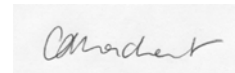|			
Contract Ref : **EUM/CO/13/4600001208/AOC**		**Document Ref** :	**EUM-BC-ATBD-003**
Issue : **Final Issue 1**		**Issue Date** :	**20 Nov 2014**

Project   :   **Preparation for the Implementation of a Bayesian Cloud Detection Scheme**

Title   :   **Bayesian Cloud Detection Algorithm Theoretical Basis Document**

Author   :   *Pearson, Embury, Bulgin*

K. Pearson
O. Embury
C. Bulgin
(University of Reading)

Checked   :

C. J. Merchant

(University of Reading)

Accepted by EUMETSAT:

Anne O'Carroll

Acceptance date: 20 November 2014

AMENDMENT RECORD

This document shall be amended by releasing a new edition of the document in its entirety. The Amendment Record Sheet below records the history and issue status of this document.

AMENDMENT RECORD SHEET

ISSUE	DATE	REASON FOR CHANGE
Draft 1	2 Dec 2013	Initial Version
Draft 2	6 Nov 2014	Section 8 Desirable Future Developments added
1	20 Nov 2014	Final comments addressed

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Purpose and Scope

This Algorithm Theoretical Basis Document (ATBD) describes and justifies the algorithms used in the Bayesian Cloud Detection scheme. The mathematical basis is outlined in Section 3 and the scheme itself is described in Section 4 within the context of its original application to Along-Track Scanning Radiometer (ATSR) data. The adaptations to this scheme that are required for other instruments are discussed in section 5, with particular application to the Advanced Very High Resolution Radiometer on Metop-A (AVHRR-A).

1.2 References

The following documents are referenced in this document:

[1]   Pearson K., Embury O. and Merchant C., Cloud Detection Algorithm Input Output Data Description. Issue 1. Date: 3/12/2013

[2]   Merchant C.J., Embury O., Rayner N.A., Berry D.I., Corlett G.K., Lean K., Veal K.L., Kent E.C., Llewellyn-Jones D.T., Remedios J.J. and Saunders, R. (2012). A 20 year independent record of sea surface temperature for climate from Along-Track Scanning Radiometers, J. Geophys. Res., **117**, C12013: doi:10.1029/2012JC008400

[3]   Embury O., Merchant C.J. and Corlett G.K. (2012). A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: Initial validation, accounting for skin and diurnal variability effects, Rem. Sens. Env., **116**, 62-78: doi:10.1016/j.rse.2011.02.028

[4]   Hocking J., Rayer P., Rundle D., Saunders R., Matricardi M., Geer A., Brunel P. and Vidot J. RTTOV v11 Users Guide. NWPSAF-MO-UD-028. Version 1.2. Date: 24/05/2013. Available from:
http://research.metoffice.gov.uk/research/interproj/nwpsaf/rtm/docs_rttov11/users_guide_11_v1.2.pdf

[5]   Frederic Chevallier, Sabatino Di Michele and Anthony P. McNally. (2006). Diverse profile datasets from the ECMWF 91-level short-range forecasts. EUMETSAT. NWP-SAF-EC-TR-010

[6]   Merchant C.J., Harris A.R., Maturi E. and Maccallum S. (2005). Probabilistic physically based cloud screening of satellite infrared imagery for operational sea surface temperature retrieval, Q. J. R. Meteorol. Soc., **131**, 2735-2755: doi: 10.1256/qj.05.15

[7]   Fox N. (2010). A guide to establish a quality indicator on a satellite sensor derived data product, Ed: Greening M-C. QA4EO, QA4EO-QAEO-GEN-DQK-001
http://qa4eo.org/docs/QA4EO-QAEO-GEN-DQK-001_v4.0.pdf

[8]   GHRSST Science Team (2010). The Recommended GHRSST Data Specification (GDS) 2.0, document revision 4, available from the GHRSST International Project Office, 2011, pp 123
https://www.ghrsst.org/documents/q/category/gds-documents/operational/GDS_2.0r5.doc

[9]   Ackerman S. A., Strabala K. I. , Menzel W. P., Frey R. A., Moeller C. C. and Gumley L. E. (1998). Discriminating clear-sky from clouds with MODIS, J. Geophys. Res., **103**, D24. 32141–32157.

[10]   Stubenrauch C.J., Cros A., Guignard A. and Lamquin N. (2010). A 6-year global cloud climatology from the Atmospheric InfraRed Sounder AIRS and a statistical analysis in synergy with CALIPSO and CloudSat, Atmos. Chem. Phys., **10**, 7197-7214: doi: 10.5194/acp-10-7197-2010

[11]   Birks A., Cox C. and Smith D. SLSTR: Algorithm Theoretical Basis Definition Document for Level 1 Observables. Issue 2.0. Date:14-Jan-2011.   S3-TN-RAL-SL-032
https://sentinel.esa.int/documents/247904/349589/SLSTR_Level-1_ATBD.pdf

[12]   Bulgin C. E., Eastwood S., Embury O., Merchant C. J., and Donlon C. (2014). The sea surface temperature climate change initiative: Alternative image classification algorithms for sea-ice affected oceans. Remote Sensing of Environment (in press).

[13]   Embury O. and Merchant C. J. (2012). A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: A new retrieval scheme. Remote Sensing of Environment, **116**, 32-46.

[14]   Merchant C.J., Harris A.R., Murray M.J. and Zavody A.M. (1999). Toward the elimination of bias in satellite retrievals of skin sea surface temperature. 1: Theory, modeling and inter-algorithm comparison, J. Geophys. Res., **104**, C10, 23565-23578.

[15]   Merchant C.J. and Harris A.R. (1999). Toward the elimination of bias in satellite retrievals of skin sea surface temperature. 2: Comparison with in situ measurements, J. Geophys. Res., **104**, C10, 23579-23590.

[16]   Embury O., Merchant C.J. and Filipiak M.J. (2012). A Reprocessing for Climate of Sea Surface Temperature from the

Along-Track Scanning Radiometers: Basis in Radiative Transfer, Rem. Sens. Env., **116**, 32 - 46, doi: 10.1016/j.rse.2010.10.016

[17]   Simpson J. J. and Gobat J. I. (1996). Improved cloud detection for daytime AVHRR scenes over land, Rem. Sens. Env., **55**, Issue 1, 21-49.

[18]   Bulgin C.E., Sembhi H., Ghent D., Remedios, J. J. and Merchant C. J. (2013). Cloud clearing techniques over land for land surface temperature retrieval from the Advanced Along Track Scanning Radiometer, Int. J. of Rem. Sens, In review.

[19]   Mecklenburg S. (2013). Sentinel-3 User Handbook. Issue 1. Date:2/9/2013
https://earth.esa.int/documents/247904/685236/Sentinel-3_User_Handbook

[20]   Mackie S. (2009). Exploiting weather forecast data for cloud detection, PhD thesis, University of Edinburgh.

1.3   Acronyms

The following acronyms are used in this document:

Term	Definition
(A)ATSR	(Advanced) Along Track Scanning Radiometer
ARC	ATSR Reprocessing for Climate
AVHRR	Advanced Very High Resolution Radiometer
BT	Brightness Temperature
ECMWF	European Centre for Medium-range Weather Forecasting
IR	Infrared
LSD	Local Standard Deviation
LUT	Look-Up Table
NWP	Numerical Weather Prediction
PDF	Probability density function
RTTOV	Radiative Transfer for the Television and Infrared Orbiting Satellite Operational Vertical Sounder
SADIST	Synthesis of ATSR Data Into Sea-surface Temperature

SRF   Spectral Response Function

SST   Sea Surface Temperature

TCWV   Total Column Water Vapour

TOA   Top Of Atmosphere

2.   SCIENTIFIC MOTIVATION

Cloud screening is a fundamental pre-processing step for sea surface temperature (SST) retrieval. Threshold based techniques have often been used to detect cloud, in which spectral and spatial tests have been applied to the data, and only those that fall within an acceptable range of values flagged as clear-sky.

Threshold-setting is something of an art, in which achieving a good true detection rate in one circumstance leads to unsatisfactory false detections in a different context. The selection of the thresholds can be a subjective decision. While they may be chosen by an expert, it is unlikely that other experts would choose identical values. The stringency or otherwise of the thresholds may also be chosen inconsistently between the different tests even by the same expert. Maintenance of expertise can be problematic, since periodic updates to the processing scheme are required or new instruments become available with different spectral characteristics.

Another issue with typical threshold approaches is that the stringency of the cloud-screening is pre-determined by the processing chain and not available to an end user. Different applications of the data will find different tolerances for cloud contamination acceptable in the final SST product. Some users, for example, may value high-accuracy over coverage and others the reverse.

More philosophically, fixed threshold approaches do not make full use of the other background (prior) information that may be available at the time the processing is performed. Rather than assuming a mean climatology in selecting the thresholds, it should in principle be more effective to use an approach that can takes into account the local conditions at the time of the retrieval.

In the Bayesian approach defined in this ATBD, these issues are addressed by calculating a probability of clear-sky to each pixel, which allows the end user to select their own tolerance for cloud contamination depending on their particular application. The Bayesian calculation that assigns this probability also provides a mechanism for including information regarding the meteorological state local to the pixel in question in a systematic and objective way. This calculation also allows for uncertainty and correlations relating to the observations to be incorporated into the scheme in a robust way.

3. MATHEMATICAL DESCRIPTION OF BAYESIAN CLOUD DETECTION SCHEME

In discussing Bayes' theorem, the notation for conditional probability is used. In this, $P(A|B,C)$ is the probability of event A occurring given the assumption that conditions/observations B and C are true. In this notation, cloud detection is the calculation of $P(c|\boldsymbol{y}^o, \boldsymbol{x}^b)$ -- i.e., the probability of clear-sky in the pixel, given the observations $\boldsymbol{y}^o$ and the prior information we have brought to the problem, $\boldsymbol{x}^b$. The Bayesian cloud detection scheme calculates this probability based on satellite observations and prior Numerical Weather Prediction (NWP) information.

Formally Bayes theorem applied to the problem of cloud detection can be expressed as:

$$P(c|\boldsymbol{y}^o, \boldsymbol{x}^b) = \frac{P(\boldsymbol{y}^o|\boldsymbol{x}^b, c)P(\boldsymbol{x}^b|c)P(c)}{P(\boldsymbol{y}^o|\boldsymbol{x}^b)P(\boldsymbol{x}^b)} \qquad (2.1)$$

where

c denotes clear-sky,

$\boldsymbol{y}^o$ is the observation vector and

$\boldsymbol{x}^b$ is the state vector (here, the meteorological conditions).

The assumption is made that the background state is independent of the clear-sky probability at the pixel scale (1x1 km). Writing $P(\boldsymbol{x}^b|c) = P(\boldsymbol{x}^b)$ then, equation (2.1) can be simplified to give:

$$P(c|\boldsymbol{y}^o, \boldsymbol{x}^b) = \frac{P(\boldsymbol{y}^o|\boldsymbol{x}^b, c)P(c)}{P(\boldsymbol{y}^o|\boldsymbol{x}^b)} \qquad (2.2)$$

The probability of the observations given the background state, $P(\boldsymbol{y}^o|\boldsymbol{x}^b)$, can be expressed as the sum of the probabilities of the two possible states (cloud $\bar{c}$ and clear c)

$$P(\boldsymbol{y}^o|\boldsymbol{x}^b) = P(c)P(\boldsymbol{y}^o|\boldsymbol{x}^b, c) + P(\bar{c})P(\boldsymbol{y}^o|\boldsymbol{x}^b, \bar{c}) \qquad (2.3)$$

where $P(\bar{c})$ is the prior probability of cloud and is equal to one minus the prior probability of clear-sky

$$P(\bar{c}) = 1 - P(c) \qquad (2.4)$$

Equations (2.2), (2.3) and (2.4) can be rearranged to give the form of the equation used in the cloud-detection algorithm

$$P(c|\boldsymbol{y}^o, \boldsymbol{x}^b) = \left[1 + \frac{(1 - P(c))P(\boldsymbol{y}^o|\boldsymbol{x}^b, \bar{c})}{P(c)P(\boldsymbol{y}^o|\boldsymbol{x}^b, c)}\right]^{-1} \qquad (2.5)$$

The term on the LHS of equation (2.5) is the required probability of clear-sky given the observations and the assumed background state. Calculating this requires us to calculate suitable values for the three different probability terms on the RHS.

4. CLOUD DETECTION FOR ALONG TRACK SCANNING RADIOMETERS

The (Advanced) Along Track Scanning Radiometer ((A)ATSR) instruments make observations at infrared and visible wavelengths at two viewing angles: the nadir view between 0-22° and the forward view between 52-55°. Both views can be exploited to give additional information for cloud detection purposes. ATSR-1 made measurements in spectral bands centred at 1.6, 3.7, 10.8 and 12 μm, whilst ATSR-2 and AATSR instruments had additional visible wavelength channels centred at 0.55, 0.66 and 0.87 μm. The 10.8 μm is often labelled simply as the 11 μm channel and, together with the 3.7 and 12 μm, described as thermal infrared channels. Although the 1.6 μm channel falls in the near-infrared part of the spectrum, it is often grouped with the shorter wavelength channels as "visible" since their radiance is dominated by reflectance rather than emission processes.

In principle, the method can be applied to any instrument with an arbitrary set of window infra-red and reflectance channels. Specifically, this includes conventional single-view meteorological imagers with channels similar to the ATSRs.

As implemented in the processor, two of the three terms on the RHS of equation (2.5) are obtained solely from pre-calculated look-up tables. The first is the prior probability of clear sky, $P(c)$. The second is the probability for the observation occurring under the given meteorological conditions assuming *cloudy* sky, $P(y^o|x^b, \bar{c})$. The only element of the prior state currently used in this look-up is the SST: no dependence on the elements of x^b describing the meteorological profile is represented in the tables. The third term is the probability for the observation occurring under the given meteorological conditions and assuming *clear* sky, $P(y^o|x^b, c)$. This is dynamically calculated making use of a forward model (here, the RTTOV radiative transfer code) in addition to a look-up table. The details regarding the determination of the three terms are dealt with sections 4.3, 4.4 and 4.5.

4.1 Algorithm Overview

The Bayesian cloud detection scheme calculates a probability of clear-sky for any given pixel based on the satellite observations, prior information about the atmosphere and surface conditions and the respective uncertainties in these variables. Figure 4.1 shows a high level overview of the classification process. The currently implemented Bayesian cloud detection scheme takes European Centre for Medium-range Weather Forecasting (ECMWF) NWP reanalysis data as input in order to simulate clear sky brightness temperatures and top of the atmosphere reflectances. The other inputs are (A)ATSR satellite observations and cloudy PDF

LUTs. The Bayesian cloud detection scheme provides the probability of clear-sky as output on a per pixel basis.

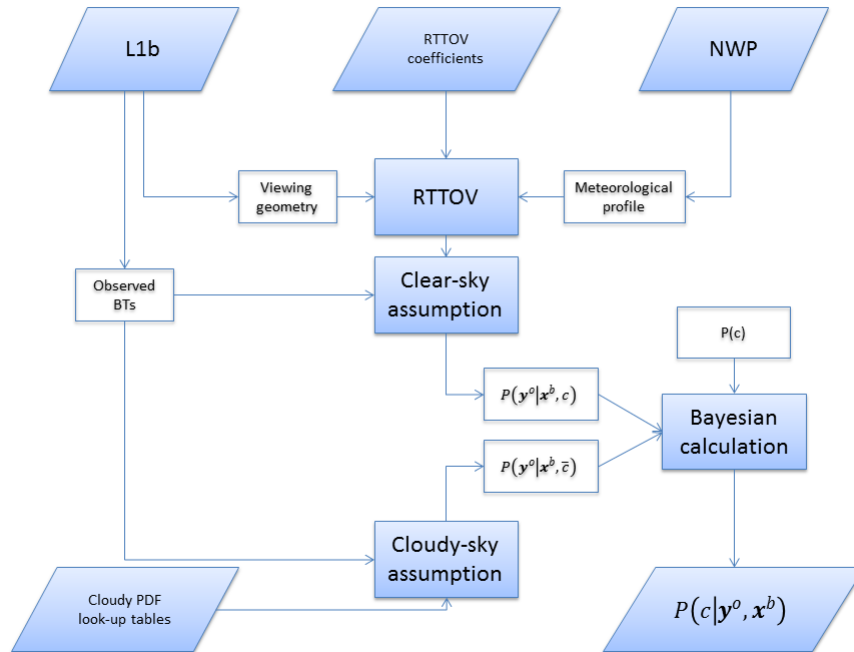**Figure 4.1. Flow chart showing an overview of the cloud detection processing scheme**.

A more detailed overview of the steps involved in the Bayesian classification is given in Figure 4.2. The rest of this section describes the methodology for carrying out the required calculations in detail. The inputs, auxiliary data and outputs are described fully in the corresponding Input Output Data Description document [1].

```
                          ┌─────────────────────────────────┐
                          │ Pixel by pixel image processing │
                          └─────────────────────────────────┘
                                          │
                          ┌─────────────────────────────────────────┐
                          │ Select pixel, allocate and zero pixel structure │◄──────┐
                          └─────────────────────────────────────────┘       │
                                          │                                  │
                          ┌─────────────────────────────────┐               │
                          │    Get image pixel coordinates   │               │
                          └─────────────────────────────────┘               │
                                          │                                  │
   ┌──────┐      Yes      ◇ Is pixel obviously cloud, land, space or ◇       │
   │ Flag │◄──────────────◇            limb?                         ◇       │
   │ pixel│               ◇                                          ◇       │
   └──────┘                               │                                  │
       │                  ┌─────────────────────────────────┐               │
       │                  │ Get pixel's latitude, longitude, solar │         │
       │                  │        & viewing geometry        │               │
       │                  └─────────────────────────────────┘               │
       │                                  │                                  │
       │                  ┌─────────────────────────────────┐               │
       │                  │  Load observation error / NEδT   │               │
       │                  └─────────────────────────────────┘               │
       │                                  │                                  │
       │                  ┌─────────────────────────────────┐               │
       │                  │       Set day/night flag         │               │
       │                  └─────────────────────────────────┘               │
       │                                  │                                  │
       │                  ┌─────────────────────────────────┐               │
       │                  │   Get coordinates of four        │               │
       │                  │   neighbouring NWP tie-points    │               │
       │                  └─────────────────────────────────┘               │
       │                                  │                                  │
       │                  ┌─────────────────────────────────┐               │
       │                  │      Run RTTOV at tie-points     │               │
       │                  └─────────────────────────────────┘               │
       │                                  │                                  │
       │                  ┌─────────────────────────────────┐               │
       │                  │ Find simulated clear-sky BTs and H-matrix at tie-points │
       │                  └─────────────────────────────────┘               │
       │                                  │                                  │
       │         ┌───────────────────────────────────────────────────┐      │
       │         │ Interpolate simulated clear-sky values and H-matrix to pixel location │
       │         └───────────────────────────────────────────────────┘      │
       │                                  │                                  │
       │                  ┌─────────────────────────────────┐               │
       │                  │       Get B and R matrices       │               │
       │                  └─────────────────────────────────┘               │
```

Calculate obs-model differences $(BT^o - BT^b)$

Calculate 3x3 local standard deviation (texture) of 11µm BT

Look up $P(c)$

$P(y_s^o|x^b,\hat{c})P(y_t^o|x^b,\hat{c}),P(y_t^o|x^b,c)$ → Get PDF values from LUTs. Spectral and textural probabilities for cloudy conditions and texture for clear-sky

Calculate $P(y_s^o|x^b,c)$ from simulated values

Calculate $P(c|y^o,x^b)$

Store results in product structure
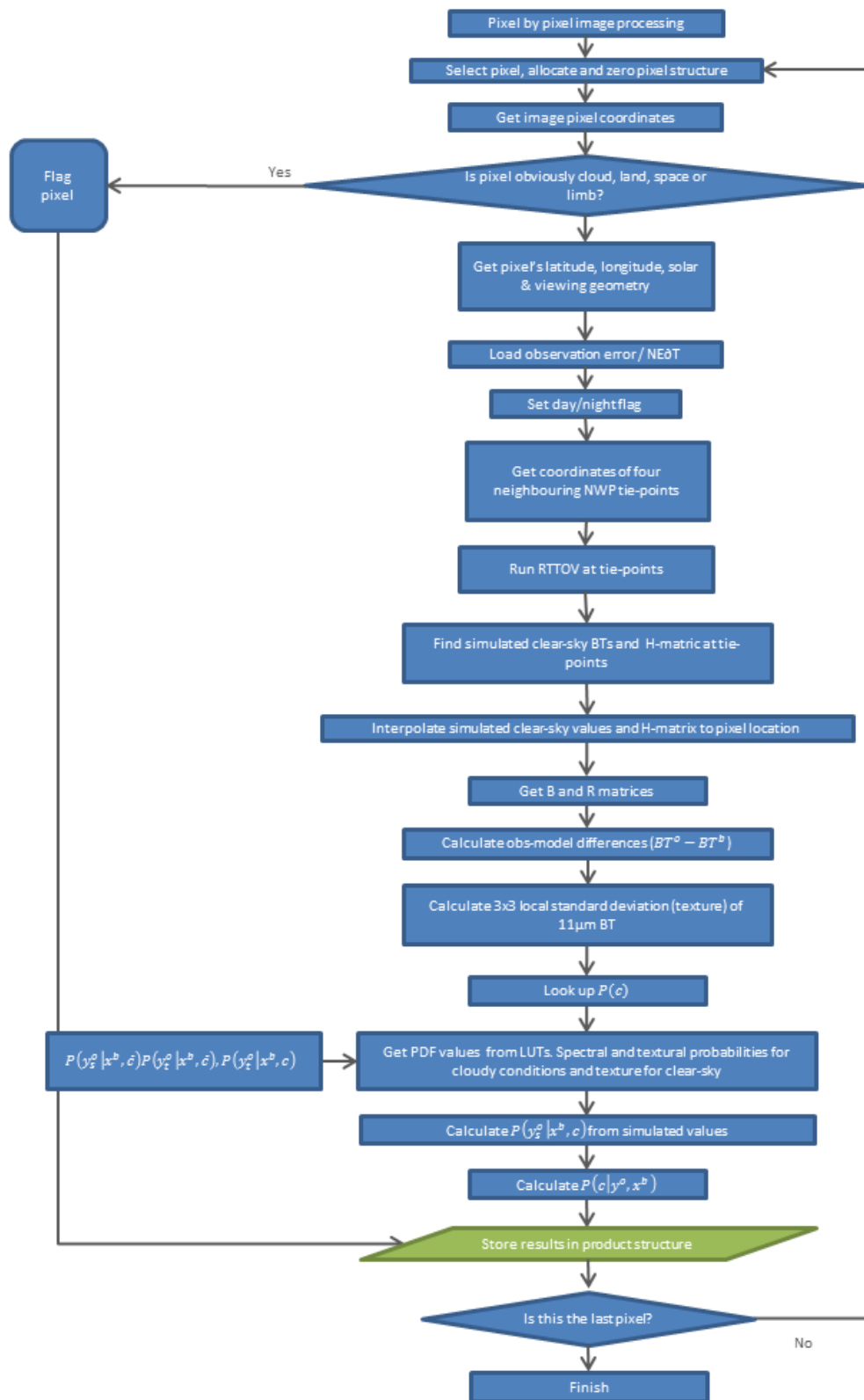
◇ Is this the last pixel? ◇ — No

Finish

Figure 4.2 Bayesian classification steps for calculating clear-sky probability prior to SST retrieval. Blue rectangles denote processing steps, blue diamonds decision making steps and green parallelograms data storage.

4.2   Sensor Data - Brightness Temperature, Reflectance

The (A)ATSR observations used in the Bayesian cloud detection scheme form the observation vector, $\boldsymbol{y}^o$. The subset of channels used in the observation vector is dependent on time of day.

The channels used in the cloud detection algorithm are present on all ATSR sensors and give consistency over the dataset time series. The cloud detection algorithm can use data for the specified channels from both the nadir and forward views when they are available. The observation vector, $\boldsymbol{y}^o$, is defined under day conditions as,

$$\boldsymbol{y}^o = \begin{bmatrix} RE_{1.6} \\ BT_{11} \\ BT_{12.0} \\ LSD_{3X3}(BT_{11}) \end{bmatrix} \tag{4.1}$$

and under night conditions as,

$$\boldsymbol{y}^o = \begin{bmatrix} BT_{3.7} \\ BT_{11} \\ BT_{12.0} \\ LSD_{3X3}(BT_{11}) \end{bmatrix} \tag{4.2}$$

where

BT   denotes brightness temperature,

LSD_{3x3} is the local standard deviation over a 3x3 pixel box,

RE   denotes reflectance and

1.6, 3.7, 11, 12.0   subscripts define the (A)ATSR channel.

As equations (4.1) and (4.2) show, the 11 and 12 μm brightness temperatures are used under all conditions whereas the 1.6 and 3.7 μm channels are switched depending on whether it is a daytime or nighttime observation. Daytime conditions are defined by a solar zenith angle less than 90º[1], and nighttime conditions by solar zenith angles above 90º. These single channel radiances are collectively termed "spectral" information. In addition, the local standard deviation in the 11 μm brightness temperature is used as a "textural" measure of the observational data.

---

[1] *Use of forward modelled visible radiances is challenging at very high solar zenith angles because of the long path length. In pre-RTTOV-11 versions of the Bayesian cloud detection algorithm within the ARC project, a twilight zone of +/-5° was defined where only a minimum thermal channel set was used (i.e., no visible channels and no 3.7 μm channel either). Now that updated reflectance channel modelling is available, the definition and details of a twilight zone should be revisited.*

$$LSD_i = \sqrt{\frac{1}{9} \sum_{\text{9 pixel box}} (y_i^o - \langle y_i^o \rangle)^2} \qquad (4.3)$$

where

y_i^o is the 11 μm brightness temperature for a given observation

$< y_i^o >$ is the mean 11 μm brightness temperature across the 3x3 pixel box.

4.3 Prior Probabilities of Clear-sky - LUT

A LUT was developed for the prior clear-sky probability, which is a function of latitude and longitude. The prior clear-sky probability was calculated by counting the number of clear relative to total number of AATSR pixels for each location, i.e., using the following equation:

$$P(c) = \frac{N_c}{N_c + N_{\bar{c}}} \qquad (4.4)$$

where:

N is the number of occurrences of the given classification at the pixel location

c denotes clear sky

$\bar{c}$ denotes cloud

$P(c)$ was calculated at 1º×1º resolution and the global distribution is shown in Figure 4.3. The maximum prior probability of clear sky is 0.5, seen in the mid-Pacific and south-east Atlantic and Indian Oceans. The Pacific and Atlantic oceans off the west-coasts of South America and Africa are typically cloudier.
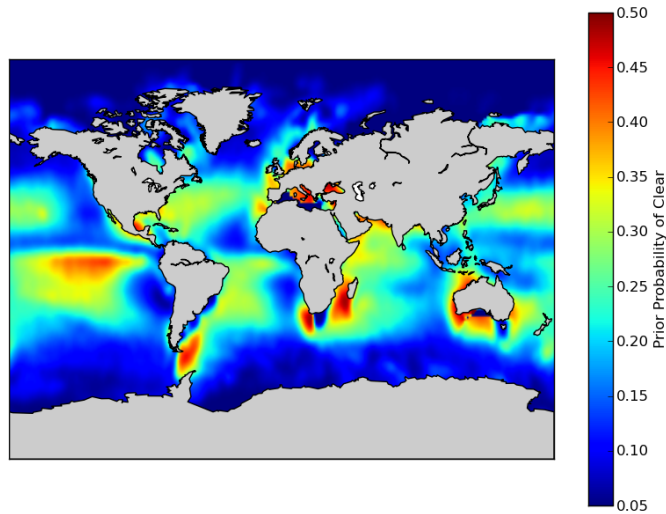
Figure 4.3. Global map of the "prior probability of clear-sky at 1 km resolution" generated from ARC processing of ATSR data.

4.4   Probability of the Observation Given Cloudy Conditions - LUT

The values used to calculate $P(\boldsymbol{y}^o|\boldsymbol{x}^b,\bar{c})$ in equation (2.5) are taken from various LUTs depending on the channels used and the availability or otherwise of both forward and nadir views. The technical specifications of the files storing the LUTs are given in the corresponding Input Output Data Description [1] and the logic identifying which LUTs should be used in what circumstances is set out below. Each LUT contains an empirically generated PDF for a subset of the observation vector under a variety of prior conditions. The PDFs are generated using the entire (A)ATSR time series of observations initially bootstrapped using the operational SADIST cloud mask to identify cloudy pixels. These were subsequently iterated once using the Bayesian cloud detection scheme as part of the ATSR Reprocessing for Climate (ARC) project [2]. As the Bayesian cloud detection scheme performed better than the operational SADIST cloud mask (see [3]), this iteration resulted in less incorrectly classified data contributing to the LUTs. Furthermore, it allowed refinement of the LUT dimensions and the treatment of the ATSR dual-view which complicates cloud detection since it is possible for cloud to be present in either neither, one or both of the (A)ATSR views of the surface.

In all cases the probability is decomposed into spectral and textural components denoted by subscripts 's' and 't' that are assumed to be independent so that

$$P(\boldsymbol{y}^o|\boldsymbol{x}^b,\bar{c}) = P(\boldsymbol{y}_s^o|\boldsymbol{x}^b,\bar{c})P(\boldsymbol{y}_t^o|\boldsymbol{x}^b,\bar{c}) \qquad (4.5)$$

Figure 4.4 shows the textural PDF (using the 11 μm channel) under day- and night-time conditions for the single-view nadir view data. The cloudy PDF is much broader than the clear-sky PDF as cloud surfaces are more

heterogeneous than the underlying sea surface over a 3x3 pixel [3x3 km] surface area. Both the clear-sky and cloudy-sky texture PDFs appear different between the day and night cases. This is mainly due to the difficulty of classifying certain types of scenes in night-time imagery. For example, sea fog has a temperature close to that of the sea surface and is uniform with a low local standard deviation; hence it is easier to detect during the day when the reflectance channels are available. This can been seen in Figure 4.4 (left), where the daytime, cloudy-sky texture PDF has a stronger peak for low values of local standard deviation than the night-time PDF. Similarly, ocean fronts have a high local standard deviation and in night-time imagery they are often misclassified as clouds by automated detection systems. This is seen in Figure 4.4 (right) where the night-time, clear-sky texture PDF is narrower than the day-time one. In order to improve the classifier performance for night-time scenes, the textural PDF LUT generated from daytime data is also used at night-time.

The textural probabilities are extracted from LUTs. If only a single view is available then the textural probability is extracted directly from an array described in Table 4.1. The relevant view LUT is selected on the basis of the satellite zenith angle. If both views are available then the probability for both views is extracted from a joint probability LUT as described in Table 4.2.

Table 4.1. 11 µm single-view, textural PDF array. Note: separate PDFs are generated for clear-sky and cloudy conditions.

Dimension	Unit	Upper limit	Lower limit	Bin size	Number of bins
11 µm texture	*K*	*2.0*	*0.0*	*0.005*	*400*
Satellite zenith angle	*°*	*60.0*	*0.0*	*30.0*	*2*

Table 4.2. 11 µm dual-view, textural PDF array. Note: separate PDFs are generated for clear-sky and cloudy conditions.

Dimension	Unit	Upper limit	Lower limit	Bin size	Number of bins
11 µm nadir texture	*K*	*2.0*	*0.0*	*0.005*	*400*
11 µm forward texture	*K*	*2.0*	*0.0*	*0.005*	*400*
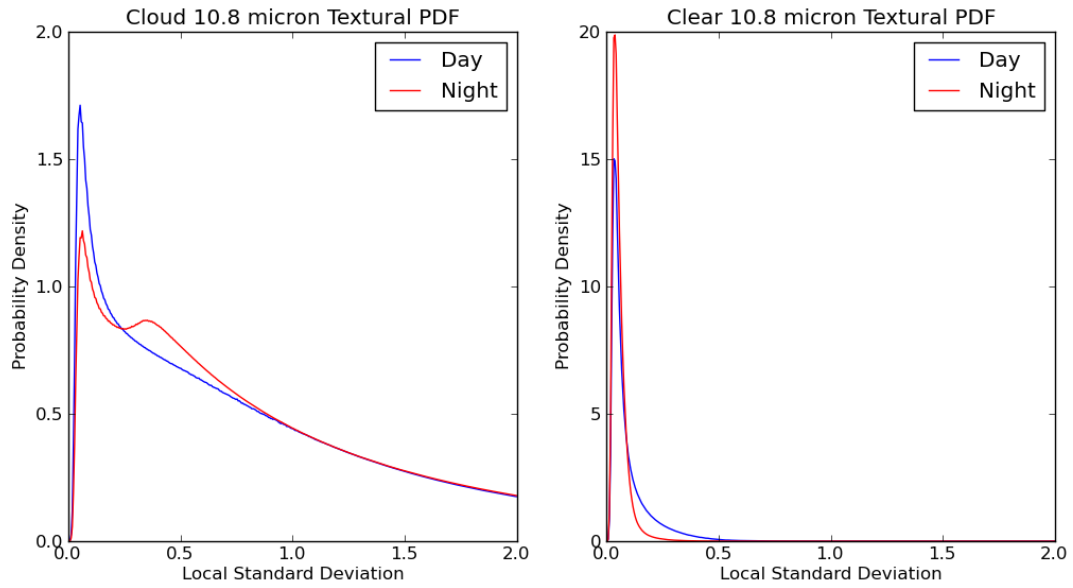
Figure 4.4. Comparison of empirical 11 μm textural PDFs for both cloud and clear-sky observations (nadir view only, as derived from both night time and day time data). Note: day-time derived data are considered more reliable, and are used for detection for day and night.

Obtaining the spectral probability follows a similar process to the textural one of using LUTs appropriate to the available view but it also consults different tables according to whether it is a day or night-time observation. The multiple channels used for the spectral observations makes the detailed process more complex and the four combinations of single/dual view and daytime/night-time are dealt with individually in the following subsections.

4.4.1 Daytime single-view

For single view daytime observations the spectral probability is split into reflectance and thermal components that are assumed to be independent. Thus,

$$P(\boldsymbol{y}_s^o|\boldsymbol{x}^b,\bar{c}) = P(\boldsymbol{y}_{1.6}^o|\boldsymbol{x}^b,\bar{c})P(\boldsymbol{y}_{11,12}^o|\boldsymbol{x}^b,\bar{c}) \tag{4.6}$$

The 1.6 μm probability is obtained from an array described in Table 4.3. The appropriate LUT in which to locate the observation is chosen according to the view selected via the satellite zenith angle and the time of day indicated by the solar zenith angle. The 11 and 12 μm observations are located simultaneously in a joint probability LUT. The specification for this array is given in Table 4.4. The appropriate PDF is selected based on the SST contained in the supplied NWP data, whether it is forward or nadir view. Although not utilised in the current implementation, the array also has a day/night flag from which the appropriate LUT is selected. In order to

maximize the resolution of the LUT in regions of interest, the axes are transformed from $[BT_{11}, BT_{12}]$ to $[(BT_{11} - SST), (BT_{11} - BT_{12})]$.

Table 4.3. Daytime, reflectance, cloudy, single-view, spectral PDF array

Dimension	Unit	Upper limit	Lower limit	Bin size	Number of bins
$RE_{1.6}$		1.00	0.0	0.01	100
Solar zenith angle	°	95.00	20.0	2.5	30
Satellite zenith angle	°	60.00	0.0	30.0	2

Table 4.4. Daytime, two-channel, thermal, cloudy, single-view, spectral PDF array

Dimension	Unit	Upper limit	Lower limit	Bin size	Number of bins
$BT_{11} - SST$	K	10.00	-20.00	1.0	30
$BT_{11} - BT_{12}$	K	9.0	-1.00	0.2	50
NWP SST	K	304.00	271.00	1.0	33
Satellite zenith angle	°	60.00	0.0	30.0	2
Day/Night	°	180.00	0.0	90.0	2

4.4.2 Night-time single-view

For single view night-time observations, the spectral probability $P(\boldsymbol{y}_s^o | \boldsymbol{x}^b, \bar{c})$ comes from a joint probability LUT for the three TIR channels contained in an array described in Table 4.5. The appropriate PDF is selected based on the SST contained in the supplied NWP data, whether it is forward or nadir view and, as the table is used elsewhere, according to the appropriate day/night flag. In order to maximize the resolution of the LUT in regions of interest, the axes are transformed from $[BT_{3.7}, BT_{11}, BT_{12}]$ to $[(BT_{11} - SST), (BT_{11} - BT_{12}), (BT_{3.7} - BT_{11})]$.

Table 4.5 Night-time, cloudy, spectral PDF array

Dimension	Unit	Upper limit	Lower limit	Bin size	Number of bins
$BT_{11} - SST$	K	10.00	-20.00	2.0	15
$BT_{11} - BT_{12}$	K	9.0	-1.00	0.2	50
$BT_{3.7} - BT_{11}$	K	10.00	-6.00	0.2	80
NWP SST	K	305.00	270.00	2.5	14
Satellite zenith angle	°	60.00	0.0	30.0	2

Figure 4.5 shows some slices of the spectral PDF described in Table 4.5 The PDFs are three-dimensional but for visualization purposes have been collapsed along one of these dimensions. They are presented for nadir only data for two different NWP SST values. The 11–12 μm BT is plotted as a function of 11 μm BT–NWP SST in the top panel, and as a function of the 11-3.7 μm BT in the bottom panel. The PDF shape and orientation

shifts significantly between the two NWP SSTs in the slices presented indicating the importance of constraining the PDF as a function of all the constituent dimensions.
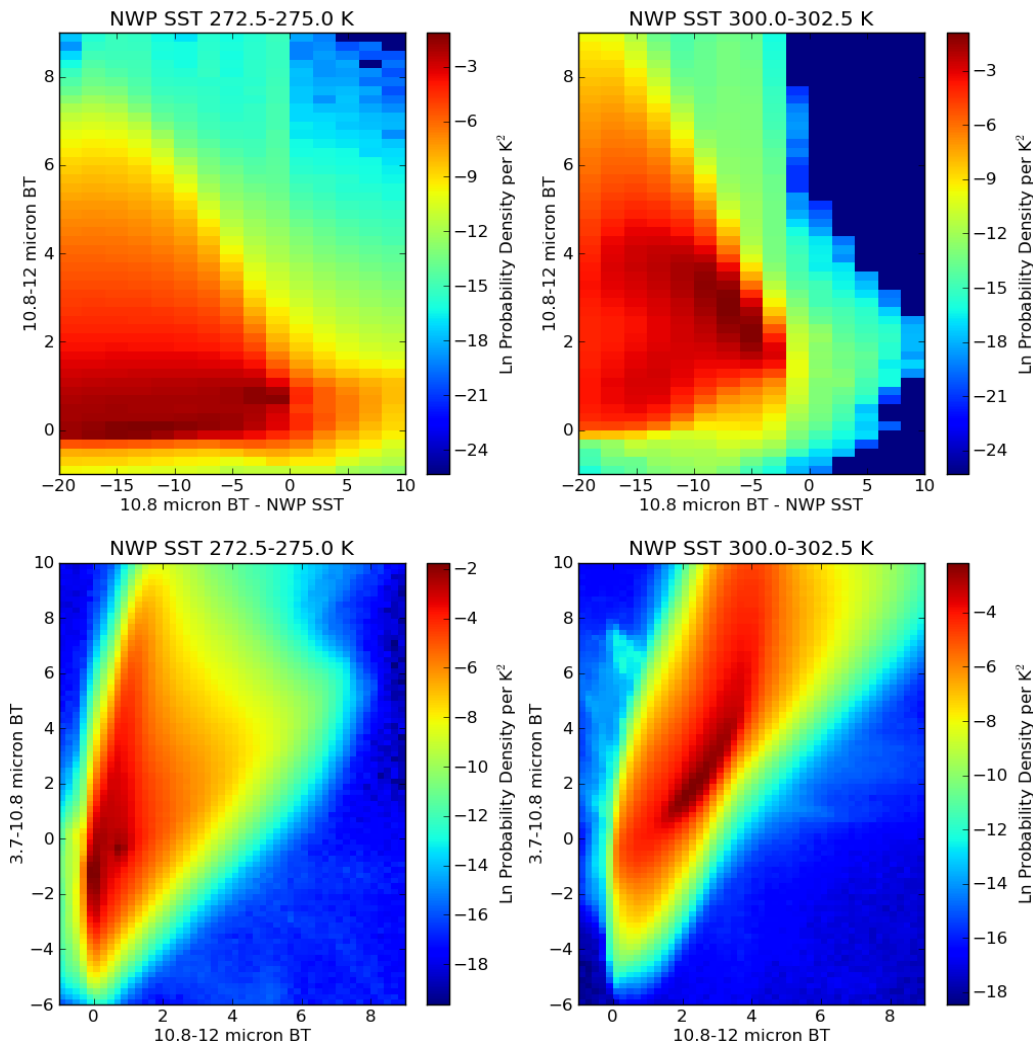
Figure 4.5. Example thermal spectral PDFs for two NWP SST values. The top panel shows the 11 µm BT minus the NWP SST against the 11 minus 12 µm BT. The lower panel shows the 3.7 minus 11 µm BT against the 11 minus 12 µm BT. The PDF shapes show significant variation as a function of SST.

4.4.3 Daytime dual-view

As for the single view case (section 4.4.1), the daytime dual view spectral probability is split into two components, assumed independent, relating to the reflectance and TIR channels (see eqn. 4.6). For the 1.6 µm reflectance channel, the probability of the observations from both views is

obtained from a joint probability LUT. This is contained in an array described in Table 4.6. The appropriate LUT is selected on the basis of the time of day indicated by the solar zenith angle. Figure 4.6 shows a graphical representation of the dual view PDF for the 1.6 μm channel with different solar zenith angles in the two panels. As the solar zenith angle increases the PDF becomes more spread out with a tendency towards higher reflectance in the nadir view. At lower solar zenith angles the 1.6 μm nadir versus forward view PDF is closer to the 1:1 line. The reflectance peak for relatively dark clouds may be the result of partially filled pixels flagged as cloud where some of the darker underlying ocean surface is also visible.

Table 4.6. Daytime, reflectance, cloudy, dual-view, spectral PDF array

Dimension	Unit	Upper limit	Lower limit	Bin size	Number of bins
$RE_{1.6}$ *forward*		*1.00*	*0.0*	*0.01*	*100*
$RE_{1.6}$ *nadir*		*1.00*	*0.0*	*0.01*	*100*
Solar zenith angle	°	*90.0*	*20.0*	*2.5*	*28*
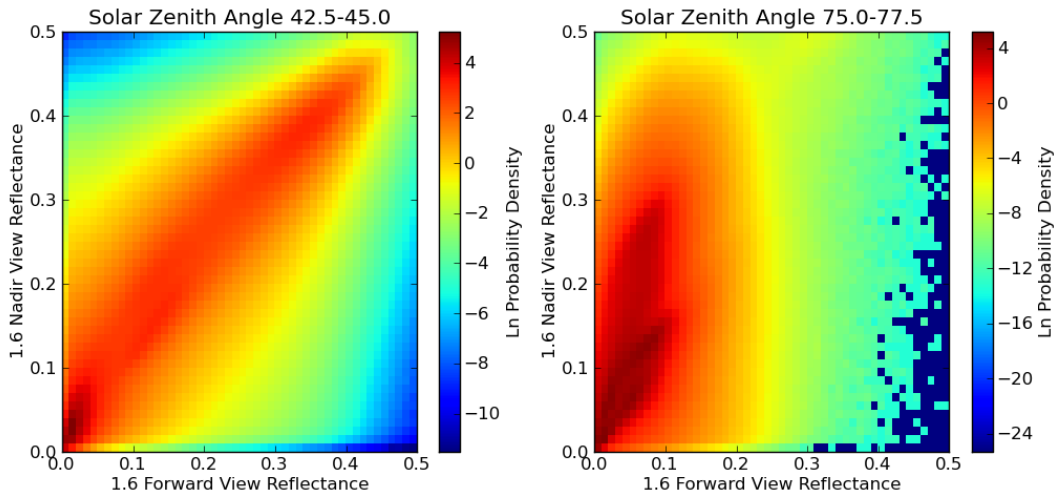
Figure 4.6. Visible spectral dual view PDFs showing the 1.6 μm nadir versus forward view reflectance given cloudy conditions.

The probability of the observations in the two TIR channels in both views are obtained from a 4 dimensional joint probability LUT contained in an array described in Table 4.7. The appropriate LUT is selected based on the SST contained in the supplied NWP data. In order to maximize the resolution of the LUT in regions of interest, the axes are transformed from $[BT_{11,\text{nad}}, BT_{12,\text{nad}}, BT_{11,\text{for}}, BT_{12,\text{for}}]$ to $[(BT_{11,\text{nad}} - SST), (BT_{11,\text{nad}} - BT_{12,\text{nad}}), (BT_{11,\text{for}} - BT_{12,\text{for}}), (BT_{11,\text{nad}} - BT_{11,\text{for}})]$.

Table 4.7. Daytime, two-channel, thermal, dual-view, cloudy, spectral PDF array

Dimension	Unit	Upper limit	Lower limit	Bin size	Number of bins
$BT_{11,\text{nad}} - SST$	K	10.00	-20.00	2.0	15
$BT_{11,\text{nad}} - BT_{12,\text{nad}}$	K	7.0	-1.00	0.4	20
$BT_{11,\text{for}} - BT_{12,\text{for}}$	K	7.0	-1.00	0.4	20
$BT_{11,\text{nad}} - BT_{11,\text{for}}$	K	7.0	-1.00	0.4	20
NWP SST	K	305.00	270.00	2.5	14

4.4.4 Night-time dual-view

The dual view night-time observations make use of the three TIR channels in both forward and nadir views. In principle, we could obtain the spectral probability directly from a 6 dimensional LUT. However, this would require a prohibitively large array to store the information. Instead the probability for each of the two views is obtained separately from the LUT described in Table 4.5. These are then combined according to

$$P(y_s^o|x^b,\bar{c}) = P(y_{s,\text{nad}}^o|x^b,\bar{c})P(y_{s,\text{fwd}}^o|x^b,\bar{c})$$

$$+P(y_{s,\text{nad}}^o|x^b,c)P(y_{s,\text{fwd}}^o|x^b,\bar{c}) \quad (4.7)$$

$$+P(y_{s,\text{nad}}^o|x^b,\bar{c})P(y_{s,\text{fwd}}^o|x^b,c)$$

where the clear sky probabilities $P(y_{s,\text{nad}}^o|x^b,c)$ and $P(y_{s,\text{fwd}}^o|x^b,c)$ are obtained from the radiative transfer modelling as described in the following section.

4.5 Probability of the Observation Given Clear Conditions – Model and LUT

The values used to calculate $P(y^o|x^b,c)$ in equation (2.5) are taken from a combination of a LUT and calculations using the forward model. As in the previous section, the probability is split into spectral and textural components so that

$$P(y^o|x^b,c) = P(y_s^o|x^b,c)P(y_t^o|x^b,c) \quad (4.8)$$

The texture probability is obtained from a LUT of the same specification (but containing different values) to that described previously in Table 4.2. The spectral probability comes from forward modelling carried out using RTTOV 11.

RTTOV 11 is the most recent edition of a fast forward model developed at the EUMETSAT NWP Satellite Application Facility to calculate atmospheric radiative transfer at infrared wavelengths [4]. The model is run at the (A)ATSR geolocation tie-points at a resolution of 25 x 32 km using NWP data from the ECMWF as inputs. The full background state

vector x^b contains all the surface and atmospheric variables that can, in principle, influence the calculated radiance. As represented by the required inputs to RTTOV this is:

$$x^b = \begin{bmatrix} T^b(Z) \\ SH^b(Z) \\ SST^b \\ P^b_{surf} \\ u^b_{10} \end{bmatrix} \tag{4.9}$$

where

T(Z)   is the atmospheric temperature profile,

SH(Z)   is the atmospheric specific humidity profile,

SST   is sea surface temperature

P$_{surf}$ is the surface pressure and

u_{10} is the 10 m horizontal wind vector.

The spectral probability for clear-sky is expressed using a Gaussian distribution as

$$P(y^o_s|x^b,c) = \frac{e^{\left(-\frac{1}{2}\Delta y^T (H^T BH + R)^{-1} \Delta y\right)}}{(2\pi)^{\frac{n}{2}}|H^T BH + R|^{0.5}} \tag{4.10}$$

where *n* is the total number of channels used and $\Delta y = y^b - y^o$ ie. the difference between the true observation vector and that derived from the forward model using the background state vector. The expression $H^T BH$ represents the error covariance in the background state vector propagated through the forward model. The H matrix contains the tangent linear of the forward model

$$H = \frac{\partial y^b_s}{\partial x^b} \tag{4.11}$$

This is an expression of the sensitivity of the forward model to changes in the state vector. RTTOV calculates the tangent linear with respect to the elements of the state vector for the calculated brightness temperature or reflectance. Although generated at all levels for all profile variables only the dominant terms are considered in the subsequent matrix operations and a reduced state vector is used:

$$x^b = \begin{bmatrix} SST^b \\ TCWV^b \\ u^b_{10} \end{bmatrix} \tag{4.12}$$

As an example, for the selection of channels used in the spectral calculation at daytime, the tangent linear matrix is

$$H = \begin{bmatrix} \dfrac{\partial RE_{1.6}}{\partial SST^b} & \dfrac{\partial BT_{11}}{\partial SST^b} & \dfrac{\partial BT_{12.0}}{\partial SST^b} \\[3ex] \dfrac{\partial RE_{1.6}}{\partial TCWV^b} & \dfrac{\partial BT_{11}}{\partial TCWV^b} & \dfrac{\partial BT_{12.0}}{\partial TCWV^b} \\[3ex] \dfrac{\partial RE_{1.6}}{\partial \boldsymbol{u}_{10}^b} & \dfrac{\partial BT_{11}}{\partial \boldsymbol{u}_{10}^b} & \dfrac{\partial BT_{12.0}}{\partial \boldsymbol{u}_{10}^b} \end{bmatrix} \tag{4.13}$$

Under night-time conditions, as would be expected, the 1.6 μm reflectance is replaced by the 3.7 μm brightness temperature. For both day and night, the derivatives with respect to wind speed are set to zero for all channels except the 1.6 μm. (There is an interaction of wind speed and thermal radiance via surface emissivity, but it is negligible compared to the other terms.) The value of $\frac{\partial RE_{1.6}}{\partial SST^b}$ is also zero.

The background error covariance matrix $\boldsymbol{B}$ contains the errors in each of the components in the reduced state vector. These are assumed to be uncorrelated so that the off-diagonal terms are zero and it is thus given by

$$\boldsymbol{B} = \begin{bmatrix} \left(\varepsilon_{SST}^b\right)^2 & 0.0 & 0.0 \\[2ex] 0.0 & \left(\varepsilon_{TCWV}^b\right)^2 & 0.0 \\[2ex] 0.0 & 0.0 & \left(\varepsilon_{\boldsymbol{u}_{10}}^b\right)^2 \end{bmatrix} \tag{4.14}$$

In the current implementation, the ε_{SST}^b component varies with location on a 0.5º×0.5º grid based on an a comparative analysis of ERA-40 and ATSR SST values. The ε_{TCWV}^b component varies with TCWV to represent a fractional error in TCWV.

The matrix $\boldsymbol{R}$ is the error covariance matrix of the differences between the model and observed values. It can be decomposed into two components $\boldsymbol{R}^m$ and $\boldsymbol{R}^o$. The model component, $\boldsymbol{R}^m$, can be expressed as:

$$\boldsymbol{R}^m = \begin{bmatrix} (\varepsilon_i^m)^2 & r^2(\varepsilon_i^m)(\varepsilon_j^m) \\[2ex] r^2(\varepsilon_i^m)(\varepsilon_j^m) & \left(\varepsilon_j^m\right)^2 \end{bmatrix} \tag{4.15}$$

where the diagonal terms describe the forward model error in the given channel and the off-diagonal terms the co-variance in that error. The observational component of this error, $\boldsymbol{R}^o$, is defined as the 'noise' in the observations or noise-equivalent delta brightness temperature (NEdT) in the thermal channels. This is assumed to be diagonal:

$$R^o = \begin{bmatrix} (\varepsilon_i^o)^2 & 0.0 \\ 0.0 & \left(\varepsilon_j^o\right)^2 \end{bmatrix} \quad\quad (4.16)$$

Ideally, NEdT levels appropriate to the observations are reported with the radiances and used; where this is not the case, the values may be fixed to the design or nominal values for the mission. We assume that r^2 is equal to zero for the model error (ie. that the model errors are also uncorrelated between channels) and that therefore the off-diagonal terms of this matrix are zero. Hence,

$$R = R^m + R^o = \begin{bmatrix} (\varepsilon_i^m)^2 + (\varepsilon_i^o)^2 & 0.0 \\ 0.0 & \left(\varepsilon_j^o\right)^2 + \left(\varepsilon_j^m\right)^2 \end{bmatrix} \quad\quad (4.17)$$

4.6 Assumptions Made

Within the Bayesian cloud detection independence is assumed between the infrared and visible channel probabilities of clear and cloud observations. In the context of the reduced state vector, TOA reflectance is assumed independent of prior SST. This assumption is made to simplify the forward modelling. Spectral and textural probabilities are also assumed to be independent allowing the extraction of two pieces of information from the observations.

In the R^m matrix an r^2 value of zero is assumed in the off-diagonal term giving no covariance between channels. In reality there may be a strong error covariance between the 11 and 12 μm channels (and the 3.7 μm channel at night), reflecting common fast radiative transfer approximations. Further research is needed to correctly determine the off-diagonal terms of this matrix.

In the cloud detection scheme, the assumption is made that the pixel will either be 'clear' over ocean or 'not clear'. Sea-ice pixels are unlikely to be well represented in the cloudy PDFs and therefore are more likely to be misclassified.

5.   ADAPTATIONS FOR USE WITH AVHRR-A

There are two major factors to be considered in adapting the cloud detection scheme to other instruments: differences in the spectral response functions (SRFs) of the channels and differences in the viewing geometry. These effects should modify the probability density LUTs that the processor requires to carry out the Bayesian calculation. These LUTs could be regenerated from scratch for each new instrument in an iterative process starting from an adequate existing cloud mask. A less labour and computing intensive method, however, is to reuse the existing LUTs generated from the AATSR instrument with appropriate adaptations for a new instrument. This is possible for instruments with channels sufficiently similar to AATSR's, which includes the AVHRR-series.

Given that the PDFs in the LUTs are highly non-linear, it is very difficult to resample them in an appropriate way to match the new SRFs. Instead an approach is adopted that applies a brightness temperature shift to the BTs to make them "AATSR-like" and the appropriate probability is then extracted as normal from the LUTs. The shift depends, in principle, on the meteorological conditions. The processing scheme set out in Figure 4.1 is therefore modified to the form shown in Figure 5.1.
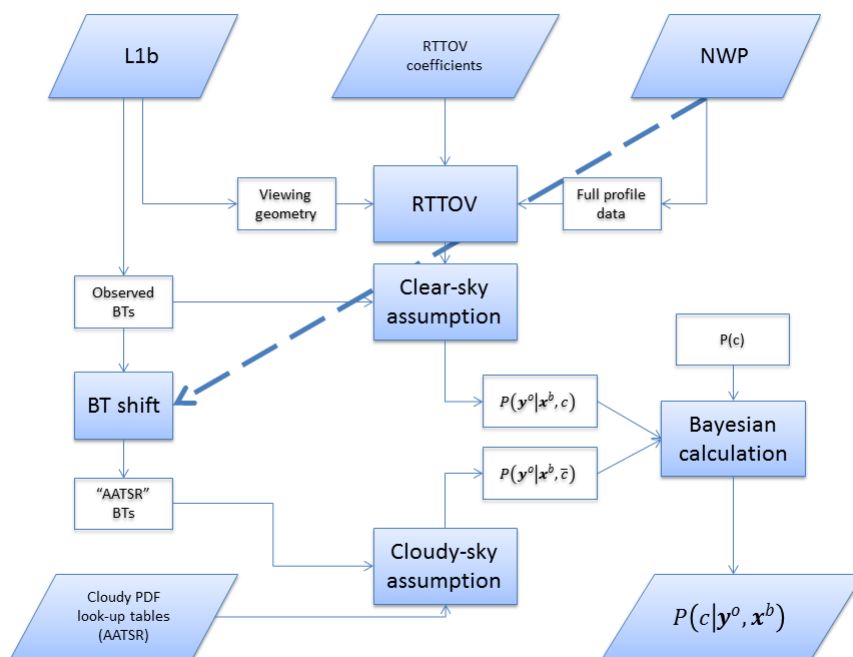
Figure 5.1 Adapted processing scheme including a Brightness Temperature shift

The details of the required adaptations are discussed in the following subsections.

5.1   Viewing Geometry

The (A)ATSR instruments have a dual-view geometry with nadir viewing angles ranging from 0º-22º and a forward view at approximately 55º. In contrast, the AVHRR-A instrument has a single, wide swath with viewing angles ranging 0º-68º. The effect of the differing geometry between the AATSR forward and nadir views is examined in Figure 5.2, Figure 5.3 and Figure 5.4.
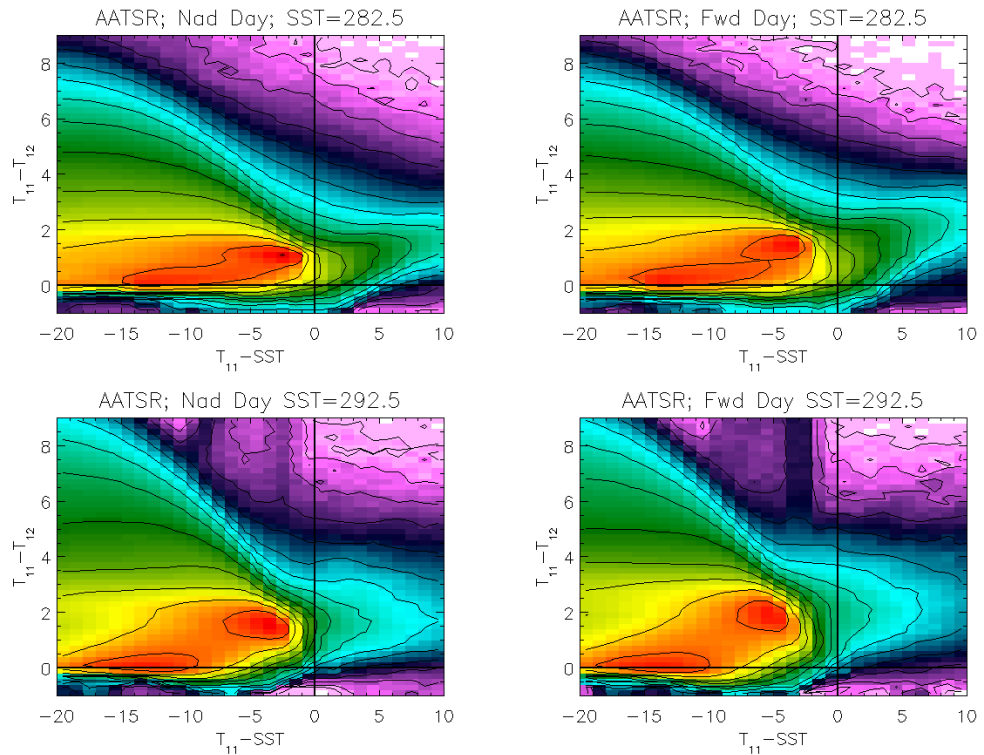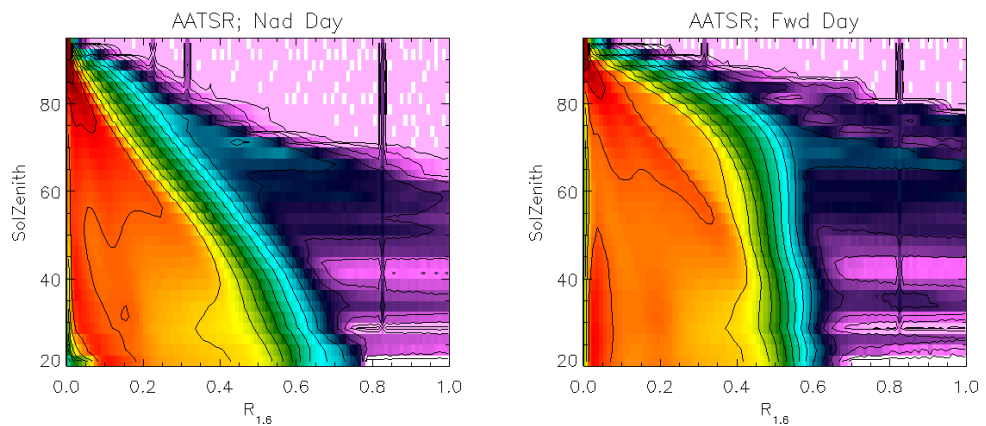

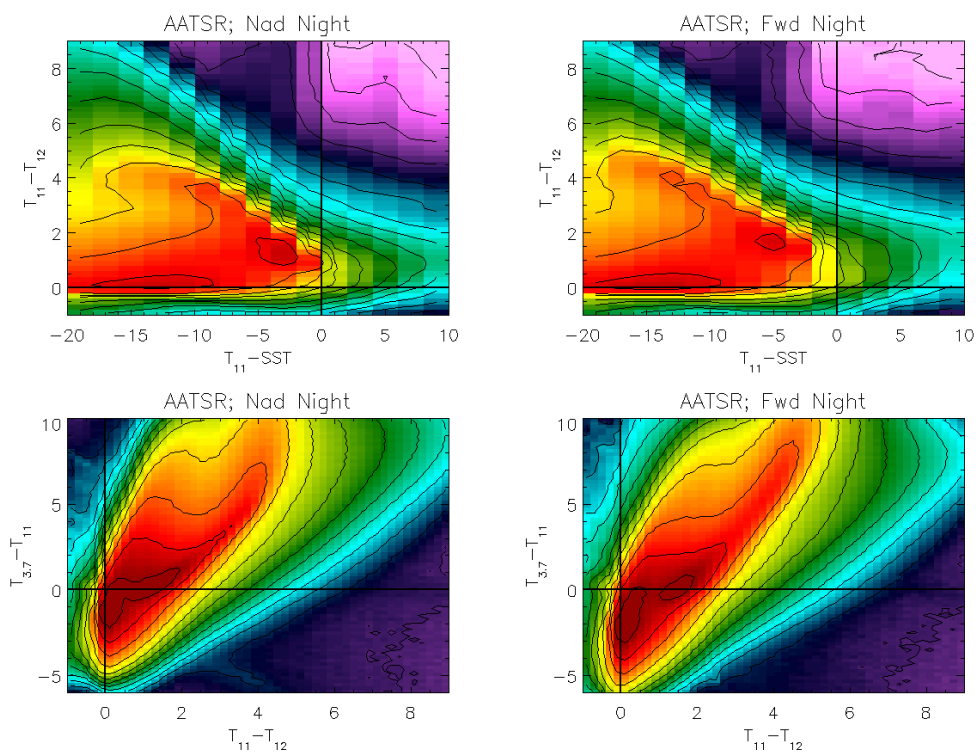**Figure 5.2 AATSR day time IR PDF LUTs**

Figure 5.3 AATSR day time Visible/Near IR PDF LUTs

Figure 5.4 AATSR night-time IR PDF LUTs

For the IR channels, both daytime and night-time PDFs show only minor differences between the two views when plotted, as here, in terms of BT differences. The different viewing geometry of AVHRR-A is therefore accounted for by linear interpolation, based on the path length for the pixel (measured as the secant of satellite zenith angle), between the values extracted from the forward and nadir single-view LUTs. For angles greater than the 55º of the (A)ATSR forward view, the forward-view LUT is used alone (although extrapolation to higher angles should be investigated in future).

5.2   Spectral Response Function

The SRFs of the 3.7 µm, 11 µm and 12 µm channels are compared for AVHRR-A and the ATSR family of instruments in Figure 5.5.
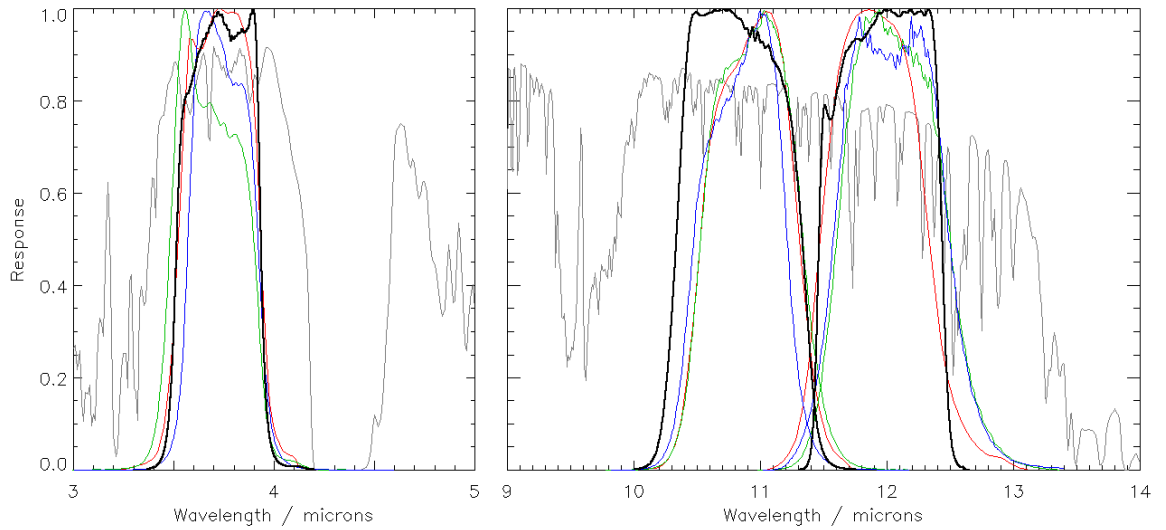
Figure 5.5 Comparison of instrumental SRFs for the 3.7 µm, 11 µm and 12 µm channels for Metop (Black), ATSR1 (Red), ATSR2 (Green) and AATSR (Blue). Light grey shows the atmospheric transmission for standard atmosphere.

The AVHRR-A 3.7 µm channel is very similar to the (A)ATSR 3.7 µm channels, the closest match being with the ATSR1 instrument (red line). For the 11 µm channel the AVHRR-A response is both wider than the (A)ATSR instruments, and weighted towards shorter wavelengths. The AVHRR-A 12 µm channel is centred close to the same wavelengths as the (A)ATSR 12 µm channels, but the AVHRR-A response is closer to a top-hat shape with lower out-of-band response on both sides of the function. For the 3.7 and 12 µm channels the difference between the AVHRR-A and (A)ATSR responses is similar to the differences between the (A)ATSR responses. The AVHRR-A response only stands out from the ATSR family for the 11 µm channel. Overall, the close agreement between the AVHRR-A and (A)ATSR SRFs mean that the AATSR PDF LUTs are usable with the AVHRR-A instrument with only minor modification

Figure 5.6, Figure 5.7 and Figure 5.8 show the difference between simulated AVHRR-A and AATSR BTs for the a range of scenes plotted against TCWV. The two sets of points represent the differences for pixels at nadir and for pixels at a swath angle equivalent to the AATSR forward view angle (55.2º). It is readily apparent that the difference between the BTs measured by the two instruments in each of the channels is a function of TCWV. The shift required by each channel is determined by a quadratic fit to the mean of the nadir and forward cases such that

$$BT_{\mathrm{AVHRR}} - BT_{AATSR} = a + bW + cW^2$$

where W is the TCWV. The aim of the fit is to leave a residual that is smaller than the noise equivalent change in temperature (NEDT). Consequently, the fit for each channel is also plotted on the figures in red along with the NEDT range about this. The fit coefficients are stored in an

ASCII LUT that includes the number of channels to be corrected followed by a line for each listing the channel identity number (6=3.7 µm, 8=11 µm and 9=12 µm) and then the three coefficients for the channel. For AVHRR-A, this contains

3

$6 \quad a_{3.7} \quad b_{3.7} \quad c_{3.7}$

$8 \quad a_{11} \quad b_{11} \quad c_{11}$

$9 \quad a_{12} \quad b_{12} \quad c_{12}$

and the current values of the coefficients are listed in Table 5.1
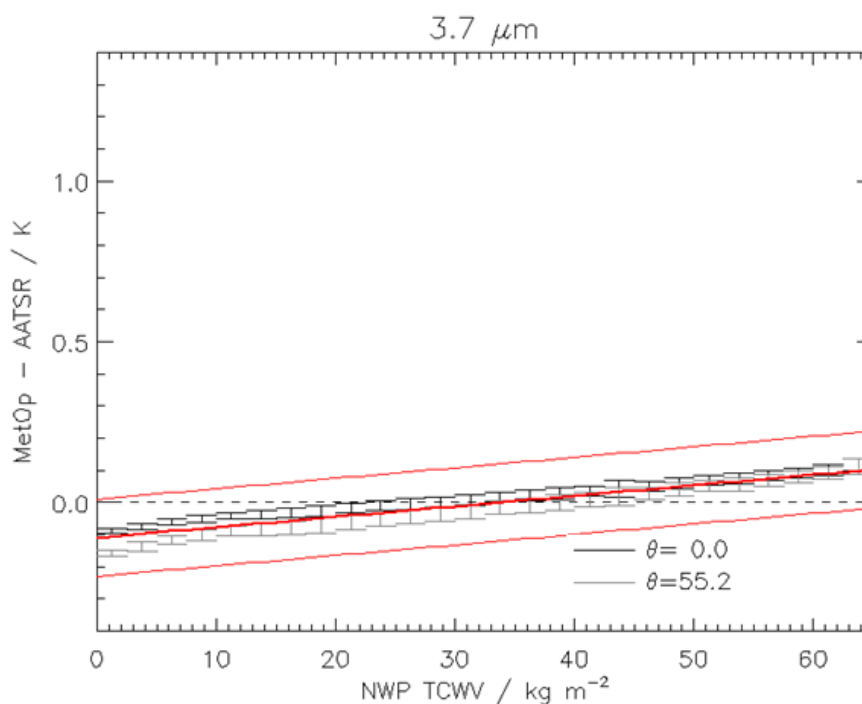
Figure 5.6 AVHRR-A - AATSR view 3.7 µm BT difference plotted against TCWV. Data are plotted for the cases of both instruments measuring at nadir and at the AATSR forward view angle. A quadratic fit to the mean of the two views and the NEDT about this are also shown.
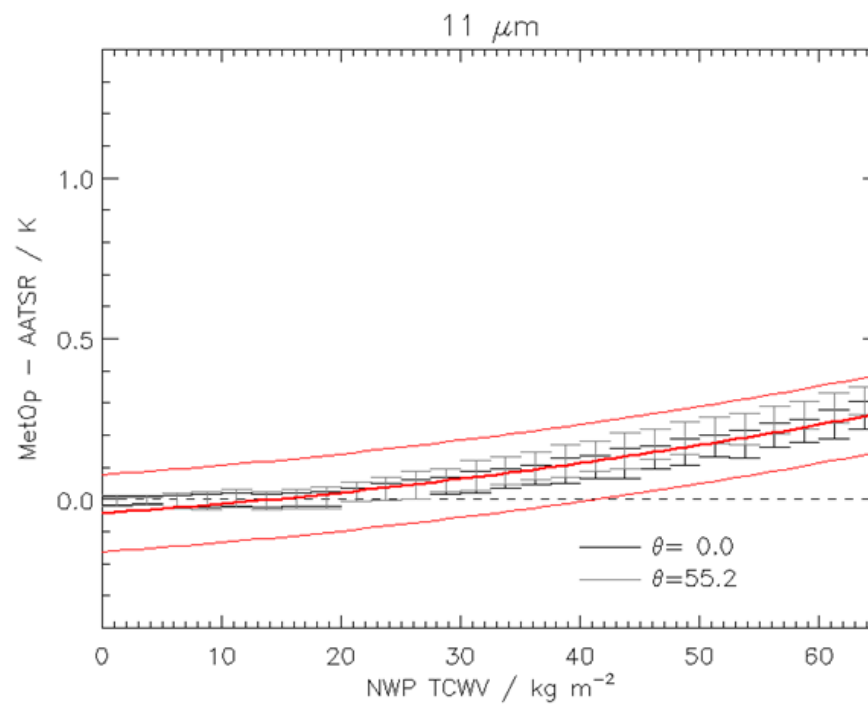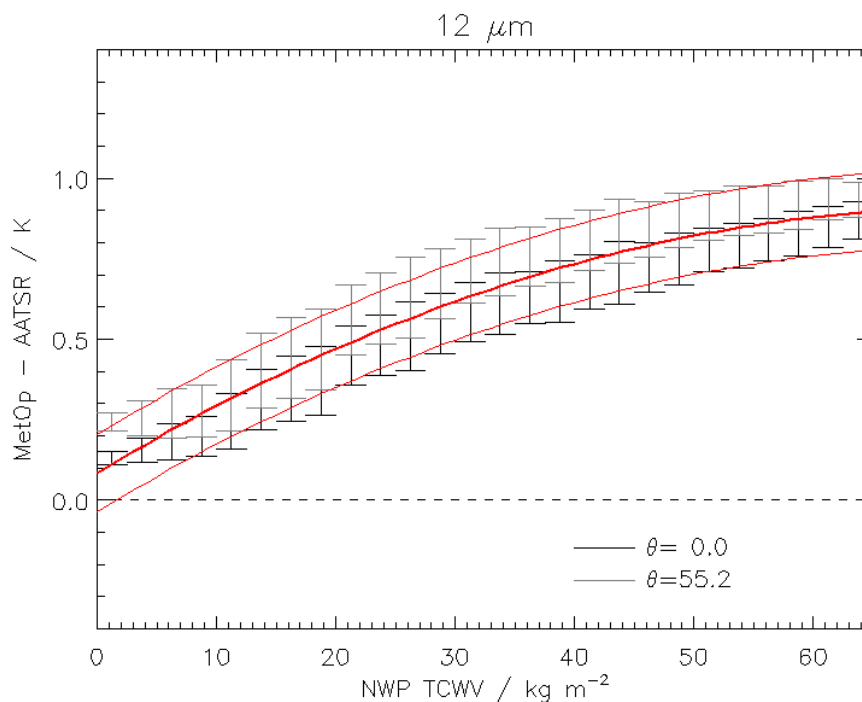
Figure 5.7 AVHRR-A - AATSR nadir view 11 µm BT difference plotted against TCWV. Data are plotted for the cases of both instruments measuring at nadir and at the AATSR forward view angle. A quadratic fit to the mean of the two views and the NEDT about this are also shown.

Figure 5.8 AVHRR-A - AATSR nadir view 12 μm BT difference plotted against TCWV. Data are plotted for the cases of both instruments measuring at nadir and at the AATSR forward view angle. A quadratic fit to the mean of the two views and the NEDT about this are also shown.

Table 5.1 Fit parameters for BT shift against TCWV, for AVHRR-A

Channel ID	Wavelength (μm)	a K	b $K/(kg\ m^{-2})$	c $K/(kg\ m^{-2})^2$
6	3.7	-0.1096	0.003215	7.559×10^{-7}
8	10.8	-0.04289	0.002531	3.439×10^{-5}
9	12	0.08537	0.02222	-1.449×10^{-4}

5.2.1 Summary of LUT shift calculation

The method to generate the required BT shift values for a non-ATSR family sensor can be summarised as follows:

Requirements:

1. NWP profile dataset such as ECMWF 91L q (see [5]) available from: http://research.metoffice.gov.uk/research/interproj/nwpsaf/rtm/profile_data sets.html

2. RTTOV11.1

3. Surface emissivity model: Gbcs Sea Surface Emissivity model (included with the Gbcs code) recommended, otherwise RTTOV built-in model

Steps

1. Filter the NWP profile data set (select ocean profiles with surface temperature > 271.15 K)

2. Run RTTOV to produce TOA BTs for each channel at zenith angles of 0° and 55°:

    i. For AATSR

    ii. For target instrument (e.g., AVHRR-A)

3. Adjust AATSR simulations to account for the 12 µm anomaly (nb. a global 0.2 K adjustment has already been applied in the AATSR LUTs so only the water vapour dependency is required:

    i. $BT_{12,\text{nad}} \rightarrow BT_{12,\text{nad}} - (-0.12 + 0.0034W)$

    ii. $BT_{12,\text{fwd}} \rightarrow BT_{12,\text{fwd}} - (-0.12 + 0.0040W)$

4. For each channel, calculate the quadratic fit to BT differences:

    $$BT_{\text{AVHRR}} - BT_{AATSR} = a + bW + cW^2$$

6.   PROBABILITY THRESHOLD SELECTION

As mentioned in section 2, one of the merits of the probabilistic approach to cloud screening is that it allows the end-user to determine the strictness level of the test they use according to their particular needs. This is done by selecting a probability threshold that balances the competing effects appropriately. Figure 6.1 is reproduced from [6] that described an early form of the Bayesian cloud detection algorithm. Although this used a different observation vector (specifically only the TIR channels and two textural measures), it serves to illustrate the trade-offs that a user must consider.
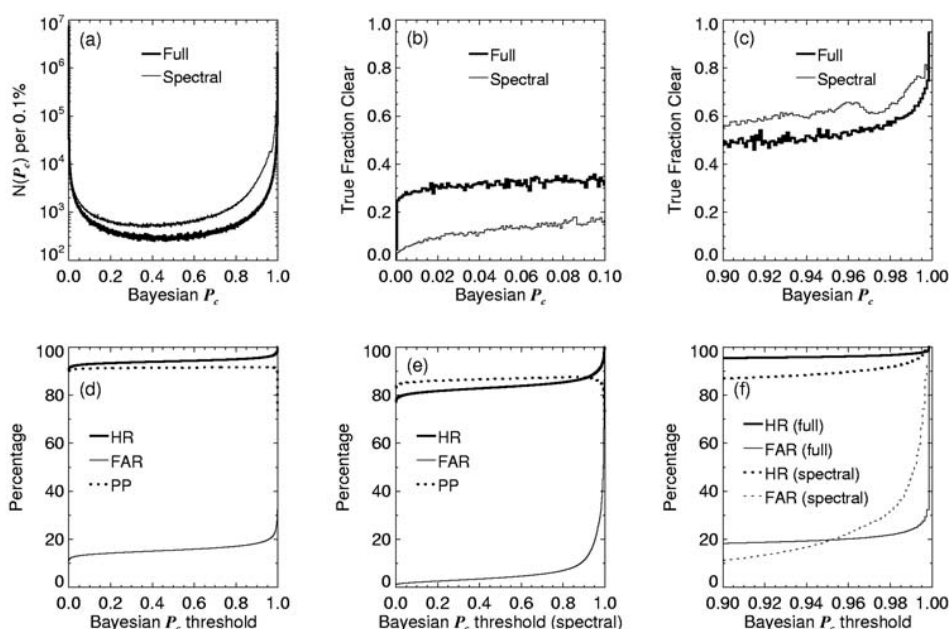
Figure 6.1 Effects of altering clear-sky probability threshold and including only spectral or full probability components from [6]: (a) histogram of the P_c in the pixels, (b) and (c) for each probability bin the fraction of the pixels falling into the bin that truly were clear, (d), (e) and (f) hit-rate (HR), false-alarm rate (FAR) and proportion perfect (PP) as functions of P_c.

Panel (a) shows a histogram of the output probability for clear sky (labelled $P_c = P(c|\boldsymbol{y}^o, \boldsymbol{x}^b)$). The distribution is clearly bimodal with the vast majority of pixels being either highly likely to be cloud or highly likely to be clear. A relatively small fraction have intermediate probabilities. The upturn at either end of the distribution is particularly sharp given the logarithmic nature of the vertical axis. Panels (b) and (c) show the fraction of pixels in each probability bin that truly were clear.

Panels (d), (e) and (f) illustrate the effect that the threshold probability has on the correct identification of cloud or clear conditions. They show the hit

rate for correctly identifying a pixel as cloudy, false alarm rate for identifying a pixel as cloudy when in reality clear and the total proportion correctly identified as clear or cloudy. Such plots give the user information regarding the degree of contamination that is likely to occur in subsequent retrievals and could be repeated for new sensors. In practice however, it is very difficult to determine what the true cloud mask is against which a new scheme is to be tested. If the scheme is replacing an existing method then differences between the two serve only to highlight that they are indeed different rather than which is better or worse. One could repeat the process of [6] by determining a true cloud mask for test scenes using expert human analysis. In practice, however, a more practical solution is to examine retrieval validation statistics as shown schematically in Figure 6.2. Here statistics for the difference between the retrieved SSTs and buoy measurements SSTs are plotted against threshold probability. At low P_c, the mean difference is large and negative as many cloudy pixels remain in the retrieval. As P_c increases, the magnitude of the difference initially reduces sharply as most of the cloud pixels are rapidly removed .There is then a steady decrease as the difference approaches an irreducible bias value characteristic of the instrument in question. The standard deviation similarly starts out large, initially reducing rapidly but then asymptotes towards a minimum value. The initial rapid removal of cloud pixels is reflected in an initial sharp reduction in the area available for SST retrievals (expressed as coverage). A steady decrease in coverage continues until a second period of rapid removal occurs at high P_c.

The task of choosing the threshold value is a compromise between accurate and uncontaminated retrievals and having sufficient data for the end purpose. Choosing a very high value for P_c may be attractive to a user seeking a climate use for example but placing it too high will also eliminate many valid observations from consideration. By contrast, a forecasting system may be more tolerant of modest biases arising from modest cloud contamination, and prefer greater coverage to greater certainty of cloud-free conditions. Experience suggests that a suitable range for the threshold is $0.5 < P_c < 0.95$. Given the similarities between the channels used by (A)ATSR and AVHRR highlighted in section 5 and the reuse of existing LUTs in the current implementation, a value of $P_c = 0.9$ as used in the ARC project is perhaps a reasonable threshold for climate purposes.
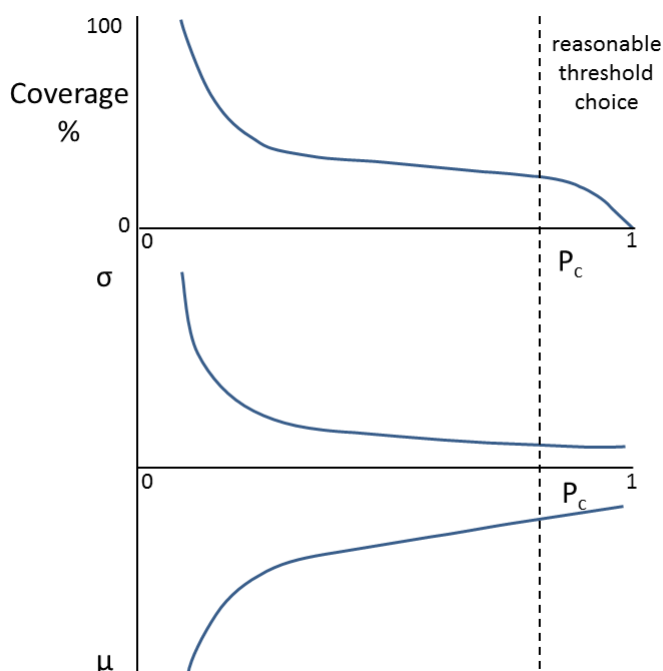
Figure 6.2 Schematic illustration of the evolution of the mean difference between SST retrieval and observed buoy SST values (μ) and the standard deviation (σ), as a function of selected clear-sky probability threshold (P_c). Also shown is the resulting effect on the area covered by the SST retrieval and an indication of a reasonable regime for selecting a rejection threshold for climate purposes.

The recommended default value of probability threshold for application with a new sensor is 0.5. During an initial period "cloudy" and "clear-sky" matches with validation data should be collected. Several months of (pre-)operational application should allow a sufficient set of matches with validation to be built up. Validation statistics and coverage should be determined as a function of the threshold, to inform a judgement about the trade-off of coverage against validation statistics most suitable for applications of the product.

7.   COMPUTATIONAL CONSIDERATIONS

The underlying Generalized Bayesian Cloud Screening (GBCS) code that implements the algorithm set out above was originally intended for reprocessing activities. Therefore, the code is designed to be single threaded with modest memory requirements so that multiple instances can easily be run in parallel (eg. via a script). Running on a Dual-CPU Intel Xeon E5-2650L system, the GBCS processor takes approximately 13 minutes to process one orbit of AVHRR data. However, the same system would be capable of processing 16 orbits in the same 13 minutes assuming adequate I/O bandwidth.

Of the 13 minutes processing per-orbit, 1 minute is used for pre-processing including: staging the input L1b file to the processing area, and extracting and interpolating ERA-interim data. The remaining 12 minutes is spent executing the main Fortran program which performs the cloud screening and generates the output netCDF file.

In many cases, especially when multiple instances are operating in parallel, the processor can become I/O bound and so has several features to help alleviate this:

- The Fortran program can read gzip compressed EUMETSAT Polar System (EPS) format files directly, thereby eliminating the need to unzip them before processing. If disk I/O is not a bottleneck, then the driving Python script may be modified to unzip the files during the pre-processing step instead.

- The Python scripts create a temporary processing area in the system temporary directory (/tmp). Usually this will be mapped to a local disk drive; however, on Linux based systems, it can be redirected to the shared memory area which will be far faster than a disk-based processing area.

- The Fortran program can output data in either netCDF3 or netCDF4 formats, with netCDF4 offering the option of internal (gzip) compression. This option reduces the required disk I/O at a cost of increased CPU usage. Alternatively, the Python scripts can compress the resulting output with gzip/bzip2 during post-processing. This option is useful to reduce the required network bandwidth when generating netCDF3 format outputs.

There is scope to optimize the main Fortran program by improving the memory layout of arrays etc. which could reduce the CPU time requirement by ~50%; however, this would only be of benefit on systems which do not become I/O bound.

8. DESIRABLE FUTURE DEVELOPMENTS / RESEARCH

8.1 Use of Bayesian probability for SST quality level determination

Quality indicators in products provide users with additional guidance about use of geophysical values. Guidance on how to "establish a quality indicator on a satellite sensor derived data product" given by the QA4EO project [7] is very general. For SST products, the quality level is defined by the Group for High Resolution SST (GHRSST) Data Specification [8] as a value expressed on a scale of 0 to 5 as follows: "0 no data, 1 (bad/unusable data e.g. cloud, rain, to close to land), 2 (worst quality usable data), to 5 (best quality usable data)". Within this framework, data producers are free to define their own approach to population of the quality level field, according to their own context and insight.

SST cannot be generally retrieved from pixels in infra-red imagery where significantly affected by clouds or aerosols, since these perturb (or destroy) the relationships between BTs and SST on which the inversion to SST is based. However, when the pixel-averaged optical depth of cloud or aerosol is small, the perturbation to the radiances can be subtle, and SST retrievals are made, albeit with reduced quality. Bayesian cloud detection essentially tests the plausibility that a set of BTs corresponds to the truly clear-sky conditions under which SST retrieval should perform best. Thus, it is reasonable to explore the degree to which the Bayesian probability can inform quality level (QL) assignment.

To address this, we need to be clear about what aspect of "quality" the QL refers to. There are two possible points of view. First, we could take the view that the QL should reflect the uncertainty in the SST, with QL = 5 reflecting the SSTs with smallest total uncertainty. This implies the existence of (preferably) a context-specific means to estimate uncertainty for each SST, or the existence of (at a minimum) some heuristics to discriminate more and less certain data. The validity of the QL assignment can then be explored by demonstrating in validation that the statistics for lower QLs are inferior to those for higher QLs. However, if a context-specific uncertainty estimate is available with the SSTs in the product (which is best practice), the QL on this view is then essentially redundant information.

Therefore, a second view of the QL is perhaps more useful, in which the QL indicates the confidence we have in the SST and its associated uncertainty estimate. In this view, *an SST with relatively large uncertainty is still of good quality if we are confident that the uncertainty estimate is realistic / valid*. This would be the situation where we are confident that the pixel is truly clear sky with no unusual instrument noise, atmospheric anomalies, etc – but the retrieval is for some reason intrinsically less certain (e.g., it is at high satellite zenith angle, so that the atmospheric transmittance is low). Conversely, in this approach, the QL would be low

even if the uncertainty estimate is relatively small, but something in the context indicates that we should have less confidence that all sources of error are properly accounted for in the uncertainty estimate. This could be, for example, because the instrument is not behaving nominally, because there is aerosol in the field of view that is not accounted for, or because there is an elevated risk of cloud contamination in the pixels (while being treated in the retrieval as if they are clear sky pixels).

The rest of this discussion therefore takes the second point of view when considering QL information. The Bayesian probability relates to the chance of cloud in the observed pixel, and therefore provides one component that could be considered in assigning QL.

When using Bayesian cloud detection, a probability of clear-sky (i.e., of nominal retrieval conditions) is returned, given the observations and contextual information. The usual next step is to define a threshold on the probability, above which SSTs are generated (see section 6). Instead of this "all-or-nothing" approach, a series of thresholds could be chosen that map onto quality levels. (This doesn't preclude further steps to modify the QL because of other factors – for example, if the instrument data suggest a period of increased noise in the sensor observations.)

To explore whether this is a useful approach, the follow steps can be identified.

1. Devise a metric or metrics for assessing the effectiveness of Bayesian-based QLs. An effective QL indicator will indicate our confidence in the realism of the SST and its uncertainty estimate. For the highest QL, we would expect SSTs to be consistent with validation data to within the estimated uncertainties (of both satellite and validation SST). In other words, we would expect minimal rates of outliers in validation. As we move to lower thresholds of probability (and therefore lower quality levels) we should see increased incidence of inconsistency, i.e., more outliers. Suitable metrics will therefore quantify this for different bands of clear-sky probability.

2. Define a method, based on the above metric, for selecting a coherent set of clear-sky probability thresholds for the QL assignment. For example, this could be done on the basis of estimated outlier rates in different threshold bands.

3. Apply the above approaches to one or more specific sensors, e.g., Metop-A AVHRR. This requires a matchup dataset (MD) of satellite SSTs to validation data with known uncertainty characteristics, in order to be able to quantify the satellite SST outlier rates. To this MD, an implementation of Bayesian cloud detection needs to be applied. This in turn requires the MD to support relevant forward modeling of BTs, and, importantly to contain both cloud and partly

cloudy matches, not pre-filtered as clear-sky only. Since the cloud detection includes some spatial coherence testing, the MD content needs to include image extracts, not only single pixels. Thus, the MD may require some preparation if a suitable dataset is not pre-existing. Characterize and document the outcomes on the MD, including the conclusion on appropriate QL thresholds.

4. Apply the QL thresholds to a sample of orbit data for each sensor. Characterize the spatio-temporal distribution of QLs and ensure it is satisfactory for user applications. For example, it would be unsatisfactory if there were regions where the QL was never QL = 5: this would need to be identified and understood. Iterate step 3 if necessary and document the outcomes.

5. Progress arising from the above should be reported in the context of GHRSST, in particular to the ST-VAL working group.

8.2 Exploitation of higher resolution visible channels

Present meteorological sensors such as AVHRR have a nadir pixel resolution of order 1 km in both the infra-red and visible channels. However, other present and planned imagers have higher resolution in at least some reflectance channels. This raises the possibility that the Bayesian cloud detection could be rendered more sensitive to sub-pixel cloud within 1 km pixels by exploiting reflectance variability within those pixels.

The reflectance detectors on the future SLSTR are on a focal plane assembly which projects at nadir as shown in Figure 8.1. SLSTR channels S4 to S6 have detectors in both green and blue locations in the figure – i.e., there are 8 detectors at 500 m resolution, 4 of which nominally lie within a given pixel for channels S7 to S9. The wavelengths of the relevant channel centres are given in Table 8.1.
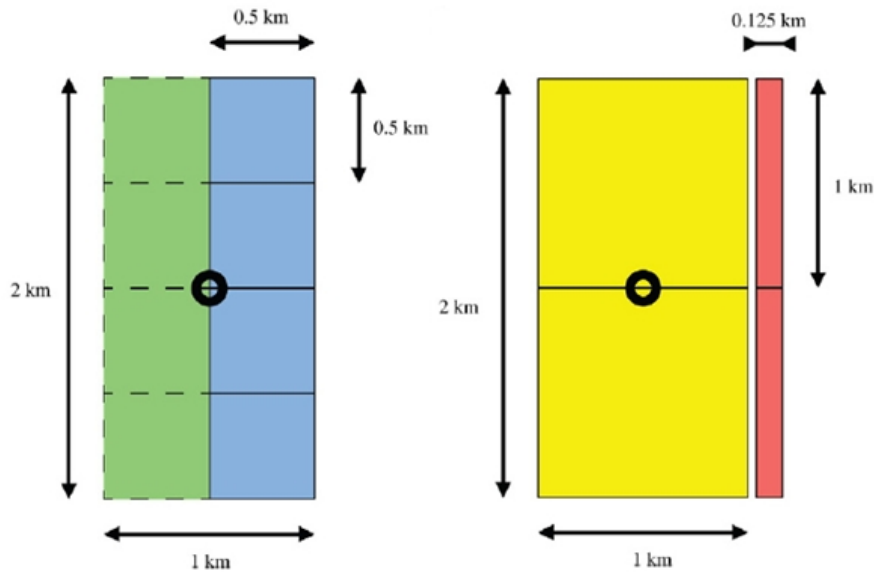
Figure 8.1 Left: Visible and shortwave ground resolution, from https://sentinel.esa.int/web/sentinel/user-guides/sentinel-3-slst, accessed 15/1/2014. Right: thermal (yellow) and fire channel (red) ground resolution.

Table 8.1 Wavelengths and stated uses of selected SLSTR channels

Channel	Wavelength (µm)	Stated use
S4	1.375	For cirrus cloud detection over land
S5	1.61	For cloud detection, ice, snow and vegetation monitoring
S6	2.25	For cloud detection and vegetation monitoring
S7	3.74	SST and LST
S8	10.85	SST and LST
S9	12	SST and LST

The Moderate-Resolution Imaging Spectroradiometer (MODIS) also has 1 km thermal channels and higher-resolution reflectance channels, whose properties are given in Table 8.2. There is therefore a reasonable correspondence between some SLSTR and MODIS channels:

SLSTR S4 $\leftrightarrow$ MODIS B5

SLSTR S5 $\leftrightarrow$ MODIS B6

SLSTR S6 $\leftrightarrow$ MODIS B7

Table 8.2 Properties of selected MODIS channels (from http://eoweb.dlr.de:8080/short_guide/D-MODIS.html, accessed 15/2/2014)

Primary Use	Band	Central wavelength *nm*	Bandwidth *nm*	Spatial resolution *m*
Land / Cloud / Aerosols / Boundaries	*B1*	*645*	*620 - 670*	*250*
	B2	*858.5*	*841 - 876*	
Land / Cloud / Aerosols Properties	*B3*	*469*	*459 - 479*	*500*
	B4	*555*	*545 - 565*	
	B5	*1240*	*1230 - 1250*	
	B6	*1640*	*1628 - 1652*	
	B7	*2130*	*2105 - 2155*	

with the closest similarity being the 1.6 μm wavelengths. This wavelength (at 1 km resolution) is also used in the current Bayesian formulation for cloud detection, being present on AVHRRs and all ATSRs. Therefore, prior to the launch of Sentinel-3, approaches using 500 m reflectance could be prototyped on MODIS B6.

Two obvious questions for exploration are:

1. Can a spatial variability measure be formulated using 4 x 500 m 1.6 μm reflectances that gives a useful additional discrimination within the Bayesian cloud detection, particularly of clouds that are sub-pixel at 1 km?

2. If yes, can this measure replace the current 3 x 3 local standard deviation measure for day-time cloud detection – or is it best used as an additional measure to inform the clear-sky probability for the central pixel of the 3 x 3?

One possible set of steps to examine this issue is identified below.

1. Review of radiometric and geometric information for 500 m reflectance channels on MODIS and SLSTR. This review should establish: the reflectance noise when viewing a uniform scene, stratified by the reflectance over the full geophysical range, if variable; the geometric overlap when viewing at off-nadir angles (and how that is regridded into products, if applicable).

2. Analysis of radiometric and geometric information obtained for simple scenarios of cloud-related reflectance variability at 500 m scales, for nadir and off-nadir situations. This is to develop enough understanding to inform choices of (the) variability measure(s) to test, and to highlight any obvious limitations on what is possible.

3. Define possible measures for testing:, e.g., standard deviation reflectance across four 500 m pixels; range; normalised range, etc.

4. Select test cases of MODIS imagery at 500 m resolution. These should include cases of: scattered fair-weather cumulus fields; marine stratocumulus; disorganised shower clouds behind a mid-latitude cold front; Saharan dust aerosol; sun-glint over a ocean with a spatially variable wind field; high zenith angles as well as near-nadir. Use local thresholds (perhaps using more than one channel, perhaps even using the 250 m imagery if available) define by expert interpretation to determine a 'truth' mask at 500 m resolution for interesting sections of the images. For these cases, calculate variability measures at 2 x 2 pixel resolution, and assess which are most discriminating of cloud identified in the reference mask.

5. Generalise from the insights gained in steps 1. to 4. to define a trial Bayesian formulation to test on full MODIS orbits. Again, a 'truth' cloud mask is required for comparison. This could be the MODIS mask, although expert inspection to validate it for the orbits of tested would be necessary. Variants in which the new measure is additional to and replaces the existing spatial variability measure should be tried and compared, using hit-rate, false-alarm-rate and true   skill   score   statistics.

6. If there are positive results: define how any successful formulation could be more formally implemented in the Bayesian processor for SLSTR, from a software engineering point of view.

8.3   Use of SLSTR channel at 1.375 μm for ocean cloud detection

As noted above, SLSTR will have a reflectance channel centred at 1.375 μm, which has no precedent on AVHRR or ATSR instruments, and is therefore not part of the channel set used by current Bayesian cloud detection code. There is a similar channel on MODIS (1.38 μm, 1 km nadir resolution, Band 26). The stated purpose of the S4 channel is thin cirrus cloud detection over land. This wavelength is useful in this regard because surface-leaving radiance can be substantially attenuated by water vapour absorption below the height of the cirrus cloud, providing a dark background against which the cirrus cloud reflectance stands out [9].

Thin cirrus cloud detection is also important over the ocean with such clouds pervasive in the inter-tropical convergence zone. One contributing mechanism is vertical wind shear in the vicinity of the tops of deep convective clouds. The anvils which form at the top of such clouds comprise ice particles, and wind shear can spread these horizontally far from the convective tower. Figure 8.2 shows the prevalence of high, semi-transparent, ice cloud as a fraction of all cloud types from a 6-year climatology derived from AIRS [10]. In areas of tropical convergence, 10%

to 40% of all cloud cover observed fell into this category, which we may take as roughly corresponding to "thin cirrus".
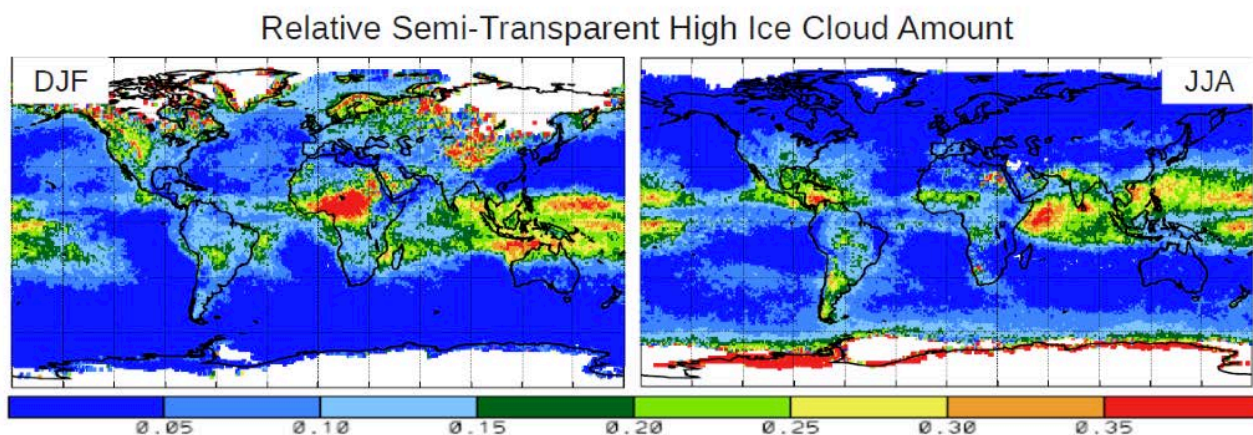

Figure 8.2 the prevalence of high, semi-transparent, ice cloud as a fraction of all cloud types from a 6-year climatology derived from AIRS [10].

It is not clear that the current Bayesian formulation is missing thin cirrus that is significant for SST retrieval in terms of causing retrieval biases – it has not been systematically assessed. To be significant for SST retrieval, any undetected cloud needs to perturb the BTs in a manner that is distinct from the effect of varying the TCWV as this is already accounted for in the retrieval. So, a possible scenario is that thin cirrus is missed by the cloud detection scheme, but leads to small errors because the radiometric impact across the three SST retrieval channels is either small in absolute terms, or is similar to the BT variability caused by TCWV. The actual situation is not currently known.

If use of the 1.375 μm channel increased the sensitivity of the Bayesian detection to thin cirrus significantly, there could be a large impact on the fraction of observations currently found to be "clear sky", particularly in convective zones where clear-sky fractions are already relatively low. Since use of reflectance information is limited to day-time cloud detection, this would likely increase the disparity between day and night-time cloud detection rates. Moreover, from a climate perspective, such an effect would introduce a significant inhomogeneity between SLSTR and the ATSR series.   If there is benefit to greater day-time sensitivity to thin cirrus, the natural approach would be to use both the 1.6 and 1.38 μm channels together as the reflectance component of the Bayesian calculation.

One possible set of steps to examine this issue is identified below.

1.   Establish whether thin cirrus cloud is missed by the present Bayesian method. There are various means by which this could be done. First, ATSR imagery should be inspected to locate thin high

cloud that is not detected. ATSR is advantageous for this, because stereo viewing of images can be used to locate clouds unambiguously as high or low, and to estimate the cloud height. Stereo viewing also increases visual sensitivity to subtle reflectance/BT features. By visually identifying high-level clouds and then switching to a view with the Bayesian cloud mask, the extent of missed thin high cloud can be assessed. Second, Bayesian detection could be applied to MODIS images (without the 1.38 μm cirrus test), and the results compared with the 1.38 μm cirrus mask. To do this, the Bayesian detection needs to be adapted to MODIS using similar techniques to that applied to Metop A in this ATBD.

2. Assuming there is evidence of undetected thin cirrus, the next step is to assess the SST retrieval impact. One way to do this is to develop local BT correlations against 1.38 μm reflectance over scenes with relatively uniform SST and TCWV. BT changes associated with the thin cirrus optical depth changes could then be inferred. These inferred BT changes would then be propagated through an SST retrieval process, to see if they have a significant SST impact. The inferred BT changes can also be compared to $\frac{\partial BT}{\partial TCWV}$ values from simulations, to ascertain the degree of similarity of BT impact between cirrus and water vapour. The following steps would be a priority only if the failure to detect thin cirrus was significant and is found to impact SST retrieval.

3. Develop a joint probability density function for 1.38 μm and 1.6 μm given (all) cloudy conditions, for example by using the MODIS cloud mask over ocean as a starting point for one or more iterations. Develop the capacity to simulate the 1.38 μm reflectance under clear-sky conditions (including in sunglint), and the joint 1.38 μm and 1.6 μm clear-sky PDF. Implement these in a variant of the Bayesian code. Run and compare results with an implementation without the additional channel. The Bayesian mask including the 1.38 μm channel can be compared to the MODIS cirrus mask from step   1.

4. Compare day and night-time cloud detection rates and SSTs regionally and globally with and without the 1.38 μm channel using Bayesian cloud detection and MODIS data. This will give a first approximation of the potential impact on the climate data record from the ATSR series in applying this cloud detection approach to SLSTR data.

8.4 Use of the 3.7 μm channel for ocean cloud detection during the day

Viewed at 3.7 μm, clouds may be colder than or warmer than the surface, according to whether the clouds reflect significant solar spectral radiance towards the sensor.

At the same time, this thermal channel is relatively transparent (compared to 11 and 12 μm) with respect to water vapour, and thus is *a priori* more tightly constrained given clear-sky conditions than the other thermal channels. It is therefore a powerful channel to use in night-time cloud detection, and it is worthwhile to consider whether it can be exploited during the day.

This possibility has become more feasible since the release of RTTOV 11, which accommodates simulation of the solar radiance stream in addition to thermal emission.

One possible set of steps to examine this issue is identified below.

1. Assess day-time simulation of 3.7 μm clear-sky observations using RTTOV 11. This needs to be done for angles remote from and close to specular reflection. The additional uncertainty in the simulation needs to be quantified and parameterised, perhaps as a function of how close the viewing geometry is to the specular geometry. Define limits (if any) beyond which the 3.7 μm BT simulation is not usable.

2. In addition to using $\frac{\partial BT}{\partial SST^b}$ and $\frac{\partial BT}{\partial TCWV^b}$ (which are used at present), explore how to exploit $\frac{\partial BT}{\partial \boldsymbol{u}_{10}^b}$ for the 3.7 μm channel in evaluating the clear-sky PDF (inclusive of solar irradiance).

3. Define and derive an empirical cloudy-sky PDF, bootstrapped using the existing Bayesian cloud detection, including the day-time 3.7 μm. This will involve changing the dimensions of the PDF compared to the night-time dimensions, to accommodate the cases of warm 3.7 μm from clouds.

4. Implement code changes and test new Bayesian processing compared to configurations not using the 3.7 μm channel during the day.

8.5 Use of SLSTR channel at 2.25 μm for ocean cloud detection

The SLSTR Level 1 ATBD [11] proposes adding to the baseline cloud detection for SLSTR a histogram test based on 2.25 μm. The details of this are not yet specified (as far as we can ascertain). The MODIS channel nearest in wavelength (B7) is not used in the standard MODIS cloud detection [9]. Ocean surface and cloud reflectance are similarly contrasting at 1.6 and 2.25 μm, so it is not obvious (over ocean) *a priori* that this

channel adds useful discrimination if the 1.6 μm is already used. Thus, there is no clear reason to make assessing use of this channel a priority.

8.6 Extension to a 3-way classifier for sea-ice areas

Cloud detection at high latitudes is complicated by the presence of sea-ice. At infrared wavelengths, ice can be difficult to distinguish from open water due to surface temperatures close to the freezing temperature. At VIS wavelengths, new or partially submerged ice has low surface reflectance, similar to open water. Ice melt in summer months can lead to the formation of melt ponds on the ice surface giving mixed pixel composition at 1 km spatial scales difficult to classify in remotely sensed data. Misclassification can lead to sea-ice pixels erroneously flagged as open water which in turn affects SST retrievals. Typically, in high latitude regions additional thresholds are placed on SST retrieval on the basis of a prior ice concentration field, which can lead to significant losses of clear-sky data.

The work of Bulgin *et al.* (2014) [12] demonstrated benefits of extending the Bayesian classifier to include a third class, using the 1.6, 11 and 12 μm channels to classify daytime imagery. Bayesian probabilities were calculated offline for clear-sky, sea-ice and cloudy observations using RTTOV 10 to simulate sea-ice conditions. The paper evaluated algorithm performance under a number of configurations and found that using a textural measure based on the standard deviation of the 1.6 μm reflectance over a 3x3 pixel box (replacing the previously used 11 μm measure under daytime conditions) gave a performance similar to the ARC algorithm in terms of the number of cloud and ice observations misclassified as clear [13], whilst significantly increasing the number of clear-sky observations correctly identified. Inclusion of the 0.6 μm channel in the Bayesian scheme increased the ice detection skill with the potential for this to be used to derive an ice surface temperature data record from AATSR and SLSTR under daytime conditions. An example of the classifier performance under the new configurations is shown in the figure below.
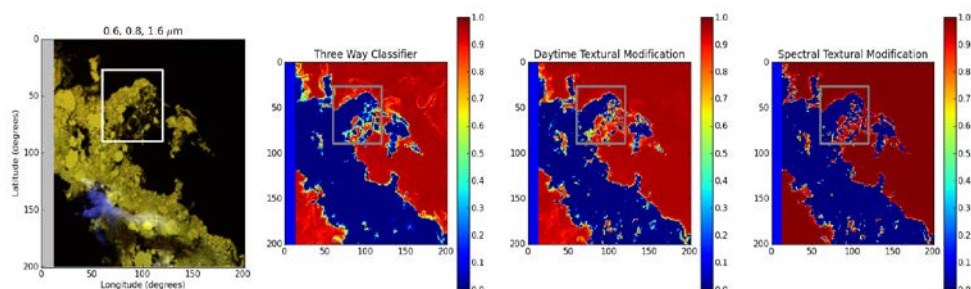

Figure 8.3 Example of Bayesian three-way classification over an area of mixed ice. The first panel shows a false colour image using the

0.6, 0.8 and 1.6 µm channels. Subsequent plots show the probability of clear-sky over ocean using a three-way classifier, including the 1.6 µm textural measure and then adding the 0.6 µm channel.

The work of Bulgin *et al.* (2014) [12] focused on nadir view observations only, and for inclusion in the ATSR climate data record processing, the following steps need to be taken to test performance 'globally', on dual-view observations and under night-time conditions.

1. In order to test the performance of the new algorithm configurations a reference cloud mask for a set of scenes is required. Within Phase 1 of the SST CCI project an extensive database of clear, cloud and ice observations including mixed and thin ice was generated by expert inspection of match-up database extracts. This needs to be updated to include forward-view imagery in order to test dual-view classification.

2. A database of expertly classified scenes under night-time conditions and both viewing angles also needs to be generated.

3. Three-way Bayesian classification can be tested on this updated database in its current offline configuration. Algorithm performance can be measured and compared to standard Bayesian and nadir-only classifications using hit rate, false alarm and true skill score metrics.

4. If the algorithm performance at high latitudes from step three is comparable to or better than the nadir only results the algorithm needs to be validated globally (to check that spurious ice is not generated outside sea-ice regions) and for the period following the Mount Pinatubo eruption (when spectral relationships may be significantly different). Several orbits of data across the climate data record should be processed and compared to standard Bayesian outputs to identify any problems occurring when using the three-way classifier and the 1.6 µm textural measure under different atmospheric conditions. In the mid-latitudes and tropics the prior probability for ice should be set to zero.

5. If the work in (4) demonstrates a benefit to applying this scheme globally, the Bayesian software needs to be developed from a two-way to a 'n'-way classifier for application to climate data record processing.

8.7 Extension to accommodate post-volcanic stratospheric aerosol events

Massive volcanic eruptions that inject material into the stratospheric occur with a return period of a few decades, with two having taken place during the era of meteorological satellites (El-Chichon (1982) and Mount Pinatubo (1991)). The sulphur injected can hydrolyse and form a layer of sulphuric acid aerosol droplets that perturbs brightness temperatures and introduces bias into retrieved SST [14]. These are long-lived (1-2 years) given the

relative stability of the stratosphere and slow-transport mechanism from the equator to the poles.

Simulation of the effects of stratospheric aerosol on ATSR-1 BTs appear to be adequate to reduce these SST biases to ~0.1 to ~0.2 K (down from ~1 to ~2 K) if used to make ATSR-1 dual-view SST retrievals "aerosol robust" [15][16]. This approach is less effective for three-channel single-view SST retrievals, and impossible for two-channel single-view SST. Simulated relative BTs impacts for ATSR are shown in Figure 8.4
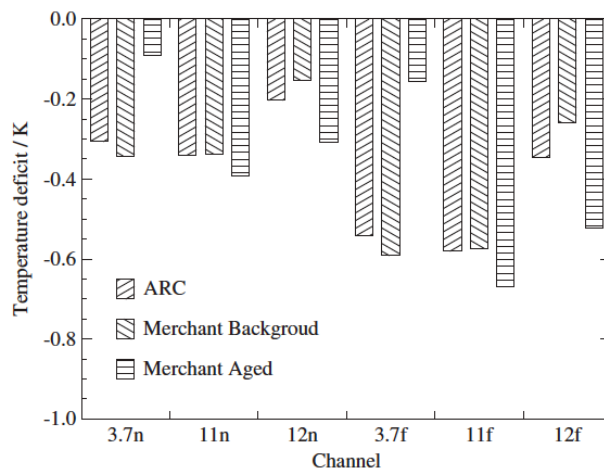

Figure 8.4 Simulated relative BTs impacts for ATSR due to stratospheric aerosol

Within the ARC [2] and SST CCI Phase 1 projects, a preliminary approach has been applied to render ATSR-1 cloud detection functional in the presence of stratospheric aerosol. The approach has been to define *a priori* a latitudinal estimate of the BT perturbation associated with the Pinatubo aerosol on a monthly basis. This dataset was inferred from the relative SST biases between aerosol-robust and non-aerosol-robust SST algorithms, which can be inverted to estimate a scaling factor (as a function of latitude and month) for the "mode" of variability shown as "ARC" in Figure 8.4. The estimated temperature deficit for each channel given that scaling was subtracted from each set of BTs, to give an estimate of the BTs expected had normal conditions prevailed. These modified BTs were then sent through the Bayesian cloud detection in the usual way.

Two approaches could be taken based upon the ARC method summarized above. First, keep the same basic approach, but work to improve the prior estimate of the aerosol impacts. Second, devise a means of making Bayesian cloud detection "aerosol tolerant" by explicitly introducing aerosol-related BT and reflectance uncertainty into (1) the simulation of the conditional probability distribution function of clear sky, and (2) a modified look-up procedure for the cloudy condition PDF.

Work will be undertaken on these possibilities within the SST CCI Phase 2 project. The work plan is reproduced below from the SST CCI Phase 2 Implementation Plan

Volcanic-aerosol-era cloud detection

- Review sources of prior volcanic aerosol products for Pinatubo and El Chichon
    - Temporal and spatial coverage
    - Relationship of observation to IR optical depth / scaling of BT aerosol mode
    - Relationship to parameters usable as aerosol parameters for BTs in RTTOV-11
    - Assess simulation for reflectance channels using prior and RTTOV-11, near and away from sunglint angles
- Develop (using aerosol modes from WP 21 and most useful aerosol products) prior mode-scaling for Pinatubo and El Chichon eras
    - Or, equivalently, prior aerosol parameters suitable for RTTOV-11
    - May need to tolerate less information for El Chichon
    - Also need associated uncertainties
- Develop means of quantitatively testing effectiveness of cloud detection alternatives
- Define alternatives
    - prior BT perturbation estimates assumed perfect
    - prior BT perturbations estimates with realistic uncertainty in scaling of aerosol mode included in clear-sky pdf calculation
    - no prior BT perturbation, but with large uncertainty in scaling of aerosol mode included in clear-sky pdf calculation ('aerosol robust' method)
- Test alternatives, first on ATSR-1 then on other relevant AVHRRs
- Determine what cloud detection to use for stratospheric aerosol eras
- Write up notes as basis of publication and ATBD contribution

8.8   Extension of Bayesian classifier to cloud detection over land

Over land, image classification into clear and cloud affected pixels is often more difficult than over ocean surfaces due to small-scale spatial structure of land cover features and temporal variations in reflectance and emissivity due to dynamic land cover change [Simpson and Gobat (1996) [17], Bulgin *et al.* (2013)[18].] The work in Bulgin *et al.* (2013) [18] demonstrates the benefit of applying a Bayesian scheme to cloud detection over land for the AATSR instrument, over previously employed threshold methods. The analysis focuses on daytime imagery, with the classifier using the 0.6, 1.6, 11 and 12 μm channels. Performance metrics are calculated in relation to a set of images with a reference cloud mask verified by expert inspection. The Bayesian scheme increases the true skill score by 5% over the SADIST threshold based tests and reduces the false alarm rate by 8%.

Three key factors were identified in affecting cloud detection performance over land: high aerosol loadings, cloud type and land surface biome. High aerosol loading can obscure the surface either by attenuation or reflection of solar and long-wave radiation. Under these conditions, the retrieved LST can be significantly perturbed from the true LST and either needs to

be corrected for aerosol presence if possible, or assigned a higher uncertainty due to the presence of aerosol.

Optically thin clouds are often difficult to detect, particularly where they over-lie cold or highly reflective surface types. Cloudy sky empirical PDFs currently used in the Bayesian classifier are based on ocean-only observations and would benefit from regeneration over land, including clouds overlying surfaces with a variety of different emissivity and reflectance characteristics.

Biome is also important in cloud detection where radiative transfer methods are used, as some surfaces are more difficult to model, for example urban areas and desert regions. These often experience significant changes in temperature over the diurnal cycle, and high surface reflectance makes distinguishing clouds using visible channels more difficult. Interpolation of data between observation tie-points may also lead to poor representation of emissivity and reflectance fields in regions of abrupt change in land surface cover. The work in Bulgin *et al.* (2013) [18] was done in the context of an algorithm comparison study and has the potential for significant development as outlined in the steps below.

1. The empirical PDFs used in Bulgin *et al.* (2013) [18] were the PDFs developed for the ARC project, using only observations over the ocean. Generation of land-based cloudy PDFs should improve classifier performance by including semi-transparent clouds over land surfaces with different surface reflectance and emissivity. These PDFs could be generated using the current Bayesian classifier and iterated if necessary. They should also include the forward view for dual-view retrievals.

2. As with the classifier development work described at high latitudes, performance analysis is dependent on the generation of reference cloud masks by expert inspection. Over land, a number of nadir view images have been manually classified, but inclusion of the forward view imagery is essential for the development of a dual-view classifier. Expert classification of nighttime images is also essential for algorithm development.

3. The new PDFs should be implemented into the processing chain and the Bayesian cloud detection performance evaluated using the extended selection of manually classified scenes.

4. Cloud detection is most difficult over regions with large diurnal cycles in temperature eg. deserts. Numerical weather prediction data used as an input for the radiative transfer modelling in the Bayesian scheme is provided at synoptic times (0000, 0600, 1200, 1800). Interpolation between these points doesn't capture the peak temperatures in the diurnal cycle (typically 1400), which affects classifier performance. The Bayesian calculation skill can be improved either by better representation of the uncertainties introduced by this process or by better representation of

changes in land surface temperature and emissivity over the diurnal cycle. Strategies for both approaches should be developed and tested.

5. Over land, classifier performance is significantly affected by the presence of aerosol. Methods should be considered as to how to make the retrieval 'aerosol robust'. The impact of aerosol on retrieved land surface temperature should be analysed by comparison of satellite and in-situ data. This should be done for a number of different land covers, aerosol types and where possible time of day. Following this review, recommendations should be made as to whether correction for aerosol biases in LST is needed/can be made during the Bayesian retrieval process in its current configuration.

8.9   Adaptation of Bayesian classifier to OLCI

The basic characteristics of the Ocean and Land Colour Instrument (OLCI) are shown in Figure 8.5.

Band #	λ center	Width	Lmin	Lref	Lsat	SNR@Lref
	nm	nm	W/(m².sr.μm)	W/(m².sr.μm)	W/(m².sr.μm)	
Oa1	400	15	21.60	62.95	413.5	2188
Oa2	412.5	10	25.93	74.14	501.3	2061
Oa3	442.5	10	23.96	65.61	466.1	1811
Oa4	490	10	19.78	51.21	483.3	1541
Oa5	510	10	17.45	44.39	449.6	1488
Oa6	560	10	12.73	31.49	524.5	1280
Oa7	620	10	8.86	21.14	397.9	997
Oa8	665	10	7.12	16.38	364.9	883
Oa9	673.75	7.5	6.87	15.70	443.1	707
Oa10	681.25	7.5	6.65	15.11	350.3	745
Oa11	708.75	10	5.66	12.73	332.4	785
Oa12	753.75	7.5	4.70	10.33	377.7	605
Oa13	761.25	2.5	2.53	6.09	369.5	232
Oa14	764.375	3.75	3.00	7.13	373.4	305
Oa15	767.5	2.5	3.27	7.58	250.0	330
Oa16	778.75	15	4.22	9.18	277.5	812
Oa17	865	20	2.88	6.17	229.5	666
Oa18	885	10	2.80	6.00	281.0	395
Oa19	900	10	2.05	4.73	237.6	308
Oa20	940	20	0.94	2.39	171.7	203
Oa21	1020	40	1.81	3.86	163.7	152

Figure 8.5 The basic characteristics of the Ocean and Land Colour Instrument (OLCI) taken from ESA's online Sentinel-3 handbook [19].

Level 1B products will be available for these channels with the following features:

- 300 m resolution for the full resolution products
- 1.2 km resolution for the reduced resolution products
- including ECMWF meteorological profiles and surface fields at tie-points

To adapt the Bayesian method to OLCI requires, in essence, a set of substitutions relative to the standard application to meteorological-type sensors such as AVHRR. These substitutions are summarised in Table 8.3

Table 8.3 Substitutions required in order to adapt the Bayesian scheme to OLCI.

Aspect	AVHRR-like configuration (day-time)	Plausible substitution for adaptation to OLCI	Comments
Channel set	1.6 μm, 11 μm & 12 μm	1020 nm, 865 nm, 620 nm	Although SNR is poorer for the longest wavelengths, both clear and turbid waters have very low (non-glint) reflectance at 1020 nm. The other channels have high atmospheric transmittance, contrasting water reflectance, and similar cloud reflectance, which is a discriminating combination. Other similar channels sets should be considered. Four or more channels could be used in principle.
Factors in model of clear-sky PDF	SST, TCWV and (1.6 μm only) U	TCWV and U	AOD of marine and/or desert dust could be added if >3 channels were to be used.
Forward model for clear sky	RTTOV 11	Requires to be explored	RTTOV 11 does have a capability to simulate reflectance channels, so the easiest first step would be to assess it for this purpose for OLCI channels. A key requirement is for the forward model to do well in conditions near sun-glint.
PDF for cloudy sky	AATSR-derived PDF, adapted to other sensors on reading	New cloudy-sky PDF required.	It may be worthwhile to explore whether the angular separation from the specular angle should be an important dimension in the PDF (not currently used as a dimension for the table).
Texture measure	Local (3x3) standard deviation of 11 μm BT	Local standard deviation of 1020 nm reflectance	Alternative measures could be explored that may be better attuned to reflectance channels
Clear-sky texture PDF	Based on 11 μm NEDT, or using empirical PDF	Based on SNR	May need additional terms to allow for surface roughness variability in sun-glint regime

Each substitution implies the following steps:

- Exploring detailed options scientifically: this may involve literature review, simulation study, etc.
- Assessing the impact on the Bayesian code of implementation for different options (i.e., the complexity and likely effort involved in refactoring/modifying the code).
- Selecting and fully defining an option for trials.
- Coding and verification against an offline implementation.
- Defining test cases (in common across all substitutions).
- Applying code to test cases and assessing test results.
- Iteration of the above if results show this to be necessary.

8.10 Extension to use of forward-modelled cloudy radiances

Clear-sky simulations are used to calculate the conditional PDF of brightness temperatures given clear sky conditions. This is done dynamically using NWP profiles describing the scene being viewed. This is a readily tractable approach because the main uncertainties in the clear-sky BTs are captured by uncertainty in the surface temperature and TCWV. These geophysical uncertainties are estimated and propagated to a multi-dimensional Gaussian distribution in BT space using the tangent linear outputs from the forward model, i.e., using $\frac{\partial BT}{\partial SST^b}$ and $\frac{\partial BT}{\partial TCWV^b}$ for all of the thermal channels considered.

Clear-sky reflectance simulations are also undertaken using a similar principle. This case is a little different, because reflectance has a geophysical limit at 0% and clear-sky reflectance can be very low over the ocean, i.e., a few percent, away from specular reflection. Thus, the calculated Gaussian distribution can include geophysically impossible negative reflectances. At present, this problem is addressed by approximately redistributing the probability associated with geophysically impossible observations, by scaling up the PDFs such that the total probability over the geophysically possible range is normalised. The main uncertainty in prior clear-sky reflectance arises from uncertainty in wind speed (which affects surface reflectivity) and from uncertainty in TCWV (absorption in the atmosphere). Thus, as for BTs, there is a restricted state space to account for, and a linear perturbation around the prior state is adequate to describe it.

In contrast, the present Bayesian formulation uses empirically derived look-up tables to describe the conditional PDFs of radiances given cloudy conditions. The reason for this approach is that the state space that affects cloudy radiances is large and cannot be linearized. The key cloud properties that affect the radiances are cloud top phase, cloud top height and cloud water/ice path. Only certain combinations of these parameters are common in nature. Nonetheless, the distribution of cloudy radiances that are possible for a given atmospheric profile is not readily formulated

as an analytic function – there is no obvious equivalent of the Gaussian distribution. Simulation of many cloud configurations is required to sample the state space adequately to give a sound numerical PDF of radiances for a given atmospheric profile of temperature and water vapour. Mackie (2009) [20] implemented a 'brute-force' approach of simulating cloudy radiances across the plausible state space and using the distribution of results in Bayesian cloud detection. In tests, there was a benefit from the "local" cloud PDF in reducing the false alarm rate for ocean cloud detection to about a fifth of the value using the global PDF (then in use). However, the hit-rate for a fixed probability threshold also reduced, which is undesirable. On balance, the results were not adequate to justify the additional computational cost.

Nonetheless, it is desirable to develop local PDFs of cloudy radiances based on forward modelling for the following reasons:

• no special steps would be required to apply the Bayesian detection to a different sensor, provided the sensor had been implemented in RTTOV
• cloud configurations that are implausible given the NWP for a particular observation do contribute to the empirical global PDF, which is sub-optimal
• the global PDF, being empirical, probably includes a fraction of contributions that were actually clear-sky, leading to higher false alarm rates than necessary

There are two ways of making the approach of Mackie more computationally efficient.

The first is to take advantage of native cloudy radiance calculations implemented in RTTOV since Mackie did her work. Many of the calculations of interest are presently not externally output from RTTOV when run in cloud mode – but this could be addressed by modification of the code. The advantage in extracting cloudy radiance information from RTTOV's native cloudy radiance calculations is that the computational overhead associated with calling RTTOV multiple times to simulate individual clouds is avoided.

The second approach is to recognize that only cloudy radiance simulations in the vicinity of the observed brightness temperatures need be evaluated. Configurations of cloud that lead to results too far from the observation don't need to be simulated: they need only be enumerated to provide the appropriate normalisation of the local cloudy PDF. This approach requires a technique to identify which of all the plausible cloud configurations are those that will give results similar to the observations and do need to be simulated. One method to explore is to implement a simple retrieval from the observations of cloud top height (CTH) to define the range of CTH over which simulations should be performed. However, this would require careful treatment of errors in the CTH estimate.

These possibilities represent research that is rather more risky than some of the other extensions discussed, but which would result in a very significant improvement in generality of the Bayesian implementation – i.e., it would become even easier to apply the Bayesian code to other processors.

8.11 Improved treatment of twilight conditions

For solar zenith angles around 90°, it may be that the basic day-time (<90°) and night-time (>90°) configuration of the Bayesian classifier described here can be improved upon.

Firstly, a fine solar-zenith-angle dependence can in principle be introduced into the new AVHRR-specific empirical cloudy-radiance look up tables. Around twilight, a resolution of ~5° in this dimension could be investigated.

Secondly, on the clear-sky side, the quality of the forward model simulations needs to be considered more carefully. At extreme solar zenith angles, the atmospheric scatter becomes very significant, even for clear low-aerosol conditions. It is not necessarily required that the simulations perform well in these conditions (particularly since the cloud screening has no auxiliary information on aerosol loadings), but it is important the the forward model error covariance is representative under these conditions. Otherwise, the Bayesian probability will not realistically capture the variability intrinsic to twilight scenes.

Thirdly, it may be that a "twilight zone" needs to be defined in terms of solar zenith angle, for which neither a day or night channel set is used. In fact, this was done within the ARC project for Along Track Scanning Radiometer cloud detection, where a minimal channel set of only 11 and 12 μm was used for cloud detection within about 5° of the terminator. This choice represents that at that time, using an older version of RTTOV as forward model, only these channels could be simulated for this region.