A method for the detection of clouds using AVHRR infrared observations

V. THIERMANN† and E. RUPRECHT

Institut für Meereskunde der Universität Kiel, 2300 Kiel 1, Germany

Abstract. A method is proposed for cloud detection over sea using a single infrared channel of a high spatial resolution satellite radiometer. The method combines a spatial coherence test and an absolute threshold test for the temperature field. The threshold for the second test is automatically determined by a histogram analysis of data which were preselected by the coherence test. The coherence test used differs from the classical standard deviation test in the way of the test value computation: the weight of the pixel to be investigated is increased. A simulation shows that under conditions of low cloud cover this makes the proposed coherence test more sensitive than the standard deviation test. When applied to real data the test was also found to be very sensitive compared with visible and IR threshold tests. Cloud cover obtained by the whole method is finally compared with that resulting from the visible and IR histogram analysis procedure of Phulpin *et al.* This comparison indicates a high reliability of the proposed method.

1. Introduction

The detection of cloud-contaminated pixels is an essential step in the application of infrared satellite data for remote sensing of the Earth and atmosphere. Even a small percentage of cloud cover within a radiometer pixel can affect the clear atmosphere radiance in such a way that the determination of sea surface temperature or temperature and humidity profiles becomes impossible. Routine processing of satellite data requires reliable automated cloud detection algorithms which have to be applicable to the wide range of possible scenes.

Presently there are two IR radiometers flying aboard NOAA satellites: the AVHRR (Advanced Very High Resolution Radiometer) and the HIRS (High Resolution Infrared Radiation Sounder). The latter provides an extensive set of narrow band radiance data for temperature and humidity profiling. The spatial resolution, however, is limited to 17·4 km diameter at nadir view. Thus, HIRS pixels are often partially cloud-filled. Compared to HIRS, the spectral information of the AVHRR is low. On the other hand, its spatial resolution of 1·1 km pixel diameter at nadir is much more suitable for cloud detection since this high spatial resolution provides a more advantageous cloud to pixel size ratio. Even for sparse cloud coverage, some of the AVHRR pixels can be significantly filled by clouds. It also enables the spatial structure of the observed field to be evaluated. These advantages make AVHRR the more suitable radiometer for cloud detection. The cloud pattern observed by the AVHRR can then be used for analysing HIRS data.

---

†Now with Forschungsinstitut für Optik, Schloss Kressbach, 7400 Tübingen, Germany

At present, cloud detection is often based on absolute thresholds for sunlight backscatter or thermal emission. The discrimination of clouds results from their higher albedo in the visible range and lower brightness temperature in the infrared compared to the cloud-free background. The results of such methods are very sensitive to the threshold values, which must be chosen from scene to scene. Uncertainty of the exact background, i.e. the radiance under assumed cloud free conditions, is the main source of error. The ocean background depends for example on the wind velocity (which affects the surface albedo) and the atmospheric aerosol content in the visible. In the infrared it depends on the surface temperature and atmospheric temperature and humidity. All these quantities vary in space and time, therefore the threshold value must be individually adapted to be optimum (for a detailed discussion see Rossow *et al.* (1985)). For operational application the threshold values must be chosen automatically by algorithms. Histogram analysis is one method to be applied. Unfortunately in many scenes, histograms cannot be unequivocally interpreted. Cluster analysis simultaneously using information from different wavelengths diminishes the ambiguity in many cases (Simmer *et al.* 1982; Phulpin *et al.* 1983), but at night data from the visible range are not available.

Another type of cloud detection evaluates the spatial coherence of the observed image. These methods do not need absolute thresholds since they look at relative structures. They are usually based on IR observations and therefore applicable day and night. The IR brightness temperature depends on the temperature and emissivity of the radiating surface (ocean, land, clouds) and also on atmospheric contributions according to the temperature and humidity fields. The atmospheric temperature and humidity fields as well as the emissivity and temperature of the ocean surface are relatively homogeneous. The emission of cloud surfaces on the other hand is in most cases spatially very variable. This is caused either by the typical topography of the cloud surface through its height related temperature as with cumulus clouds or by the variability of the cloud transmission according to the cloud's thickness and density as with cirrus or thin stratus clouds. This spatial inhomogeneity of cloud top IR radiances makes the spatial coherence of IR data a valuable measure of cloudiness above a homogeneous background like the sea.

The coherence technique most commonly used is the standard deviation test proposed by Coakley and Bretherton (1982) and Saunders (1986). In order to classify a pixel of brightness temperature T_{22} within the surrounding pixel field

$$\begin{array}{ccc} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{array}$$

the standard deviation of the nine pixels is computed. T_{22} is flagged cloud-contaminated (or cloudy) if this standard deviation exceeds a certain threshold S_L.

This method generally provides good results. However, in §3 we will show, that it does not have maximum sensitivity if only a few pixels of the field are cloud contaminated. This can be understood by considering a single cloudy pixel T_{22} in a clear environment. The effect of a reduced temperature T_{22} on the total variance is small because only one of the nine pixels differs significantly in temperature. The weight of the central pixel for computation of the test value is only as high as that of any of the surrounding pixels. In §3, we will discuss this in more detail. Before, in §2, we introduce a modified coherence method having an increased weight of the central pixel. This method will then be compared with threshold techniques in §4.

In order to increase the reliability, the coherence test will be supplemented by a threshold test in §5. A new method for the threshold determination will be demonstrated. Finally, in §6, results of an automated scheme, which combines the improved coherence test and the threshold test, is compared with results of the two-dimensional visible and IR histogram procedure of Phulpin *et al.* (1983).

2. Description of coherence test

Consider the pixel with brightness temperature T_{22} to be investigated. The following algorithm is proposed: Four test values x_1 to x_4 are computed which are (half) the sums of the absolute values of the differences between T_{22} and its neighbours in the north-south, east-west and diagonal directions, such that:

$$x_1 = (|T_{12} - T_{22}| + |T_{32} - T_{22}|)/2$$
$$x_2 = (|T_{21} - T_{22}| + |T_{23} - T_{22}|)/2$$
$$x_3 = (|T_{11} - T_{22}| + |T_{33} - T_{22}|)/2$$
$$x_4 = (|T_{13} - T_{22}| + |T_{31} - T_{22}|)/2$$

If one of the four test values exceeds a critical threshold X_L, T_{22} is called cloudy, otherwise clear.

Since always differences with respect to T_{22} are considered, the sensitivity of the test value to T_{22} is higher than the sensitivity of the whole pixel field variance to T_{22}. The consequences of this increased weight of T_{22} in our test scheme are discussed in the next section.

3. Comparison with standard deviation test

Before we can compare the proposed test with the 'classical' standard deviation test we have to find equivalent threshold values. Too large a threshold leads to too low a sensitivity, too low a threshold leads to too high a rate of false detection (i.e. flagging a clear pixel cloudy). The last effect is caused by the noise of the clear pixels (radiometric noise plus clear pixel temperature scatter).

In order to find equivalent thresholds we consider nine cloud-free pixels. In a statistical analysis we assume the noise to be normally distributed and use its standard deviation σ_n to normalize the thresholds. The probability of false detection (T_{22} is flagged 'cloudy') as a function of the normalized thresholds S_L/σ_n and X_L/σ_n is given in figure 1. Threshold values S_L and X_L which produce the same false detection rate are taken to be equivalent. If we accept a false detection rate of 1 per cent we find a pair of corresponding thresholds of about 1·6 (standard deviation test) and 3·6 (proposed coherence test). We can now renormalize the curve for the proposed coherence test in a way that both tests have the same threshold value at a 1 per cent false detection rate. The result is plotted as a dashed curve. This curve can directly be compared with the solid curve of the standard deviation test, as done in the next paragraph.

Above we regarded σ_n to be the noise of the clear background. We can also regard σ_n to be the scatter of a field of nine cloud-contaminated pixels. This means, for example, if the scatter due to clouds is double the clear background noise, the normalized threshold becomes half the clear background value. Now the curves express the increase of the number of detected cloudy pixels with increasing cloud caused temperature scatter (or decreasing normalized thresholds). For a 1 per cent false detection rate we can compare the dashed curve of the proposed test with the solid curve of the standard deviation test. For example if the scatter is double as high
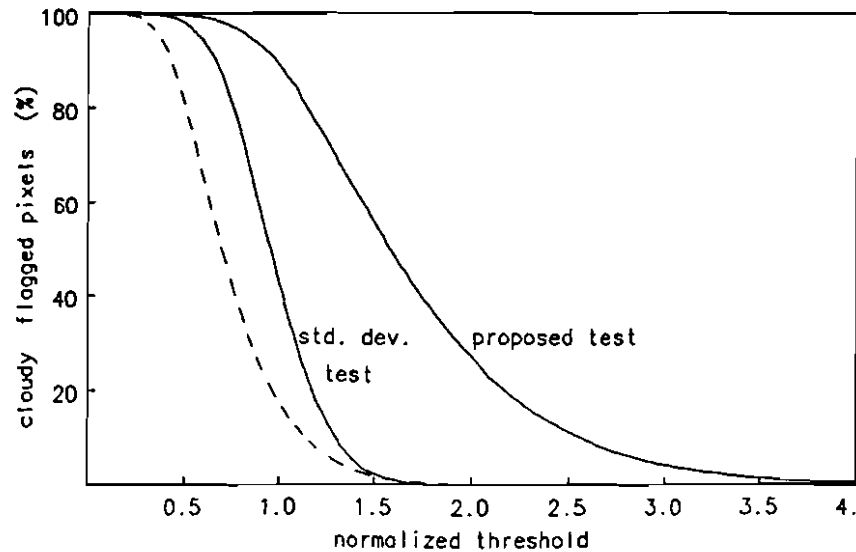
Figure 1. The probability that a pixel is flagged cloudy as a function of the normalized thresholds S_L/σ_n for the standard deviation test and X_L/σ_n for the proposed coherence test. σ_n is the noise of the pixel field. The dashed curve is a renormalization of the solid curve for the proposed coherence test to equal the solid curve for the standard deviation test at a probability of 1 per cent.

as the noise leading to a 1 per cent false detection rate (normalized threshold of about 0·8), roughly 75 per cent of the pixels are removed by the standard deviation test whereas only 35 per cent are removed by the proposed coherence test. At first view this result seems to demonstrate that the standard deviation test is generally more sensitive than the proposed coherence test. But we have to keep in mind that we considered the special case of a field of nine cloud-contaminated pixels (unimodal temperature distribution).

We will now consider a single cloud-contaminated pixel T_{22} in a clear environment. In a statistical sense, the cloud-contaminated pixel has a reduced (or possibly also increased) temperature compared to the adjacent ones. The probability that this pixel is detected as being cloud-contaminated is given in figure 2 as a function of the difference between T_{22} and the expected temperature of the clear pixels, normalized by the background noise value σ_n. The two pairs of curves correspond to threshold values which lead to 1 per cent and 10 per cent false detection under cloud free conditions (values are read off the solid curves of figure 1). As figure 2 shows, for both false detection rates the probability that the cloudy pixel is detected is up to about two times higher for the proposed test than for the standard deviation test.

With increasing cloud cover it becomes more likely that more than the central pixels of the test field are cloud-contaminated also. Such situations have been simulated as follows. A large field was produced with a noise of 0·06 K. (This noise value was found to be typical for our AVHRR data even though it is much lower than the commonly used value of 0·12 K just for the radiometric noise.) A certain fraction of these pixels was then simulated to be cloud-contaminated. This was done by reducing their temperature by a random value between 0·2 K and 2 K (uniform

distribution). The spatial pattern of these 'cloud–contaminated' pixels was arbitrary without any tendency of grouping them into clusters.

The two tests were applied to this data field. Threshold values of 0·1 K for S_L and 0·22 K for X_L have been used. These values correspond to a 1 per cent false detection rate for clear sky. Figure 3 shows how many of the cloud-free pixels were left 'cloud-free' by the respective methods as a function of the fraction of cloud-contaminated pixels. We will call the fraction of the cloud-contaminated pixels in the following simple 'cloud cover' even though the real cloud cover, taking into account that pixels are often only partially filled by clouds, might be much less.

For a zero cloud cover the definition of the threshold causes 99 per cent of the cloud-free pixels to be correctly identified. With increasing cloud cover, this fraction becomes quickly smaller. All coherence tests suffer from the fact that false detection is likely for clear pixels directly adjacent to cloud-contaminated pixels. The high degree of false detection in this example especially follows from the spatially unorganized location of the cloudy pixels. However we see that the proposed method is slightly less susceptible to false detection. And even though the absolute difference is small, in cases of medium cloud cover our method might leave significantly more clear pixels for subsequent use: e.g. for a 40 per cent cloud cover in the graph the number of 'cloud free' pixels left by our method is more than one third higher.

In figure 4 the percentage of cloudy pixels which are not detected is given as a function of cloud cover. It can be seen that the sensitivity of the proposed method is significantly higher for cloud covers less than about 30 per cent. If the cloud cover increases the difference of both methods decreases. (For a 100 per cent cloud cover we finally find the situation which was described by means of figure 1, where the standard
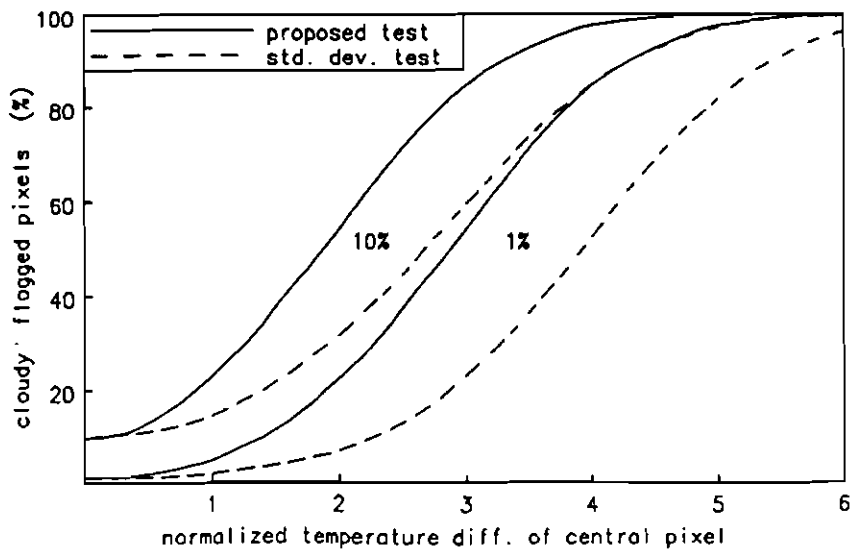

Figure 2. The probability that a cloudy pixel is flagged cloudy by the standard deviation test and the proposed coherence test, respectively, as a function of the difference of the pixel temperature to the expected temperature of cloud-free pixels. This difference was normalized by the noise of the clear pixels σ_n. The two sets of curves correspond to thresholds determined from figure 1 for a 1 per cent and a 10 per cent false detection rate.
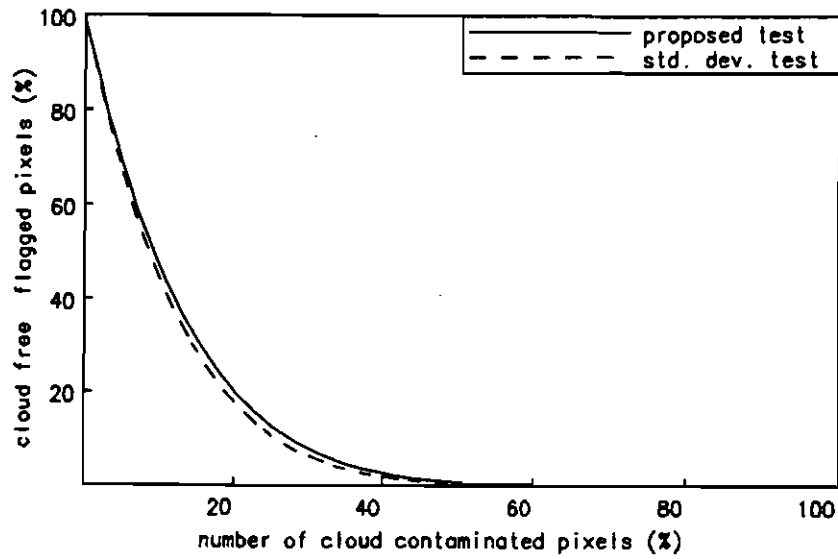
Figure 3. The probability that a clear pixel is flagged clear as a function of the fraction of cloud-contaminated pixels for an example pixel temperature distribution (see text). Thresholds are chosen to produce 1 per cent false detection in the limit of small cloud-cover.
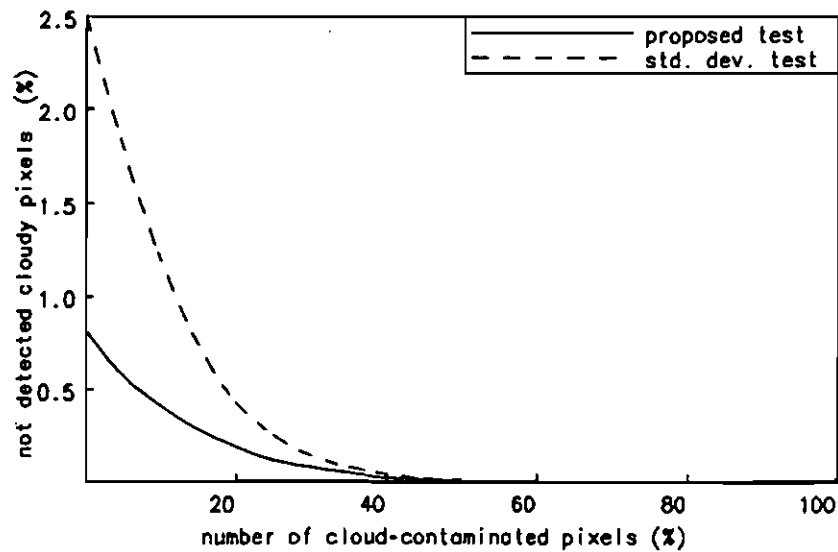

Figure 4. The probability that a cloudy pixel is flagged clear as a function of the fraction of cloud-contaminated pixels for the pixel temperature distribution and threshold of figure 3.

deviation test is more advantageous. For the given scales this cannot be seen in figure 4.)

We can summarize that the proposed test is more sensitive than the standard deviation test if the cloud cover is small, leaving at the same time more clear pixels flagged 'cloud free'. On the other hand it is less sensitive if the cloud cover is large. In real situations, however, the case of large cloud cover is generally much less critical for cloud detection. If the majority of pixels is cloud-contaminated, it is very likely that also significant portions of the pixels are cloud-filled. This means that usually a large pixel temperature scatter is found as well as a significantly reduced mean pixel temperature, which enables these pixels to be detected by simple threshold methods. Since the proposed technique is more sensitive in situations where other tests often fail, we conclude that it is more valuable for cloud detection. This conclusion was empirically supported by case studies not shown here.

4. Comparison with threshold tests

For a comparison of our coherence test with threshold tests we change from simulation to a real scene: an AVHRR image of 512 by 512 pixels from 27 October 1984, 16·20 GMT, over the North Atlantic. This scene contains partly broken low lying clouds (convective) in the southern part and higher clouds due to the arrival of a frontal zone in the northern part.

The coherence test is compared with two threshold test: an albedo threshold test and a sea surface temperature SST threshold test. For the albedo test we used AVHRR channel 1 (0·58–0·68 μm). The SST (in °C) was computed from a split-window-algorithm. We used the coefficients given by Schlüssel *et al.* (1987) for a scan angle of 20°:

$$\text{SST} = -1·05 + 1·018\,T_4 - 2·756\,T_5$$

where $T_{4,5}$ are the brightness temperatures of channel 4, 10·3–11·3 μm, and channel 5, 11·5–12·5 μm, in °C. It is preferable to use the SST instead of a single channel brightness temperature in a threshold test because the computed SST is corrected for scan angle and water vapour dependent absorber mass variations.

We have chosen the threshold values very carefully to provide maximum possible sensitivity without having obvious false detection. This was done by subjectively comparing obtained cloud patterns for different thresholds. Finally, values of 3 per cent for the albedo test and 21·5°C for the SST test were chosen. For the coherence test we have selected a value of $X_L = 0·25$ K to be reasonable for this scene and many others. The number of 'cloudy' flagged pixels as a function of X_L is given in figure 5. A value of $X_L = 0·25$ K is slightly above the steep slope region of false detection due to the background noise.

A pixel-by-pixel comparison of the test results is shown in figures 6 and 7. Generally the three methods agree well. The albedo test exhibits a relatively high detectivity near the right border due to the larger scan angle and the corresponding higher air mass related background reflectivity. Striking are the violet parts of figure 7 within the large cloud field in the north. They consist of thin clouds not detectable by the albedo test but by both other tests. The SST test did not detect the red, orange or yellow coloured pixels, especially in the lower right part, which belong to relatively warm low level clouds. On the other hand it may have led to some cases of false
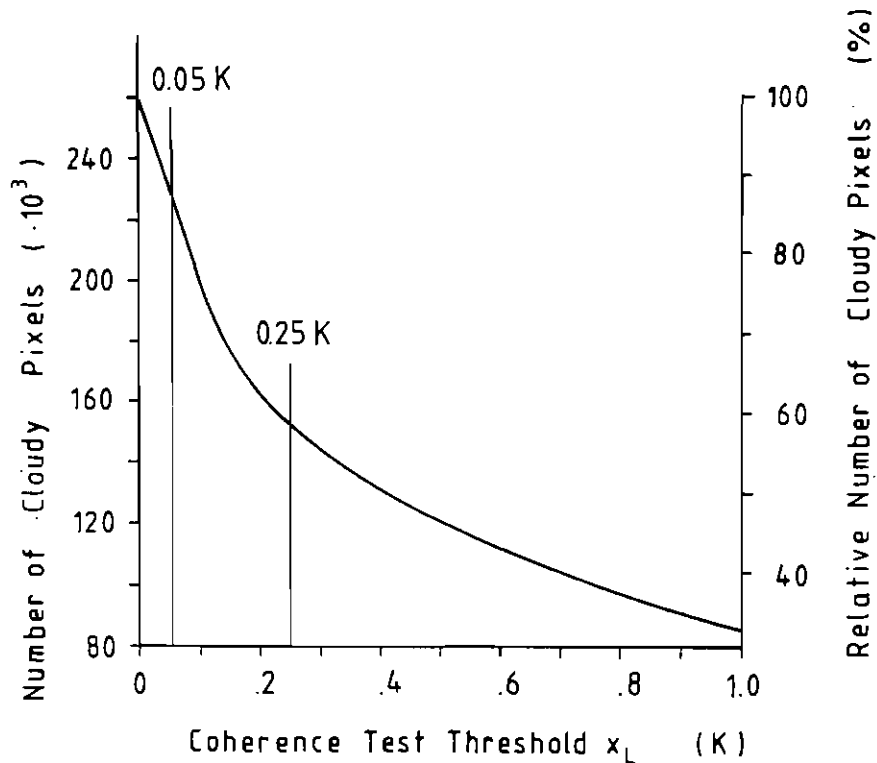
Figure 5. Dependence of the number of cloudy flagged pixels on the threshold of the proposed coherence test for an example scene of 512×512 AVHRR pixels (see text). Chosen thresholds for cloud detection (0.25 deg K) and for preselection of data for histogram computation (0.05 deg K) are marked by vertical bars.

detection (blue) in the upper part due to a cooler ocean surface in partly cloud covered areas.

The location of the high number of pixels flagged cloudy by the coherence test (red, orange, violet) is plausible, and not many pixels are designated cloudy only by other tests (yellow, green or blue). So the coherence test seems to be very sensitive in almost all situations, even in the case of high thin clouds (cirrus). However we see that within large flat clouds with low temperature variance, some obviously cloudy pixels remained undetected (green).

5. Combination with threshold test, the automated scheme

Since the coherence test sometimes misses homogeneous clouds with low temperature variance, it should be supplemented by another test. Because a homogeneous cloud cover is usually found within the central areas of extended clouds filling the whole pixel, the pixels brightness temperature is significantly reduced in most cases and a simple IR threshold test, being also applicable at night, is expected to be very effective. (An exception might be the case of low stratus clouds, see §7.) Thus the combination of a coherence test and an IR threshold test should be ideal. How the coherence test can also be used to find a reasonable threshold value for the threshold test is demonstrated in the following.

We assume the cloud top temperature generally to be lower than the SST. In principle the threshold can be defined in two ways. If the threshold equals the temperature of the warmest cloud contaminated pixel, all cloudy pixels would be rejected. However false detection would occur in many cases. Since the coherence test is already very effective at cloud edges and in order to avoid additional false detection, we define the threshold in the inverse way as the temperature of the coldest clear pixel. To find this value we apply the coherence test to the considered scene with a strongly reduced value of, say, $X_L = 0.05$ K. About 20 per cent of the clear pixels survive this procedure (see figure 5) whereas cloudy pixels are practically all eliminated. For the test scene, the histogram of the pixels which were selected in this way together with the histogram of all pixels is given in figure 8. We see that from the histogram of preselected pixels the threshold can easily be determined as the low temperature footpoint of the distribution. This would not unambiguously be possible from the original histogram.

In order to find automatically the threshold we used the following simple algorithm, which gave good results for many investigated scenes. One should keep in mind that even after application of the strong coherence test there might remain few scattered cloudy pixels in the histogram (not visible in figure 8). Firstly, eliminating all pixels having temperatures below 0°C clears the histogram from remaining pixels, which are evidently cloudy due to their very low temperatures. (For the polar regions a slightly lower value has possibly to be chosen.) Secondly, separated clusters in the histogram found at the lower temperature side of the main maximum are considered to be cloudy and are eliminated if their total frequencies count less than 5 per cent of all remaining pixels. Thirdly, from the resulting histogram the brightness temperature which marks the 95 per cent cumulative frequency (starting at the warmest pixel) is computed and a value of 2 K subtracted to define the final threshold. The value of 2 K was found to be reasonable for typical temperature distributions within 512 by 512 pixel scenes. The threshold derived for the histogram given in figure 8 is 16°C, for example. If we apply this threshold to our test scene, all cloudy pixels which were not detected by the coherence test but by the SST threshold test are rejected.

6. Comparison of automated scheme with the bispectral method of Phulpin *et al.* (1983)

A comparison of our whole method (coherence test plus IR threshold test with automatic threshold determination) was carried out with the bispectral method of Phulpin *et al.* (1983). This method is based on a two-dimensional, visible and IR histogram analysis. For part of the NOAA-7 orbit No. 3587 (ALPEX scene) from 4 March 1982, we compared the fraction of cloudy AVHRR pixels within about 600 HIRS pixel related field of views. The data are from the western North Atlantic, the North Sea and the Mediterranean.

Table 1 compares the number of HIRS pixels containing a certain fraction of cloudy AVHRR pixels as determined by the method of Phulpin *et al.* (set 1) and our method (set 2): The percentage of cloudy pixels found by our method is often significantly higher than that obtained by the bispectral method. The bispectral method never found a cloud in a HIRS pixel when our method described it cloud-free. The cloud cover determined by our method was never less than that determined by the bispectral one. Therefore our method seems to be very sensitive and suitable for providing a relatively safe basis of clear radiance pixels. On the other hand, there is no indication of overdetection within widely clear areas. This can be inferred from a
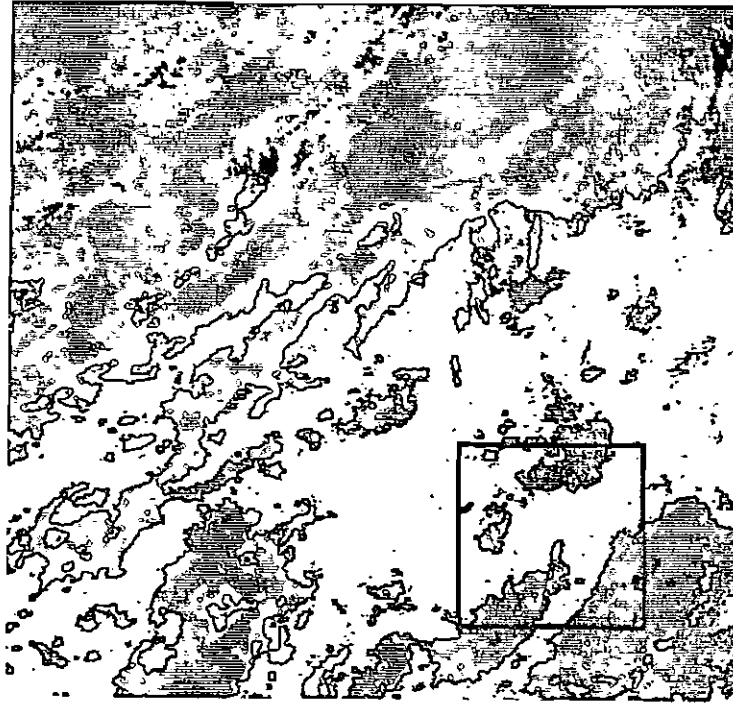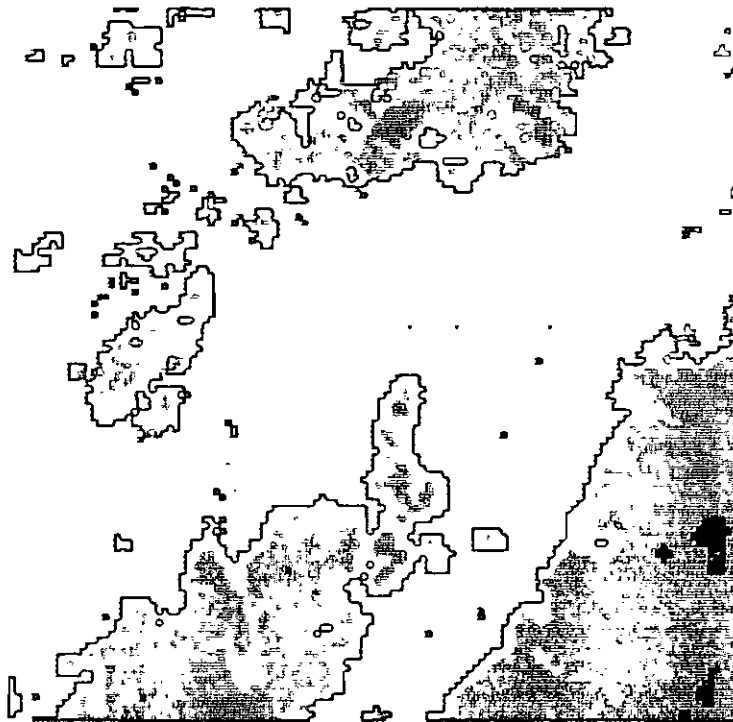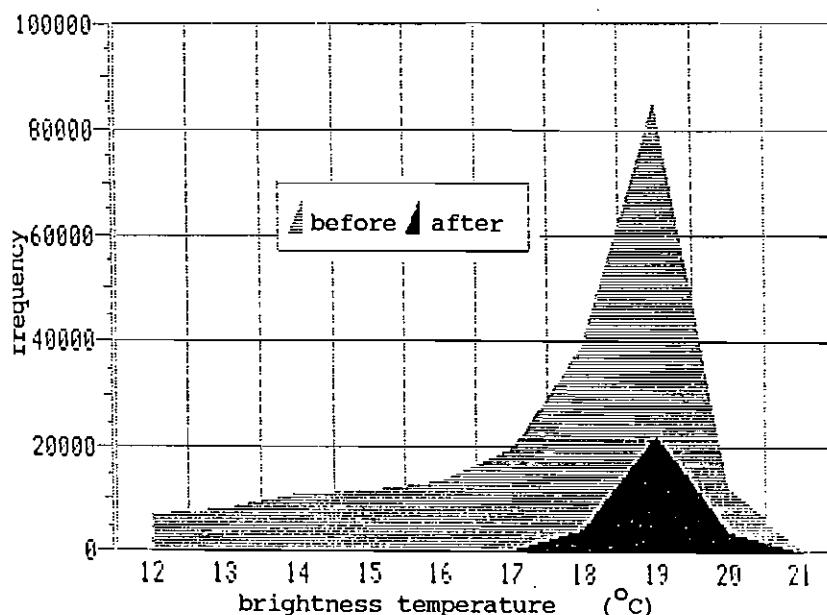
Figure 6.

Figure 7.

Figure 8. Histogram of AVHRR channel 4 brightness temperatures for the example scene before and after application of the coherence test with $X_L = 0.05$ K.

comparison of our cloud cover results in the rows of the bispectral 0 per cent and 1–9 per cent classes: Our method flags much less pixels cloudy in the 0 per cent class than in the 1–9 per cent class of set 1.

7. Concluding remarks

Cloud contamination yields large errors in the remote sensing of earth and atmosphere using visible and IR satellite observations. A reliable cloud detection method is therefore of great importance for the accuracy of the derived results.

In this study a new method was developed for application over sea. It is based on a single IR channel of a high resolution radiometer (AVHRR) and therefore applicable day and night. It combines a coherence and a threshold test. The used coherence test has an improved sensitivity compared to the standard deviation test under conditions of low cloud cover, which is the most critical situation. A threshold test supplements the coherence test ideally since in most cases it easily detects the central regions of extended clouds, which are sometimes too homogeneous to be detected by the coherence test. The determination of the IR threshold value for the threshold test is

Figure 6. Colour coded cloud detection results for the example scene. Red: A pixel is flagged cloudy by the coherence test only, yellow: albedo test only, blue: SST test only, orange: coherence and albedo test, green: albedo and SST test, violet: coherence and SST test, grey: all tests effective.

Figure 7. Enlarged view of the inset area of figure 6.

V. Thiermann and E. Ruprecht

Table 1. Comparison of cloud detection results.

	Cases	0	1 to 9	10 to 19	20 to 29	30 to 39	40 to 49	50 to 59	60 to 69	70 to 79	80 to 89	90 to 99	100
	0	13	19	1	0	2	0	0	0	0	0	0	0
	1 to 9	0	4	12	7	7	6	5	3	1	0	0	0
	9 to 19	0	0	1	1	7	9	10	3	6	5	0	0
	20 to 29	0	0	0	0	1	2	2	5	7	12	1	0
	30 to 39	0	0	0	0	0	0	3	7	5	9	3	0
	40 to 49	0	0	0	0	0	0	0	5	5	11	11	0
	50 to 59	0	0	0	0	0	0	0	0	2	9	9	0
	60 to 69	0	0	0	0	0	0	0	0	2	3	27	3
	70 to 79	0	0	0	0	0	0	0	0	0	2	14	5
	80 to 89	0	0	0	0	0	0	0	0	0	0	12	8
	90 to 100	0	0	0	0	0	0	0	0	0	0	0	299

The header of the upper columns reads "Cloudy flagged pixels (per cent), set 2". The left side label reads "Cloudy flagged pixels per cent, set 1".

Set 1: bispectral histogram analysis of Phulpin *et al.* (1983); set 2: our IR method. Given are the number of HIRS pixels with a certain cloud cover (percentage of cloudy pixels) as determined by both methods. Data from NOAA-7 orbit No. 3587, 4 March 1982 (western North Atlantic, North Sea and Mediterranean Sea).

performed by histogram analysis of data which passed the coherence test with a very strong criterion. It was demonstrated how this preselection procedure considerably diminishes the ambiguity in the histogram analysis for the threshold definition.

The proposed test is simple, very fast, does not require any *a priori* information about the field to be tested and is therefore very suitable for operational application. A comparison with other methods reveals that it responds very sensitively to cloud-contamination. As with all coherence tests, there may be an overdetection at cloud edges. However our coherence test was shown to be less susceptible to false detection than the standard deviation test.

Problems are possible in the case of low and warm clouds having a temperature very close to the SST. All IR tests, which are applicable at night, have this problem in common. However the coherence test will improve the threshold test results in many cases, since the coherence test responds to temperature differences as small as 0·25 K whereas the absolute threshold typically requires a safety distance to SST of about 1 K.

It shall finally be noted that our method was also found to be very effective for many other scenes than the one discussed here in detail. These scenes include those with high surface temperature gradients at the Californian coast characterized by upwelling. Another important aspect is the scan angle dependence. We found subjectively only at very large scan angles (the outer 10 per cent of the swath) a tendency to oversensitivity of the coherence test, probably due to increased grid point distances on the earth's surface.

Acknowledgments

References

COAKLEY, J. A., and BRETHERTON, F. P., 1982, Cloud cover from high resolution scanner data: Detecting and allowing for partially filled fields of view. *Journal of Geophysical Research*, **87**, 4917–4932.

PHULPIN, T., DERRIEN, M., and BRARD, A., 1983, A two dimensional histogram procedure to analyze cloud cover from NOAA satellite high-resolution imagery. *Journal of Climate and Applied Meteorology*, **22**, 1332–1345.

ROSSOW, W. B., MOSHER, F., KINSELLA, E., ARKING, A., DESBOIS, M., HARRISON, E., MINNIS, P., RUPRECHT, E., SEZE, G., SIMMER C., and SMITH, E., 1985, ISCCP Cloud Algorithm Intercomparison. *Journal of Climate and Applied Meteorology*, **24**, 877–903.

SAUNDERS, R. W., 1986, An automated scheme for the removal of cloud contamination from AVHRR radiances over Western Europe. *International Journal of Remote Sensing*, **7**, 867–886.

SCHLÜSSEL, P., SHIN, H.-Y., EMERY, W. J., and GRASSL, H., 1987, Comparison of satellite derived sea-surface temperatures with *in situ* skin measurements, *Journal of Geophysical Research*, **92**, 2859–2874.

SIMMER, C., RASCHKE, E., and RUPRECHT, E., 1982, A method of determination of cloud properties from two-dimensional histograms. *Annalen der Meteorologie*, **18**, 130–132.