

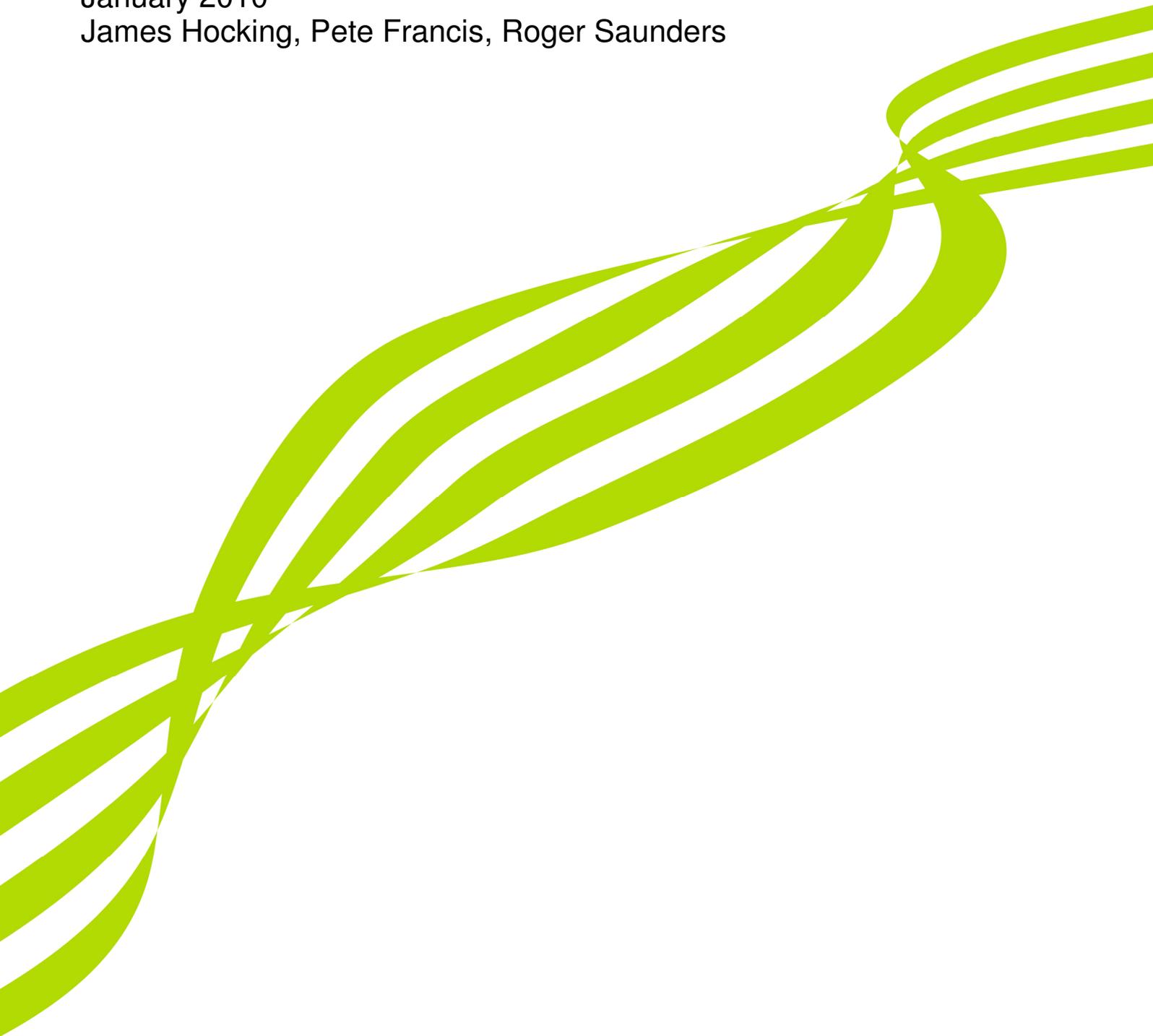


**Met Office**

Cloud Detection in Meteosat Second Generation  
Imagery at the Met Office  
Forecasting R&D Technical Report No. 540

January 2010

James Hocking, Pete Francis, Roger Saunders



## Document review history

<b>Date</b>	<b>Version</b>	<b>Author/changed by</b>	<b>Action/comments</b>	<b>Approval</b>
April 2009	0.1	Pete Francis & James Hocking	Draft version for comments.	-
Oct 2009	0.2	Roger Saunders	Changes to paper.	-
Nov 2009	0.3	James Hocking	Incorporated comments from RWS. Further additions.	-
Dec 2009	0.4	James Hocking	Amended following comments from RWS.	-
Jan 2010	1.0	James Hocking	Amended following comments from JRE.	JRE

## Contents

Abstract .....	2
1. Introduction .....	3
2. Summary of the SEVIRI instrument.....	3
3. Cloud mask determination .....	4
4. Effectiveness of cloud detection tests .....	17
5. Comparison against SAFNWC SEVIRI cloud mask.....	19
6. Comparison against the MODIS cloud mask.....	21
7. Summary .....	23
8. Acknowledgements .....	24
9. References.....	25
10. Annex – Test Thresholds, Offsets and Coefficients .....	27

## **Abstract**

The Met Office has been using data from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard the Meteosat Second Generation (MSG) satellites for nowcasting, assimilation in numerical weather prediction (NWP) and for providing guidance to forecasters since September 2004. This paper describes the cloud mask derived from MSG imagery which is used by all downstream products. The Met Office SEVIRI mask is compared with that of the Nowcasting Satellite Applications Facility (SAFNWC), and with the Moderate Resolution Imaging Spectro-Radiometer (MODIS) cloud mask. These comparisons show good agreement between the cloud masks: the Met Office and SAFNWC masks are in accord with the MODIS mask for 90.65% and 89.69% of SEVIRI pixels tested, respectively.

## **1. Introduction**

The Met Office has used imagery from the European Meteosat geostationary satellites for many years initially for forecaster guidance but increasingly for quantitative use of the cloud products in the nowcasting system, Nimrod (Golding, 1998). One of the key requirements is to make the products available within minutes of the measurement. More recently the direct assimilation of the cloud products and radiances from the geostationary satellites has been developed (e.g. Munro *et al.*, 2004; Kelly and Francis, 2008) requiring the data are accurately pre-processed. The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) provides some products in near real time (see [http://www.eumetsat.int/Home/Main/Image\\_Gallery/Real\\_Time\\_Imagery/index.htm](http://www.eumetsat.int/Home/Main/Image_Gallery/Real_Time_Imagery/index.htm)), but in order to meet the timeliness requirements the locally received data are processed at the Met Office to allow flexibility to tailor any products for the UK and other areas of interest to customers of the Met Office.

A significant milestone in the provision of geostationary imagery over the European and African regions was the launch of the second generation Meteosat (MSG) satellites in 2002 by EUMETSAT. This satellite contained new technology sensors including the Spinning Enhanced Visible and Infrared Imager (SEVIRI) radiometer. Data from SEVIRI have been available since 2004, and will continue to play a major role in the nowcasting and short-range forecasting capabilities at the Met Office until at least 2018, by which time the third generation of Meteosat satellites is scheduled to have started operations. Compared to the first generation of Meteosat imagers, SEVIRI can provide imagery at significantly higher spatial and temporal resolution, and senses radiation in 12 different spectral channels in the solar and infrared spectra, compared with the 3 channels on the earlier satellites. There is scope, therefore, for significant improvement in our ability to derive quantitative products for use in nowcasting systems, numerical weather prediction (NWP) applications and as imagery for forecasters by fully exploiting the SEVIRI data.

Section 2 of this paper briefly introduces the SEVIRI instrument. The methodology to identify cloud-contaminated pixels in the SEVIRI data is documented in Section 3. Section 4 examines the relative effectiveness of the cloud tests applied to the imagery. Sections 5 and 6 compare the Met Office SEVIRI mask with that of the Nowcasting Satellite Applications Facility (SAFNWC), and with the cloud mask derived from Moderate Resolution Imaging Spectro-radiometer (MODIS) imagery. Finally, Section 7 summarises the results. The Annex supplies values of thresholds and offsets and other details omitted from Section 3 for the sake of clarity.

## **2. Summary of the SEVIRI instrument**

SEVIRI is a radiometer on MSG that measures radiation reflected from and emitted by the earth's surface and its atmosphere at more wavelengths, greater temporal resolution and greater spatial resolution than any other geostationary meteorological satellite. It is located above approximately 0° latitude, 0° longitude, covering a region including Europe, Africa,

Middle East and the North and South Atlantic. MSG-1 was launched in August 2002, and became operational (as Meteosat-8) in January 2004. MSG-2 was launched in December 2005, and, as Meteosat-9, was declared the prime 0° satellite in April 2007. The other two satellites in the series, MSG-3 and MSG-4, are due to be launched in 2011 and 2013 respectively.

A summary of the 12 different spectral channels is given in Table I. Four of the channels (i.e. those centred at 0.6  $\mu\text{m}$ , 0.8  $\mu\text{m}$  and 1.6  $\mu\text{m}$  plus the High Resolution Visible (HRV) channel) are located in the visible/near-infrared part of the spectrum (i.e. day-time only use), with the 3.9  $\mu\text{m}$  channel acting as a solar + thermal infrared channel during the day and a thermal infrared-only channel during night-time. The remaining 7 channels are thermal infrared only, and therefore behave in the same way by day or night. The spatial resolution of the data is 3 km (at sub-satellite point) for 11 of the channels, with the HRV channel having a spatial resolution of 1 km though not with full disk coverage. The full disc is viewed once every 15 minutes, enabling monitoring of rapidly evolving phenomena, and potentially aiding the weather forecaster in the swift recognition and prediction of dangerous weather phenomena such as intense convection, heavy rain, fog and rapid cyclogenesis. More details on the SEVIRI instrument can be found in Schmetz *et al.* (2002).

The MSG image data are received at the Met Office continuously via the EUMETCast system (EUMETSAT, 2006), and are routinely processed by the operational processing system to generate single channel images (e.g. visible and infrared), false-colour (RGB) images and a variety of other derived products. Some of the imagery is available on the Met Office web site at [www.metoffice.gov.uk/weather/satellite/index.html](http://www.metoffice.gov.uk/weather/satellite/index.html).

### **3. Cloud mask determination**

The nowcasting system requirement at the Met Office is for cloud products sampled every 30 minutes and generated within 3 minutes of the nominal end of the measurement period. MSG has a set of channels designed to provide accurate cloud top pressure and amount estimates. The first step in deriving cloud parameters is to produce a cloud mask which identifies clear, partly cloudy and fully cloudy fields of view using techniques previously developed for the Advanced Very High Resolution Radiometer, AVHRR, (e.g. Saunders and Kriebel, 1988). However an important development for processing MSG radiances is to include clear-sky simulated radiances produced using NWP model profiles input to a fast radiative transfer model. This allows the differences between observed and simulated radiances to be compared which should be small for clear sky radiances. A summary of the cloud detection tests applied to SEVIRI radiances is given in Table II. They encompass spectral differencing using various channel combinations and comparing the measured differences with the simulated differences. Spatial coherence and visible reflectance information is also used.

In the following, the symbols  $R_{(0.6)}^{ob}$ ,  $R_{(0.8)}^{ob}$ ,  $R_{(1.6)}^{ob}$  and  $R_{(HRV)}^{ob}$  represent the observed reflectances at 0.6  $\mu\text{m}$ , 0.8  $\mu\text{m}$ , 1.6  $\mu\text{m}$  and in the HRV channel respectively,  $T_{B(3.9)}^{ob}$ ,  $T_{B(8.7)}^{ob}$ ,  $T_{B(10.8)}^{ob}$  and  $T_{B(12.0)}^{ob}$  are the observed brightness temperatures (BTs) for the 3.9  $\mu\text{m}$ , 8.7  $\mu\text{m}$ , 10.8  $\mu\text{m}$  and 12.0  $\mu\text{m}$  channels respectively, and  $T_{B(3.9)}^b$ ,  $T_{B(8.7)}^b$ ,  $T_{B(10.8)}^b$  and  $T_{B(12.0)}^b$  are corresponding background clear-sky brightness temperatures, as calculated from a recent forecast profile using the RTTOV fast radiative transfer model (Saunders *et al.* 1999). Hourly forecast fields are available and these are interpolated onto 15-minute intervals for use by RTTOV. EUMETSAT have derived a SEVIRI pixel-resolution land surface emissivity (LSE) map (Lutz and König, 2008) from the University of Wisconsin-Madison/Cooperative Institute for Meteorological Satellite Studies (UW/CIMSS) high resolution MODIS LSE atlas (Seemann *et al.*, 2008). The pixel LSE values from the EUMETSAT atlas are averaged over NWP model grid points and the resulting emissivities are input to RTTOV. Currently RTTOV version 9 is being used which allows the radiative transfer calculations to be carried out on the background model levels, rather than having to interpolate onto the 43 levels standard to previous versions of RTTOV. Throughout the cloud processing algorithms, simulated brightness temperatures and model data (where they are used explicitly) are linearly interpolated onto pixel locations. The interpolation takes account of surface type so that wherever possible the land/sea classification of the pixel matches that of all the interpolated model data points.

The observed reflectance values in each visible channel have been normalised by the following factor:

$$\begin{cases} \sec \theta_{sol} & \theta_{sol} \leq 85^\circ \\ \sec 85^\circ + 2.29(\theta_{sol} - 85) & 85^\circ < \theta_{sol} \leq 90^\circ \end{cases} \quad (1)$$

where  $\theta_{sol}$  is the solar zenith angle.  $\sec \theta_{sol}$  is the commonly applied multiplier which accounts for the varying local solar zenith angle at the surface. It is modified at high solar zenith angles in order to avoid excessively high reflectance values near the terminator: for  $\theta_{sol} > 85^\circ$  the factor increases linearly with  $\theta_{sol}$  at the gradient of  $\sec \theta_{sol}$  at  $\theta_{sol} = 85^\circ$ .

The tests carried out on each pixel depend upon the time of day and the surface type. Day time is defined by  $\theta_{sol} < 80^\circ$ , twilight by  $80^\circ \leq \theta_{sol} < 90^\circ$  and night time by  $\theta_{sol} \geq 90^\circ$ . A pixel land/sea mask distinguishes between land, sea and coastal pixels, the latter being those pixels which contain both land and sea surfaces. There are 8 gradations of coastal pixels representing the ratio of land to sea within each pixel. A further pixel land surface type map derived from the International Geosphere-Biosphere Programme (IGBP) surface atlas (Belward, 1996; Francis, 2004) is used to identify specific surface types (e.g. "barren" surfaces) and inland water bodies.

Cloud tests which employ the visible channels can incorrectly identify cloud, referred to as false positives hereafter, near sun glint. This mostly affects water surfaces, but desert regions at high viewing angles such as the Saudi Peninsula can also be affected. Two parameters are calculated to mitigate this:  $\theta_{sat}^{specular}$  is the satellite zenith angle at the centre

of the sun glint area (the solar specular reflection point) and  $\theta_{\text{glint}}$  is a proxy for the intensity of the sun-glint for a given pixel based on geometrical considerations, with  $\theta_{\text{glint}} = 0^\circ$  indicating the solar specular reflection point.  $\theta_{\text{glint}}$  is defined by:

$$\cos(\theta_{\text{glint}}) \equiv 2\cos(\theta_{\text{sol}})\cos(\theta_{\text{sat}}) - \cos(\theta_{\text{scat}}) \quad (2)$$

where  $\theta_{\text{sat}}$  is the satellite zenith angle, and  $\theta_{\text{scat}}$  is the scattering angle defined for the range  $[0^\circ, 180^\circ]$  from backward scattering ( $0^\circ$ ) to forward scattering ( $180^\circ$ ) directions. Sea and inland water pixels are tagged as being affected by sun glint if  $\theta_{\text{glint}} < 25^\circ$ . Some cloud tests employ further conditions to identify sun-glint affected pixels.

There follows a description of each of the tests applied to identify cloud-contaminated pixels. The tests are also tabulated in Table II. The values of offsets and thresholds are given in the Annex. Figures 1 and 2 demonstrate the mask produced by each test for day-time and night-time cases.

### 3.1 Snow test

This test is based on the SAFNWC snow test developed by Derrien and Le Gléau (2005) and is applied over land during the day time and twilight (for  $\theta_{\text{sol}} \leq 85^\circ$ ). It requires the observed reflectances at 0.6, 0.8 and 1.6  $\mu\text{m}$ , the observed brightness temperatures at 3.9, 10.8 and 12.0  $\mu\text{m}$ , and the RTTOV-calculated background clear-sky brightness temperature at 10.8  $\mu\text{m}$ .

In this test, a pixel is flagged as being snow-contaminated if all the following conditions hold true:

$$\frac{R_{(0.6)}^{ob} - R_{(1.6)}^{ob}}{R_{(0.6)}^{ob} + R_{(1.6)}^{ob}} > 0.3 + 0.15(\cos\theta_{\text{scat}} - 1)^2 \quad (3)$$

$$\frac{T_{B(3.9)}^{ob} - T_{B(10.8)}^{ob}}{\cos\theta_{\text{sol}}} < \Delta T_{B(3.9-10.8)}^{snow} \quad (4)$$

$$T_{B(10.8)}^{ob} > T_{B(10.8)}^b - \Delta T_{B(10.8)}^{gross} - \Delta T_{B(10.8)}^{snow} \quad (5)$$

$$T_{B(10.8)}^{ob} < T_{B(10.8)}^{snow} \quad (6)$$

$$T_{B(10.8)}^{ob} - T_{B(12.0)}^{ob} < \Delta T_{B(10.8-12.0)}^{snow} \quad (7)$$

$$R_{(0.8)}^{ob} > R_{(0.8)}^{snow} \quad (8)$$

where  $\Delta T_{B(3.9-10.8)}^{snow}$ ,  $\Delta T_{B(10.8)}^{snow}$ ,  $T_{B(10.8)}^{snow}$ ,  $\Delta T_{B(10.8-12.0)}^{snow}$  and  $R_{(0.8)}^{snow}$  are pre-defined snow test thresholds and offsets, and  $\Delta T_{B(10.8)}^{gross}$  is an offset relating to the gross test (see below). The first two tests exploit the lower reflection of snow and ice compared to water clouds at 1.6

$\mu\text{m}$  and  $3.9 \mu\text{m}$ . Conditions (5) and (6) are based on the fact that snow and ice are expected to have slightly lower temperatures than snow- and ice-free land surfaces beneath a clear sky. Condition (7) discriminates between thin cirrus cloud and snow and ice (see the thin cirrus test below). The final test attempts to prevent shadows being mistaken for snow or ice.

### 3.2 Gross test

This is the most general cloud test, and is applied at all times and for all surface types. It has been used in many previous cloud detection schemes, such as the APOLLO scheme (Kriebel *et al.* 2003). The basic premise of the test is that if the observed brightness temperature at  $10.8 \mu\text{m}$  is colder than that which one would expect given the background cloud-free model profile, then this is likely to be due to the presence of cloud. Hence, a pixel is flagged as being cloud-contaminated if:

$$T_{B(10.8)}^{ob} < T_{B(10.8)}^b - \Delta T_{B(10.8)}^{gross} \quad (9)$$

where  $\Delta T_{B(10.8)}^{gross}$  is a pre-defined offset which can vary with time of day and underlying surface type.

### 3.3 Thin cirrus test

This is another widely used test (e.g. Inoue 1985, Kriebel *et al.* 2003, Derrien and Le Gléau 2005) which exploits the higher absorption by ice clouds at  $12.0 \mu\text{m}$  compared to  $10.8 \mu\text{m}$ : the attenuation by high, thin clouds of radiation emitted by the surface is greater at  $12.0 \mu\text{m}$  than at  $10.8 \mu\text{m}$ . Therefore, the difference in observed brightness temperatures in these channels is expected to be larger than the difference in the corresponding simulated brightness temperatures for pixels containing such cloud. The test is applied at all times and for all surface types. A pixel is flagged as being cloud-contaminated if:

$$T_{B(10.8)}^{ob} - T_{B(12.0)}^{ob} > T_{B(10.8)}^b - T_{B(12.0)}^b + \Delta T_{B(10.8-12.0)}^{thin-cirrus} \quad (10)$$

where  $\Delta T_{B(10.8-12.0)}^{thin-cirrus}$  is a pre-defined offset which is larger over barren surfaces to account for spectral variation in surface emissivity. An additional condition must be satisfied over land surfaces:

$$T_{B(10.8)}^{ob} < 303.15K \quad (11)$$

This reduces the risk of false positives over very warm, moist, cloud-free areas as noted by Derrien and Le Gléau (2005).

### 3.4 Fog/low-cloud test

This test is also common to other cloud masks (e.g. Eyre *et al.* 1984, Saunders and Kriebel 1988, Derrien and Le Gléau 2005). It is applied to night-time and twilight pixels only over all surface types. The emissivity of liquid water clouds is larger at  $10.8 \mu\text{m}$  than at  $3.9 \mu\text{m}$  so the observed brightness temperatures are expected to differ by a greater amount than the

simulated clear-sky brightness temperatures for pixels containing liquid water clouds. A pixel is flagged as being cloud-contaminated if:

$$T_{B(10.8)}^{ob} - T_{B(3.9)}^{ob} > T_{B(10.8)}^b - T_{B(3.9)}^b + \Delta T_{B(10.8-3.9)}^{fog} \quad (12)$$

and:

$$T_{B(10.8)}^{ob} > T_{B(10.8)}^{fog} \quad (13)$$

where  $\Delta T_{B(10.8-3.9)}^{fog}$  is a pre-defined offset and  $T_{B(10.8)}^{fog}$  is a pre-defined threshold. Barren surface types also exhibit higher emissivity at 10.8  $\mu\text{m}$  so a larger offset is used for these surface types to reduce the risk of confusing such surfaces with cloud.

False positives are observed to occur over certain surface types such as barren regions and tropical rainforest, particularly over central Africa and tropical South America. In order to reduce the number of false positives, the test is not applied over these particular surface types if:

$$T_{B(8.7)}^{ob} - T_{B(3.9)}^{ob} < 0.3 \quad (14)$$

Liquid water cloud has a much higher emissivity at 8.7  $\mu\text{m}$  than at 3.9  $\mu\text{m}$  so this condition rarely rejects genuine cloud, whereas the land surfaces in question have more similar emissivities in the case of tropical forest or lower emissivities at 8.7  $\mu\text{m}$  for arid surfaces.

Water clouds also exhibit lower emissivity at 8.7  $\mu\text{m}$  than at 10.8  $\mu\text{m}$  which can be exploited over vegetated (i.e. non-barren) surfaces at high latitudes (for example, Derrien and Le Gleau, 2005) at all times using the 8.7  $\mu\text{m}$  channel. A pixel is flagged as being cloud-contaminated if:

$$T_{B(10.8)}^{ob} - T_{B(8.7)}^{ob} > T_{B(10.8)}^b - T_{B(8.7)}^b + \Delta T_{B(10.8-8.7)}^{fog8.7} \quad (15)$$

and:

$$T_{B(10.8)}^{ob} > T_{B(10.8)}^{fog8.7} \quad (16)$$

where  $\Delta T_{B(10.8-8.7)}^{fog8.7}$  is a pre-defined offset and  $T_{B(10.8)}^{fog8.7}$  is a pre-defined threshold.

### 3.5 Mixed scenes test

The mixed scenes test is similar in principle to the thin cirrus test, and again has a precedent in other cloud masks (e.g. Saunders and Kriebel 1988, Derrien and Le Gléau 2005). It detects thin, high clouds over all surface types at night-time only using the fact that ice clouds absorb more readily at 12.0  $\mu\text{m}$  than at 3.9  $\mu\text{m}$ . A pixel is flagged as being cloud-contaminated if:

$$T_{B(3.9)}^{ob} - T_{B(12.0)}^{ob} > T_{B(3.9)}^b - T_{B(12.0)}^b + \Delta T_{B(3.9-12.0)}^{mixed-scenes} \quad (17)$$

where  $\Delta T_{B(3.9-12.0)}^{mixed-scenes}$  is a pre-defined offset.

### 3.6 Sea Surface Temperature test

This test is applied to sea pixels only. A pixel is flagged as being cloud-contaminated if:

$$SST_{ret} - T^* < \Delta T_{SST} \quad (18)$$

where  $T^*$  is the skin temperature from the background field,  $\Delta T_{SST}$  is a pre-defined offset, and  $SST_{ret}$  is the retrieved sea-surface temperature, calculated from the following regression relationship:

$$SST_{ret} = a_0 T_{B(10.8)}^{ob} + a_1 S T_{B(10.8)}^{ob} + a_2 S (T_{B(10.8)}^{ob} - T_{B(12.0)}^{ob}) + a_3 S^2 (T_{B(10.8)}^{ob} - T_{B(12.0)}^{ob}) + a_4 \quad (19)$$

where  $a_{0-4}$  are pre-defined regression coefficients,  $S = (1/\cos\theta_{sat}) - 1$  and  $\theta_{sat}$  is the satellite zenith angle. The regression coefficients were derived from radiative transfer simulations based on a reference set of atmospheric profiles (Chevallier, 1999), and are given in the Annex.

### 3.7 8.7 $\mu\text{m}$ test

This test is also applied to sea pixels only. A pixel is flagged as being cloud-contaminated if:

$$T_{B(8.7)}^{ob} - T_{B(8.7)}^{pred} > \Delta T_{B(8.7)} \quad (20)$$

where  $\Delta T_{B(8.7)}$  is a pre-defined offset, and  $T_{B(8.7)}^{pred}$  is the 8.7  $\mu\text{m}$  brightness temperature predicted from the following regression relationship:

$$T_{B(8.7)}^{pred} = b_0 T_{B(10.8)}^{ob} + b_1 T_{B(12.0)}^{ob} + b_2 \quad (21)$$

where  $b_{0-2}$  are pre-defined regression coefficients. The regression coefficients were calculated as for the sea surface temperature test.

### 3.8 10.8 $\mu\text{m}$ spatial coherence test

This test is also common to other masks including Saunders and Kriebel (1988) and Derrien and Le Gléau (2005). It makes use of the observation that broken cloud and cloud edges typically exhibit higher spatial variability in their thermal characteristics than the surface beneath, particularly over oceans. A pixel is flagged as being cloud-contaminated if:

$$\sigma[T_{B(10.8)}^{ob}] > \sigma_{cloudy}^{10.8} \quad (22)$$

where  $\sigma[T_{B(10.8)}^{ob}]$  is the standard deviation of the observed 10.8  $\mu\text{m}$  brightness temperature in a 3 x 3 box centred on the pixel in question, and  $\sigma_{cloudy}^{10.8}$  is a pre-defined threshold for this standard deviation which varies with underlying surface type. The test is not applied to land surfaces in daylight and twilight, and is not applied to coastal pixels at all.

### 3.9 HRV spatial coherence test

Each standard SEVIRI pixel corresponds precisely to a 3 x 3 block of HRV pixels. This test uses the standard deviation of reflectances in such a 3 x 3 HRV block,  $\sigma[R_{(HRV)}^{ob}]$ , to identify cloud in the corresponding SEVIRI pixel. For some pixels (defined below),  $\sigma[R_{(HRV)}^{ob}]$  is normalised by the mean of the 9 HRV reflectances,  $\mu[R_{(HRV)}^{ob}]$ . This normalisation can reduce the occurrence of false positives and allows for cloud detection under sun glint conditions. For the purposes of this test, sun glint is defined by  $\theta_{glint} < 40^\circ$ .

For all day-time sea pixels and twilight sun-glint-affected sea pixels, a SEVIRI pixel is flagged as being cloud-contaminated if:

$$\frac{\sigma[R_{(HRV)}^{ob}]}{\mu[R_{(HRV)}^{ob}]} > \sigma_{cloudy}^{HRV, norm} \quad (23)$$

where  $\sigma_{cloudy}^{HRV, norm}$  is a pre-defined threshold which varies according to whether or not pixels are affected by sun glint.

For twilight sea pixels unaffected by sun-glint, a SEVIRI pixel is flagged as being cloud-contaminated if:

$$\sigma[R_{(HRV)}^{ob}] > \begin{cases} \sigma_{cloudy}^{HRV, twi} & \theta_{sol} \leq 85^\circ \\ \sigma_{cloudy}^{HRV, twi} + 0.005 (\theta_{sol} - 85) & \theta_{sol} > 85^\circ \end{cases} \quad (24)$$

where  $\sigma_{cloudy}^{HRV, twi}$  is a pre-defined threshold.

For land pixels in day or twilight, a SEVIRI pixel is flagged as being cloud-contaminated if:

$$\sigma[R_{(HRV)}^{ob}] > \sigma_{cloudy}^{HRV, land} \quad (25)$$

and:

$$\min_{3 \times 3} [R_{(HRV)}^{ob}] > 0.1 \quad (26)$$

where  $\sigma_{cloudy}^{HRV, land}$  is a pre-defined threshold which is larger for barren surface types. For land pixels condition (26) requires that all 9 of the HRV reflectances should exceed 0.1 which reduces false positives caused by the edges of inland water bodies and by cloud shadows.

The test is not applied to snow-contaminated pixels or to coastal pixels. A further check is made on the surface types of the 3x3 block of standard resolution SEVIRI pixels around the pixel under test: they must either be all land or all water (i.e. sea or inland water) for the test to be applied. This further reduces the incidence of false positives at the edges of inland water bodies.

### 3.10 0.8 $\mu\text{m}$ spatial coherence test

This test uses the standard deviation of 0.8  $\mu\text{m}$  reflectances in the 3 x 3 block around the pixel in question,  $\sigma[R_{(0.8)}^{ob}]$ . It is similar in principle to the HRV spatial coherence test and is

useful in capturing some of the small-scale cloud in those regions where HRV data are unavailable. This test is applied only to sea pixels. To reduce false positives around cloud edges the test is not applied if:

$$R_{(0.8)}^{ob} < \mu[R_{(0.8)}^{ob}] \quad (27)$$

where  $R_{(0.8)}^{ob}$  is the reflectance of the central pixel and  $\mu[R_{(0.8)}^{ob}]$  is the mean 0.8  $\mu\text{m}$  reflectance in the 3 x 3 block of pixels.

To reduce false positives due to noise in the channel the test is not applied if:

$$R_{(0.8)}^{ob} < 0.04 \quad (28)$$

Reflectances are calculated under the assumption that the underlying surface (or cloud) behaves approximately as a Lambertian reflector. Under certain bi-directional configurations (notably at high illumination and viewing angles), this assumption can be violated sufficiently that calculated reflectance values may exceed 1.0. To reduce false positives due to these cases the test is not applied if:

$$\max_{3 \times 3} [R_{(0.8)}^{ob}] > 1.0 \quad (29)$$

where the maximum is taken over the reflectances of the 3 x 3 block of pixels.

To reduce false positives due to noise in the channel near the terminator the test is not applied if:

$$\theta_{sol} > 85^\circ \quad (30)$$

To reduce false positives due to cloud shadows in twilight conditions the test is not applied if:

$$\theta_{sol} > 80^\circ \text{ and } R_{(0.8)}^{ob} < 0.1 \quad (31)$$

The severity of sun-glint increases as the solar specular reflection point approaches the edge of the disc. The parameter  $\theta_{sat}^{specular}$  described previously is used to distinguish cases of “moderate” and “severe” sun glint. Moderate sun glint is defined by:

$$35^\circ < \theta_{sat}^{specular} < 65^\circ \text{ and } \theta_{glint} < 25^\circ \quad (32)$$

and severe sun glint is defined by:

$$\theta_{sat}^{specular} > 65^\circ \text{ and } \theta_{glint} < 40^\circ \quad (33)$$

In cases of moderate sun-glint, the test is not applied if:

$$\min_{3 \times 3} [R_{(0.8)}^{ob}] < 0.07 \quad (34)$$

where the minimum is taken over the reflectances of the 3 x 3 block of pixels. This reduces the occurrence of false positives due to cloud shadows caused by high illumination angles.

For day-time pixels unaffected by sun glint, a pixel is flagged as being cloud-contaminated if:

$$\sigma[R_{(0.8)}^{ob}] > \sigma_{cloudy}^{0.8,day} \quad (35)$$

where  $\sigma_{cloudy}^{0.8,day}$  is a pre-defined threshold.

For all twilight pixels, and for day-time pixels affected by moderate sun glint, a pixel is flagged as being cloud-contaminated if:

$$\frac{\sigma[R_{(0.8)}^{ob}]}{\mu[R_{(0.8)}^{ob}]} > \sigma_{cloudy}^{0.8,norm} \quad (36)$$

where  $\sigma_{cloudy}^{0.8,norm}$  is a pre-defined threshold for the normalised standard deviation which varies according to the time of day and whether or not the pixel is affected by sun glint. The test is not applied to pixels affected by severe sun glint.

### 3.11 Visible threshold test

The observation that clouds frequently have higher reflectances than the underlying surface is commonly used for day-time cloud detection (eg Saunders and Kriebel 1988). Over the sea, a pixel is flagged as being cloud-contaminated if:

$$R_{(0.8)}^{ob} > \frac{R_{(0.8)}^{sea}}{(\cos \theta_{sol})^{0.35}} \quad (37)$$

where  $R_{(0.8)}^{sea}$  is a pre-defined reflectance threshold.

Over coasts, a pixel is flagged as being cloud-contaminated if:

$$R_{(0.6)}^{ob} > \frac{R_{(0.6)}^{coast}}{(\cos \theta_{sol})^{0.35}} \quad (38)$$

where  $R_{(0.6)}^{coast}$  is a pre-defined reflectance threshold. In both (37) and (38), the  $\theta_{sol}$ -dependence was empirically derived to account for increased clear-sky reflectances at higher solar zenith angles.

Over land, the EUMETSAT clear-sky reflectance map (CRM) product (EUMETSAT, 2008) is used to generate dynamic thresholds. A pixel is flagged as being cloud-contaminated if:

$$R_{(0.6)}^{ob} > R_{(0.6)}^{CRM} + \Delta R_{(0.6)}^{land} \quad (39)$$

where  $R_{(0.6)}^{CRM}$  is the clear-sky 0.6  $\mu\text{m}$  reflectance and  $\Delta R_{(0.6)}^{land}$  is a pre-defined offset. The nominal CRM product time is 1200 UTC: although these data may be used effectively over much of the disc throughout many slots each day, some restrictions are required to avoid

false positives in the cloud mask due to varying illumination conditions. The CRM data are used only for slots between 0500 UTC (0400 in the northern hemisphere winter) and 2000 UTC. A higher offset is used for slots before 0600 and after 1800, and also at high solar zenith angles to mitigate false positives caused by enhanced forward scattering. In addition, certain barren surfaces (for example the Middle East region) can similarly exhibit strong forward scattering and so fixed thresholds (independent of the CRM data) are employed over these regions under problematic illumination conditions. Finally, if the 0.6  $\mu\text{m}$  CRM reflectance exceeds 0.60 for a pixel, the fixed thresholds are used as such high CRM values are often symptomatic of either snow or cloud contamination in the CRM product. In the cases where the CRM data are not used, a pixel is flagged as cloud-contaminated if:

$$R_{(0.6)}^{ob} > R_{(0.6)}^{land} \quad (40)$$

where  $R_{(0.6)}^{land}$  is a pre-defined threshold. The details of the conditions under which the CRM data are used are given in the Annex.

The test is not applied to any pixels contaminated by snow or ice, or to sea or inland water pixels affected by sun glint.

### 3.12 Visible/near-infrared ratio test

This is another test used by Saunders and Kriebel (1988). It is applied to day-time and twilight pixels, over both sea and land (but not coasts). It exploits the fact that water typically has a smaller reflectance at 0.8  $\mu\text{m}$  than at 0.6  $\mu\text{m}$ , while the opposite is true for vegetated land surfaces. In contrast, clouds have a similar reflectance in the two channels. Over the sea or over inland water, a pixel is flagged as being cloud-contaminated if:

$$\frac{R_{(0.8)}^{ob}}{R_{(0.6)}^{ob}} > R_{(0.8/0.6)}^{sea} \quad (41)$$

where  $R_{(0.8/0.6)}^{sea}$  is given by:

$$R_{(0.8/0.6)}^{sea} = \begin{cases} 0.85 & \theta_{sol} \leq 60^\circ \\ 1.05 - 0.40 \cos \theta_{sol} & 60^\circ < \theta_{sol} \leq 80^\circ \\ 1.23 - 1.4366 \cos \theta_{sol} & 80^\circ < \theta_{sol} \leq 89^\circ \end{cases} \quad (42)$$

Over the land, a pixel is flagged as being cloud-contaminated if:

$$\frac{R_{(0.8)}^{ob}}{R_{(0.6)}^{ob}} < R_{(0.8/0.6)}^{land} \quad (43)$$

where  $R_{(0.8/0.6)}^{land}$  is given by:

$$R_{(0.8/0.6)}^{land} = \begin{cases} 1.60 & \theta_{sol} \leq 60^\circ \\ 1.10 + 1.00 \cos \theta_{sol} & 60^\circ < \theta_{sol} \leq 78^\circ \\ 0.85 + 2.2024 \cos \theta_{sol} & 78^\circ < \theta_{sol} \leq 89^\circ \end{cases} \quad (44)$$

The  $\theta_{sol}$ -dependencies in (42) and (44) were derived empirically. The test is restricted to specific land surface types (see Annex) as it can introduce false positives particularly over water-logged or sparsely vegetated surfaces. False positives can also occur around rivers and lake edges: these are reduced by requiring the pixels in the 3 x 3 block around the pixel under test to be either all land or all water (i.e. sea or inland water). The test is not applied at all to sea or inland water pixels affected by sun glint. In addition, the test is not applied to sea pixels with  $\theta_{glint} < 40^\circ$  and for which HRV data are available on the grounds that the HRV test captures essentially the same cloud with much lower risk of false positives due to sun glint. The test is also not applied to snow-contaminated pixels or to coastal pixels.

### 3.13 Twilight low-cloud temporal differencing algorithm

Derrien and Le Gléau (2007) devised a temporal differencing algorithm which is designed to capture low cloud and fog in twilight conditions. Day-time detection of low cloud is efficiently performed by visible tests. At night-time the differing emissivity of liquid water clouds at 3.9  $\mu\text{m}$  compared to 10.8  $\mu\text{m}$  is exploited. In twilight conditions detection of low clouds is difficult because the visible channels become noisy approaching the terminator and the 3.9  $\mu\text{m}$  channel is contaminated by solar radiation.

The temporal differencing algorithm is applied in two parts: first a seeding process takes place in which twilight pixels in the current slot that were classified as cloudy in the cloud mask from one hour ago and whose thermal characteristics have not changed significantly are flagged as seed pixels. Once all pixels have been examined a region growing process occurs in which seed groups are expanded to encompass adjacent pixels which share the thermal and visible characteristics of the seed group. Some modifications to the original algorithm were made to account for differences between the SAFNWC and Met Office cloud masks.

The seeding process in the Met Office mask requires the pixel to have been flagged as cloudy both in the T-45 minute and T-60 minute cloud masks. Any pixels flagged by the Thin Cirrus, Mixed Scenes or 8.7  $\mu\text{m}$  cloud tests in the T-60 minute mask are rejected from the seeding process since these tests typically detect high-level cloud. Any pixels flagged *only* by one or more of the spatial coherence tests in the T-60 minute mask are also rejected since these tests often flag sub-pixel cloud and cloud edges. Such pixels will often pass the conditions of the seeding process regardless of whether cloud actually remains in the pixel in the current mask. An exception to this last rule is made for sea pixels where the 0.8  $\mu\text{m}$  reflectance exceeds 0.30 in the current mask since this is a good indication that the pixel still contains cloud.

The seeding process is vulnerable to false positives in the prior cloud mask. For this reason snow-contaminated pixels are excluded from the seeding process, as are barren land surface types and coastal pixels. False positives are more common for such pixels than for others.

Sea pixels are flagged as seeds if they satisfy the above conditions and if:

$$\left| T_{B(10.8)}^{ob} - T_{B(10.8)}^{T-60} \right| < \Delta T_{B(10.8)}^{T-60,sea} \quad (45)$$

and:

$$\left| (T_{B(10.8)}^{ob} - T_{B(12.0)}^{ob}) - (T_{B(10.8)}^{T-60} - T_{B(12.0)}^{T-60}) \right| < \Delta T_{B(10.8-12.0)}^{T-60} \quad (46)$$

where  $T_{B(10.8)}^{T-60}$  and  $T_{B(12.0)}^{T-60}$  are the 10.8  $\mu\text{m}$  and 12.0  $\mu\text{m}$  brightness temperatures of the pixel one hour earlier, and  $\Delta T_{B(10.8)}^{T-60,sea}$  and  $\Delta T_{B(10.8-12.0)}^{T-60}$  are pre-defined thresholds.

Land pixels are flagged as seeds if, in addition to the above conditions, they satisfy:

$$\left| T_{B(10.8)}^{ob} - T_{B(10.8)}^{T-60} \right| < \Delta T_{B(10.8)}^{T-60,land} \quad (47)$$

and:

$$\left| (T_{B(10.8)}^{ob} - T_{B(8.7)}^{ob}) - (T_{B(10.8)}^{T-60} - T_{B(8.7)}^{T-60}) \right| < \Delta T_{B(10.8-8.7)}^{T-60} \quad (48)$$

where  $T_{B(8.7)}^{T-60}$  is the 8.7  $\mu\text{m}$  brightness temperature of the pixel one hour earlier, and  $\Delta T_{B(10.8)}^{T-60,land}$  and  $\Delta T_{B(10.8-8.7)}^{T-60}$  are pre-defined thresholds.

All seed pixels are then flagged as cloud-contaminated in the cloud mask.

Once all pixels have been through the seeding process, the region growing algorithm may be invoked. This relies heavily on information in the 0.6  $\mu\text{m}$  channel. Some combinations of viewing and illumination geometry result in this information becoming less reliable as an indicator of cloud. These are typically cases of enhanced forward or backward scattering and generally occur when the terminator is near the edge of the disc, near sun glint, and at high viewing angles. Steps are therefore taken to avoid performing the region growing in these circumstances.

The region growing routine is only called if  $15^\circ < \theta_{sat}^{specular} < 70^\circ$ . The  $\theta_{sat}^{specular}$  parameter is used as a measure of how close the terminator is to the edge of the disc, as well as an indicator of severe sun glint.

Groups of adjacent seed pixels which number fewer than 9 are discarded from the region growing. For those seed groups which remain, the average reflectance at 0.6  $\mu\text{m}$  of the seed group,  $R_{(0.6)}^{AVG}$ , and the average 10.8  $\mu\text{m}$  BT of the seed group,  $T_{(10.8)}^{AVG}$ , are calculated. Any

seed groups for which  $R_{(0.6)}^{AVG} < 0.20$  are discarded from the region growing. This avoids some region growing on seed groups which result from false positives in the prior mask. It also helps to ensure the seed group is clearly discernible in the 0.6  $\mu\text{m}$  channel which reduces the chance of false positives resulting from the region growing. Pixels adjacent to a seed group are added to the group provided they are not snow-contaminated and the following conditions are satisfied:

$$75^\circ < \theta_{sol} < 89^\circ \quad (49)$$

$$\theta_{scat} < 150^\circ \quad (50)$$

$$\theta_{sat} < 70^\circ \text{ whenever } \theta_{sat}^{specular} < 25^\circ \text{ or } \theta_{sat}^{specular} > 50^\circ \quad (51)$$

$$\theta_{glint} > 25^\circ \quad (52)$$

$$\theta_{glint} > 40^\circ \text{ when } \theta_{sat}^{specular} > 50^\circ \quad (53)$$

(water and barren land surfaces only)

$$T_{(10.8)}^{AVG} - 5.0 < T_{B(10.8)}^{ob} < T_{(10.8)}^{AVG} + 0.5 \quad (54)$$

$$R_{(0.6)}^{ob} > \min(1.05 \cdot R_{(0.6)}^{AVG}, factor \cdot R_{(0.6)}^{min}) \quad (55)$$

Snow-contaminated pixels can lead to false positives due to their higher reflectances and so are excluded. Conditions (49) and (50) restrict the region growing to the twilight region and mitigate false positives which can result from strong forward scattering. Condition (51) avoids false positives being added due to higher observed reflectances at large viewing angles when the terminator approaches the edge of the disc. Conditions (52) and (53) avoid false positives which occur near sun glint. Condition (54) ensures that only pixels with temperatures similar to, or slightly colder than, the seed group are added in the region growing process. The final condition exploits the visible information available in the twilight region, looking for pixels which have higher reflectances than the average of the seed group. In this last condition,  $R_{(0.6)}^{min}$  is the smallest allowed reflectance for pixels to be added to the seed group, and is set to 0.45 for barren surfaces and 0.30 over other surface types. *factor* is an empirically-derived multiplier greater than or equal to 1.0 which increases linearly with the solar zenith angle and is applied only when the terminator approaches the edge of the disc (see Annex). This multiplier reduces the risk of false positives being introduced to the mask due to strong forward or backward scattering, or noise in the 0.6  $\mu\text{m}$  channel, close to the terminator. When a new pixel is added to the seed group, the pixels adjacent to it are also processed. In this way the region growing algorithm “fills in” missing cloud and accounts

for cloud motion between the prior slot and the current one. Finally, the pixels added by the region growing process are flagged as cloud-contaminated in the cloud mask.

### 3.14 Partial cloud

Pixels that are flagged as cloudy by one or more of the spatial coherence tests (10.8  $\mu\text{m}$ , HRV and 0.8  $\mu\text{m}$ ) and are not flagged by any other test are designated partially cloudy. These pixels are excluded from the downstream cloud processing which generates products such as cloud top height.

## 4. Effectiveness of cloud detection tests

To illustrate the relative effectiveness of each test Table III lists the proportion of all cloud flagged by each test over all slots on 16 November 2009. The table also lists the amount of cloud flagged *uniquely* by each test as a proportion of the total cloud flagged by the test. In the following discussion this quantity will be referred to as the relative unique cloud fraction (relative in the sense that it is measured as a proportion of the cloud flagged by the test rather than *all* cloud in the mask). It is useful to know what proportion of all cloud in an image is captured by each test. However, some cloud tests are only applied to a subset of all pixels (for example the SST and 8.7  $\mu\text{m}$  tests are only applied to sea pixels), and some tests are designed to detect specific types of cloud (such as the fog/low cloud test). Therefore the relative unique cloud fraction provides a fairer measure of the comparative usefulness of each test.

In the interests of clarity Table III does not break the test statistics down by time of day or surface type. However, since more cloud tests are applied to day time pixels than to night time pixels (as a result of the visible data being used during the day), it is generally true that more pixels are detected by tests uniquely at night than in the day time. Similarly, all of the cloud tests are applied to sea pixels whereas only a subset are applied to land pixels, so the chances of a cloudy pixel being detected by only one test is generally lower over sea than land. The relative homogeneity of sea surfaces compared to land also means that cloud test thresholds may be set to allow more cloud to be flagged over sea compared to land, while still avoiding false positives. This also increases the likelihood that more than one test will detect a given cloud-contaminated pixel over sea compared to land. It follows that individual tests become more important to the mask at night time and over land.

The gross test, being the most general in principle and being applied to all surface types at all times of day, captures the largest proportion of the cloud (82.55%) and a relatively large fraction of this cloud (10.23%) is flagged uniquely by this test. It flags 85.91% of all night time cloud, and 78.55% of all day time cloud. Over land this difference is likely to result from the NWP model failing to represent land surface temperature extremes well in all cases: at night model surface temperatures may be too warm in some regions which can result in false positives appearing in the cloud mask, while in the daytime the model surface temperatures may sometimes be too cold resulting in cloud being missed from the mask.

The thin cirrus test is useful for capturing semi-transparent ice cloud which may often be missed by other tests. It is most useful in the day time over land (it captures 28.08% of all cloud in this category, 22.15% of which is uniquely flagged by this test). The mixed scenes test detects 25.80% of all night time cloud, but most of this cloud is also captured by other tests (in particular the thin cirrus and gross tests). Over sea, high cloud is often captured by the 8.7  $\mu\text{m}$  and SST tests. The result is that the thin cirrus and mixed scenes tests each flag very little cloud uniquely over ocean.

The fog/low cloud test has the largest relative unique cloud fraction of all tests (16.86%). This is not surprising since low cloud is not well captured by any other test at night time, particularly over land. The twilight low-cloud temporal differencing algorithm complements the fog/low cloud test. Being designed to detect a very specific type of cloud, this only flags 1.78% of all cloudy pixels. The impact of the test on the cloud mask is dependent on synoptic conditions: the test can be very beneficial in cases of extensive low stratus which are often not fully captured by the other cloud tests under twilight conditions: Figure 3 shows an example of this.

The HRV spatial coherence test has a comparatively high unique cloud fraction of 9.48%. This test captures 30.94% of day time cloud over sea with a relative unique cloud fraction of 10.26%. This compares with 13.58% of day time cloud over land, 7.33% of which is uniquely identified by this test. The HRV channel is able to resolve sub-pixel cloud which the other channels cannot detect, and, as with other tests, the test thresholds may be set rather lower over ocean than land due to the greater homogeneity of the background resulting in more cloud being flagged over ocean.

The 0.8  $\mu\text{m}$  spatial coherence test does not generally add much cloud to the mask beyond that flagged by the HRV test where HRV data are available. However, in areas where the HRV data are unavailable the 0.8  $\mu\text{m}$  test is particularly useful over sea in capturing some of the small cumulus missed by other tests.

The 10.8  $\mu\text{m}$  spatial coherence test flags a large proportion of all cloud (60.56%), with 11.35% of this being flagged by this test alone. This test is particularly effective for sea pixels, flagging 77.26% of all cloud over ocean, and 11.92% of this uniquely. These latter pixels mostly contain broken cloud and cloud edges.

The SST test also flags a large proportion of cloud over oceans (68.84%), but only 0.34% of this is flagged by this test alone for the reasons given above. Similarly, the 8.7  $\mu\text{m}$  test also flags only a small amount of cloud uniquely.

The use of the clear-sky reflectance product is beneficial for the visible threshold test, allowing for particularly efficient thresholds over land while admitting few false positives (see Sections 5 and 6). The test flags 59.94% of all day time cloud over land, and 9.09% of these

---

pixels are detected by this test alone. Over sea this test flags 49.33% of day time cloud, but very little of this is flagged only by this test.

The thresholds in the visible/NIR ratio test admit less cloud over land surfaces than sea, and as a result the test flags more cloud over sea than land (57.92% compared to 34.30%). In fact the test is not applied over certain land surface types as areas which are sparsely vegetated or which are covered by substantial amounts of surface water can result in false positives. The relative unique cloud fraction for day time land pixels is 5.88% compared to 0.25% for day time sea.

An important caveat for this analysis is that it does not take false positives into account: pixels wrongly flagged by a test will often only be flagged by that one test, and so false positives in the mask will tend to increase the count of uniquely flagged pixels erroneously. The following sections compare the Met Office SEVIRI cloud mask with that of the SAFNWC and with the MODIS cloud mask. These comparisons indicate those tests most likely to introduce false positives.

## **5. Comparison against SAFNWC SEVIRI cloud mask**

The Nowcasting Satellite Applications Facility (SAFNWC) also generates a SEVIRI cloud mask based on a series of threshold tests (Derrien and Le Gléau, 2005). A number of cloud tests are common to the SAFNWC and Met Office masks. However, a significant way in which the masks differ is in how test thresholds are determined: the SAFNWC mask employs pre-calculated look-up tables derived from large numbers of off-line radiative transfer simulations on a wide range of model profiles. The look-up tables are indexed by geometrical variables such as solar zenith angle and satellite zenith angle, NWP fields such as surface temperature and total column water vapour, and other ancillary data such as elevation and climatological data. This contrasts with the Met Office cloud mask in which thresholds are derived from clear-sky radiative transfer simulations based on recent NWP forecast fields, ancillary datasets, and empirically derived relationships with variables such as solar zenith angle, as detailed in Section 3.

Table IV compares the Met Office and SAFNWC masks for all slots on 16 November 2009. “Cloud matches” refers to the percentage of pixels flagged by each test which were cloudy in the SAFNWC mask. This is a measure of the skill of the test against the SAFNWC mask. “Cloud matches as % of all cloud” gives the proportion of all the SAFNWC cloud which is flagged by the test, and likewise “cloud mismatches as % of all clear” gives the percentage of SAFNWC clear pixels flagged as cloudy by the test. In the following discussion it should be borne in mind that the SAFNWC mask does not represent the truth in terms of cloudy/clear classification. Agreement between the masks does not guarantee that both masks are correct. Likewise, “cloud mismatches” do not always represent false positives in the Met Office mask. The comparison of the two masks serves to highlight those cloud tests which *may* be introducing false positives into the cloud mask.

The gross test flags the largest proportion of SAFNWC clear pixels (5.09%). For land pixels 92.79% of the flagged pixels agree with the SAFNWC mask compared to 98.61% over sea. The difference is due to the reliance of the gross test on accurate model surface temperature, and, in general, land surface temperatures are not characterised as well as sea surface temperatures in NWP models. If the model surface temperatures are too warm, the thresholds will be set too high thus increasing the risk of false positives. This situation is not uncommon over parts of Africa at night.

The twilight temporal differencing algorithm exhibits comparatively low skill among the tests, with 92.34% of the pixels flagged by the test being cloudy in the SAFNWC mask. False positives (due to other tests) in the prior cloud mask are often perpetuated by the seeding process. This is the most common source of false positives introduced by the temporal differencing scheme. At high viewing and solar zenith angles, atmospheric dust and aerosol can produce higher reflectances in the 0.6  $\mu\text{m}$  channel and this can cause the region growing process to add cloud-free pixels to the mask. For example, false positives occur with greater frequency around the Middle East region where the atmosphere is frequently laden with sufficient quantities of dust to cause problems for the algorithm.

The comparison suggests that the fog/low cloud test has the lowest skill of all the tests: of all pixels flagged by the test, 85.74% are cloudy in the SAFNWC mask, though there is a large difference between the value for sea pixels – 89.18% – and that for land pixels – 75.31%. In fact, while false positives do occur, there are occasions in which the fog/low cloud test flags cloud missed by the SAFNWC mask. Figure 4 shows such an example. Over sea, the majority of false positives introduced by this test occur at high viewing angles. This may be a result of the CO<sub>2</sub> “limb-cooling” effect in the 3.9  $\mu\text{m}$  channel which is not captured well by the radiative transfer simulations since RTTOV is being used at the limits of its specifications towards the edge of the Earth disc.

The HRV test also appears to have relatively low skill, but the 2009 version of the SAFNWC mask used in this comparison does not make use of HRV data. On the basis of comparisons with visible imagery, it is believed that the majority of the SAFNWC clear pixels flagged by this test are in fact cloud-contaminated. Figure 5 shows an example. The 2010 release of the SAFNWC cloud mask will exploit HRV spatial coherence information (Derrien *et al.*, 2009).

Many pixels flagged by the 0.8  $\mu\text{m}$  spatial coherence test which are clear in the SAFNWC mask are also in fact cloud-contaminated (see for example Figure 6). Thus the skill of this test is somewhat higher than indicated in Table IV.

The 10.8  $\mu\text{m}$  spatial coherence test also flags a relatively large proportion (2.56%) of clear SAFNWC pixels. The SAFNWC mask also employs a version of the 10.8  $\mu\text{m}$  test and the discrepancies are mostly located around cloud edges and areas of small cumulus since it is

these areas where small differences between the thresholds used by each mask have the greatest impact.

Overall, 95.06% of the pixels flagged by the thin cirrus test are cloudy in the SAFNWC mask. There is extremely good agreement over sea where 99.18% of flagged pixels are cloudy in the SAFNWC mask. It is somewhat worse over land (92.75%), where warm, clear regions with high water vapour content can result in false positives.

The visible threshold test agrees very well with the SAFNWC over both land and sea, with 97.38% and 99.28% of flagged pixels matching the SAFNWC mask respectively. Before the CRM data were used, fixed thresholds were applied over land surfaces: it was found that the CRM data allowed a significant amount of extra cloud to be detected with very few additional false positives introduced (Hocking *et al.*, 2009). The good agreement between this test and the SAFNWC mask over land suggests the SAFNWC approach compares very favourably with the use of the clear-sky reflectance data.

There is also good agreement between the visible/NIR ratio test and the SAFNWC mask, with 91.53% and 99.31% of flagged pixels being cloudy in the SAFNWC for land and sea respectively. The lower value for land pixels is due to the greater chance of false positives with this test over land surfaces mentioned earlier.

The mixed scenes and SST tests agree with the SAFNWC mask to a very high degree. As noted above, they add only a small amount of cloud to the mask that is not flagged by any other test: the result of the comparison is not then surprising, since most of the pixels they flag as cloudy are corroborated by other cloud tests.

Overall the Met Office and SAFNWC masks give the same clear/cloudy classification to 92.57% of the pixels tested.

## **6. Comparison against the MODIS cloud mask**

The Moderate Resolution Imaging Spectro-radiometer (MODIS) carried onboard the polar orbiting Terra and Aqua satellites has 36 channels in visible, near-infrared and thermal infrared wavelengths. All channels have a spatial resolution of at least 1 km at the sub-satellite point. The Level 2 MODIS cloud mask product (Ackerman *et al.*, 2006) is generated from a series of threshold tests, some of which have analogues in the Met Office SEVIRI mask. MODIS pixels which are processed successfully by the cloud detection algorithm are flagged as high confidence clear (>99% confidence level for clear), confidently clear (>95%), uncertain (>66%), or cloudy. The MODIS mask does not represent the “truth” in terms of cloud-contamination, but nevertheless is a useful product for the purposes of validating the SEVIRI cloud mask.

The Met Office and SAFNWC SEVIRI cloud masks were compared with the MODIS cloud mask for all MODIS granules which fell within the MSG field-of-view on 16 November 2009. Each MODIS granule is scanned over a 5 minute period, while each SEVIRI full-disc scan takes (nominally) 15 minutes. The 3 granules scanned during a given 15-minute SEVIRI scan interval were matched up with that SEVIRI slot. The maximum temporal difference between the MODIS and SEVIRI data was therefore 15 minutes.

To compare the SEVIRI and MODIS masks, each MODIS pixel was mapped onto the nearest SEVIRI pixel. Each SEVIRI pixel for which at least 10% of the mapped MODIS pixels were flagged cloudy was flagged as cloud-contaminated. Each SEVIRI pixel for which at least 90% of the mapped MODIS pixels were confidently clear (i.e. with greater than 95% confidence) was flagged as clear. SEVIRI pixels for which neither condition was true were excluded from the comparison: this excludes SEVIRI pixels with too many MODIS pixels mapped to them being classified as uncertain. The cut-off for cloudy pixels was set at 10% since the SEVIRI HRV channel has similar resolution to MODIS at the sub-satellite point (1km) so it would be expected that both instruments can resolve clouds on this scale (where HRV data exist), and HRV pixels are approximately 10% the size of a normal SEVIRI pixel. The comparison was restricted to SEVIRI pixels with MSG satellite zenith angles of 75° or less.

The resulting MODIS mask was compared on a pixel-by-pixel basis with the SEVIRI masks. Some differences between the masks can be ascribed to errors in registration between MODIS and SEVIRI pixels which are largely due to parallax effects, and also to the difference in scan times.

Table V shows the results of the comparison for each Met Office cloud test with the MODIS mask. The terms “cloud hits”, “clear hits”, “false positives” and “false negatives” in Tables V and VI are used for convenience, and are in respect to the MODIS cloud mask rather than the “truth”. Note that due to Terra and Aqua over-pass times there are very few match-ups covering twilight illumination conditions and so there are insufficient data to compare the output of the twilight temporal differencing algorithm with the MODIS mask. The results are consistent with those in Table IV for most of the tests. One notable exception is the fog/low cloud test which shows much higher skill against the MODIS mask supporting the idea that the true skill of the test is higher than indicated by the comparison with the SAFNWC mask. The HRV spatial coherence test also appears more skillful when compared to MODIS, though still only 92.76% of pixels flagged by the test were cloudy in the MODIS mask. This may be partly a result of errors in registration between the MODIS and SEVIRI pixels: much of the cloud flagged by this test is fragmented which will be more susceptible to registration errors than more extensive areas of contiguous cloud.

The 10.8  $\mu\text{m}$  spatial coherence test shows lower skill against the MODIS mask. As with the HRV test, many discrepancies between the masks occur with fragmented cloud and around

cloud edges due to registration errors. However the 10.8  $\mu\text{m}$  test does introduce some false positives around cloud edges.

The thin cirrus and mixed scenes tests also show lower skill against the MODIS mask, especially over land. However, the MODIS cloud detection algorithm identifies some semi-transparent cloud as clear which the Met Office and SAFNWC masks flag as cloud. The MODIS mask has an additional flag for thin cirrus cloud detected under day time conditions using its 1.38  $\mu\text{m}$  channel. This flag indicates high cloud which is sufficiently thin that it may be corrected for when retrieving surface properties using visible channels. By considering MODIS pixels with this flag set to be cloudy, the thin cirrus hit rate against MODIS rises to 94.54%. (With this extra thin cirrus flag, the proportion of all pixels for which the Met Office and MODIS masks agree drops by 0.41% and for the SAFNWC mask it drops by 0.36%). Some of the remaining discrepancies appear to be due to thin cirrus not captured by the MODIS mask at night (see Figure 7). Over sea the thin cirrus and mixed scenes tests agree well with MODIS (cloud hit rates of 98.59% and 99.33% respectively).

Table VI shows the overall statistics for the comparison between the Met Office and MODIS masks, and the SAFNWC and MODIS masks. The Met Office mask agrees with the MODIS mask for 90.65% of all pixels tested. The SAFNWC mask agrees with the MODIS mask for 89.69% of pixels tested. The results indicate that the Met Office mask is flagging slightly more cloud correctly (as compared to MODIS) than the SAFNWC mask while introducing a similar proportion of false positives. This demonstrates that overall the Met Office SEVIRI mask is performing as well as or better than the SAFNWC mask when compared to the MODIS cloud mask. The SAFNWC mask sometimes captures semi-transparent cloud more effectively than the Met Office cloud mask (as seen in Figure 7), while the Met Office mask can be more efficient at detecting low cloud at night (see Figure 4, for example).

## 7. Summary

Data from MSG are used by the Met Office for the production of imagery products, nowcasting and assimilation in NWP. Processing of the raw data received via EUMETCast is undertaken to produce a range of quantitative products for the nowcasting system. A fundamental pre-processing step is the identification of cloud-contaminated pixels. This paper has described the threshold tests applied to MSG imagery in real-time to distinguish cloudy, partially cloudy, and clear pixels.

An analysis of the relative effectiveness of the tests has been presented which indicates those tests that are most valuable to the mask. The results show that the gross test flags the most cloud overall and a substantial proportion of this is not captured by any other test over both land and sea. The 10.8  $\mu\text{m}$  spatial coherence and SST tests also flag a large proportion of cloud over sea, though while the 10.8  $\mu\text{m}$  test is useful for detecting broken cloud and cloud edges (particularly at night), the SST test flags only a small amount of cloud uniquely. The HRV spatial coherence test captures small-scale cloud missed by other tests (again

especially useful over sea), and where HRV data are unavailable the 0.8  $\mu\text{m}$  spatial coherence test captures some of this cloud. Most cloud tests applied over land detect significant amounts of cloud which no other test captures due to the greater difficulties in cloud detection over relatively heterogeneous land surfaces (as compared to sea). This is especially true of the gross and thin cirrus tests, and the visible threshold and visible/near-infrared ratio tests. The visible tests also flag a large amount of cloud over sea, but much of this is also captured by other tests. The fog/low cloud test and twilight low-cloud temporal differencing algorithm are both valuable for capturing low cloud under night time and twilight conditions. The 8.7  $\mu\text{m}$  and mixed scenes tests add relatively little cloud to the mask beyond the other tests, but are retained in the interests of detecting as much cloud as possible and function as a “back-up” in cases where other channel data are missing from an image.

Comparisons with the SAFNWC SEVIRI cloud mask and the MODIS cloud mask have been used to validate the cloud mask and to indicate which tests are most likely to introduce false positives. In general, false positives are more common over land surfaces. The gross test sometimes flags too much cloud over land, especially at night, as a result of errors in model surface temperatures. The thin cirrus test is vulnerable to false positives over land in regions of high water vapour content. The visible/near-IR ratio test also introduces some false positives over land surfaces: this is more common at high solar zenith angles. The 10.8  $\mu\text{m}$  spatial coherence test can flag too many pixels around cloud edges. The fog/low cloud test introduces some false positives, but not as many as the comparison against the SAFNWC mask would suggest. Finally, the twilight temporal differencing algorithm is vulnerable to false positives from other tests in the prior cloud mask. Despite these differences, the Met Office and SAFNWC masks show good agreement, especially over ocean. The two masks identify 90.65% and 89.69% of pixels correctly as compared to the MODIS cloud mask.

An important aspect of the Met Office cloud detection is the use of simulated clear-sky BTs based on recent NWP forecast fields. By using the difference between observed and simulated BTs in the cloud tests, atmospheric effects are accounted for (notably  $\text{CO}_2$  absorption in the 3.9  $\mu\text{m}$  channel), enabling relatively tight thresholds to be used. The benefit of this approach is seen where these tests detect cloud missed by alternative cloud detection schemes. This is particularly true of the fog/low cloud test, which captures low cloud missed by the SAFNWC mask as noted above. Furthermore, improvements to NWP models over time are expected to have a positive impact on those tests which make use of the simulated data.

## **8. Acknowledgements**

The authors thank Jean-Pierre Olry at Météo-France for providing the SAFNWC cloud mask data. The SAFNWC mask images shown are copyright SATMOS (INSU/CNRS-Météo-France). The MODIS data were obtained from the Level 1 and Atmosphere Archive and Distribution System (LAADS).

---

## 9. References

Ackerman S, Strabala K, Menzel W, Frey R, Moeller C, Gumley L, Baum B, Seemann SW, Zhang H. 2006. *Discriminating Clear-Sky from Cloud with MODIS Algorithm Theoretical Basis Document (MOD35)*. [[http://modis-atmos.gsfc.nasa.gov/MOD35\\_L2/atbd.html](http://modis-atmos.gsfc.nasa.gov/MOD35_L2/atbd.html)]

Belward AS (ed.). 1996. The IGBP-DIS global 1 km land cover data set (DISCover) - proposal and implementation plans. IGBP-DIS Working Paper No. 13, Toulouse, France, 61 pp.

Chevallier F. 1999. *TIGR-like sampled databases of atmospheric profiles from the ECMWF 50-level forecast model*. NWP SAF Research Report No. 1, 18pp. [Available from the librarian at ECMWF].

Derrien M, Le Gléau H. 2005. MSG/SEVIRI cloud mask and type from SAFNWC. *International Journal of Remote Sensing*, **26**: 4707-4732.

Derrien M, Le Gléau H. 2007. Temporal-differencing and region-growing techniques to improve twilight low cloud detection from SEVIRI data. In *Proceedings of Joint 2007 EUMETSAT Meteorological Satellite Conference and the 15th Satellite Meteorology & Oceanography Conference of the American Meteorological Society, Amsterdam, The Netherlands, 24-28 September 2007*. EUMETSAT P.50, ISBN 92-9110-079-X.

Derrien M, Le Gléau H, Raoul M-P. 2009. SAFNWC/MSG cloud product recent enhancements. In *Proceedings of the 2009 EUMETSAT Meteorological Satellite Conference, Bath, UK, 21-25 September 2009*. EUMETSAT P.55, ISBN 978-92-9110-086-6.

EUMETSAT. 2006. *EUMETCast - EUMETSAT's Broadcast System for Environmental Data*. EUM TD 15. [[http://www.eumetsat.int/Home/Main/Documentation/Technical\\_and\\_Scientific\\_Documentation/Technical\\_Descriptions/index.htm](http://www.eumetsat.int/Home/Main/Documentation/Technical_and_Scientific_Documentation/Technical_Descriptions/index.htm)]

EUMETSAT. 2008. *MSG Meteorological Products Extraction Facility Algorithm Specification Document*. EUM.MSG.SPE.022. [[http://www.eumetsat.int/Home/Main/Documentation/Technical\\_and\\_Scientific\\_Documentation/Technical\\_Notes/SP\\_1124282585834?!=en](http://www.eumetsat.int/Home/Main/Documentation/Technical_and_Scientific_Documentation/Technical_Notes/SP_1124282585834?!=en)]

Eyre JR, Brownscombe JL, Allam, RJ. 1984. Detection of fog at night using Advanced Very High Resolution Radiometer (AVHRR) imagery. *Meteorological Magazine*, **113**: 266.

Francis PF. 2004. *The Development of an Infrared Land Surface Emissivity Atlas, and its Comparison with MODIS/Terra Products*. Met Office Forecasting Research Technical Report No. 405.

---

Golding BW. 1998. Nimrod: A system for generating automated very short range forecasts. *Meteorological Applications*, **5**: 1-16.

Hocking J, Francis PN, Saunders RW. 2009. The Use of Ancillary Surface Data Products in the Met Office SEVIRI Cloud Mask. In *Proceedings of the 2009 EUMETSAT Meteorological Satellite Conference, Bath, UK, 21-25 September 2009*. EUMETSAT P.55, ISBN 978-92-9110-086-6.

Inoue T. 1985. On the Temperature and Effective Emissivity Determination of Semi-Transparent Cirrus Clouds by Bi-Spectral Measurements in the 10  $\mu\text{m}$  Window Region. *Journal of the Meteorological Society of Japan*, **63**: 88-99.

Kelly G, Francis PN. 2008. Use of clear SEVIRI radiances in the Met Office Meso-scale Assimilation system. In *Proceedings of the 2008 EUMETSAT Meteorological Satellite Conference, Darmstadt, Germany, 8-12 September 2008*. EUMETSAT P.52, ISBN 978-92-9110-082-8.

Kriebel KT, Gesell G, Kästner M, Mannstein H. 2003. The cloud analysis tool APOLLO: improvements and validations. *International Journal of Remote Sensing*, **24**: 2389-2408.

Lutz HJ, König M. 2008. The Use of Surface Emissivity Information within the Meteosat Second Generation Meteorological Product Generation. In *Proceedings of the 2008 EUMETSAT Meteorological Satellite Conference, Darmstadt, Germany, 8-12 September 2008*. EUMETSAT P.52, ISBN 978-92-9110-082-8.

Munro R, Köpken C, Kelly G, Thépaut J-N, Saunders RW. 2004. Assimilation of Meteosat radiance data within the 4D-Var system at ECMWF: Data quality monitoring, bias correction and single-cycle experiments. *Quarterly Journal of the Royal Meteorological Society*, **130**: 2293-2313.

Saunders RW, Kriebel KT. 1988. An improved method for detecting clear sky and cloudy radiances from AVHRR data. *International Journal of Remote Sensing*, **9**: 123-150.

Saunders RW, Matricardi M, Brunel P. 1999. An improved fast radiative model for assimilation of satellite radiance observation. *Quarterly Journal of the Royal Meteorological Society*, **125**: 1407-1425.

Schmetz J, Pili P, Tjemkes S, Just D, Kerkmann J, Rota S, Ratier A. 2002. An introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological Society*, **83**: 977-992.

Seemann SW, Borbas EE, Knuteson RO, Stephenson GR, Huang H-L. 2008. Development of a Global Infrared Land Surface Emissivity Database for Application to Clear Sky Sounding

Retrievals from Multi-spectral Satellite Radiance Measurements. *Journal of Applied Meteorology and Climatology*, **47**: 108-123.

## 10. Annex – Test Thresholds, Offsets and Coefficients

This annex lists the values of the thresholds, offsets and regression coefficients for the tests described in Section 3. It also provides details of the conditions under which the tests are applied where these were omitted in Section 3 for the sake of clarity.

### 10.1 Snow test

The following table gives the values of the offsets and thresholds used in the snow test.

Parameter	Value	Equation reference
$\Delta T_{B(3.9-10.8)}^{snow}$	10.00 K	(4)
$\Delta T_{B(10.8)}^{snow}$	8.00 K	(5)
$T_{B(10.8)}^{snow}$	283.15 K	(6)
$\Delta T_{B(10.8-12.0)}^{snow}$	2.00 K	(7)
$R_{(0.8)}^{snow}$	0.20	(8)

### 10.2 Gross test

The value of  $\Delta T_{B(10.8)}^{gross}$  in equation (9) varies with surface type and time-of-day as follows:

Surface	Time-of-day	Value
Sea	Day/twilight/night	2.50 K
Land, snow test success	Day/twilight only	9.00 K
Land/coast model snow	Day/twilight/night	4.00 K
Land, barren	Day	10.00 K
Land, barren	Twilight	6.00 K
Land, barren	Night	4.00 K
All other land/coast	Day/twilight/night	3.50 K

Barren surfaces are those with IGBP surface types of 7, 8, 9, 10, 16, or 22.

### 10.3 Thin cirrus test

The value of  $\Delta T_{B(10.8-12.0)}^{thin-cirrus}$  in equation (10) is as follows:

Surface	Time-of-day	Value
Sea	Day/twilight/night	1.4 K
Land/coast, non-barren	Day/twilight/night	1.4 K
Land/coast, barren	Day/twilight/night	1.9 K

Barren surfaces are those with IGBP surface types of 6, 7, 8, 9, 10, 16, or 22.

## 10.4 Fog/low cloud test

For the 3.9  $\mu\text{m}$  fog/low cloud test, the value of  $\Delta T_{B(10.8-3.9)}^{fog}$  in equation (12) is as follows:

Surface	Time-of-day	Value
Sea	Twilight/night	3.8 K
Land/coast, non-barren	Twilight/night	3.5 K
Land/coast, barren	Twilight/night	6.0 K

Barren surfaces are those with IGBP surface types of 7, 8, 9, 10, 16, or 22. The value of  $T_{B(10.8)}^{fog}$  in equation (13) is 258.0 K. The extra check specified in equation (14) is applied to barren surfaces and also to pixels with IGBP surface type 2.

The 8.7  $\mu\text{m}$  fog/low cloud test is only applied at high latitudes, defined by pixel lines lower than 700 and greater than 3200. The test is not applied over barren surfaces, defined by IGBP surface types 6, 7, 8, 9, 10, and 16. The value of  $\Delta T_{B(10.8-8.7)}^{fog8.7}$  in equation (15) is given by:

Surface	Time-of-day	Value
Sea	Day/twilight/night	2.5 K
Land/coast	Day/twilight/night	2.2 K

The value of  $T_{B(8.7)}^{fog8.7}$  in equation (16) is 253.0 K.

## 10.5 Mixed scenes test

The value of  $\Delta T_{B(3.9-12.0)}^{mixed-scenes}$  in equation (17) is given by:

Surface	Time-of-day	Value
Sea	Day/twilight/night	4.0 K
Land/coast	Day/twilight/night	3.0 K

## 10.6 Sea Surface Temperature test

The regression coefficients in equation (19) are as follows:

$$a_0 = 1.01248, a_1 = 0.010237, a_2 = 0.08866, a_3 = -0.013593, a_4 = -2.96384.$$

The value of  $\Delta T_{SST}$  in equation (18) is -7.5 K.

## 10.7 8.7 $\mu\text{m}$ test

The regression coefficients in equation (21) are as follows:

$$b_0 = 0.5373, b_1 = 0.4500, b_2 = 2.6106.$$

The value of  $\Delta T_{B(8.7)}$  in equation (20) is 0.89 K.

### 10.8 10.8 $\mu\text{m}$ spatial coherence test

The value of  $\sigma_{cloudy}^{10.8}$  in equation (22) is as follows:

Surface	Time-of-day	Value
Sea	Day/twilight/night	0.5 K
Land/coast	Night	2.5 K

### 10.9 HRV spatial coherence test

The thresholds for the HRV spatial coherence test are as follows:

Parameter	Surface	Time-of-day	Value	Equation reference
$\sigma_{cloudy}^{HRV,norm}$	Sea, no sun-glint	Day	0.08	(23)
$\sigma_{cloudy}^{HRV,norm}$	Sea, sun-glint	Day/twilight	0.06	(23)
$\sigma_{cloudy}^{HRV,twi}$	Sea, no sun-glint	Twilight	0.018	(24)
$\sigma_{cloudy}^{HRV,land}$	Land, non-barren	Day/twilight	0.06	(25)
$\sigma_{cloudy}^{HRV,land}$	Land, barren	Day/twilight	0.08	(25)

Barren land surfaces are defined by IGBP surface types 7, 8, 9, 10, 16, or 22.

### 10.10 0.8 $\mu\text{m}$ spatial coherence test

The thresholds for the 0.8  $\mu\text{m}$  spatial coherence test are as follows:

Parameter	Surface	Time-of-day	Value	Equation reference
$\sigma_{cloudy}^{0.8,day}$	Sea, no sun-glint	Day	0.01	(35)
$\sigma_{cloudy}^{0.8,norm}$	Sea, sun-glint	Day	0.12	(36)
$\sigma_{cloudy}^{0.8,norm}$	Sea	Twilight	0.15	(36)

### 10.11 Visible threshold test

The value of  $R_{(0.8)}^{sea}$  in equation (37) is 0.20, and the value of  $R_{(0.6)}^{coast}$  in equation (38) is 0.40.

For land pixels the choice of threshold depends on a number of conditions. The value of  $R_{(0.6)}^{land}$  in equation (40) is 0.65. This static (CRM-independent) threshold is applied under the following conditions:

- For slots before 0500 UTC (0400 in winter) and slots after (and including) 2000 UTC.

- In the autumn, winter, and spring, for pixels with X coordinates above 2900 and Y coordinates below 2300 and with  $\theta_{sol} > 65^\circ$  (this mitigates against false positives over South America at high illumination angles).
- In the spring, summer, and autumn, for barren pixels (IGBP surface types 6, 7, 8, 9, 10, or 16) with  $\theta_{sat} > 50^\circ$  and Y coordinates above 2300 (this avoids persistent false positives over the Middle East).
- In the spring, summer, and autumn, for pixels with X coordinates below 1050 and Y coordinates above 2050 and *either*  $\theta_{sol} > 70^\circ$  *or* if the slot is before 0600 UTC (this reduces false positives over the Middle East and north-east Africa at high solar zenith angles).
- If the 0.6  $\mu\text{m}$  CRM reflectance exceeds 0.60.

Winter is defined as November, December, and January. Summer is defined as May, June, and July. These definitions are based on the maximal solar elevation. Pixel coordinates are measured from the south-east corner of the MSG imagery.

The value of  $\Delta R_{(0.6)}^{land}$  in equation (39) takes the value 0.08, except under the following conditions, when it is set to 0.16:

- For slots before 0600 UTC and slots after (and including) 1800 UTC.
- In winter, if  $\theta_{sol} > 70^\circ$ .
- In spring, summer, and autumn, if  $\theta_{sol} > 60^\circ$ .

### 10.12 Visible/near-infrared ratio test

The thresholds for this test are given in section 3 above. However, there are some restrictions on the application of the test to land surfaces:

- The test is called for all land surfaces whose IGBP surface types are *not* in the following list: 6, 7, 8, 9, 10, 11, 16, 19, 20, 21, and 22.
- For IGBP surface type 19, the test is only applied to pixels with Y coordinate greater than 3230, and to pixels with X coordinate less than 1790 and Y coordinate greater than 3020.
- For IGBP surface type 20, the test is only applied to pixels with Y coordinate greater than 3220, and to pixels with X coordinate less than 1790 and Y coordinate greater than 3175.
- For IGBP surface type 21, the test is only applied to pixels with Y coordinate greater than 1420.

Pixel coordinates are measured from the south-east corner of the MSG imagery.

### 10.13 Twilight low-cloud temporal differencing algorithm

The following table lists the thresholds used in the seeding process:

Parameter	Value	Equation reference
$\Delta T_{B(10.8)}^{T-60, sea}$	1.00 K	(45)
$\Delta T_{B(10.8-12.0)}^{T-60}$	0.60 K	(46)
$\Delta T_{B(10.8)}^{T-60, land}$	1.00 K	(47)
$\Delta T_{B(10.8-8.7)}^{T-60}$	0.50 K	(48)

The term *factor* in equation (55) is set to 1.0 if  $25^\circ < \theta_{sat}^{specular} < 60^\circ$  (in this range the terminator is not too near the edge of the disc). For  $\theta_{sat}^{specular}$  outside this range, *factor* is defined by:

$$factor = \begin{cases} 1.0 & \theta_{sol} \leq 80^\circ \\ 1.0 + 0.1 (\theta_{sol} - 80) & \theta_{sol} > 80^\circ \end{cases} \quad (56)$$

For the purposes of the temporal differencing algorithm, barren land surfaces are defined by IGBP surface types 7, 8, 9, 10, and 16.

Channel	Descriptor	Central wavelength (µm)	Band (µm)	Comments
1	VIS006	0.6	0.56-0.71	
2	VIS008	0.8	0.74-0.88	
3	IR_016	1.6	1.50-1.78	
4	IR_039	3.9	3.48-4.36	IR + solar
5	WV_062	6.2	5.35-7.15	H <sub>2</sub> O absorption
6	WV_073	7.3	6.85-7.85	H <sub>2</sub> O absorption
7	IR_087	8.7	8.30-9.10	
8	IR_097	9.7	9.38-9.94	O <sub>3</sub> absorption
9	IR_108	10.8	9.80-11.8	Window
10	IR_120	12.0	11.0-13.0	Window
11	IR_134	13.4	12.4-14.4	CO <sub>2</sub> absorption
12	HRV	0.75	0.60-0.90	

Table I. Summary of SEVIRI channels.

Cloud test	Day/Night/ Twilight?	Land/Sea/ Coast?	Observations required	Ancillary data required
<b>Snow</b>	Day time and twilight only ( $\theta_{sol} \leq 85^\circ$ )	Land only	$R_{(0.6)}^{ob}, R_{(0.8)}^{ob}, R_{(1.6)}^{ob},$ $T_{(3.9)}^{ob}, T_{(10.8)}^{ob}$	$T_{(10.8)}^b$
<b>Gross</b> $T_{B(10.8)}^{ob} < thresh_{gross}$	All	All	$T_{(10.8)}^{ob}$	$T_{(10.8)}^b$
<b>Thin cirrus</b> $T_{B(10.8)}^{ob} - T_{B(12.0)}^{ob} > thresh_{thincirrus}$	All	All	$T_{(10.8)}^{ob}, T_{(12.0)}^{ob}$	$T_{(10.8)}^b, T_{(12.0)}^b$
<b>Fog/low-cloud</b>	(i) Night and twilight (ii) All	(i) All (ii) All	$T_{(3.9)}^{ob}, T_{(8.7)}^{ob}, T_{(10.8)}^{ob}$	$T_{(3.9)}^b, T_{(8.7)}^b, T_{(10.8)}^b$
<b>Mixed scenes</b> $T_{B(3.9)}^{ob} - T_{B(12.0)}^{ob} > thresh_{mixedscenes}$	Night only	All	$T_{(3.9)}^{ob}, T_{(12.0)}^{ob}$	$T_{(3.9)}^b, T_{(12.0)}^b$
<b>SST</b> $SST_{retrieved} - T^* < thresh_{SST}$	All	Sea only	$T_{(10.8)}^{ob}, T_{(12.0)}^{ob}$	Pre-calculated regression coefficients, background skin temperature ( $T^*$ ) Pre-calculated regression coefficients
<b>8.7 <math>\mu\text{m}</math></b> $T_{B(8.7)}^{ob} - T_{B(8.7)}^{pred} > thresh_{8.7}$	All	Sea only	$T_{(8.7)}^{ob}, T_{(10.8)}^{ob}, T_{(12.0)}^{ob}$	
<b>10.8 <math>\mu\text{m}</math> spatial coherence</b>	All	Land (night only) and sea only	$T_{(10.8)}^{ob}$ in surrounding 3x3 block of pixels	
<b>HRV spatial coherence</b>	Day time and twilight only	Land and sea only	$R_{(HRV)}^{ob}$ in 3x3 block of high resolution pixels	
<b>0.8 <math>\mu\text{m}</math> spatial coherence</b>	Day time and twilight only ( $\theta_{sol} \leq 85^\circ$ )	Sea only	$R_{(0.8)}^{ob}$ in surrounding 3x3 block of pixels	
<b>Visible threshold</b> $R_{(0.8)}^{ob} > thresh_{visible}^{sea}$ (sea) $R_{(0.6)}^{ob} > thresh_{visible}^{land}$ (land)	Day only	All	$R_{(0.6)}^{ob}, R_{(0.8)}^{ob}$	EUMETSAT clear sky reflectance data
<b>Visible/NIR ratio</b> $R_{(0.8)}^{ob} / R_{(0.6)}^{ob} > thresh_{visnir}^{sea}$ (sea) $R_{(0.8)}^{ob} / R_{(0.6)}^{ob} < thresh_{visnir}^{land}$ (land)	Day time and twilight only ( $\theta_{sol} \leq 89^\circ$ )	Land and sea only	$R_{(0.6)}^{ob}, R_{(0.8)}^{ob}$	
<b>Twilight temporal differencing</b>	Twilight only ( $75^\circ < \theta_{sol} < 90^\circ$ )	All	$R_{(0.6)}^{ob}, T_{(8.7)}^{ob}, T_{(10.8)}^{ob},$ $T_{(12.0)}^{ob}, T_{(8.7)}^{T-60}, T_{(10.8)}^{T-60},$ $T_{(12.0)}^{T-60}$	Cloud mask from T-60 min and T- 45 min prior slots

Table II. Summary of cloud mask tests.

<b>Cloud Test</b>	<b>Pixels flagged as % of total cloud</b>	<b>Pixels flagged uniquely as % of all pixels flagged by test</b>
<b>Gross</b>	82.55	10.23
<b>Thin cirrus</b>	12.22	7.39
<b>Fog/low cloud</b>	5.39	16.86
<b>Mixed scenes</b>	12.67	0.46
<b>SST</b>	48.86	0.34
<b>8.7 <math>\mu\text{m}</math></b>	14.57	1.05
<b>10.8 <math>\mu\text{m}</math> spatial coh.</b>	60.56	11.35
<b>HRV spatial coh.</b>	13.40	9.48
<b>0.8 <math>\mu\text{m}</math> spatial coh.</b>	12.95	2.60
<b>Visible threshold</b>	22.40	3.06
<b>Visible/NIR ratio</b>	24.67	1.30
<b>Twilight temp. diff.</b>	1.78	3.88

Table III. Illustrating the relative effectiveness of cloud tests. Statistics based on all slots from 16 November 2009.

<b>Cloud Test</b>	<b>Cloud matches (%)</b>	<b>Cloud matches as % of all cloud</b>	<b>Cloud mismatches as % of all clear</b>
<b>Gross</b>	96.79	82.01	5.09
<b>Thin cirrus</b>	95.06	11.93	1.16
<b>Fog/low cloud</b>	85.74	4.82	1.50
<b>Mixed scenes</b>	99.21	12.91	0.19
<b>SST</b>	99.00	49.66	0.93
<b>8.7 <math>\mu\text{m}</math></b>	97.24	14.54	0.77
<b>10.8 <math>\mu\text{m}</math> spatial coh.</b>	97.80	60.79	2.56
<b>HRV spatial coh.</b>	90.29	12.42	2.50
<b>0.8 <math>\mu\text{m}</math> spatial coh.</b>	96.59	12.84	0.85
<b>Visible threshold</b>	98.69	22.69	0.56
<b>Visible/NIR ratio</b>	98.03	24.82	0.93
<b>Twilight temp. diff.</b>	92.34	1.68	0.26

Table IV. Comparison of the Met Office cloud tests with the SAFNWC cloud mask over all slots on 16 November 2009. “Cloud matches” refers to the percentage of pixels flagged by each test which were cloudy in the SAFNWC mask. “Cloud matches as % of all cloud” gives the proportion of all the SAFNWC cloud which is flagged by the test, and likewise “cloud mismatches as % of all clear” gives the percentage of SAFNWC clear pixels flagged as cloudy by the test.

Cloud Test	Cloud hits (%)	Cloud hits as % of all cloud	False positives as % of all clear
<b>Gross</b>	97.24	80.72	5.33
<b>Thin cirrus</b>	93.11	9.72	1.68
<b>Fog/low cloud</b>	96.51	4.82	0.41
<b>Mixed scenes</b>	95.92	9.59	0.95
<b>SST</b>	99.00	51.74	1.22
<b>8.7 <math>\mu</math>m</b>	98.63	13.98	0.45
<b>10.8 <math>\mu</math>m spatial coh.</b>	94.20	58.09	8.34
<b>HRV spatial coh.</b>	92.76	11.16	2.03
<b>0.8 <math>\mu</math>m spatial coh.</b>	97.44	14.73	0.90
<b>Visible threshold</b>	99.28	27.29	0.46
<b>Visible/NIR ratio</b>	98.67	28.48	0.89

Table V. Comparison of the Met Office cloud tests with the MODIS cloud mask for all MODIS granules falling within an MSG satellite zenith angle of 75° on 16 November 2009.

	Cloud hits (%)	Clear hits (%)	False positives (%)	False negatives (%)
<b>Met Office</b>	64.79	25.86	4.16	5.18
<b>SAFNWC</b>	63.88	25.81	4.21	6.09

Table VI. Summary of the comparison of the Met Office and SAFNWC SEVIRI masks against the MODIS cloud mask. Cloud and clear hits show the proportion of cloudy/clear SEVIRI pixels which agree with the MODIS mask. False positives shows the proportion of pixels flagged as cloudy which were clear in the MODIS mask, and false negatives shows the proportion of pixels flagged as clear which were cloudy in the MODIS mask.

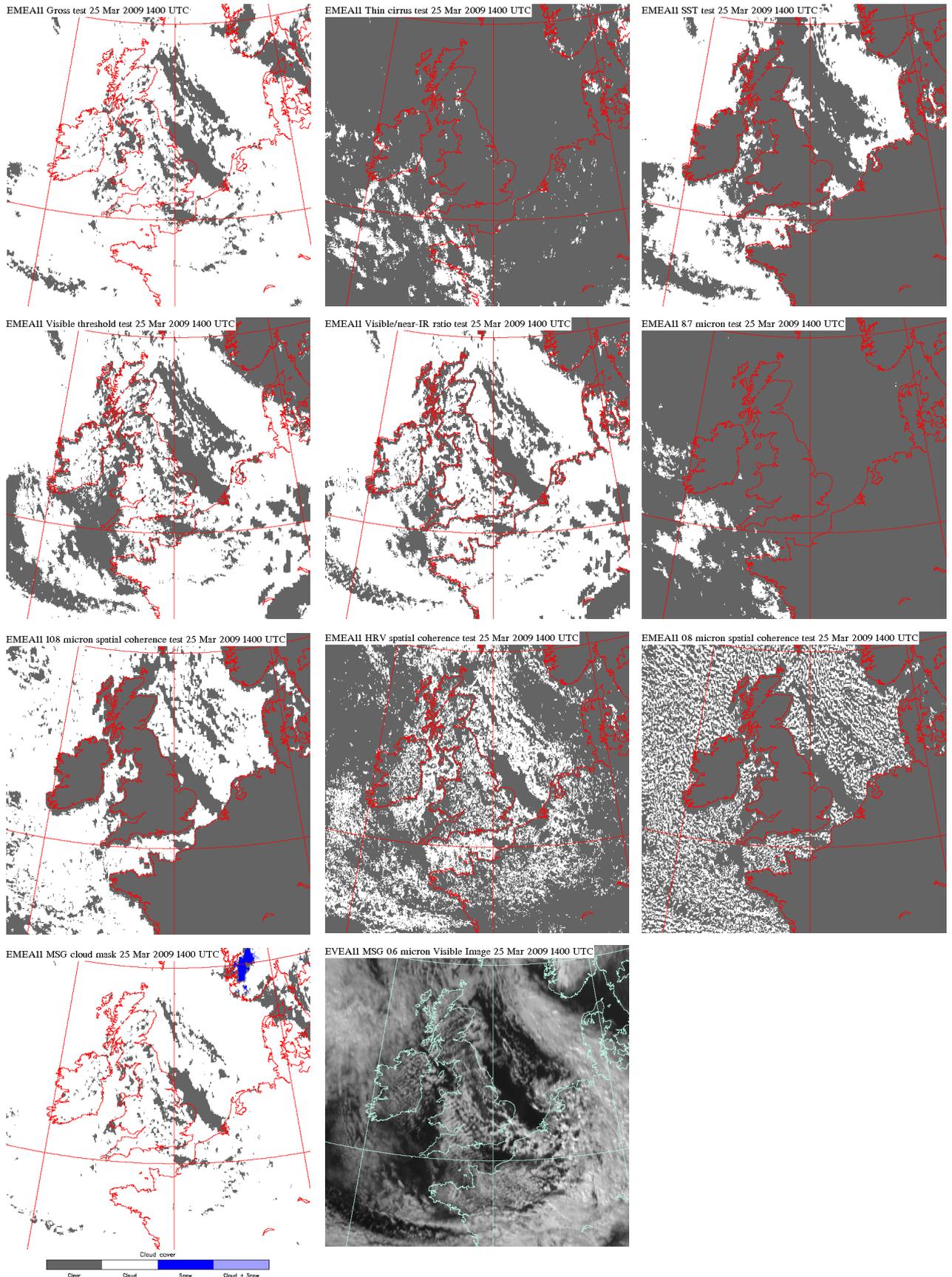


Figure 1. Illustrating cloud flagged by each day time cloud test on a case from 25 March 2009 at 1400 UTC.

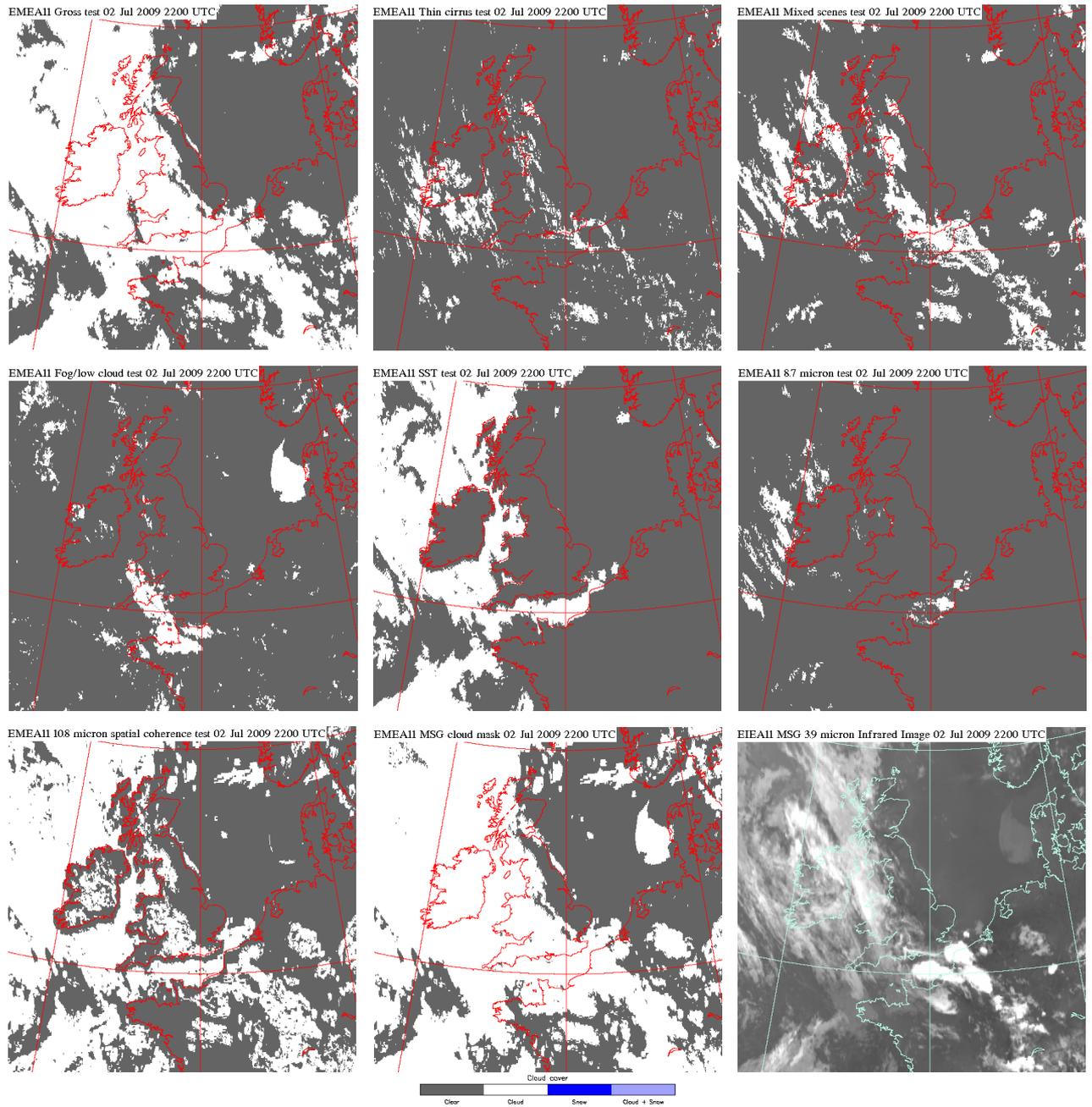


Figure 2. Illustrating cloud flagged by each night time cloud test on a case from 2 July 2009 at 2200 UTC.

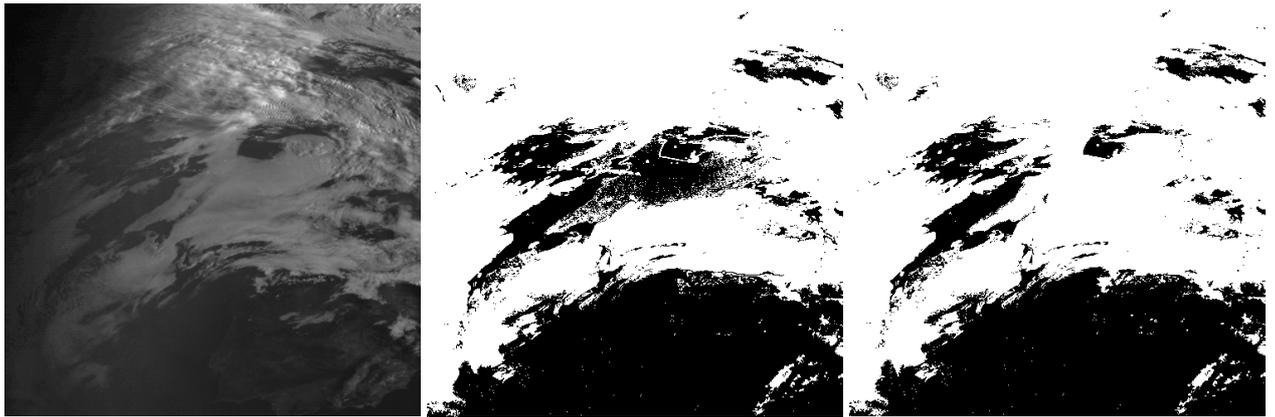


Figure 3. A case from 0700 UTC on 2 April 2008 which demonstrates the impact of the twilight low-cloud temporal differencing algorithm. The image on the left shows the 0.6  $\mu\text{m}$  channel. The central image shows the mask without the temporal differencing algorithm. The image on the right shows the full mask.

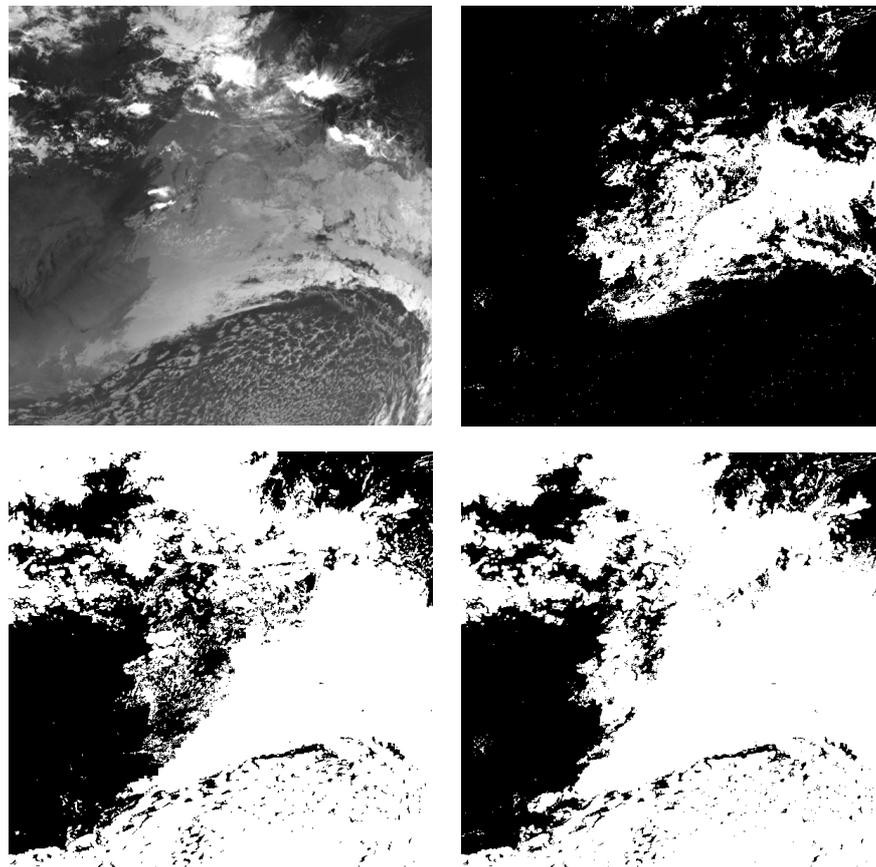


Figure 4. A case from 2100 UTC on 16 November 2009. The top left image shows the 3.9  $\mu\text{m}$  channel in which an area of low stratus is clearly visible over South Africa and neighbouring countries. The top right image shows the corresponding mask from the fog/low cloud test. The bottom right image shows the full Met Office cloud mask for the scene. The bottom left image is the SAFNWC cloud mask. Although the Met Office mask does not capture the cloud fully, some of the cloud flagged will appear as false positives when compared to the SAFNWC mask.

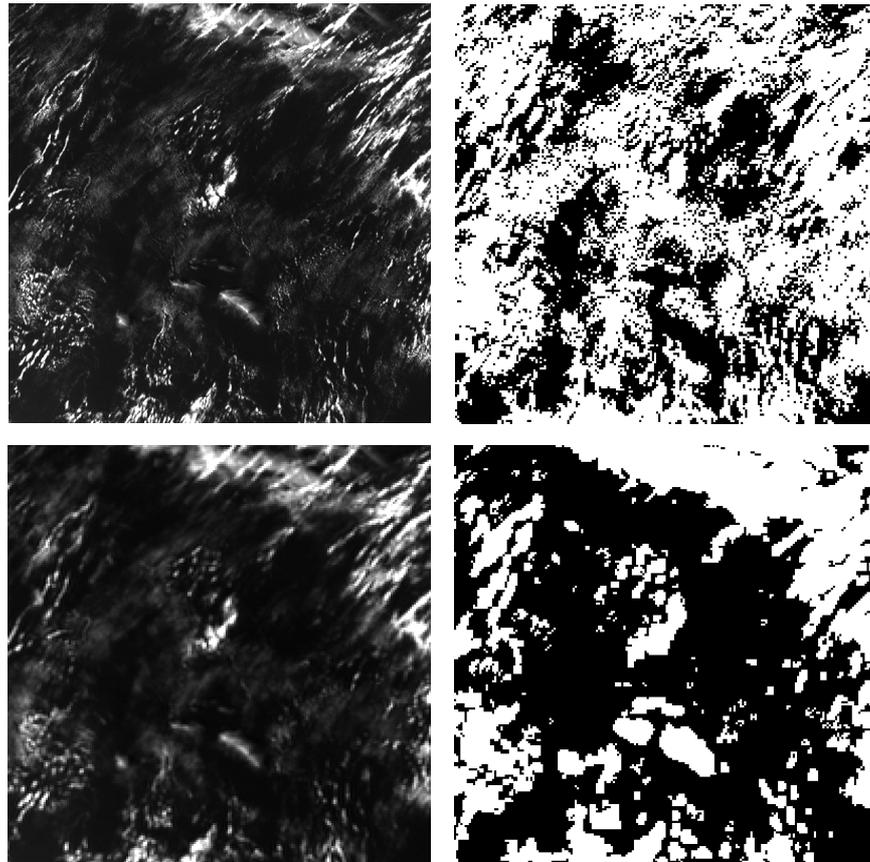


Figure 5. A case from 1500 UTC on 16 November 2009. The images on the left are from the HRV (top) and 0.6  $\mu\text{m}$  (bottom) channels with contrast stretching to highlight tenuous and broken cloud. The top right image shows the mask due to the HRV spatial coherence test. The bottom right image shows the SAFNWC mask: this version of the mask does not exploit the HRV data.

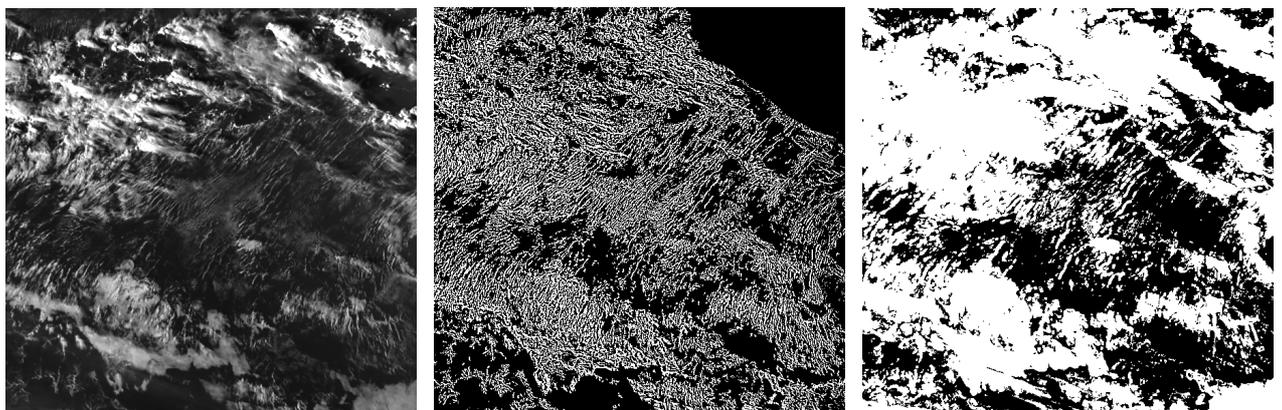


Figure 6. A case from 1730 UTC on 16 November 2009. The image on the left shows the 0.6  $\mu\text{m}$  visible channel. The centre image shows the corresponding mask from the 0.8  $\mu\text{m}$  spatial coherence cloud test. The image on the right shows the SAFNWC cloud mask. The 0.8  $\mu\text{m}$  test captures some of the broken cloud missed by the SAFNWC mask.

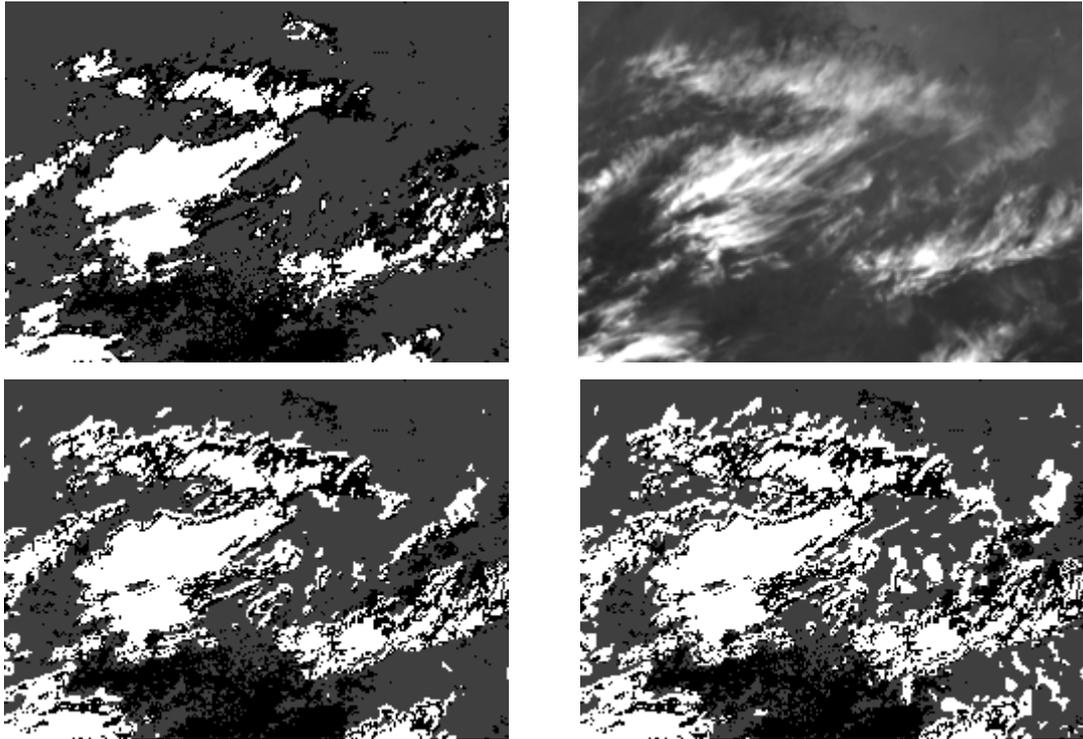


Figure 7. This example from the 2140 UTC Terra MODIS granule on 16 November 2009 shows the MODIS cloud mask mapped to SEVIRI pixels on the top left. The corresponding Met Office and SAFNWC masks are shown on the lower left and right respectively. The 10.8  $\mu\text{m}$  SEVIRI channel is shown on the top right. In the masks, cloud-flagged pixels are white, and clear-flagged pixels are grey. The black pixels have been excluded from the MODIS comparison (see text). The MODIS mask misses rather a lot of the thin cirrus cloud. The SAFNWC mask captures more of this cloud than the Met Office mask.

**Met Office**  
FitzRoy Road, Exeter  
Devon EX1 3PB  
United Kingdom

Tel: 0870 900 0100  
Fax: 0870 900 5050  
[enquiries@metoffice.gov.uk](mailto:enquiries@metoffice.gov.uk)  
[www.metoffice.gov.uk](http://www.metoffice.gov.uk)