (a) SDI and (b) the 2-class probability of clear, for pixels in each of the 3 classes (as identified by the highest posterior probability) for the case study. Solid line: dust; dashed line: cloud; and dot-dashed: clear. The 3-class algorithm was run using a path length-dependent spectral PDF for dust.



(a) (b)

Figure 7.22: Histograms of (a) SDI and (b) the 2-class probability of clear, for pixels in each of the 3 classes for the case study. Solid line: dust; dashed line: cloud; and dot-dashed: clear. The 3-class algorithm was run using a spectral PDF for dust with no path length-dependence.



(a) (b)

Figure 7.23: The calculated posterior probability with which each pixel belongs to the class to which it is assigned: (a) using a path length-dependent spectral PDF for dust, and (b) using a spectral PDF for dust independent of path length. Solid line: dust; dashed line: cloud; and dot-dashed: clear.

7.5.1 Discussion of Case Study Results

The final classification results in figure 7.19 both appear plausible, although 7.19a indicates some dust at high latitudes, which is unlikely to reflect reality. The dust plume coming from the Saharan desert, seen off the North-East coast of Africa in figure 7.5 is captured in both the classifications, which also distinguish what are probably small clouds within the dust. A large number of pixels off the coast of Namibia are classed as dust, which agrees with the SDI image, and can be explained by sand blowing from the Namib desert, which lies along the coast and is separated from the interior by a mountain range, or possibly from the Kalahari desert which lies beyond the mountain range, but could be lifted high into the atmosphere and transported across the mountains - dust from the Sahara is known to be transported high in the atmosphere (see some of the references cited at the start of this chapter) and it is possible this occurs in the Kalahari too. Other studies have identified low altitude dust along the Namibian coast, for example Eckardt and Kuring (2005), which is missing from datasets compiled using sensors such as TOMS which cannot reliably observe dust at low altitudes. The results here do not differentiate between high and low altitude dust, but it is nonetheless encouraging that dust at low altitudes, generally difficult to detect in satellite data, McPeters et al. (1998), may have been detected. This is particularly encouraging if the dust is low dust from the Sahara over the Red Sea and Mediterranean is also captured in both the SDI image, and in the classification - although more dust pixels are identified in the Mediterranean region when no path length-dependence is used for the spectral PDF for dust.

It is not surprising that the histograms in figures 7.21a and 7.22a show a broader peak for dust than for either the cloud or clear classes, as there are far fewer pixels in the dust class than in the other two, making the shape of the distribution less well defined. Clear-classed pixels form a well-defined peak in the SDI histogram, at a low SDI value as expected. Cloud-classed pixels also form a reasonably well defined peak, at a higher SDI value than the clear peak. This can be explained, as cloud and dust are not expected to be easily distinguishable using SDI, which is a measure for separating clear and dust observations, after cloud observations have been removed. True cloud observations can therefore be anticipated to appear anywhere in the SDI space. The SDI distribution for dust-classed pixels is more difficult to interpret. In the path-length-dependent case, a significant fraction of dust-classed pixels exhibit properties which seem more representative of clear observations (i.e. the peak at SDIs near 0). While a peak in the SDI distribution is apparent in the expected position for dust, these possibly mis-classed pixels form a second, distinct peak which contains more pixels. When no path length-dependence is assumed for the spectral PDF for dust, a more plausible distribution of SDI values is seen, which, although broad, peaks at a higher SDI value than that for cloud-classed pixels. The SDI histograms indicate a successful classification for all 3 classes, although the path length-dependency does seem to result in some misclassification of dust pixels. It should be remembered, however, that SDI itself cannot be considered a 'truth', rather it is an index of dust that, operationally, is taken to be a reliable indicator of the presence of dust for SDI >0.25 in regions away from the satellite limb view.

The histograms in figures 7.21b and 7.22b show cloud- and clear-classed pixels to have a

high- and low- probability of cloud respectively when calculated using the 2-class algorithm. This is not surprising, as the algorithms are of course very similar, however it is reassuring that the introduction of a third class only results in pixels previously classed as either cloud or clear being now classed as dust, and not in pixels changing from a clear to a cloud classification or vice-versa. Figures 7.21b and 7.22b show that pixels now classed as dust were mostly classed as cloud using the two class method, as expected.

A distribution of the posterior probability of membership to the class to which each pixel is assigned is plotted in figure 7.23. Pixels assigned to the clear and cloud classes are associated with less ambiguity than those assigned to the dust class. This may reflect the spectral similarity between dust and cloud observations. A more definite peak is seen in the dust classification calculated with no path length-dependency for the spectral PDF for dust, suggesting a less ambiguous classification. This is surprising, as the path length dependency of the spectral PDF for dust was expected to result in its having higher values, which could be associated with a more certain classification. Figure 7.16 shows that the spectral PDF values for dust, calculated with and without the atmospheric path length dependence, is similar. The highest values are seen in mostly the same areas in both images, although the values are higher when no path length-dependence is used, which is surprising. This is especially noticeable in the Mediterranean and Red Sea regions, where figure 7.16b indicates a high probability of the observations corresponding to dust, which is not seen in figure 7.16a. The path length-dependency also increases the spectral PDF at higher latitudes and around the edge of the disk (which correspond to longer path lengths when viewed from this sensor), where dust is unlikely to be present. This is not necessarily a problem, as the prior probability for dust $P\left(c_{dust}\right)$ is very low in these regions, however it does suggest that the technique may lead to non-dust features being identified as dust in other regions, which may explain the second peak in figure 7.21a.

The posterior probabilities for dust in figures 7.17a and 7.18a are very similar - notably in the Mediterranean and Red Sea regions, where the spectral PDFs for dust are quite different. This suggests that the textural component of the PDF for dust dominates in this region, an explanation supported by the high values in figure 7.8 for this area.

Conclusions drawn from this case study should be considered with the caveat that the results may be not be representative of a larger dataset. The absence of a 'truth' means that some judgment is required on the part of the investigator, in determining what the 'true' class of a pixel is likely to be (through examination of SDI data and the results of the 2-class probability of clear). No practical alternative to the qualitative comparison presented here could be found, however, and so, notwithstanding the criticisms above, it was concluded from this case study that the path length-dependence of the spectral PDF for dust does not improve the accuracy of the classification, and in fact introduces some artifacts into the results at high latitudes. It was therefore decided not to include a path length-dependency in calculation of the spectral PDF for dust. It was also concluded that the classification was successful, at least for this image, although

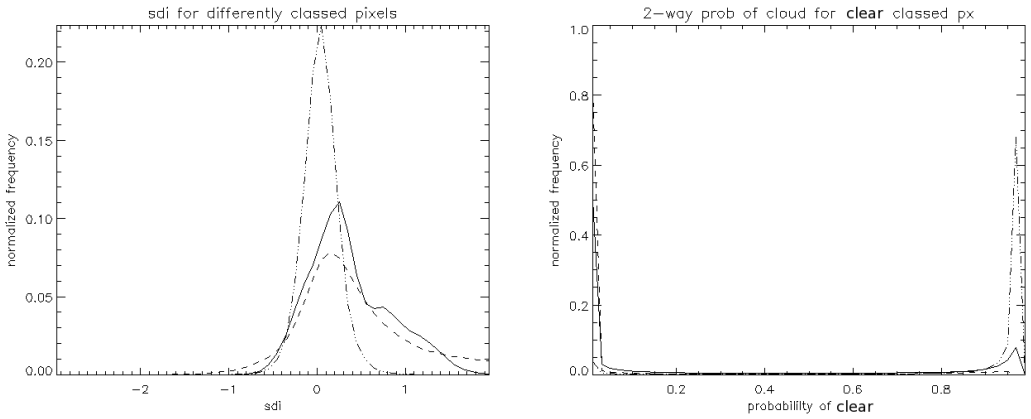
here it is important to note that the SDI image, against which comparisons have been made and the success of the classification judged, is not a 'truth'. There may be shared biases between the results, meaning agreement with the SDI image could occur even in cases of mis-classification, and similarly, the SDI itself, being only appropriate for cloud-screened imagery, may indicate some truly cloud pixels to be dust.

7.6 Application of Classifier to a Larger Dataset

In section 7.5, the classification algorithm was successfully applied to one image from the SEVIRI sensor. To demonstrate the technique more fully, 22 night-time images, acquired by the same SEVIRI sensor between 2004 and 2006 were used. All the images were acquired between 18:15 and 06:00 UTC, and only those ocean pixels which correspond to a solar zenith angle greater than 90^o were investigated. The SDI for each pixel was calculated, in the same way as for the case study, using the method presented in Merchant et al. (2006a). The 2-class algorithm was run for the images, producing a posterior 2-class probability of clear. The 3-class algorithm was run, using a spectral PDF for dust with no dependency on atmospheric path length, and pixels were assigned to the class for which their posterior probability of membership was highest.

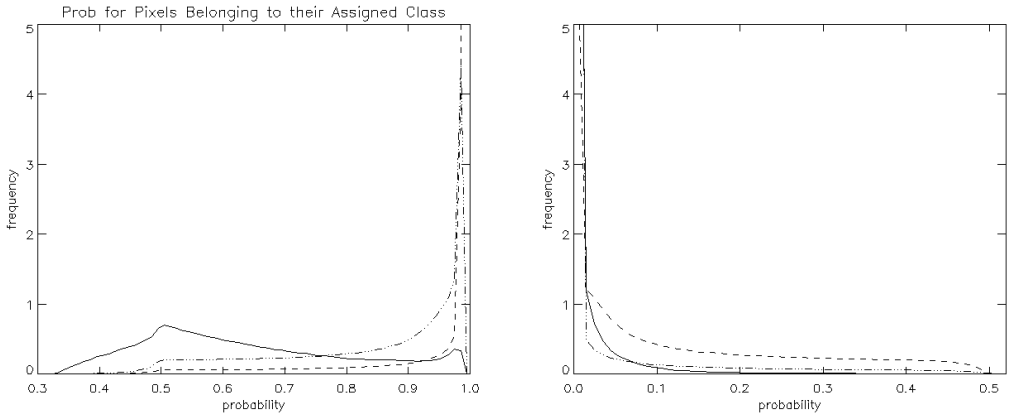
The distribution of SDI values for pixels assigned to each class is shown in figure 7.24a. Similarly, the distribution of the 2-class probability of cloud for pixels in each class is shown in figure 7.24b. To measure the certainty of the classification, figure 7.25a shows distributions of the probability with which pixels belong to the class to which they are assigned. This should be compared to figure 7.25b, which shows the distribution of calculated probabilities for pixels which are not assigned to the class, i.e. the distribution of calculated probabilities for dust for all pixels which are not classified as dust in the result, and likewise for cloud and clear.

Ambiguities in the 3-class classification can be looked at in more detail in figure 7.26, which shows the distribution of probabilities with which pixels assigned to each specific class are associated with each of the other two classes. The histograms were calculated and normalized for each image before being summed to create the histograms in the figures. For reference, figure 7.27 shows the distribution of posterior probabilities for clear calculated using the 2-class method for all pixels - only one histogram is necessary, as pixels are assigned to the class corresponding to whichever side of the distribution they fall, i.e. a calculated probability of clear greater than 0.5 corresponds to a clear pixel. The ascribed ‘certainty’ of the 2- and 3- class methods can then be compared.



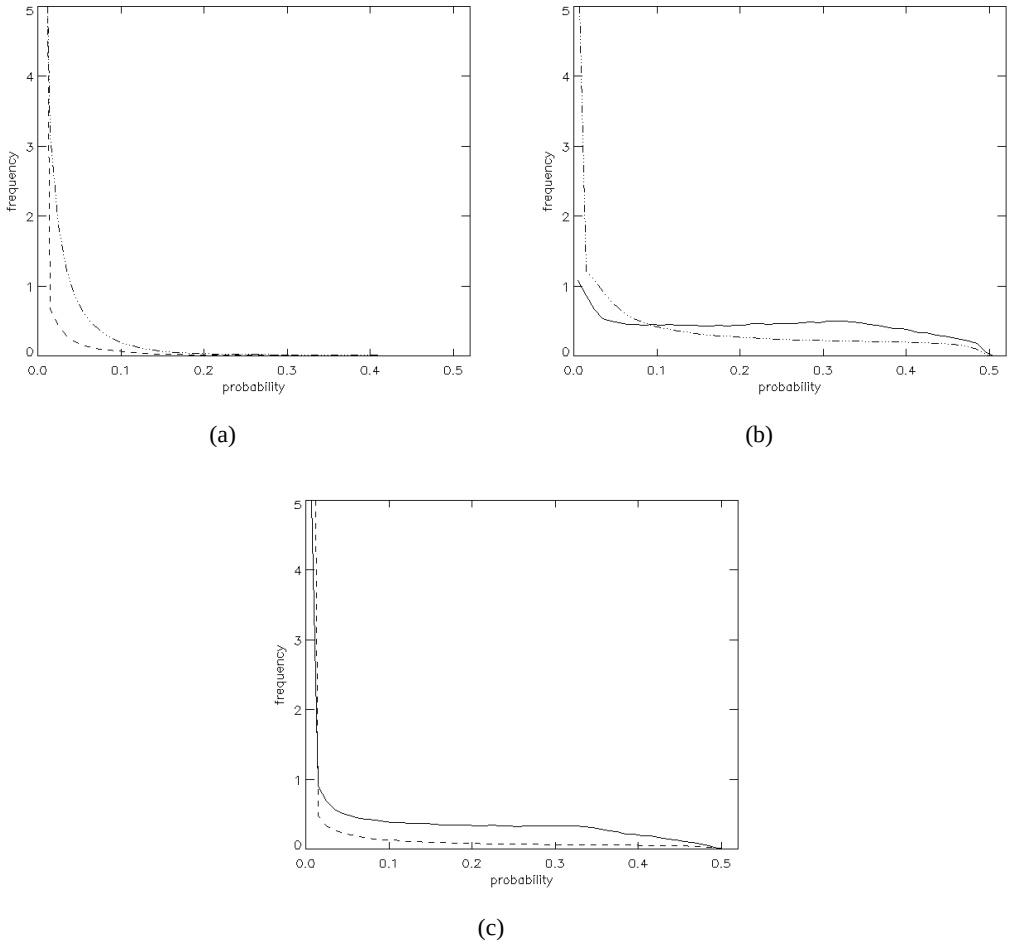
(a) (b)

Figure 7.24: Histograms of (a) SDI and (b) the 2-class probability of clear, for pixels in each of the 3 classes. Solid line: dust; dashed line: cloud; and dot-dashed: clear.



(a) (b)

Figure 7.25: (a) Histograms of the probability with which pixels are associated with the class to which they are assigned; (b) Histograms of the calculated probability of each class for pixels not assigned to that class. Solid line: dust class; dashed line: cloud class; and dot-dashed line: clear class.



(a) (b)

(c)

Figure 7.26: Histogram of the probability with which (a) pixels assigned to the clear and cloud classes are associated with the dust class; (b) pixels assigned to the clear and dust classes are associated with the cloud class; (c) pixels assigned to the dust and cloud classes are associated with the clear class. The solid line represents pixels assigned to the dust class, the dashed line represents pixels assigned to the cloud class, and the dotdotdashed line represents pixels assigned to the clear class.

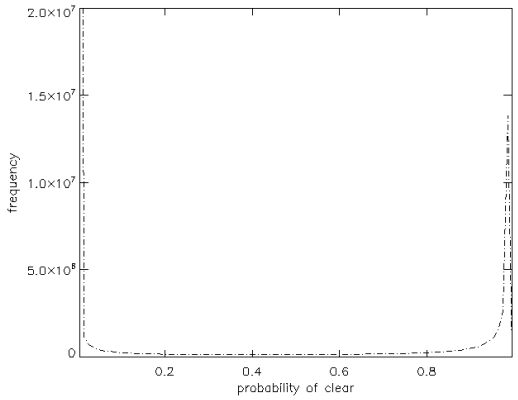


Figure 7.27: Distribution of calculated posterior probabilities for clear using 2-class method.

7.6.1 Discussion of Results

The histograms in figure 7.24 indicate a successful classification. Clear-classified pixels are associated with lower SDI values, and a narrower distribution, than either dust- or cloud-classified pixels. Cloud and dust pixels are anticipated to correspond to broader SDI distributions than clear pixels, reflecting the larger number of atmospheric states that such observations encompass. The peaks for clear, cloud and dust SDI distributions appear in the expected relative positions, with the peak for dust-classified pixels being at a higher SDI than either clear- or cloud-classified pixels. This suggests that the pixel-classification does reflect reality - pixels assigned to each class correspond to the SDI distributions that would be expected for observations of each class. Figure 7.24b also shows the expected distribution for a successful classification. If the 2-class algorithms is treated as an accurate discriminator between clear and cloud pixels, then pixels classed as cloud in the 3-way classification should correspond to a high probability of cloud in the 2-class results. Likewise, clear-classified pixels should correspond to a high probability of clear (equivalent to a low probability of cloud) in the 2-class results. This is not surprising, as the algorithms are of course very similar, but it is reassuring that the introduction of a third class does not appear to reduce the accuracy of the clear-cloud discrimination. Dust-classified pixels are seen in this figure to have been mostly associated with a cloud class in the 2-class algorithm, which is also not surprising, as dust and clouds often appear spectrally similar.

The certainty of the classification can be described by the probability with which pixels are associated with the class to which they are assigned. Figure 7.25a indicates a very high level of certainty for clear- and cloud-classed pixels. It is likely that the lower probabilities with which dust-classed pixels are associated with dust can be attributed to some dust-classed pixels corresponding to a high probability for both cloud and dust. A similar ambiguity could be anticipated for some cloud-classed observations, but would not necessarily be as noticeable in the plotted distribution as for dust-classed pixels. Ii is possible that *all* dust observations correspond to high probabilities for both dust and cloud, while the same is true for only *some* cloud observations (those in the region for which the prior probability of dust is high, and which correspond to clouds that are spectrally and/or texturally similar to dust). This would mean that the effect would only be seen in the distribution for dust-classed pixels, as the distribution of dust probabilities for cloud-classed pixels would be dominated by cloud pixels which are more distinct from dust, which are almost certainly the majority. The slightly higher probabilities with which non-cloud-classed pixels are associated with cloud in figure 7.25b are likely to correspond to dust-classed pixels. These two plots suggest that the posterior probability for dust for dust-classed pixels is often only slightly higher than that calculated for cloud. In general, however, pixels are seen to have a low probability of belonging to a class other than that to which they are assigned, indicating a relatively certain classification.

The ambiguity of the classification can be investigated using figure 7.26. Cloud- and clear-classed pixels are both associated with very low probabilities of dust, which is likely to be an effect of the prior probability for dust constraining the probability for dust to low values except in

a specific geographical region (unless the calculated PDF for dust is very high). More ambiguity can be seen in figures 7.26b and c, which show clear- and dust-classed pixels to be associated with slightly higher probabilities for cloud, and cloud- and dust-classed pixels to be associated with slightly higher probabilities for clear. In both 7.26b and c, the suggested ambiguity is higher for dust-classed pixels, but still small enough to indicate a high certainty for the classification.

7.7 Conclusions

The results presented in the preceding sections show the dust-cloud-clear classifier to work well. Cloud and clear observations continue to be discriminated accurately, while dust-classed observations, although associated with a slightly lower certainty, correspond well with the SDI. In the absence of a 'true' dust-cloud-clear classification for comparison, the problem of assessing the skill of the discrimination can only be approached qualitatively. Previous work supports the accuracy of SDI as a tool for discrimination of clear and dust pixels (Merchant et al., 2006a), but no work has been done to thoroughly investigate how cloud observations are represented in SDI. Similarly, the 2-class cloud detection method has been demonstrated to perform with high level of skill over the ocean at night (Merchant et al., 2005), but the method it uses relies on the implicit assumption that cloud and clear are the only possible atmospheric states. Validation against a 'true' dataset of classified imagery would be a worthwhile extension of this work. The most detailed studies into Saharan dust in the literature tend to be targeted case studies, see for example Gobbi et al. (2000); Pinker et al. (2001); Highwood et al. (2003). A comparative study of this method with the results of such localized studies would be interesting.

The technique presented here could be useful for applications that require the identification of Saharan dust in global satellite imagery, for example for NWP. Removing suspected cloud pixels prior to retrieving dust observations, as is generally done at present, necessarily biases sets of recorded dust observations towards those which appear spectrally more similar to clear sky than to cloud. The results presented here show that, at least using the Bayesian cloud detection technique, most dust would, in a 2-class scheme, be detected as cloud, and so is likely to be removed prior to any dust-discrimination processing. This technique avoids that bias and so may be able to contribute to solving some of the issues surrounding the quantity and behavior of Saharan dust in the atmosphere, which are currently compounded by the lack of a reliable estimate of the location of dust in the atmosphere at any given time (Tegen and Fung, 1994).

Dust discrimination techniques such as SDI require spectral data at four wavelengths (3.7, 8.7, 11 and 12 μm), which are not available from all sensors. A reliable dust detection technique such as that presented here, where NWP information compensates for the lack of spectral information, allows for data from such sensors to be exploited more fully. This potentially increased dust monitoring could significantly improve estimates of the volume of dust in the atmosphere, with positive implications for chemical transport and climate modelling.

Chapter 8

Conclusions and Future Work

Discriminating cloudy from clear pixels is an essential step in deducing much useful information from satellite imagery. Records of both surface observations, and of observed atmospheric variables are likely to be affected if retrieval of these parameters, almost always carried out post cloud-screening, relies on an inaccurate method of cloud detection. This has implications for any application which exploits parameters derived from satellite data, for example NWP and climate research. Accurate cloud detection is therefore incredibly important if the full potential of satellite-borne sensors to record a global picture of the state of the Earth and its atmosphere at a given moment in time is to be realized and exploited.

This project has contributed to a new, accurate and reliable method of cloud detection which is suitable for real-time applications such as NWP. As an approach, the Bayesian method has the advantage of a sound physical basis and of quickly calculating a product in which the certainty of individual pixel results is clear. For applications which require the processing of a large volume of data recored by many different satellite sensors, the flexibility of the method, making it simple to adapt to imagery from different sensors, and appropriate for imagery acquired at any location and time means it has the potential to be an extremely useful cloud detection tool. Such applications include NWP, which provided the motivation for this project, which was partly funded by the U.K. Met. Office. As the accuracy of NWP data continues to improve, through the application of research such as that presented here, the accuracy of results calculated by the Bayesian approach can also be expected to increase, and so the method also has the potential strength of ‘self-improvement’.

The Bayesian method for cloud detection, which has been advanced and improved by this work, does have some limitations which should be noted, notwithstanding the results of the technique demonstrated in chapter 5. These weaknesses underlie all uses of the technique investigated here, and so should be considered alongside all the findings of this work.

In order to perform computationally efficient classification, which is essential to applications such as NWP, several approximations are made. The RTM used to predict clear sky observations (and cloud observations in chapter 6) and adapted to predict observations of desert dust in chapter 7, is designed to run quickly, and is not expected to make predictions as accurate as those which would be produced by a full line-by-line radiative transfer model. The uncertainties accounted for

in the $\mathbf{H}$-, $\mathbf{B}$- and $\mathbf{R}$- matrices, for simulations of both clear-sky observations (described in chapter 2), and cloud observations (described in chapter 6) are approximated as Gaussian distributions. In reality, this is unlikely to reflect the true shape of the error distributions. Assimilating errors in the form of more appropriate distributions would be computationally expensive, increasing processing time and making the algorithm impractical for the real-time applications for which it is intended. The inaccuracy introduced by this approximation is considered both necessary, on practical grounds, and small enough to be justified.

Uncertainty in NWP data is likely to vary in space and time, but must be approximated as constant for the fields considered in the Bayesian cloud detection technique (these errors are considered in the $\mathbf{B}$-matrix described in equation 2.4, on page 18). This is to some extent justified by the difficulty of quantifying uncertainties in NWP data. It is anticipated that, as advances are made in NWP, the uncertainties will both decrease and be quantified more accurately, but at present this method of accounting for uncertainties in the NWP data is considered as practical as any available, and is therefore considered appropriate.

In this work, the Bayesian technique has been extended and demonstrated to be suitable for day- and night-time cloud detection applications for land and sea imagery, and for simultaneous retrievals of dust, cloud and clear-sky observations. A method of spectral clustering is also presented and shown to have the potential to reduce ambiguity in the classification of some pixels.

In addition to improving the algorithm to a standard suitable for release under public license, several features have been added and validation work carried out. The technique is no longer optimized solely for applications requiring cloud detection for night-time ocean imagery, and has been shown to be suitable for data from more than the single satellite-borne sensor for which previous validation work had been done. A clustering method which may increase the skill of the detection around ocean fronts was investigated, as was a technique for the fast forward modelling of cloudy atmospheric states. It was demonstrated that this latter technique could be included in the Bayesian cloud detection scheme, thereby increasing its dependence on spatially- and temporally- specific information for cloud detection. In this way, the technique can be viewed as being ‘tailored’ to individual pixels. A one step classification of pixels into cloud, clear and dust classes was also developed from the Bayesian cloud detection method, thereby removing the need for cloud screening of imagery prior to retrieving satellite observations of dust.

8.1 Cloud Detection with Pre-Clustering of Imagery

One of the strengths of the Bayesian approach to cloud detection is that, rather than producing a binary mask whereby each pixel is classified as cloudy or clear, a probability of clear is calculated for each pixel. This allows application-specific tolerances to cloud contamination to be set and makes it easy to identify ambiguously classed pixels which may not be suitable for further

processing, i.e. those with a probability of clear around 50%. Ideally a probability of clear very close to either 100% or 0% would be calculated for all pixels, and there would be no ambiguous classifications.

In practice some pixels, such as those imaging ocean fronts, can be ambiguously classified, see for example figure 4.1d in chapter 4. Pixels on different sides of an ocean front can be associated with different thermal regimes. When considered together, these pixels can be associated with a variability in surface temperature that is more usually associated with clouds than with ocean surfaces. Textural considerations in the algorithm therefore raise the probability that such observations correspond to cloud, rather than to clear sky, leading to such pixels sometimes being falsely classified as cloud. This may be overcome if pixels are first separated into thermal regimes within which the variation in surface temperature is more likely to represent that of a clear ocean surface (in a non-frontal region) than a cloud top, and then processed together for cloud detection. It is shown in chapter 4 that separating pixels into clusters prior to processing for cloud detection, so as to consider only pixels from the same thermal regime together, can reduce the incidence of false classifications of this type.

Many methods exist for clustering. Section 4.1 in chapter 4 presents investigations into Split and Merge Clustering (SMC), Clustering Large Applications (CLARA) and Fixed Point Clustering (FPC), and argues their relative suitability to this application. FPC is judged the most suitable and further investigations into the appropriateness of implementing clustering in the cloud detection scheme are carried out using FPC in sections 4.1.5 and 4.2.

One potential problem associated with clustering is the loss of spectral information for individual pixels. If all pixels are each assigned to a specific cluster, and the average spectral properties of the pixels in each cluster are processed for cloud detection, then the same probability of clear will be calculated for all pixels in a cluster. This is not necessarily appropriate for all the individual pixels in the cluster.

This problem can be overcome if pixels are unrestricted in the number of clusters they may belong to, and if the relative strength of their association to each one is quantified, as occurs using the FPC clustering method. The spectral properties of the cluster then become the weighted average of the properties of member pixels, and each pixel is assigned spectral properties for the cloud detection processing based on combining the spectral properties of the clusters to which it belongs, weighted by the strength of its association to each one.

A more bimodal distribution of calculated probabilities of clear was found to result from including clustering of the imagery as a pre-processing step, see section 4.2. This indicates a more certain cloud detection result. The loss of individual pixel information was shown not to have been entirely overcome, as some single pixel clouds were missed, but were detected successfully when imagery was not pre-clustered, see figures 4.2 and 4.4, and the discussion in section 4.3.

Further study could quantify the extent to which single pixel clouds are missed as a result of pre-clustering, and so assess the overall improvement, or otherwise, brought by pre-clustering to the skill of the cloud detection.

FPC having been identified as a suitable clustering technique in this work, more work could ascertain the most appropriate parameters to use for the clustering. A necessary precursor to any further work, however, is significant optimization of the clustering algorithm, which at present is too slow (requiring several days for a ∼200x200 pixel region to be clustered) for a larger case study, or for further investigation, to be practical.

Advances in NWP, and more accurate quantification of the uncertainties associated with NWP data, should lead to a reduction in the number of ambiguous pixel classifications resulting from the Bayesian cloud detection algorithm. This would be a result of the spectral PDF calculated for clear sky conditions becoming narrower and steeper, leading to more extreme values being found for $P\left(\mathbf{y_s}\left|\mathbf{x^b}, c\right.\right)$, making it more often the dominant term in $P\left(\mathbf{y}\left|\mathbf{x^b}, c\right.\right)$, see section 2.4 in chapter 2. The textural PDF having been identified as the source of ambiguity in regions of ocean fronts, then, if such regions could be identified, it may be appropriate to perform the cloud detection on the basis of spectral considerations only, i.e. to set $P\left(\mathbf{y_t}\left|\mathbf{x^b}, c\right.\right) = P\left(\mathbf{y_t}\left|\mathbf{x^b}, \bar{c}\right.\right) = 1$. This may reduce, or remove altogether, the problems that the work here seeks to address through clustering.

8.2 Cloud Detection for Day and Land Imagery

The range of applications for which the Bayesian cloud detection method was suitable prior to this project was limited to those using night-time imagery acquired over ocean surfaces, see the introduction to chapter 5. An advantage claimed for this cloud detection method at that time, was its flexibility - in principle, application of the same technique to imagery from any sensor for which an RTM exists should be relatively straightforward. Validation of the method had only previously appeared in the literature for imagery from one sensor, and a part of this work involved restructuring the algorithm to be more flexible, and so easier to adapt to imagery from other sensors. Chapter 5 describes the work done to extend the algorithm to exploit surface emissivity, in section 5.1, and reflectivity data, in section 5.2, and so to operate over land for both day- and night-time imagery, significantly widening the range of applications for which it is suitable. A quantitative validation of the technique with these extensions was carried out using imagery acquired from another sensor. The results are compared to those of operational techniques for the same externally supplied database, and a higher, or comparable, level of skill to the operational techniques is seen for the Bayesian method in every case.

Section 5.3.3 in chapter 5 shows that, for land targets imaged at night-time, the Bayesian technique achieved a True Skills Score (TSS) almost 10% higher than the only available operational result for these targets (see section 5.3.1 on page 64 for the definition of the skills

scores used). This is a result of the Hit Rate (HR) being more than 32% higher than that of the operational method. For sea targets imaged at night, the Bayesian HR is at least 9% higher than both operational techniques, giving it a TSS of 87.3%, in comparison to the TSS of 79.3% and 74.0% achieved by the operational techniques of the U.K. Met. Office and Météo-France respectively. The Bayesian technique is shown to significantly outperform operational techniques for both night-land and night-sea targets, showing that it is now a suitable tool for any application requiring cloud detection for any night-time-acquired satellite imagery.

Cloud detection for day-time imagery is generally more straightforward, image data at visible wavelengths generally being available to aid detection, in addition to the thermal-wavelength data which is exploited for night-time imagery. All approaches to cloud detection are therefore likely to be more successful when applied to imagery acquired during the day. Despite the high level of skill demonstrated by the operational techniques for day targets, the Bayesian method achieved a 2% higher TSS than the operational skill of Météo-France for land targets, see section 5.3.3. For sea targets, the Bayesian approach achieved a 3% higher TSS than the U.K. Met. Office, and was within 1% of the TSS achieved by Météo-France. This is encouraging, as both operational techniques perform well for daytime imagery, with TSS of 86.4% and 90.4% for the U.K. Met. Office and Météo-France respectively. This shows the Bayesian cloud detection scheme to be at least as suitable as current operational schemes for cloud detection in both day and night satellite imagery, acquired over both land and sea.

The results in section 5.3.3 demonstrate the successful validation of the Bayesian cloud detection method for imagery from a satellite sensor it had not been validated for previously, indicating the flexibility of the technique, which is not sensor-specific. The high skill scores achieved, relative to those of operational approaches, show it to be suitable to the full range of imagery acquired by a satellite, and so to be suitable for implementation in a wide range of applications. It is therefore hoped that the algorithm could contribute to improvements for the many applications that rely on successful cloud detection for the exploitation of data from within satellite-imagery.

8.3 Local NWP-Dependent PDF for Cloud

The Bayesian method for cloud detection provides accurate results partly because it has a physical basis and relies on a spatially- and temporally-specific calculation, in contrast to more traditional cloud detection techniques where thresholds for differentiating between clear and cloud observations are assumed to be more or less constant. An RTM is used to simulate pixel-specific observations of clear sky from NWP data. Uncertainties in the NWP data and in the forward-modelling are used to infer from the simulation the probability distribution in observation space of clear sky. The value corresponding to an individual pixel observation is read as the NWP-conditional probability that the observation corresponds to clear sky. An analogous

number is needed for the probability that the observation corresponds to cloud. At present this is read from a normalized distribution of cloud observations, compiled from a global dataset. It is unlikely that such a global distribution will represent all possible cloudy atmospheric states, and it is impossible for it to represent only those cloudy states which are realistic for the time and place imaged by the pixel. This can be overcome by the use of an NWP data to calculate a temporally- and spatially- specific PDF for cloud. This is is more challenging than the clear sky case, as clouds of different optical depths, with different phases, filling different fractions of the imaged pixel and at different altitudes must all be represented. Simulating enough cloud observations to cover this range of atmospheric states with an RTM would be too computationally expensive to be practical. An alternative method of calculating a time- and space-specific PDF for cloud is therefore required in order to make the technique fully physically robust.

Section 6.1 in chapter 6 describes how relationships between simulated cloud radiances and liquid- and ice-water paths, and cloud-top altitudes and cloud fractions, are used to limit the number of cloudy atmospheric states requiring forward-modelling with the RTM. The presented method limits the cloudy states represented in the PDF to those which are realistic for the pixel and includes some consideration of the relative likelihood of the various different cloud states, given the NWP data. This makes the local PDF for cloud more dependent on the NWP data and the uncertainties associated with it than the clear sky PDF. In addition to the approximations made in the RTM for the clear sky case, further approximations are made in the cloudy case for simulation of cloud observations, and these are discussed in section 6.4.1. The method relies on these uncertainties, and those for the NWP data, being well estimated. It also relies on their not being unreasonably large, which would create a broad, flat cloud PDF, in which the conditional probability of cloud for almost all observations would be very similar, and which would therefore not be useful in discriminating cloud observations. The results of the case study in section 6.5 show the feasibility of calculating a local PDF for cloud within the algorithm, and suggest the potential of the local PDF for cloud for improving the skill of the cloud detection. When applied to a larger dataset in section 6.6, however, the local cloud PDF was seen to decrease the overall skill of the detection. A decrease in FAR was achieved for day, night, land and sea targets, but this is countered by a large decrease in HR, meaning the only improvement in TSS was for day sea targets, for which the local cloud PDF effected a rise in TSS of 4.5%. For day land, night sea and night land, the change in TSS resulting from replacing the global PDF with the local PDF for cloud was a decrease of 15.2%, 8.4% and 12.8%.

These results are probably a consequence of biases in the RTM simulations and in the NWP data which are unknown or not fully accounted for. NWP data is used in the simulations of cloud observations, and also in the relative weighting of the representation within the PDF. In this way, the cloud PDF is more sensitive to the NWP data than the global PDF. The local PDF for cloud is steeper and narrower than the global PDF it was substituted for, making its position more significant, particularly to clouds which are represented at the edges of the distribution. Inaccurate NWP data, or uncertainties in the RTM, could lead to the PDF being centred inappropriately.

Similarly, if uncertainties in both the RTM and the NWP data are not sufficiently accounted for, the PDF will be narrower than is appropriate, leading to cloud observations which would have been at the edge of the distribution being missed. NWP data is thought to be less reliable over land, see section 5.1.3, so it is not surprising that the greatest decrease in TSS is seen over land. The scattering approximations made in the RTM for simulation of cloudy radiances, described in section 6.4.1, are likely to have a greater effect on simulations at 3.9μm than on 11 and 12μm, and this may explain why the effect of the local PDF for cloud is so much worse for nighttime imagery.

Another possible explanation for the poor results seen for the technique in chapter 6 is the application of the RTM sensitivity to specific NWP fields (represented in the **H'**-matrix) for *clear sky* simulations to be used for the calculations involving the simulations of cloudy sky radiances. The simulation of cloudy radiances is likely to be sensitive to more fields in the NWP data than the simulation of clear sky, as highlighted in section 6.7. For example, cloud simulations might be sensitive to the temperature at a particular altitude, whereas the sensitivity for simulations of clear sky may be negligible. Expanding the **H'**- and **B**-matrices to include terms accounting for sensitivities not considered for clear sky simulations, or increasing the sensitivities in the **H'**-matrix to those fields already considered is discussed in section 6.7. As the algorithm is currently structured, neither of these options would be computationally straightforward to implement without increasing processing time by an unreasonable amount. If further NWP fields were considered, suitable values for the **B**-matrix would have to be found, and quantifying the uncertainty in NWP fields is problematic, mainly because it is likely to be non-linear, and to vary in space and time. Although it is at present computationally infeasible to consider a separate **H'**-matrix for each simulated cloud observation, and it is accepted that the sensitivity is likely to depend on the cloud parameters such as cwp, and so to vary between the simulated cloud radiances, it may be possible to increase the values calculated for the clear sky **H'**-matrix before implementing it in the calculations for the cloud PDF. This would require some work to identify the appropriate value for the increase, as the same **h'**-matrix would still have to be used for all represented clouds, but would be computationally straightforward to implement, and may have some affect on the skill of the technique.

Advances in NWP, and in fast forward modelling, mean that these effects are anticipated to be less significant in the future, and the skill of the detection using the local PDF should increase, possibly to beyond that currently achieved using the global PDF. While these results are disappointing, the dataset used is comprised of 'difficult' targets for cloud detection, and so may contain cloud targets whose spectral properties that are not allowed for by the approximations of the RTM, making these targets more sensitive to the approximations than would generally be the case. A comparison of the skill of the detection with and without the local PDF for cloud using a more general dataset could allow a more appropriate assessment of the effect on the detection's skill.

8.4 Dust-Cloud-Clear Classifier

Cloud detection can be viewed as classification of an image into two classes. Under the Bayesian approach, any number of classes for which the necessary prior information is available can be considered. Chapter 7 describes how this feature of the technique was exploited to address the problem of observing desert dust in the atmosphere. Difficulties associated with the retrieval of dust observations from satellite imagery mean that the volume and distribution of dust in the atmosphere at any given time is highly uncertain, see for example Tegen and Fung (1994), affecting models of climate and chemical transport, as well as creating problems for dust storm forecasting in NWP.

One method of dust detection which is used in NWP relies on imagery having been recorded at four infra-red wavelengths (Merchant et al., 2006a), while some meteorological satellite-borne sensors only record imagery at three infra-red wavelengths. A method by which this lack of spectral information can be compensated for by exploiting NWP data in addition to the available spectral information, is presented in 7.1. Most schemes for the retrieval of dust observations are only suitable for application to pixels that have already passed an initial cloud-screening step, which means that only pixels already deemed to be cloud-free can be classified as dust. This necessarily results in a record of dust observations which is biased towards those dust observations which appear more similar to observations of clear sky than to observations of cloud. This uncertainty affects historical records of dust and their interpretation, for example within climate and chemical transport models.

The case study in section 7.5 demonstrates that the Bayesian approach can be used to simultaneously assign pixels a probability of belonging to classes of dust, cloud and clear sky observations in one step, removing the above bias.

No ‘true’ dust-cloud-clear classification data was available against which to quantitatively assess the results of the Bayesian classifier. To give a qualitative analysis of the classification, the SDI was calculated for pixels in the test data using the method shown in Merchant et al. (2006a) to be successful in discriminating dust and clear sky observations (see section 7.1 for a brief description of SDI). The Bayesian cloud detection algorithm, shown in section 5.3.3 to be a reliable technique for the discrimination of cloud from clear observations, was also run for the test data. Following the implementation of the classifier for the test data, each pixel was assigned to the class for which the calculated probability of its membership was highest. Figures 7.23 and 7.25 show that there were few ambiguously classifications and almost all pixels were associated strongly with only one class, indicating that the result can be associated with relatively high certainty.

The distribution of SDI values, and of the probability of clear found in the cloud detection, were then plotted for each class, see sections 7.5 and 7.6. Pixels classified as dust were seen to correspond to an SDI distribution which would generally be associated with dust, and to a flat

distribution of probabilities of clear, with a small peak at the cloud end of the distribution. It is notable that observations classified as dust using this technique correspond to probabilities associated with both clear and cloud under the cloud detection algorithm. This suggests that some dust observations would be removed in cloud screening prior to implementation of a dust retrieval scheme, such as SDI, and so not be recorded. This one-step classification method has the potential to avoid this. Clear classified pixels were not surprisingly found to correspond to high probabilities of clear in the cloud detection results, and to low SDI values, indicating that the discrimination of clear observations using both methods agrees with the classification results. Observations classified as cloud corresponded to low probabilities of clear, and to a distribution of SDI values similar to that which would be anticipated for dust, but peaking at a slightly lower value.

The results of the case study in section 7.5, and of the larger-scale study in section 7.6, indicate that the Bayesian approach can be successfully used to simultaneously discriminate dust, cloud and clear observations. It has been shown that the problem of retrieving dust observations in the absence of a fourth spectral channel can be overcome if NWP data is exploited to compensate for the information loss. Furthermore, some evidence is seen for the bias inherent to retrievals of dust observations which require imagery to first be cloud screened. The presented method avoids this bias and so could be a means by which the accuracy of records of dust observations may be improved, and could thereby contribute to improvements in NWP, climate and chemical transport modelling applications.

8.5 Future Work

The work of this project into the feasibility of a local PDF for cloud in chapter 6 could be a means by which the method could be made to be more spatially- and temporally- specific, and therefore arguably more appropriate to the individual observations processed. Although the results do not show any increase in cloud detection skill, a more representative dataset could be used to test the technique and more appropriately assess its effect on the method's overall skill. A suitable dataset would have to contain observations centred on the wavelengths exploited in this project, NWP data and an externally-supplied 'true' cloud mask. Although substantial work has been done to make the algorithm more flexible, and so more straightforward to apply to imagery from other sensors, some restructuring work would still be required were the dataset to come from a satellite sensor for which the technique has not yet been implemented, particularly if an RTM other than RTTOV were required for the forward-modelling. A practical aspect of this possible future work would be likely to be reading the new observation data and NWP data into the algorithm from whatever format they were provided in, and converting them to a format appropriate for the RTM.

The successful extension of the algorithm from a cloud detection technique to a cloud-dust-clear classifier in chapter 7 could be built upon to create a classifier with a dust and a clear class,

and classes for different types of cloud. This could have positive implications for NWP, climate modelling, and any application which ultimately requires retrievals of different types of clouds. At present, such applications generally rely on a two-step process: firstly cloud detection, and then cloud classification. The Bayesian approach offers the potential for a one-step process. The arguments for discriminating cloud from clear observations on the basis of spatially- and temporally-specific criteria can also be applied to discrimination between different types of cloud, and so the Bayesian technique, with its physical basis, could provide an appropriate.

Clouds could be separated into classes corresponding to cloud top altitude and optical depth, possibly following the nine cloud type definitions used by the International Satellite Cloud Climatology Project (ISCCP: *http://isccp.giss.nasa.gov/*). This could have applications in NWP, and in climate research. The following paragraphs present some thoughts on how this could be achieved.

The method used to calculate a local PDF for cloud presented in chapter 6 could be used to calculate PDFs for different types of clouds. Predicted cloud observations, interpolated from RTM simulations, could be separated on the basis of the cloud properties to which they correspond and used to construct PDFs for cloud of specific types. Computationally, this should be relatively straightforward, as the algorithm has already been restructured to consider multiple classes, as described in chapter 7. Although structurally straightforward, calculation of the individual PDFs is computationally expensive and it is likely that the algorithm would become considerably slower, which may limit the practicality, even of investigative work. Several non-trivial issues would have to be addressed before PDFs for different cloud types could be appropriately calculated - the most significant of these are discussed briefly below.

In chapter 6, all clouds represented in the local PDF for cloud are approximated to be single layer- and single phase. If the same method were followed to construct PDFs for different cloud types, then only PDFs for single-layer and single-phase clouds could be constructed. This raises the question of how observations of multi-layer, or mixed-phase, cloud could, or should, be classified.

An appropriate value for the prior probability for each cloud class would have to be found. This could come from a global climatology, or from a dataset from which global mean prevalence could be calculated for different cloud types. Alternatively, it could be calculated from the weights which are currently used to weight the representation of clouds within the PDF by the relative likelihood of their being present. These options would have to be explored to find the most suitable value.

While the development of a cloud type classifier would be a useful progression from the work of this project, several non-trivial issues would have to be addressed carefully. It is anticipated that answering some of these issues would require several small-scale studies to be carried out

and the results analyzed in order to assess the most appropriate solutions.

8.6 Summary

This project has usefully advanced and extended the range of applications for which the Bayesian approach to cloud detection has been demonstrated to be successful and useful. The cloud detection scheme can now be implemented for land- and sea-imagery, acquired during day- or night-time, and has been applied to imagery from more than one sensor, demonstrating that it is now more flexible. The skill of the cloud detection has been compared to current operational techniques, and been shown to perform with comparable (a difference in TSS of less than 1%), or greater, skill, justifying the claim that it is suitable and appropriate for a wide range of applications. The approach has also been developed to classify imagery into multiple classes - an extension which is shown to achieve successful discrimination of dust, cloud and clear observations. This is anticipated to be useful in reducing some bias in retrieved dust observations, and to widen the range of imagery from which dust observations can be retrieved, a fourth infrared channel of spectral data no longer being required. This has implications for NWP, and for climate and chemical transport modelling. In addition, it may be possible in the future to extend this aspect of the work to classify observations into classes of cloud type, which could have benefits for NWP and climate research. A method for the fast forward modelling of cloudy atmospheric states is presented and demonstrated to be computationally feasable, although the results of its implementation in the cloud detection scheme are disappointing. Further work could improve its effect on the cloud detection scheme.

It is hoped that the findings of this project could contribute to improvements in NWP. This in turn should lead to further improvement in the results of the presented techniques, which rely on the exploitation of NWP data. In addition to NWP applications, it is anticipated that the work may be of interest to any application of satellite imagery requiring cloud or dust detection. As cloud-detection is a necessary pre-processing step for many applications, the methods and results could also benefit those in many fields of climate modelling and earth observation.

Bibliography

S. Ackerman, K. Strabala, W. P. Menzel, A. Frey, C. Moeller, and L. Gumley. Discriminating clear sky from clouds with modis. *Journal of Geophysical Research*, 103(D24):32141–32157, 1998.

Z. Ameur, S. Ameur, A. Adane, H. Sauvageot, and K. Bara. Cloud classification using the textural features of meteosat images. *International Journal of Remote Sensing*, 25(21):4491–4503, 2004.

A. Baran. The dependence of cirrus infrared radiative properties on ice crystal geometry and shape of the size-distribution function. *Quarterly Journal - Royal Meteorological Society*, 131: 1129–1142, 2005.

A. J. Baran, S. Haveman, P. N. Francis, and P. D. Watts. A consistent set of single-scattering properties for cirrus cloud: tests using radiance measurements from a dual-viewing multi-wavelength satellite-based instrument. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 70-80:549–567, 2003.

M. J. Barnsley, D. Allison, and P. Lewis. On the information content of multiple view angle (mva) images. *International Journal of Remote Sensing*, 18:1937–1960, 1997.

Bayes and Price. An essay towards solving a problem in the doctrine of chances. by the late rev mr bayes, f.r.s. communicated by mr. price, in a letter to john canton, a.m.f.r.s. *Philisophical Transactions*, 53:370–418, 1763.

R. Borde, P. Dubuisson, R. Santer, C. Schmechtig, and D. Dassailly. Atbd: Cirrus cloud detection. Technical Report PO-TN-MEL-GS-0005, Laboratoire Interdisciplinaire des Sciences de l'Environment, Maison de la Recherche, 32 Av. Foch BP 59, 62930 Wimereux-France, February 2000.

F. S. Boudala, G. A. Isaac, Q. Fu, and S. G. Cober. Parameterization of effective ice particle size for high-latitude clouds. *International Journal of Climatology*, 22:1267–1284, 2002.

K. N. Bower and T. W. Choularton. A paramerisation of the effective radius of ice free clouds for use in global climate models. *Atmospheric Research*, 27:305–339, 1992.

F. Bréon and S. Colzy. Cloud detection from the spacebourne polder instrument and validation against surface synoptic observations. *Journal of Applied Meteorology*, 38:777–785, 1998.

R. D. Cess, G. L. Potter, J. P. Blanchet, G. J. Boer, A. D. del Genio, M. Déqué, V. Dymnikov, V. Galin, W. L. Gates, S. J. Ghan, J. T. Kiehl, A. A. Lacis, H. le Treut, Z. Li, X. Liang, B. J.

McAvaney, V. P. Meleschko, J. F. B. Mitchell, J. Morcrette, D. A. Randall, L. Rikus, E. Roeckner, J. F. Royer, U. Schlese, D. A. Sheinin, A. Slingo, A. P. Sokolov, K. E. Taylor, w. M. Washington, R. T. Wetherald, I. Yagai, and M. Zhang. Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *Journal of Geophysical Research*, 95:16601–16615, 1990.

P. Chen, R. Srinivasan, and G. Fedosejevs. An automated cloud detection method for daily noaa 16 advanced very high resolution radiometer data over texas and mexico. *Journal of Geophysical Research*, 108(D23):4742, 2003.

F. Chevalier. Sampled databases of 60-level atmospheric profiles from the ecmwf analyses. Research Report 4, EUMETSAT - NWP SAF, December 2001 2001.

I. Chiapello and C. Moulin. Toms and meteosat satellite records of the variability of saharan dust transport over the atlantic during the last two decades. *Geophysical Research Letters*, 29(8): 17–1 – 17–4, 2002.

CMS/Météo-France. Validation report for pge01-02-03 of saf/nwc/msg. *Météo-France / Centre de Mtorologie Spatiale Report SAF/NWC/IOP/MFL/SCI/VAL/01*, version 1.0:87, 2005.

R. Cossu, S. Chaudhuri, and L. Bruzzone. A context-sensitive bayesian technique for partially supervised classification of multitemporal images. *IEEE Geoscience and Remote Sensing Letters*, 2(3):352–356, 2005.

C. Cox and W. Munk. Measurement of the roughness of the sea surface from photographs of the sun's glitter. *Journal of the Optical Society of America*, 44(11):838–850, 1954.

M. F. Cronin, N. A. Bond, C. W. Fairall, and R. A. Weller. Surface cloud forcing in the east pacific stratus deck/cold tongue/itcz complex. *Journal of Climate*, 19:392–409, 2005.

P. H. Daum and Y. Liu. Generalized expressions for effective radius, cloud radiative properties and their application to studies of the first indirect aerosol effect. In *Fifteenth ARM Science Team Meeting Proceedings*, Daytona Beach, Florida, 2005.

M. N. Deeter and K. F. Evans. A hybrid eddington single scatter radiative transfer model for computing radiances from thermally emitting atmospheres. *Journal of Quantified Spectroscopy and Radiative Transfer*, 60:635–648, 1998.

M. Derrien, B. Farki, L. Harang, H. LeGleau, D. Pochic, and A. Sairouni. Automatic cloud detection applied to noaa-11/avhrr imagery. *Remote Sensing of Environment*, 43(3):246–267, 1993.

A. Dybbroe, K. Karlsson, and A. Thoss. Nwcsaf avhrr cloud detection and analysis using dynamic thresholds and radiative transfer modelling. part i: Algorithm description. *Journal of Applied Meteorology*, 44(1):39–54, 2005.

F. D. Eckardt and N. Kuring. Seawifs identifies dust sources in the namib desert. *International Journal of Remote Sensing*, 26(19):4159–4167, 2005.

O. E. Embury. Atsr re-analysis for climate (http://arc.geos.ed.ac.uk), work package 1.1.5, task 8, 2006.

S. J. English, J. R. Eyre, and J. A. Smith. A cloud-detection scheme for use with satellite sounding radiances in the context of data assimilation for numerical weather prediction. *Quarterly Journal - Royal Meteorological Society*, 125(A):2359–2378, 1999.

P. Faugere, P. Le Borgne, and H. Roquet. Realisation d'une climatologie mondialede la temperature de surface de la mer a echelle fine. *La Meteorologie*, 35:24–35, 2001.

M. J. Filipiak and C. J. Merchant. Technical note - sea surface model for reflectance and emissivity. *International Journal of Remote Sensing*, submitted 2008.

N. H. Fletcher. *The Physics of Rainclouds*. Cambridge University Press, 1962.

A. Fouilloux and J. Iaquinta. Assessment of cloud characteristics from satellite observations by means of self-organised neural networks. *Remote Sensing of the Environment*, 66:101–109, 1998.

C. Francoise, A. Brisson, P. Le Borgne, and A. Marsouin. Definition of a radiosounding database for sea surface brightness temperature simulations: application to sea surface temperature retrieval algorithm determination. *Remote Sensing of Environment*, 81:309–326, 2002.

Q. Fu. An accurate parameterization of the solar radiative properties of cirrus clouds for climate models. *Journal of Climate*, 9:2058–2082, 1996.

G. P. Gobbi, F. Barnaba, R. Giorgi, and A. Santacasa. Altitude-resolved properties of a saharan dust event over the mediterranean. *Atmospheric Environment*, 34:5119–5127, 2000.

D. Hall, J. Key, K. Casey, G. Riggs, and D. Cavalieri. Sea ice surface temperature product from modis. *IEEE Transactions on Geoscience and Remote Sensing*, 42(5):1076–1087, 2004.

Q. Han, W. B. Rossow, and A. A. Lacis. Near-global survey of effective droplet radii in liquid water clouds using isccp data. *Journal of Climate*, 7:465–497, 1994.

C. Hennig. Clusters, outliers and regression: fixed point clusters. *Journal of Multivariate Analysis*, 86:183–212, 2003.

E. J. Highwood, J. M. Haywood, M. D. Silverstone, S. M. Newman, and J. P. Taylor. Radiative properties and direct effect of saharan dust measured by the c-130 aircraft during saharan dust experiment (shade): 2. terrestrial spectrum. *Journal of Geophysical Research*, 108(D18), 2003.

R. J. Hogen, C. Jakob, and A. J. Illingworth. Comparison of ecmwf winter-season cloud fraction with radar-derived values. *Journal of Applied Meteorology*, 40(3):513–525, 2001.

B. N. Holben, D. Tanré, A. Smirnov, T. F. Eck, I. Slutsker, N. Abuhassan, W. W. Newcomb, J. S. Schafer, B. Chatenet, F. Lavenu, Y. J. Kaufman, J. Vande Castle, A. Setzer, B. Markham, D. Clark, R. Frouin, R. Halthore, A. Karneli, N. T. O'Neill, C. Pietras, R. T. Pinker, K. Voss,

and G. Zibordi. An emerging ground-based aerosol climatology: aerosol optical depth from aeronet. *Journal of Geophysical Research*, 106(D11):12067–12097, 2001.

Y. X. Hu and K. Stamnes. An accurate parameterization of radiative properties of water clouds suitable for use in climate models. *Journal of Climate*, 6:728–743, 1993.

K. D. Hutchison, J. K. Roskovensky, J. M. Jackson, A. K. Heidinger, T. J. Kopp, M. J. Pavolonis, and R. Frey. Automated cloud detection and classification of data collected by the visible infra-red imager radiometer suite (viirs). *International Journal of Remote Sensing*, 26(21): 4681–4706, 2005.

K. Ide, P. Courtier, M. Ghil, and A. C. Lorenc. Unified notation for data assimilation: operational, sequential and variational. *Journal of the Meteorological Society of Japan*, 75(1B):181–189, 1997.

C. Jakob. Cloud cover in the ecmwf reanalysis. *Journal of Climate*, 12(4):947–959, 1999.

J.-D. Jang, A. A. Viau, F. Anctil, and E. Bartholomé. Neural network application for cloud detection in spot vegetation images. *International Journal of Remote Sensing*, 27(4):719–736, 2006.

I. Jankowiak and D. Tanré. Satellite climatology of saharan dust outbreaks: method and preliminary results. *Journal of Climate*, 5(6):646–656, 1992.

C. E. Junge, E. Robinson, and F. L. Ludwig. A study of aerosols in pacific air masses. *Journal of Applied Meteorology*, 8:340–347, 1969.

D. L. B. Jupp. Directional radiance and emissivity measurement models for remote sensing of the surface energy balance. *Environmental Modelling and Software*, 13:341–351, 1998.

K. Kanani, L. Poutier, F. Nerry, and M. Stoll. Directional effects consideration to improve outdoors emissivity retrieval in the 3-13um domain. *Optics Express*, 15(19), 2007.

L. Kaufman and P. J. Rousseeuw. *Finding groups in data: an introduction to cluster analysis*. John Wiley and Sons, Inc., New York, 1989.

C. A. Knight. Ice nucleation in the atmosphere. *Advanced Colliod Interfacial Science*, 10:369–395, 1979.

A. Kokhanovsky and T. Nauss. Satellite-based retrieval of ice cloud properties using a semianalytical algorithm. *Journal of Geophysical Research*, 110(D19206), 2005.

A. Kokhanovsky, V. Rozanov, J. Burrows, K. Eichmann, W. Lotz, and M. Vountas. The sciamachy cloud products: algorithms and examples from envisat. *Advances in Space Research*, 36:789–799, 2005.

K. T. Kriebel. Albedo of vegetated surfaces: Its variability with differing irradiances. *Remote Sensing of Environment*, 8:283–290, 1979.

J. Li, D. Waliser, J. Jiang, D. Wu, W. Read, J. Waters, A. Tompkis, L. Donner, J. Chern, W. Tao, R. Atlas, Y. Gu, K. Liou, A. Del Genio, M. Khairoutdinov, and A. Gettelman. Comparison of eos mils cloud ice measurements with ecmwf analyses and gcm simulations: initial results. *Geophysical Research Letters*, 32(L18710), 2005.

X. Li and A. H. Strahler. Geometric-optical bidirectional reflectance modelling of the discrete crown vegetation canopy: effect of crown shape and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2):276–292, 1992.

X. Li and A. H. Strahler. A conceptual model for effective directional emissivity from nonisothermal surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2508–2517, 1999.

S. Liang, H. Fang, M. Chen, C. J. Shuey, C. Walthall, C. Daughtry, J. Morisette, C. B. Schaaf, and A. H. Strahler. Validating modis land surface reflectance and albedo products: methods and preliminary results. *Remote Sensing of Environment*, 83:149–162, 2002.

K. N. Liou. Influence of cirrus clouds on weather and climate processes: a global perspective. *Monthly Weather Review*, 114:1167–1199, 1986.

W. T. Liu and C. Gautier. Thermal forcing on the tropical pacific from satellite data. *Journal of Geophysical Research*, 95:13209–13217, 1990.

W. Lucht. Expected retrieval accuarcies of bidirectional reflectance and albedo from eos-modis and misr angular sampling. *Journal of Geophysical Research*, 103(D8):8763–8778, 1998.

S. Mackie, O. Embury, C. J. Merchant, C. Old, and P. N. Francis. Generalised bayesian cloud detection for satellite imagery. part 1: Technique and validation for night-time imagery over land and sea. *International Journal of Remote Sensing*, submitted 2008a.

S. Mackie, O. Embury, C. J. Merchant, C. Old, and P. N. Francis. Generalised bayesian cloud detection for satellite imagery. part 2: Technique and validation for day-time imagery. *International Journal of Remote Sensing*, submitted 2008b.

S. Mackie, C. J. Merchant, and P. N. Francis. Bayesian cloud detection with pre-clustering of imagery. In *The 2006 EUMETSAT Meteorological Satellite Conference*, Helsinki, Finland, 2006.

G. M. Martin, D. W. Johnson, and A. Spice. The measurement and parameterization of effective radius of droplets in warm stratocumulous clouds. *Journal of the Atmospheric Sciences*, 51(13): 1823–1842, 1994.

M. Massons, D. Domingo, and J. Lorente. Seasonal cycle of cloud cover analyzed using meteosat images. *Annales Geophysicae*, 16:331–341, 1998.

K. Masuda, T. Takashima, and Y. Takayama. Emissivity of pure and sea waters for the model sea surface in the infrared window region. *Remote Sensing of Environment*, 24:313–329, 1988.

N. A. McFarlane, G. J. Boer, J. P. Blanchet, and M. Lazare. The canadian climate centre second generation general circulation model and its equilibrium climate. *Journal of Climate*, 5:1013–1044, 1992.

G. M. McFarquhar, S. Iacobellis, and R. C. J. Somerville. Scm simulations of tropical ice clouds using observationally based parameterizations of microphysics. *Journal of Climate*, 16(11): 1643–1664, 2003.

T. McIntire and J. Simpson. Arctic sea ice, cloud, water and lead classification using neural networks and 1.6um data. *IEEE Transactions on Geoscience and Remote Sensing*, 40(9):1956–1972, 2002.

R. D. McPeters, P. K. Bhartia, A. J. Krueger, J. R. Herman, C. G. Wellemeyer, C. J. Seftor, G. Jaross, O. Torres, L. Moy, G. Labow, W. Byerly, S. L. Taylor, T. Swissler, and R. P. Cebula. Earth probe total mapping spectrometer (toms) data products user's guide. NASA Technical Publication 1998-206895, National Aeronautics and Space Administration, Goddard Space Flight Centre, 1998.

C. Merchant, P. Le Borgne, A. Marsouin, and H. Roquet. Optimal estimation of sea surface temperature from split-window observations. *Remote Sensing of Environment*, 112(5):2469–2484, 2008.

C. J. Merchant, O. Embury, P. Le Borgne, and B. Bellec. Saharan dust in nighttime thermal imagery: Detection and reduction of related biases in retrieved sea surface temperature. *Remote Sensing of Environment*, 104:15–30, 2006a.

C. J. Merchant, O. E. Embury, and C. P. Old. Rttovcld version 8: Independent assessment. Technical report, The University of Edinburgh, Institute for Atmos. & Environ. Science, Crew Building, King's Buildings, Edinburgh EH9 3JN, 2006b.

C. J. Merchant, A. R. Harris, E. Maturi, and S. MacCallum. Probabilistic physically-based cloud-screening of satellite infra-red imagery for operational sea surface temperature retrieval. *Quarterly Journal - Royal Meteorological Society*, 131(611):2735–2756, 2005.

C. J. Merchant, L. A. Horrocks, J. R. Eyre, and A. G. O'Carroll. Retrievals of sea surface temperature from infrared imagery: origin and form of systematic errors. *Quarterly Journal - Royal Meteorological Society*, 132:1205–1223, 2006c.

MétéoFrance. Ocean and sea ice saf atlantic sst product manual, 2006.

P. Minnis, D. P. Garber, D. E. Young, R. E. Arduini, and Y. Takano. Parameterizations of reflectance and effective emittance for satellite remote sensing of cloud properties. *Journal of the Atmospheric Sciences*, 55(22):3313–3339, 1998.

J. M. Moran and M. D. Morgan. *Meteorology - The Atmosphere and the Science of Weather*. Prentice Hall, Upper Sadlle River, New Jersey, fifth edition, 1997.

J. J. Morcrette. Impact of changes to the radiation transfer parameterisations plus cloud optical properties in the ecmwf model. *Monthly Weather Review*, 118:847–873, 1990.

S. C. Mossop. The origin and concentration of ice crystals in clouds. *Bulletin of the American Meteorological Society*, 66(3):264–273, 1985.

S. M. Newman, J. A. Smith, M. D. Glew, S. M. Rogers, and J. P. Taylor. Temperature and salinity dependence of sea surface emissivity in the thermal infrared. *Quarterly Journal - Royal Meteorological Society*, 131:2539–2558, 2005.

P. J. R. North. Estimation of aerosol opacity and land surface bidirectional reflectance from atsr-2 dual-angle imagery: operational method and validation. *Journal of Geophysical Research*, 107 (D12), 2002.

C. O'Dowd, M. H. Smith, I. E. Consterdine, and J. A. Lowe. Marine aerosol, sea-salt and the marine sulphur cycle: a short review. *Atmospheric Environment*, 13(1):73–80, 1996.

K. O. Ogunjobi, Z. He, K. W. Kim, and Y. J. Kim. Aerosol optical depth during episodes of asian dust storms and biomass burning at kwangju, south korea. *Atmospheric Environment*, 38: 1313–1323, 2004.

S. Ou and K. Liou. Ice microphysics and climate temperature feedback. *Atmospheric Research*, 35:127–138, 1995.

R. T. Pinker, G. Pandithurai, B. N. Holben, O. Dubovnik, and T. O. Aro. A dust outbreak in sub-sahel west africa. *Journal of Geophysical Research*, 106(D19):22923–22930, 2001.

L. W. Pinkley, P. P. Sethna, and D. Williams. Optical constants of water in the infrared: influence of temperature. *Journal of the Optical Society of America*, 67(494-499), 1977.

j. L. Privette, W. J. Emery, and D. S. Schimel. Inversion of a vegetation reflectance model with noaa avhrr data. *Remote Sensing of Environment*, 58:187–200, 1998.

J. L. Privette, M. Mukelabai, H. Zhang, and C. B. Schaaf. Characterization of modis land albedo (mod43) accuracy with atmospheric conditions in africa. In *Geoscience and Remote Sensing Symposium, 2004*. IGARSS'04, 2004.

R. W. Reynolds, N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang. An improved in situ and satellite sea surface temperature analysis fro climate. *Journal of Climate*, 15:1609–1625, 2002.

E. Ricciardelli, F. Romano, and V. Cuomo. Physical and statistical approaches for cloud identification using meteosat second generation-spinning enhanced visible and infrared imager data. *Remote Sensing of Environment*, 112:2741–2760, 2008.

C. D. Rodgers. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Reviews of Geophysics and Space Physics*, 14(4):609–624, 1976.

R. R. Rogers and M. K. Yau. *A Short Course in Cloud Physics*, volume 113 of *International Series in Natural Philosophy*. Pergamon Press plc, Oxford, England, third edition, 1989.

J. K. Ross. *The radiation regime and architechture of plant stands.* Junk, W., Norwell, Massachuessets, 1981.

W. Rossow, F. Mosher, E. Kinsella, A. Arking, M. Desbois, E. Harrison, P. Minnis, E. Ruprecht, G. Seze, C. Simmer, and E. Smith. Isccp cloud algorithm intercomparison. *Journal of Climate and Applied Meteorology*, 24(9):877–903, 1985.

W. Rossow and R. Schiffer. Isccp cloud data products. *Bulletin, American Meteorological Society*, 72(1):2–20, 1991.

W. B. Rossow, A. W. Walker, and L. C. Garder. Comparison of isccp and other cloud amounts. *Journal of Climate*, 6:2394–2418, 1993.

J. Roujean, M. Leroy, and P. Deschamps. A bidirectional reflectance model of the earth's surface for the correction of remote sensing data. *Journal of Geophysical Research*, 97(D18):20455–20468, 1992.

R. Saunders. An automated scheme for the removal of cloud contamination from avhrr radiances over western europe. *International Journal of Remote Sensing*, 7(7):867–886, 1986.

R. Saunders. Rttov-7 users guide. eumetsat - nwp saf. available online at *http://www.metoffice.gov.uk/research/interproj/nwpsaf/rtm/rttov7_ug.pdf*, 2002.

R. Saunders, P. Brunel, S. English, P. Bauer, U. O'Keeffe, P. Francis, and P. Rayer. Rttov-8 - science and validation report. *EUMETSAT - NWP SAF*, NWPSAF-MO-TV-007 1.6, 2005.

R. Saunders and K. Kriebel. An improved method for detecting clear sky and cloudy radiances from avhrr data. *International Journal of Remote Sensing*, 9(1):123–150, 1988.

J. Schmetz, P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier. Radiometric performance of seviri. *Bulletin of the American Meteorological Society*, pages ES50–ES51, 2002.

Y. Shuai, C. B. Schaaf, A. H. Strahler, J. Liu, and Z. Jiao. Quality assessment of brdf/albedo retrievals in modis operational system. *Geophysical Research Letters*, 35(L05407), 2008.

J. Simpson and T. McIntire. A recurrent neural network classifier for improved retrievals of areal extent of snow cover. *IEEE Transactions on Geoscience and Remote Sensing*, 39(10):2135–2147, 2001.

J. Simpson, A. Schmidt, and A. R. Harris. Improved cloud detection in along track scanning radiometer (atsr) data over the ocean. *Remote Sensing of the Environment*, 65:1–24, 1998.

R. C. J. Somerville, S. F. Iacobellis, and D. E. Lane. Testing cloud-radiation schemes with single-column models and arm observations. In *Ninth ARM Science Team Meeting Proceedings*, San Antonio, Texas, 1999.

R. Spang, J. Remedios, S. Tilmes, and M. Riese. Mipas observation of polar stratospheric clouds in the arctic 2002/2003 and antarctic 2003 winters. *Advances in Space Research*, 36:868–878, 2005.

A. H. Strahler and J. P. Muller. Modis brdf/albedo product: Algorithm theoretical basis document. *Version 5.0*, http://modis.gsfc.nasa.gov/data/atbd/atbd_mod09.pdf, 1999.

J. Stroeve, J. E. Box, F. Gao, S. Liang, A. Nolin, and C. Schaaf. Accuracy assessment of the modis 16-day albedo product for snow: comparisons with greenland in situ measurements. *Remote Sensing of Environment*, 94:45–60, 2004.

I. Tegen and I. Fung. Modelling of mineral dust in the atmosphere: sources, transport and optical thickness. *Journal of Geophysical Research*, 99(D11):22897–22914, 1994.

O. Torres, P. K. Bhartia, J. R. Herman, A. Sinyuk, P. Ginoux, and B. Holben. A long-term record of aerosol optical depth from toms observations and comparison to aeronet measurements. *Journal of the Atmospheric Sciences*, 59:398–413, 2002.

M. . J. Uddstrom, W. R. Gray, R. Murphy, N. A. Oien, and T. Murray. A bayesian cloud mask for sea surface temperature retrieval. *Journal of Atmospheric and Oceanic Technology*, 16(1): 117–132, 1999.

Z. Wan. Modis land-surface temperature algorithm theoretical basis document (lst atbd). version 3.3. technical document NAS5-31370, Institute for Computational Earth System Science, University of California, April 1999 1999.

W. Wanner, X. Li, and A. H. Strahler. On the derivation of kernels for kernel-driven models of bidirectional reflectance. *Journal of Geophysical Research*, 100(D10):21077–21089, 1995.

P. D. Watts, M. R. Allen, and T. J. Nightingale. Wind speed effects on sea surface emission and reflection for the along track scanning radiometer. *Journal of Atmospheric and Oceanic Technology*, 13:126–141, 1996.

B. A. Wielicki, B. R. Barkstrom, E. F. Harrison, R. B. Lee, L. Smith, and J. E. Cooper. Clouds and the earth's radiant energy system (ceres): An earth observing system experiment. *Bulletin of the American Meteorological Society*, 77:853–868, 1996.

A. H. Woodcock. Salt nuclei in marine air as a function of altitude and windforce. *Journal of Meteorology*, 10:362–371, 1953.

X. Wu and W. L. Smith. Emissivity of rough sea surface for 8-13um: Modelling and validation. *Applied Optics*, 36(12):2609–2619, 1997.

K. Wyser. The effective radius in cirrus clouds. *Journal of Climate*, 11:1793–1802, 1998.

Y. Yu, L. Privette, and A. Pinheiro. Analysis of the npoess viirs land surface temperature algorithm using modis data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(10):2340–2350, 2005.

Appendix A

Tests Used to Construct a Cloud Mask 'Truth'

The dynamic threshold testing method used in the construction of both visible- and thermal-cloud 'truths' cloud mask uses 5 tests. For the visible 'truth', thresholds are set for the maximum reflectance for clear pixels at 0.67μm, 0.87μm and 1.6μm (pixels with reflectances above these thresholds are classed as cloud). Using these threshold tests to clear the scene of cloudy or ambiguous pixels, a 1.6μm 'background field' is calculated (through linear interpolation of the clear reflectance values). The image can be magnified, and the background pixel field adjusted over small (10x10 pixels) or large (50x50 pixels) areas, as the user specifies. Pixel classifications from the threshold tests can be reset and a new classification, based on the difference between pixel reflectance at 1.6μm and the background field value, can be made. Similarly, a 0.87μm 'median field' is calculated using only pixels classified by other tests as clear. In the median field, each pixel is assigned the median value of all the clear pixels in the box surrounding it (the box is set by the user to be either 3x3, 5x5 or 7x7 pixels). A 'difference from median field' threshold test can then be applied and used to classify pixels as cloud if they fall above it. Once the user is satisfied with the result of the above-described tests, an 'eye test' can be done. The scene is 201x201 pixels, and is subdivided into 16 smaller scenes, in which the user is able to change the classification of individual pixels. The earlier tests can then be repeated if necessary, but pixels classified explicitly in the 'eye test' will retain the classification from that test, regardless of which side of the thresholds they fall on. The thermal cloud 'truth' mask is made using the same tests, but with 3.7μm, 11μm and 12μm used to set the thresholds for the initial tests, 3.7μm for the background field, and 11μm for the median field. Single channel images from the appropriate channels (visible for the visible truth and vice-versa) were displayed for reference throughout the 'truth' construction, and a false colour composite of the region made from visible imagery was also displayed for the construction of both 'truths'.