A cloud-detection scheme for use with satellite sounding radiances in the context of data assimilation for numerical weather prediction

By S. J. ENGLISH*, J. R. EYRE and J. A. SMITH
Meteorological Office, UK

(Received 20 July 1998; revised 23 December 1998)

SUMMARY

A scheme for detecting cloud-affected radiances is described. The method is used to determine the probability of cloud-free conditions given the observations and the prior knowledge we have about the atmosphere from a numerical weather prediction (NWP) model. This is achieved using a likelihood method. It combines the strengths of some alternative methods (e.g. comparison of infra-red and microwave channels sounding the lower troposphere and comparison of infra-red window channels with sea surface temperature) in a powerful and flexible method. It is powerful because it uses different types of information simultaneously. It is flexible because it makes no assumption about which instrument is being processed, or what type of prior information (NWP, climatology etc.) is used. Therefore, it can readily be extended to new situations and data types (e.g. Advanced TIROS Operational Vertical Sounder (ATOVS)). It is suitable for use on general cloud-detection problems, using combined microwave and infra-red data. It has been tested using TIROS Operational Vertical Sounder (TOVS) radiances. The new method has been compared with an alternative cloud-detection method tailored specifically for TOVS and has been developed to a level of robustness adequate for operational use. The new method gave very similar results to the alternative method, especially over the ocean. The differences that did occur have been investigated by comparing with cloud information derived from the Advanced Very High Resolution Radiometer (AVHRR). Both the alternative method and the new scheme were found to have deficiencies when dealing with very low cloud. Some cloud missed by the existing scheme is identified by the new scheme. Over land, cloud detection is more difficult. The two schemes disagree more often, but validation using AVHRR is also more difficult because of increased surface heterogeneity and more variable emissivity and surface temperature errors. The new method is therefore shown to perform at least as well as an alternative method in operational use, whilst gaining the flexibility required for future systems. The implications for ATOVS are discussed.

KEYWORDS: Advanced TIROS Operational Vertical Sounder (ATOVS) Bayesian Cloud detection Data assimilation Numerical weather prediction Radiances TIROS Operational Vertical Sounder (TOVS)

1. INTRODUCTION

The assimilation of radiances measured by satellite sounding-instruments into numerical weather prediction (NWP) models using a variational approach is now well established at a number of operational NWP centres.

The variational technique can be applied either in one dimension (the vertical) and the resulting atmospheric profiles assimilated into a three-dimensional NWP analysis (see, for example, Eyre et al. (1993) and Gadd et al. (1995)), or in three dimensions, with the radiance information directly assimilated into the NWP model’s three-dimensional fields (Andersson et al. 1994; Derber and Wu 1997). These schemes have been developed for assimilation of TOVS (TIROS Operational Vertical Sounder) data from the United States’ National Oceanographic and Atmospheric Administration (NOAA) series of polar-orbiting satellites (see Smith et al. (1979)), but they are equally applicable

* Corresponding author: NWP Division, Meteorological Office, London Road, Bracknell, Berkshire RG12 2SZ, UK.
to radiances from the Advanced TOVS (ATOVS) instruments (see Saunders (1993)) on
the next generation of satellites, starting with NOAA-15 launched in May 1998. Both
TOVS and ATOVS are combined systems of complementary infra-red and microwave
radiometers.

One of the main scientific problems in the assimilation of such radiances, particu-
larly infra-red radiances, concerns the treatment of the effects of cloud. To date, all
the variational schemes mentioned above have been implemented on ‘clear-column’
radiances, i.e. radiances that have undergone pre-processing either to establish that they
are ‘cloud-free’ or to estimate ‘cloud-cleared’ radiances—the values that would have
been measured from the same temperature/humidity profile in the absence of cloud. De-
ficiencies in these pre-processing steps are still a major source of error in the radiances
presented for assimilation, and this represents a significant quality-control problem for
the data assimilation system as a whole.

The variational technique can be applied directly to cloud-affected radiances (see
Eyre (1989)), thus avoiding the need for prior cloud detection and cloud clearing.
However, this technique is susceptible to deficiencies in the forward modelling of cloud.
Although it has potential for the future, it has not yet been developed to a stage where it
is sufficiently robust for operational use. There still remains, therefore, a need for skilful
and robust methods for cloud detection and/or clearing.

Within the TOVS system, information on tropospheric temperature and humidity
comes predominantly from the 20-channel High-resolution Infra-red Radiation Sounder
(HIRS). The 4-channel Microwave Sounding Unit (MSU) has much lower horizontal
and vertical resolution, and so its role is to provide temperature information of lower
quality in very cloudy areas, and to give valuable information to assist the cloud
detection/clearing processes for HIRS data in cloud-free and partly cloudy areas. Many
schemes have been developed for cloud detection and clearing of HIRS data. They have
been reviewed by Eyre and Watts (1987).

Of particular interest are the schemes developed at the United States’ National
Environmental Satellite Data and Information Service (NOAA/NESDIS), as these have
been used in operational TOVS processing systems from which global clear-column
radiances are made available for use by global NWP centres. The initial scheme used for
TOVS was described by Smith et al. (1979). This was replaced in 1982 by the scheme
described by McMillin and Dean (1982), which has been used operationally, with minor
modifications only (see McMillin 1993), until the present. This scheme (and others)
makes use of the fact that cloud affects channels sensing the same atmospheric layer in
a differential way. In particular, clouds have substantial effects on tropospheric-sounding
infra-red channels, whereas microwave channels are comparatively little affected. The
mutual consistency (or lack of it) between infra-red and microwave measurements is
therefore a useful tool in assessing the presence of cloud.

In the transition from TOVS to ATOVS, the infra-red component (HIRS) is changed
only in minor detail but the microwave component is greatly enhanced, with two new in-
struments which together comprise the Advanced Microwave Sounding Unit (AMSU).
AMSU provides tropospheric temperature and humidity information comparable to
HIRS, and it does so in many (though not all) cloudy conditions. In cloud-free areas,
HIRS still provides superior vertical resolution for temperature in the lower troposphere.
It is therefore desirable to detect these areas and to make use of cloud-free HIRS in-
formation where possible. However, the addition of AMSU makes it less important to strive
to use infra-red information in difficult, cloud-affected conditions. Therefore the focus
of the work described here is to improve the detection of cloud-affected infra-red data,
but not to address the issue of cloud-clearing for such data.
TOVS cloud-detection schemes developed to date have been somewhat ad hoc in character. They have applied a series of tests—threshold, inter-channel and others—designed to detect different types of cloud. Although these schemes have been developed to levels of skill and robustness adequate for operational use, we suggest two areas of potential improvement. Firstly, the different cloud tests have previously been applied separately in series, whereas the available information should be more powerful if used simultaneously. Secondly, there is a tendency for cloud detection schemes to acquire increasing layers of complexity. Whilst this may lead to improved results in research and development systems, it poses problems for maintainability in operational systems. Also it tends to make them specific to one set of instruments, and thus inhibits transition to new instruments (e.g. from TOVS to ATOVS).

In the present paper, we describe a scheme developed using the concepts of Bayesian probability theory. It seeks to estimate the probability that a given infra-red measurement is affected by cloud, given all the other information available. This information, in its present implementation for TOVS, comes from two sources, interchannel consistency and NWP-model fields.

Interchannel consistency is used implicitly, by assessing the probability that measurements in a set of channels are consistent with the assumption of no cloud. The most powerful of such information comes from infra-red—microwave consistency, although consistency between infra-red channels in different spectral regions may also contribute. This is the same as the information used by McMillin and Dean (1982) and others, except that here the information is used simultaneously rather than serially.

NWP-model information enters through an assessment of the probability that radiances computed from a shortrange forecast of temperature and humidity profiles (with their expected errors) are consistent with the measured radiances and the assumption that they are cloud-free. Such information has been used previously within systems for processing and assimilating radiance information, but only in a limited and ad hoc manner as part of the quality-control procedures. Here, we use it systematically within a single framework. In addition to the potential theoretical advantages of this approach (which may in practice be small), this approach has the important advantage of facilitating the incorporation of the cloud-detection step into the same variational framework as the retrieval and assimilation; the new cloud-detection test can be included as part of the first step of an iterative (nonlinear) variational scheme. It thus removes the need for a separate processing module for cloud detection, with implications for the simplicity and maintainability of an operational system.

Section 2 sets out the theoretical basis of the cloud-detection method and shows how it can be incorporated within a variational framework. Section 3 describes the implementation of the method within a one-dimensional variational analysis (1D-Var) scheme for TOVS data. Sections 4 and 5 present results of testing the new method, firstly against a NWP-independent cloud-detection scheme, and secondly against cloud detection based on AVHRR data. Section 6 presents conclusions from the present work with TOVS data and comments on its application to ATOVS data.

2. Theory

(a) Bayes theorem

The probability \( P(c|y^o, x^b) \) that an infra-red sounding is cloud-free (denoted by \( c \)), given a set of observations \( y^o \), and background profile \( x^b \), can be written in terms of
joint and conditional probabilities using Bayes theorem as

$$P(c|y^o, x^b) = P(y^o, x^b|c)P(c)/P(y^o, x^b). \quad (1)$$

We develop a method to estimate $P(c|y^o, x^b)$ from known quantities in order to identify fields of view (fovs) which are likely to be cloudy. That is, we wish to identify when $P(c|y^o, x^b) < P_c$ where $P_c$ is a threshold for cloud. The joint probabilities can be written in terms of the conditional probabilities,

$$P(y^o, x^b) = P(y^o|x^b)P(x^b) \quad (2)$$

$$P(y^o, x^b|c) = P(y^o|x^b, c)P(x^b|c). \quad (3)$$

Substituting these into Eq. (1),

$$P(c|y^o, x^b) = P(y^o|x^b, c)P(x^b|c)P(c)/\{P(y^o|x^b)P(x^b)\}. \quad (4)$$

(b) Discussion of terms

Taking natural logarithms of Eq. (4) leads to a cost function

$$J = -\ln\{P(c|y^o, x^b)\}$$

$$= -\ln\{P(y^o|x^b, c)\} - \ln\{P(x^b|c)\} - \ln\{P(c)\} + \ln\{P(y^o|x^b)\} + \ln\{P(x^b)\}. \quad (5)$$

The first term on the right-hand side is the conditional probability of the observations, given the background and cloud-free conditions. Under assumptions of Gaussian and uncorrelated errors in $x^b$ and $y^o$, this can be shown (see Rodgers (1976)) to be

$$- \ln\{P(y^o|x^b, c)\} = -\frac{1}{2}(\Delta y)^T(H'(x^b)^T(BH'(x^b) + R)^{-1}(\Delta y), \quad (6)$$

where the notation of Ide et al. (1997) is used (i.e. $H'$ is the linearized observation Jacobian, $y = H(x^b)$ is the estimated observation vector calculated from $x^b$ using observation operator $H$, $x^b$ is the model background profile, $B$ is the error covariance associated with $x^b$, and $R$ is the total observational error covariance associated with $H$ and the observations $y^o$). In addition, we define the quantity $\Delta y = (y^o - y)$.

The probability $P(x^b|c)$ is assumed to be independent of the probability of cloud-free conditions such that

$$P(x^b|c) = P(x^b).$$

The global prior probability of a cloud-free sounding $P(c)$ is a constant. Denoting the constant term as $C$, we may write Eq. (5) as

$$J = -\ln\{P(c|y^o, x^b)\}$$

$$= -\frac{1}{2}(\Delta y)^T(H'(x^b)^T(BH'(x^b) + R)^{-1}(\Delta y) + \ln\{P(y^o|x^b)\} + C. \quad (7)$$

The second term on the right-hand side of Eq. (7) cannot easily be quantified but its form can be inferred. The term $P(y^o, x^b)$ is the probability of measured (potentially cloudy) brightness temperatures, given $x^b$ which contains no information on cloud. Compared with the probability density function (PDF) for the cloud-free case (Eq. (6)), the effect of cloud is to take density out of the cloud-free peak and to put it elsewhere, predominantly on the cold side. The magnitude of this effect will be greatest for the window channels and will be progressively reduced for channels with weighting functions having maxima higher in the atmosphere. For microwave channels, it will be very small compared with
infra-red channels, and usually on the warm side. The effect of the second term on the right-hand side of Eq. (7) is, therefore, to restrict high probabilities of cloud to soundings where the infra-red measurements are colder than the estimated observations from the background. So, the condition of cloudiness can be defined as occurring when

$$J(\Delta y) = -\frac{1}{2}(\Delta y)^T[H'(x^b)^TBH'(x^b) + R]^{-1}(\Delta y) + C > J_c$$  \hspace{1cm} (8)$$

and

$$\Delta y(\text{infra-red window}) < 0,$$  \hspace{1cm} (9)$$

where $J_c$ is a threshold on the cost function such that above $J_c$ the fov is considered cloudy. Because $C$ is a constant, it can be absorbed into the threshold of Eq. (8). Equation 8 alone is a likelihood function and is not specific to one instrument, although the precise value of $J_c$ depends on the channels and instruments used and will need to be reassessed for other channel combinations. The addition of Eq. (9) is a refinement which depends on the availability of an infra-red window channel. Rodgers (1970) used a Bayesian approach to evaluate $P(x, c|y^0)$ where $x$ is the atmospheric profile, except that instead of finding the likelihood of clear air (i.e. a cloud-detection method) he found the likelihood for cloud-top pressure and cloud fraction as part of a retrieval method. Rodgers suggested that, in an operational context, his method could be enhanced by replacing $P(x, c|y^0)$ with $P(x, c|y^0, x^b)$, taking $x^b$ from an NWP model. The approach presented here differs from his in that we have implemented his suggested change and use it exclusively for cloud detection.

(c) Discussion of the form of the cost function

The discussion in subsection 2(b) showed that, given the observations and a background profile, the probability of cloud-free conditions has two significant terms. One restricts high-cloud probabilities to measurements where $\Delta y < 0$ in the infra-red window channels, the other which gives a quadratic cost-function defined in Eq. (6). The quadratic form of the cost function is shown in Fig. 1, where real observations have been superimposed. The cost function has been contoured as a function of $\Delta y$(infra-red) and $\Delta y$(microwave) for selected channels of the TOVS instruments. The shape of the minimum in the cost function changes with the choice of channels. The top left-hand figure shows the cost function when microwave and infra-red channels are selected which sense different parts of the atmosphere—in the present case MSU channel 2 (mid-tropospheric, with a peak at about 650 hPa) and HIRS channel 8 (an infra-red window). The contours show that the shape of the minimum is a horizontally orientated ellipsoid and that the spread of the observations is perpendicular to the axis of this ellipsoid. As a result, any threshold on the value of the cost function is simply a threshold on HIRS channel 8, which effectively checks consistency between the measured radiance and the forecast surface-temperature. If we use HIRS channel 6 rather than HIRS channel 8, the axis of the ellipsoid becomes orientated closer to 45° such that its value is lowest when $\Delta y$ is similar for both channels. For this case, a large value of $\Delta y$ in HIRS channel 6 can have a smaller cost than a small $\Delta y$ in HIRS channel 6, depending on the value of MSU channel 2. A similar result is found for HIRS channel 15.

As more infra-red channels are added, the minimum becomes sharper. The shape of the cost function is interpreted as follows: values of $\Delta y$ with moderate amplitude and the same sign (low cost) are consistent with cloud-free conditions and a forecast temperature error of reasonable magnitude, whereas those of different sign (high cost) are inconsistent with any reasonable forecast error and show the presence of cloud.
Figure 1. Contours of quadratic cost function $\Delta y$ (as defined in Eq. (6) of the text) for different HIRS channels, given $\Delta y$ for MSU channel 2: (a) HIRS channel 8; (b) HIRS channel 6; (c) HIRS channel 15; (d) average for HIRS channels 4–8 and 13–15. Contour interval is 1. Superimposed points show corresponding real individual observations from TOVS data over the United Kingdom and north-east Atlantic, except in (d) where they show averages for HIRS channels 4–8 and 13–15.

When HIRS channels 4 to 8 and 13 to 15 are all included, a very sharp minimum results which is orientated close to the line $\Delta y$(microwave) $\approx \Delta y$(infra-red). Here $\Delta y$ refers to an average overall the infra-red channels used. The spread of observations now can be interpreted as two superimposed PDFs. One has a Gaussian nature, centred on the minimum of the cost function. The second has significant negative values of $\Delta y$(infra-red) where $\Delta y$(microwave) $\approx 0$. The test determines to which of these PDFs the observed radiance vector is most likely to belong. If the error covariance matrices were exact (rather than estimates), the best possible test would occur when all channels were used. In practice, stratospheric and humidity channels are not included because of uncertainty in the error correlation terms between temperature and humidity elements, and between tropospheric and stratospheric elements. A combination of the long-wave and short-wave tropospheric infra-red channels (HIRS channels 4 to 8 and 13 to 15) and MSU channel 2 is used.

3. IMPLEMENTATION

The new cloud-detection method uses the error covariance and observation Jacobian matrices that are required in a 1D-Var retrieval and are derived in the same way as those used by Gadd et al. (1995). The cost function can be calculated in the first iteration of a nonlinear iterative retrieval, or as a pre-processing step to the direct use of radiances in an assimilation scheme.
(a) Sources of data

TOVS radiances for Europe and the North Atlantic, received and processed locally in the UK using the method described by Dibben and Chapman (1991), have been used to test the new cloud-detection method. These are calibrated and the microwave radiances mapped to the grid of the HIRS instrument. The data are not adjusted for the effects of either emissivity or scan angle and are available at full resolution (40 km), although they are processed by selecting only the clearest views in 3 by 3 boxes. Coincident AVHRR data are also processed and mapped to the HIRS grid. In addition, AVHRR imagery has been used for qualitative assessment.

(b) Estimation of $x^b$

The Meteorological Office Unified Model (Cullen 1993) generates values of temperature and humidity at levels inherent to the model. These are then interpolated to a fixed set of 40 pressure levels which are the vertical levels required by the radiative transfer model. These interpolated values are also the first guess of the state vector, $x$, required by 1D-Var. The set of pressure levels ranges from 1000 hPa to 0.1 hPa, whereas model data are only reliable to 30 hPa. Between 30 hPa and 0.1 hPa an extrapolation based on the observations from the Stratospheric Sounding Unit (Smith et al. 1979) are used to augment the profile, as described by Gadd et al. (1995). Emissivities are assumed to be fixed and equal to unity at infra-red wavelengths, whereas for microwaves, a value of 0.6 is assumed over the sea and of 0.95 over land and sea ice.

(c) Calculation of observation operator $H$ and its Jacobian $H'$

The observation operator $H$ and its gradient $H' = \nabla_x H(x^b)$ are calculated using a fast radiative-transfer model (RTTOV—see Eyre (1991)) and $H$ is assumed to be linear in the vicinity of the background $x^b$. No modelling of cloud is required by the new method.

(d) Estimation of observation and background-error covariance matrices $R$ and $B$

The observation and forward model-error covariance matrices, $O$ and $F$ (note $R = O + F$), are used exactly as by Gadd et al. (1995). They are assumed to be diagonal matrices. Experience has shown that, after bias correction, the random component of the error is around 0.4 K, varying slightly from channel to channel. The background-error covariance matrix $B$ is not diagonal, and the correlation terms are very important as they determine how the new method interprets differences between $\Delta y$ in infra-red and microwave observations. $B$ is a well-conditioned positive definite matrix, with diagonal terms derived by comparison with the NWP model and correlation terms derived from a set of radiosonde collocations (Gadd et al. 1995). It should be noted that the matrices used to represent the error covariance matrices are only estimates and themselves have errors. An improved estimate of $B$ would be useful both to this cloud detection method and the temperature and humidity retrieval.

4. Comparison with NWP-independent cloud-detection using TOVS

(a) Overall statistics

The new cloud-detection method has been compared with the method developed by Eyre and Watts (1987), based largely on earlier work by McMillin and Dean (1982). This method has been developed to the point at which it has achieved operational robustness
Figure 2. Variation of percentage frequency of clear or cloudy fobs with cloud cost threshold $J_c$: (a) Land LASS = cloud denotes the number of fobs classed as cloudy by the tests based on a paper by Eyre and Watts (1987) which have a cloud cost less than the value on the x-axis (so no fobs are classed as cloudy by LASS which have a cloud cost less than 5); (b) Sea LASS = cloud (no sea fobs are classed as cloudy when the cost is less than 17); (c) Land LASS = clear denotes the number of fobs where the LASS class is clear with a cost above the threshold on the x-axis; (d) Sea LASS = clear denotes the number of fobs where classed as clear by LASS, which have a cost greater than the threshold.

(see Dibben and Chapman (1991)) and is used in the operational processing of locally-received TOVS data at the Meteorological Office. This processing system is called the Local Area Sounding System, so here after the operational version of the Eyre and Watts method will be referred to as the LASS method. During a four month period (January–April 1997) the number of cloudy and clear flags using the LASS cloud-flag and the likelihood cost function were calculated for TOVS radiances for the North Atlantic and Europe. Both day and night passes of the satellite were used. If required, visible information could be added as another element of $y^v$, but this has not been done in the present study. The number of clear and not-clear (i.e. cloudy and partly cloudy) fobs for sea and land points is shown in Fig. 2. The cloud cost is usually consistent with the LASS cloud-flag, i.e. the cost falls as the number of clear soundings rises and the number of cloudy soundings falls. The LASS method gives a small number of clear soundings for costs as high as 100, whereas, over the sea, no cloudy soundings are detected when the cost is less than 20. Applying a threshold of 20 for sea fobs would therefore lead to a slightly higher rejection rate, but no additional soundings would be accepted as clear (in this sample). Over land, we have less consistency, as the overlap between the clear curve and the cloudy curve is larger. It is, therefore, less straightforward to choose a threshold, but for simplicity we shall use 20 everywhere. It is not possible to determine from Fig. 2 whether the new detection-method is providing additional skill, or less skill, or whether the differences occur because, for some situations, neither scheme has a high level of skill.
Figure 3. Cloud cost \( J \) for all fongs over the sea, plotted against position in the dataset, showing the clear separation between the clear and cloudy populations. Dots denote fongs where the new and LASS methods give the same result. Open triangles denote cases where the cost \( J > 20 \), but the LASS class is clear. The single asterisk denotes the only occasion when the LASS class was cloud, and the cost \( J < 20 \). See Fig. 4 for corresponding data for fongs over the land.

Figure 3 shows the cost for 40 000 soundings over sea. Using a cost threshold of 20 to define cloud, fongs where the LASS cloud-test and the cloud cost are in agreement are denoted by dots. Diamonds represent points where the cloud cost is high but the LASS flag indicates clear air. Stars denote points where the cloud cost is low but the LASS flag indicates cloud. It is noticeable that the populations of fongs where the two cloud-tests agree form quite separable populations. The fongs where the new method gives cloud but the LASS method does not have costs ranging from 20 to 100. Only one of these 40 000 soundings gives a low cost when the LASS flag indicated the presence of cloud. Figure 4 shows that, over land, fongs classed as clear and cloudy by the LASS method can have very similar costs. A threshold of 20 gives roughly equal numbers of fongs where one test indicates cloud and the other cloud-free to those where the opposite is true i.e. the two populations of fongs where the tests agree are not as separable as they are for sea surfaces. However, it is still true that, for values of cloud cost less than seven, the LASS method always indicates clear and for values greater than 100 it almost always indicates clouds, but the 'overlap' zone is wider. Overall, the two methods agree 98.4% of the time, and agree for 99.4% of fongs over the sea, and for 97.1% of fongs over the land. The number of cloudy fongs increases slightly from 73.3% using the LASS method to 73.8% using the new method, with a threshold of 20; splitting this by sea and land we find that the number of cloudy fongs over sea rose from 69.1% to 69.7% and over land from 78.8% to 79.0%, the higher figure corresponding to the new method for both sea and land.

Few differences were found between day and night passes. In this sample of data, both methods give 5% less cloud in night-time passes. Differences between the new and LASS methods changed insignificantly from day to night, the number of differences
rising by less than 0.1%. Only night-time passes are used in the detailed studies in section 5, to ensure that the information used by both methods is the same, and because there is no evidence of substantially higher skill by day using the LASS method.

(b) Discussion of differences

In subsection 4(a), we demonstrated that for sea points there is very little difference in results between the two methods, but that a small population of fields of view detected as clear by the LASS method have a high cost. When the differences were examined in more detail it was noticed that they most frequently occurred in fields of view adjacent to those classed as cloudy. The TOVS data-processing system described by Dibben and Chapman (1991) selects the clearest field of view from a 3 by 3 box of observations—and the cases classed as clear but with high cost usually occurred where the other observations in the 3 by 3 box were classed as cloudy. This does not, however, prove that the new method is more successful—an alternative interpretation is that, in regions of broken cloud, the B matrix describes forecast errors inadequately.

The differences over land are more random in nature, and more frequent. The matrix B is modified slightly over the land, to allow for larger errors in the skin temperature (Gadd et al. 1995). This has to be done since real skin temperature errors are larger, and the bias correction scheme is tuned to sea fows, but the greater random error in skin temperature reduces the ability to identify cloud. Over land, other sources of error in H also increase (e.g. emissivity), but we do not have a good estimate of F and this issue is not addressed further in the present paper.
5. Comparison with AVHRR Imagery

(a) Qualitative

In section 4, we showed that the LASS and new cloud-detection methods give comparable results, especially over the sea. Comparison with AVHRR images shows that over the sea both methods give cloud masks which correspond well to representations of low temperatures in the AVHRR image. An example, for 14 February 1997, is given in Fig. 5. The corresponding cloud-cost image is given in Fig. 6. The HIRS cloud-map corresponds very closely to the cloud in the AVHRR image, and the brighter (colder) the AVHRR representation the higher the cost. Together, these results show that the method is functioning as expected (i.e. the likelihood function correlates well with clear regions) but in many situations cloud detection is straightforward (HIRS channel 8 can be lowered by 100 K in the presence of optically thick high-cloud) so this alone does not demonstrate that the method is useful. More detailed examination shows a high level of skill. For example, the narrow break in the low-level cloud extending southwest from Brittany (48°N, 5°W) towards the Azores (40°N, 20°W) is clearly observed in the cloud-cost map, with costs below 20, whereas the surrounding cloud has costs in excess of 100. Where the cloud is broken, the cost remains well above the threshold value of 20. Figure 7 maps the different results from the two methods, distinguishing between occasions when both methods give a cloud flag, both give a clear flag, and occasions when the new method reveals a cloudy fov but the LASS method a clear fov. Where the opposite is true (i.e. LASS = cloudy, new = clear), triangles are used. In this example, there are no triangles over the sea and only two over the land. This is because we have set the threshold at a low value (20). Overall, there are thirteen sea fovs where the two methods give different results. Comparison with Fig. 6 shows that all of these have cloud cost in the range 35–50. This is well above our threshold of 20, so that the cloud cost indicates cloud. Eight of these fovs are to the west of southern Portugal. Normally, every third fov on every third scan is processed, forming a regular grid at nominal 120 km resolution (sub-satellite). However when the LASS method indicates that this ‘on-grid’ fov is cloudy, the clearest ‘off-grid’ fov from the surrounding 3 by 3 box of fovs is processed. Seven of the eight are ‘off-grid’; that is they have been selected as clear when the on-grid fovs are flagged as cloudy. Figure 5 shows that, although these fovs are near the edge of an area of low cloud, they are within the cloudy area. For these fovs, the low and possibly broken cloud is on the limit of detectability, and the additional background information used by the new method enables us to determine that these fovs are cloudy. The other five fovs where disagreement occurs are for broken cumulus to the west of Norway, and it is therefore very difficult to determine qualitatively which method gives the correct result.

(b) Quantitative

The values of $\Delta y$ averaged over the fovs where disagreement occurred are shown in Table 1, together with the corresponding typical standard deviations in clear air.

By contrast, $\Delta y$ for MSU channel 2 varied only from $-0.1$ K to $0.1$ K (compared to a typical standard deviation in this channel of 0.5 K). This very good agreement for MSU channel 2 appears to rule out the possibility that NWP-model background errors are responsible for the large values of $\Delta y$ in the infra-red channels, and the lowering of the observed radiances compared to the background is therefore most probably due to cloud. The large radiance-increments in the infra-red channels will give large increments in the analysis or retrieved profiles, which, if the large value of $\Delta y$ is caused by cloud and not errors in the background profile, would give erroneous retrieved values. Quantitative
Figure 5. AVHRR channel 4 image for 0720 UTC 14 February 1997. © Dundee satellite receiving station.
Figure 6. TOVS Bayesian cloud cost \( J_c \) for the satellite pass on 14 February 1997 from which came the AVHRR channel 4 image in fig. 5.

Figure 7. Difference between LASS and new, Bayesian method, cloud results for the satellite pass on 14 February 1997 from which came Figs. 5 and 6. Open circles show where both methods class the fov as clear, and closed circles show where both class the fov as cloudy. Asterisks show where the new method gives cloud but LASS gives clear. No sea fovs occur where the opposite (open triangles) is true.
comparison with AVHRR is required to give independent confirmation that these fovs are indeed cloudy. Using AVHRR, it is possible to derive an independent estimate of the clear radiance, which can be compared with the observed radiance where the two cloud-detection methods disagree.

A cloud mask was generated on the AVHRR grid (Saunders and Kriebel 1988). The number of cloudy pixels, the mean clear and the mean cloudy AVHRR brightness temperatures for pixels within the HIRS fov were calculated, using the ATOVS and AVHRR Processing Package (Dibben et al. 1997). An example was taken from 5 March 1998, when extensive cloud was present over the United Kingdom and north-east Atlantic. An unusually large number of disagreements occurred between the new test and the LASS method. In using AVHRR to validate a HIRS cloud detection scheme, it is important to take full account of the different resolutions. The AVHRR cloud-tests are designed to generate a cloud mask on the AVHRR grid. Only those AVHRR fovs which are completely cloud-free will contribute to the ‘percentage clear’ on the HIRS grid. If a significant number of AVHRR fovs are slightly cloud-contaminated this will lead to a high percentage of cloudy AVHRR fovs within the HIRS fov, but does not necessarily imply that the HIRS radiances are significantly affected by cloud. For example, if 10% of AVHRR fovs in a HIRS fov themselves have a fractional cloud cover of only 10%, then the fractional cloud cover in the HIRS fov is just 1%. The effect of this is to give a misleading suggestion that AVHRR overestimates cloud fraction on the HIRS grid. It is, therefore, more useful to compare the HIRS channel 8 measurement with the average of the clear AVHRR fovs. If the two values agree to within the instrument errors, then the HIRS fov can be assumed to be insignificantly cloud-contaminated.

Despite a weak correlation between percentage cloudiness from AVHRR and the new cloud-cost, a strong correlation was found for the difference between HIRS channel 8 and the mean clear AVHRR brightness temperature. Figure 8(a) shows the difference between the estimated clear-sky radiance from AVHRR channel 4 (denoted by $y^c$ AVHRR) with the observation from HIRS channel 8 (denoted by $y^o$ HIRS), for sea fovs only for the 5 March 1998 pass. Fovs where the LASS method gives cloud are denoted by circles and LASS clear fovs are denoted by triangles. At very low cost (<10) almost all fovs have a very small difference (within $\pm 1$ K), although there are a few where larger differences are observed (3–4 K). Above a cost of 50, all but one of the HIRS fovs contained 100% cloudy AVHRR pixels, so a mean clear brightness temperature could not be calculated. For costs between 10 and 50, an increasing cold bias in the HIRS channel 8 brightness temperature compared to AWRR ‘clear’ channel 4 was found, which rose from 2 K at a cost of 20 to 6 K at a cost of 50. Figure 8(b) shows the two estimates of the clear-sky brightness temperature (i.e. that calculated from the background and that derived from AVHRR). For costs between 10 and 50, the differences were usually less than $\pm 1$ K, once a small bias of 1.5 K was removed, confirming that the two methods of estimating clear brightness temperature

<table>
<thead>
<tr>
<th>HIRS channel</th>
<th>$\Delta y$ (K)</th>
<th>Typical standard deviation in clear air (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>-8</td>
<td>1.6</td>
</tr>
<tr>
<td>7</td>
<td>-3</td>
<td>1.1</td>
</tr>
<tr>
<td>6</td>
<td>-0.5</td>
<td>0.7</td>
</tr>
</tbody>
</table>

$\Delta y = y^o - y$, where $y^o$ is a set of observations and $y$ is the estimated observation vector. See subsection 2(a) of the text.
were consistent. Figure 8(c) shows that the mean of all AVHRR pixels in each HIRS fov is close to the HIRS channel 8 value, all falling within ±0.05 K with a 0.05 K bias. Figure 8(c) also shows very clearly that, for this pass, the LASS method corresponds very closely to applying a threshold of 35 in the new cost-function. Figure 8(d) shows the difference between the mean AVHRR pixel value and the background calculated brightness temperature for HIRS channel 8. Again, this clearly shows that the new and LASS methods agree well for cost below 10 and above 35, that they are well supported by the AVHRR data, and that, for intermediate cost (10–35), a small number of fovs identified as clear by the LASS method have a cold bias likely to be caused by small amounts of low cloud. Whereas earlier calculations used a threshold of 20, the quantitative comparison with AVHRR shows that a lower threshold of 10 would be required to eliminate all cloud contamination. However, this would then be at the expense of rejecting some data where background errors were very large (i.e., small background-error and slight cloud-contamination could give a result comparable to that resulting from very large background-errors). For these reasons, an optimal value for the threshold is difficult to define, but we can say that a value of 20 implies a higher level of risk, and gives very similar results to the alternative method tested, whereas a lower value of 10 is safer for cloud detection but may lead to some very useful information being lost.
Over land, the comparison is much less clear cut. Panel for panel, Fig. 9 shows the same information as Fig. 8 but for land fovs only. Figure 9 shows that the new method detects cloudy and cloud-free fovs, but with higher noise levels, and that the mean value of AVHRR channel 4 is still very close to that of HIRS channel 8, despite increased surface heterogeneity. However, Fig. 9(a) shows that no clear correlation can be found between the cost function and the difference between the HIRS channel 8 brightness temperature and the estimated clear brightness temperature from AVHRR channel 4. Moreover, the results are equally unconvincing for the LASS method. There are large differences between the two methods for intermediate cost (10–35), but neither validates well against the AVHRR data, and it is not possible to determine which has greater skill. Furthermore, Fig. 9(b) shows that very large differences occur between the two estimates of cloud-free brightness temperature. Particularly at high cost, the clear AVHRR brightness temperature can differ by ±15 K. This comparison shows the difficulty of cloud detection over land both for AVHRR and HIRS, which continues to be a severe limitation on the effective use of infra-red radiances over the land. Consequently, a robust validation of the method over land remains elusive.

Interestingly, for very cold background values (HIRS channel 8 brightness temperature less than 240 K) the LASS method gives some clear air fovs for which the cost is high because $\Delta y$ is large and positive. This is seen most clearly in Fig. 9(d), where cloud gives large negative differences, but, at high cost, large positive differences are
also observed. The AVHRR method returns high-percentage coverage of cloud for these cases, but the LASS method indicates clear air and the mean clear AVHRR channel 4 brightness temperature is very close to the HIRS channel 8 value. These fows were found to occur along the sea-ice edge and in Norwegian fjords (where the land is very cold but the sea relatively warm). These positive values of \( \Delta T \) are believed to result from a mismatch of the surface type used by the background and the true surface-type influencing the measured radiances. This would be likely to lead to gross errors in the retrieval. We could control the quality of the data at these points by using Eq. (9) only when the background in the appropriate infra-red window channel (in this case HIRS channel 8) is greater than 250 K. Although this increases the specificity of the method to one instrument, it demonstrates how the method can be extended to other instrument-specific quality-control issues without significantly increasing its complexity.

The study of 5 March 1998 showed that the AVHRR data need to be tuned carefully in a manner similar to bias correction of TOVS radiances (Eyre 1992). This is because the mean bias between AVHRR channel 4 and the model background was very similar to the mean bias of HIRS channel 8. Indeed the two instruments showed very small (0.05 K) bias with respect to each other. This suggests that there is a bias in either the model background or the forward model. A very large bias in the model background would be required to explain differences of around 2 K, and there is no evidence for such a bias. As surface emissivity was set to unity, and the strength of the absorption by the water-vapour continuum is still uncertain, bias in the forward model seems more likely.

6. APPLICATION TO ADVANCED TOVS

The ATOVS was launched on 13 May 1998. A major motivation for developing a more flexible cloud-detection method was to take full advantage of additional channels it would provide. ATOVS carries additional microwave channels, the Advanced Microwave Sounding Unit (AMSU), and, the most interesting channel for this application, the 52.4–53.2 GHz channel which peaks at 950 hPa, significantly lower than MSU channel 2. Consequently, for the reasons discussed in the introduction, it may enhance the scheme’s ability to detect low cloud. The scheme was run on data from one ATOVS orbit from 16 November 1998 (totalling 30,000 observations) for sea fows only, using the standard set of HIRS channels (1–8, 13–15) and AMSU channel 5, and then repeating, adding AMSU channel 4. The value of \( J_c \) was increased from 20 to \( 20 \times 10^9 \) to allow for the fact we have 10 channels rather than 9. Overall, the rejection rate increased for this dataset from 72.3% to 75.6%. Figure 10 shows the impact on differences of observed and calculated brightness temperatures for those fows accepted as clear. Figure 10(a) shows the cost calculated without AMSU channel 4 plotted against the cost calculated with channel 4. There is considerable spread, indicating that the addition of channel 4 is significantly modifying the result. Figure 10(b) shows that this reduces the standard deviation of the observed and calculated brightness-temperatures in those channels for which the weighting functions are not orthogonal to AMSU channel 4. In particular, significant reductions occur for channels 7, 8, 13 and 14, with small reductions (barely detectable on Fig. 10(b)) for HIRS channels 6 and 15. Of course, standard deviation would be reduced just by applying a threshold on the HIRS channel 8 observation minus calculation difference, so that reduced standard deviation alone does not necessarily show increased effectiveness. However, Fig. 10(c,d) shows histograms for HIRS channels 7 and 8 which demonstrate that the statistical distribution of clear fows appears more Gaussian using AMSU channels 4 and 5 in the calculation of the cost. Acceptance rate
for small departures has stayed the same for HIRS channel 7 and actually risen for HIRS channel 8. The acceptance rate for larger departures has decreased. This demonstrates that AMSU channel 4 is adding real skill to the cloud detection for HIRS.

7. CONCLUSIONS

A new cloud-detection method, based on Bayesian probability theory, has been described. When applied to four months of locally received TOVS data, it has given very similar results to the method used in operational processing of locally received TOVS radiances at the Meteorological Office for 99% of cases over sea, and over 97% over land using MSU channel 2, HIRS channels 4–8 and 13–15 and a cost threshold of 20 to define cloud. (See the LASS system described by Dibben and Chapman (1991) using cloud detection and clearing based on the work of Eyre and Watts (1987).) In addition, no significant difference is found in the number of disagreements between the methods by day and night. The differences that did occur have been studied using AVHRR as a reference in both a qualitative and quantitative sense. The comparisons showed that many of the differences between the two methods were indeed cloudy fovs which had been classed as clear by the LASS method. The study indicates that some residual cloud
effects are observed for cloud costs as low as 10 over sea. Over land, the two cloud-detection methods studied give many more differences, and validation using AVHRR is more difficult. However, at costs of 10 or below it seems likely that residual cloud effects are very small.

Both methods using TOVS miss some very low cloud. This is cloud which is much lower than the height of the peak of the MSU channel 2 weighting function, and is often misinterpreted as a warming or moisture deficit in the background. What is required is better microwave representation of the lower boundary layer, in particular the lower tropospheric channels. Using the framework described in section 2, new channels can easily and optimally be incorporated. Consequently, this scheme will readily allow full advantage to be taken of the new channels on ATOVS, in particular the channels at 52.8 and 150 GHz, which are sensitive to temperature at 950 hPa and to water vapour at a similar height in mid-latitude conditions. In tropical conditions, the 89 GHz channel will have maximum sensitivity to water vapour at this altitude, and in polar and subpolar conditions the 183 ± 7 GHz will take on the same role. So if these channels are all included in the scheme it will be applicable at all latitudes and to lower altitude. A major advantage of the new method over its predecessors is, therefore, not its accuracy for TOVS (which is similar) but its ease of transition to future instruments, which may have an enhanced capability for cloud detection. The impact of adding the 52.8 GHz ATOVS channel was tested and found to give significant benefits for the lowest altitude channels. In addition, the new method is versatile and can easily be used to address quality-control issues other than cloud (e.g. incorrect surface type).

In summary, we have moved nearer to optimum use of information available for cloud detection: microwave radiances, longwave and shortwave infra-red radiances, and background (forecast) information.

ACKNOWLEDGEMENTS

The authors acknowledge the help of Paul Dibben and Dave Chapman for providing the locally received TOVS and AVHRR data and are grateful to Richard Renshaw who did much of the coding work for the calculation of the cloud cost and commented on the manuscript. The Dundee Satellite Receiving Station, University of Dundee, Scotland, is acknowledged for allowing the reproduction of one of their quick look AVHRR images. Roger Saunders (ECMWF) gave useful guidance on the use of AVHRR cloud-detection. We thank two anonymous reviewers for their helpful comments.

REFERENCES


