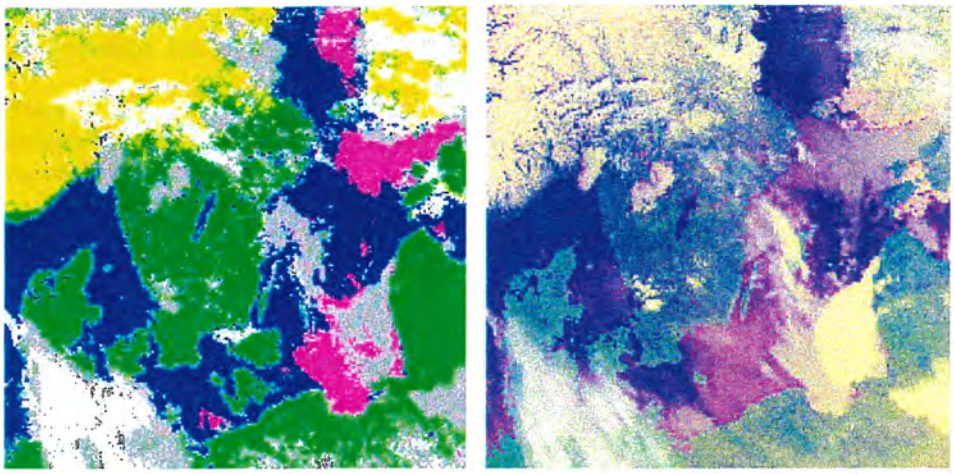
Nr 92, 1999

SMHI
Meteorologi
Finnish Meteorological Institute
Swedish Meteorological and Hydrological Institute

Nowcasting SAF

Investigations of NOAA AVHRR/3 1.6 μm imagery for snow, cloud and sunglint discrimination



Otto Hyvarinen
Karl-Göran Karlsson
Adam Dybbroe

Visiting scientist report

July 1999

ISSN 0283-7730

Contents

Front cover: The two images on the front cover are from a NOAA 15 pass **08:09 UTC April 3, 1999**. *The satellite data have been rectified to a polar stereographic map projection in 1-km resolution, and covers an area of Southern Scandinavia with Norrköping approximately in the center. The left most image is the result of the 6-class Maximum-Likelihood classification described in the report, and the right most is the corresponding RGB using channel 1,2 and 4. The classes are Land (green); Sea (blue); Snow and sea ice (yellow); Cirrus = Ice particle clouds (white); Water clouds (grey); and sunglint (purple). The chosen scene is rather complex having strong sunglint in combination with low level cloudiness in the Baltic sea, and a large homogenous area in the south east affected by what seems to be haze (increased aerosol content). Despite these difficulties the classification does catch most of the low level cloudiness and it is doing reasonable well in destinguishing these clouds from the cloud free sea surface in sunglint areas.*

1 Visiting scientist program: Study tasks and achievements

The visit started on Tuesday 6 April and continued throughout Friday 2 July 1999.

This chapter describes the goals of the study and it includes also a summary of the main results. A deeper and more comprehensive description is given in the following chapters.

1.1 Purpose and goal of the study

The AVHRR/3 imager on the NOAA-15 satellite (launched in May 1998) includes a new spectral channel at 1.6 μm (denoted 3A) which has not been used earlier for operational monitoring of meteorological conditions. This channel has been available in research mode from the ATSR instrument onboard the ERS satellites and from the Thematic Mapper instrument of the Landsat satellites. However, the use of channel 3A for NOAA-15 will be restricted to a few test periods since the 3.7 μm channel (denoted 3B) will still be maintained as the operational AVHRR channel 3. Such a test period occured during the spring of 1999. From March 9th until April 20th NOAA-15 were transmitting AVHRR Channel 3A and 3B data according to the following schedule: Channel 3A was operational during daylight passes only, as the spacecraft crosses the terminator into daylight, with the additional constraint that the spacecraft subpoint is North of 40 degrees N. As the spacecraft crosses the terminator into darkness, Channel 3A is toggled off, and Channel 3B toggled on.

Thus 3A transmissions were not continuous, but shared with Channel 3B data transmissions. In effect all morning passes received at Norrköping during the test period were transmitting channel 3A data, while all evening passes transmitted channel 3B data.

The 3A channel is spectrally situated in an atmospheric window with only minor contributions from absorbing gases in the atmosphere. In this spectral region, interesting reflectance differences between clouds and snow exist and the channel is therefore foreseen to be of great use for snow and cloud discrimination.

This study was to assess the usefulness and possibilities of channel 3A for snow and cloud detection at day time, and the changes to the current cloud mask algorithm which will be needed.

The reflectance behaviour of different objects at 1.6 μm wavelength in AVHRR/3 imagery was studied by use of data from the test period during spring 1999. Particular attention was given to the potential for discriminating snow, sunglint and clouds and comparisons were made to the present discrimination capability achieved using AVHRR/2 data. In addition, the potential for discriminating the ice and water phases of clouds was briefly investigated.

The investigations for this visiting scientist stay were broken down into the following tasks:

1. Compilation and close examination of a specific training dataset containing typical channel 3A signatures for different cloud and surface targets.
2. Qualitative understanding of the reflectance behaviour of clouds, snow-covered surfaces and sunglints in channel 3A.
3. Results from studies comparing the reflectances of channels 3A and 3B.
4. Development of a preliminary version of a daytime cloud screening algorithm utilising AVHRR channel 3A.

1.2 Achievements

The following statements and conclusions related to the above-mentioned study tasks can be reported here:

Task 1: Compilation and close examination of a specific training dataset containing typical channel 3A signatures for different cloud and surface targets.

1. A comprehensive data base of different types of phenomena was compiled by two independent analysts.

2. Because of the short length of the test period and because only morning passages could be used. the targets were clustered into two distinct sets with the restricted values of solar zenith angles and satellite viewing geometry.

3. A total of 1317 individual targets were extracted subdivided into the following main classes: water, sunglint, land, snow, water clouds and cirrus.

4. Targets were chosen for an area covering the entire European region.

Task 2: Qualitative understanding of the reflectance behaviour of clouds, snow-covered surfaces and sunglints in channel 3A.

1. The discrimation of classes using channel 3A exclusively is rather problematic since there is a great overlap between classes. Best separability is shown between snow and water clouds while the separation between thin Cirrus clouds and snow is very problematic. Furthermore, the distinction between water clouds and Cirrus clouds superposed over low-level water clouds is difficult to make. The strong reflection from cloud droplets seems to dominate over the non-reflective behaviour of ice crystals.

2. Adding channel 1 information increases the separability of classes considerably. E.g., a very good separability of water clouds and snow is achieved by using the quota between channels 3A and 1.

3. Strong sunglint can be discriminated from water clouds using the combined information of channels 3A and 1. However, there is still some overlap between weak sunglint and waterclouds.

4. Cirrus clouds are difficult to discriminate from land surface using just information from the visible channels.

5. Strong sunglint-like features were noticed for most classes in all visible channels including channel 3A, especially for water clouds. For water clouds, also a backscattering enhancement was noticed at high solar and satellite zenith angles.

6. Significantly enhanced backscattering was noticed for land surfaces at high solar and satellite zenith angles. This feature could possibly be seasonal depending on the vegetation state of land surfaces. It could cause serious problems for cloud classification applications, especially concerning the identification of thin low clouds.

Task 3: Results from studies comparing the reflectances of channels 3A and 3B.

1. To have a qualitative understanding of the differences of separability using channels 3A and 3B, an approximation of the 3B reflectance has been calculated.

2. The separability of sunglint in 3B seems to be significantly better than for 3A. Even the weak sunglint can be discriminated using the reflectance of channels 3B and 1 as a contrary to the case using channel 3A. Consequently, any cloud classification method using the foreseen distributed AVHRR/3 HRPT data set transmitted from NOAA-L and the following satellites is likely to meet worse conditions for sunglint detection compared to the present situation using AVHRR/2 data.

3. The snow separability with 3A seems to be much better than with 3B. However, forward-scattering enhancement appears ("snowglint") and may confuse results in both channels.

Task 4: Development of a preliminary version of a daytime cloud screening algorithm utilising AVHRR channel 3A.

1. Two different kinds of methods were investigated, a maximum likelihood method and a thresholding method.

2. To assist in choosing the features and to get some quantitative results for the separability of classes, the Jeffries-Matusita distance separability index was used.

3. The separability index studies did generally support conclusions made so far but some results were also both ambiguous and surprising. E.g., interesting was that indications suggested that the ch4-ch5 feature could be used to enhance separation of sunglint and waterclouds. These indications are indeed plausible due to the fact that this feature is sensitive to water vapour and that low level clouds obscure the lower part of the atmosphere where the abundance of the atmospheric moisture content is located (approximately 90% below 1 km).

4. For the cloud mask generation, information of channels 1, 3A and 4 seem to be enough, but for the more complete classification, information from all channels is probably needed.

5. In the classification tests, result scores (percentage of correctly classified targets) for the cloud mask exceeded 90% for both the two classification methods. For the cloud and surface type classification, results scores were between 80% and 90%.

6. Full scene classifications were also executed for a large part of the available scenes during the test period. Results were generally promising but several of the noticed separability problems were naturally also noticed in resulting classification images.

2 Training dataset

The training dataset for the test period was compiled mostly by Otto Hyvärinen (in the following denoted OH), the visiting scientist, and Gunhild Olofsson (in the following denoted GO) of the SMHI Forecasting Section. Thus for the study there were available two datasets made by two independent observers. The targets were selected and labeled using the Interactive Training Manager [1], an application developed for this purpose at SMHI.

In Figure 1 are shown the histogram of the azimuth angle differences and a scatterplot for the sun and satellite zenith angles for all targets. The targets are clustered into two distinct sets with very restricted values. This is understandable, because of the short length of the test period and because only morning passages could be used. These two data sets will in the following be referred to as the forward scattering (azimuth angle difference approximately 160 degrees) and the backward scattering modes (the azimuth angle difference approximately 20 degrees).

2.1 Classes

Each training sample was assigned to one of 50 possible classes as defined in the Interactive Training Manager system. For this study a coarser partition was sufficient, so the following "metaclasses" were constructed:

WATER CLOUDS Stratus, Stratocumulus, Cumulus, Altocumulus and Altostratos

ALL CIRRUS All Cirrus clouds, which is subclassed as

- **THIN CIRRUS** Thin Cirrus over surface
- **OTHER CIRRUS** Cirrus over lower clouds and Cirrosstratus (including frontal cirrus with likely sub-layer clouds)

LAND Clear (cloud free) land. No significant cloud cover.

WATER Open sea without any substantial cloud/aerosol contamination

SNOW Snow covered surface and open sea ice

SUNGLINT Sunglint on open sea

As these metaclasses do not cover all the available classes of the Training Manager, some collected targets, for example shadows, were not used. The number of targets collected during the test period for these metaclasses is shown in table 1.

Though the number of all targets collected by both observers does not differ much, there are some differences: the OH-dataset contains more clouds targets and the GO-dataset more surface targets.

The geographical distribution of targets is shown in Figures 2 and 3. Most of the targets are from Northern Europe and surrounding waters, at least one noticeable exception being the land targets. Clear land without snow was easier to find in Central and Southern Europe in this period, compared to the mostly snow covered Scandinavia and Eastern Europe. Also a significant part of the sunglint targets are from the Mediterrian region.

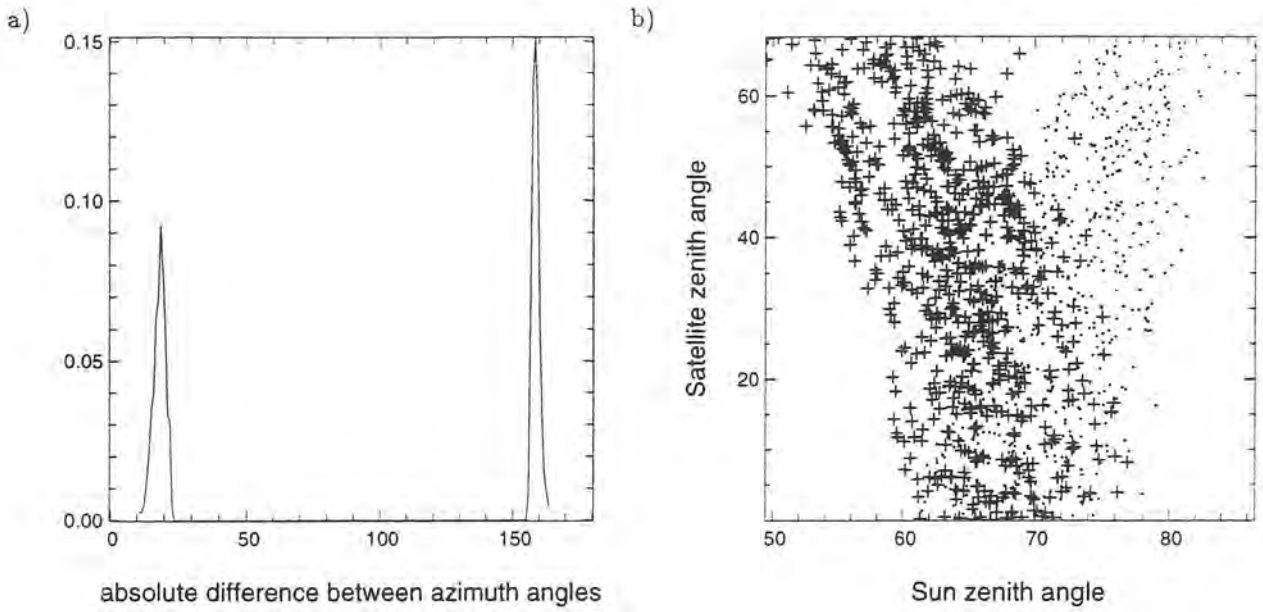


Figure 1: The histogram of the sun and satellite azimuth angle differences and the scatterplot of the sun and satellite zenith angles for all targets. "." have an azimuth angle difference less 50 degrees (the backward scattering mode) and "+" greater than 100 degrees (the forward scattering mode).

Table 1: The number of targets collected by the observers OH and GO using channel 3A data.

class	OH	GO	other
WATER	34	183	5
SUNGLINT	51	25	2
LAND	36	123	6
SNOW	92	122	15
WATER CLOUDS	247	129	11
ALL CIRRUS	140	92	2
-THIN CIRRUS	60	60	2
-OTHER CIRRUS	80	32	0
	600	674	43

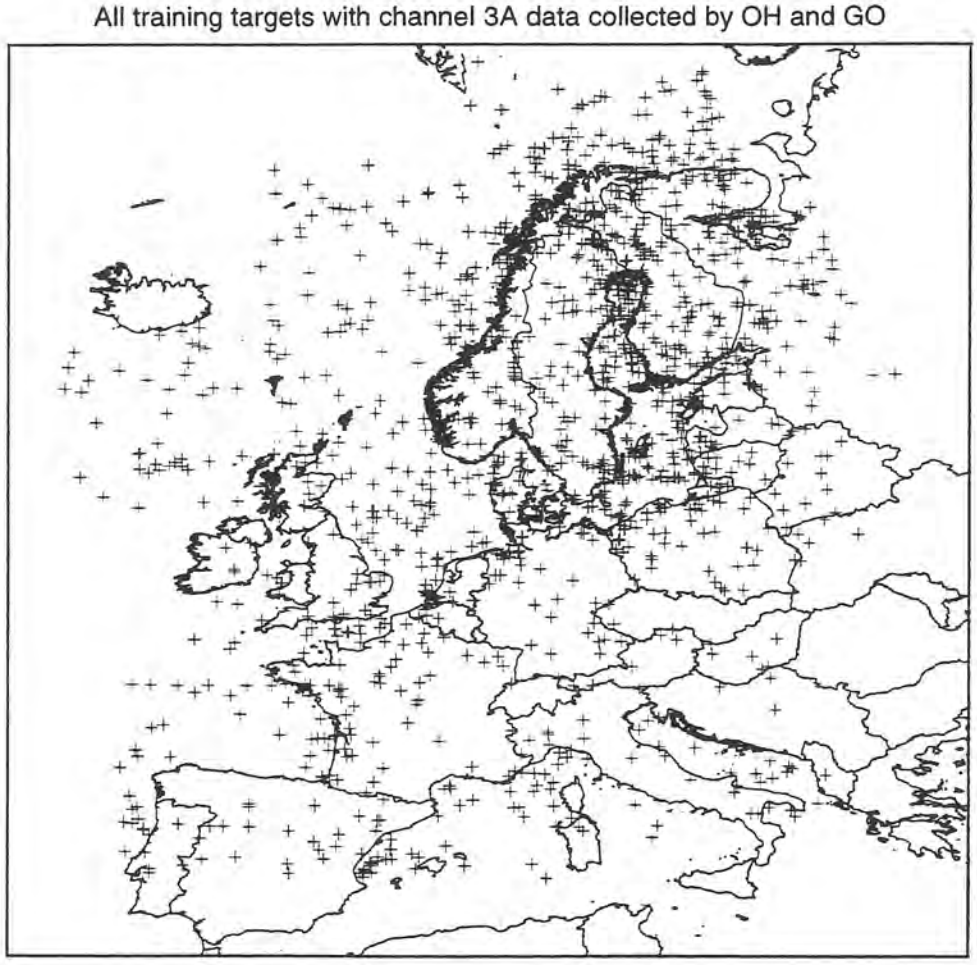


Figure 2: The geographical distribution of training targets.

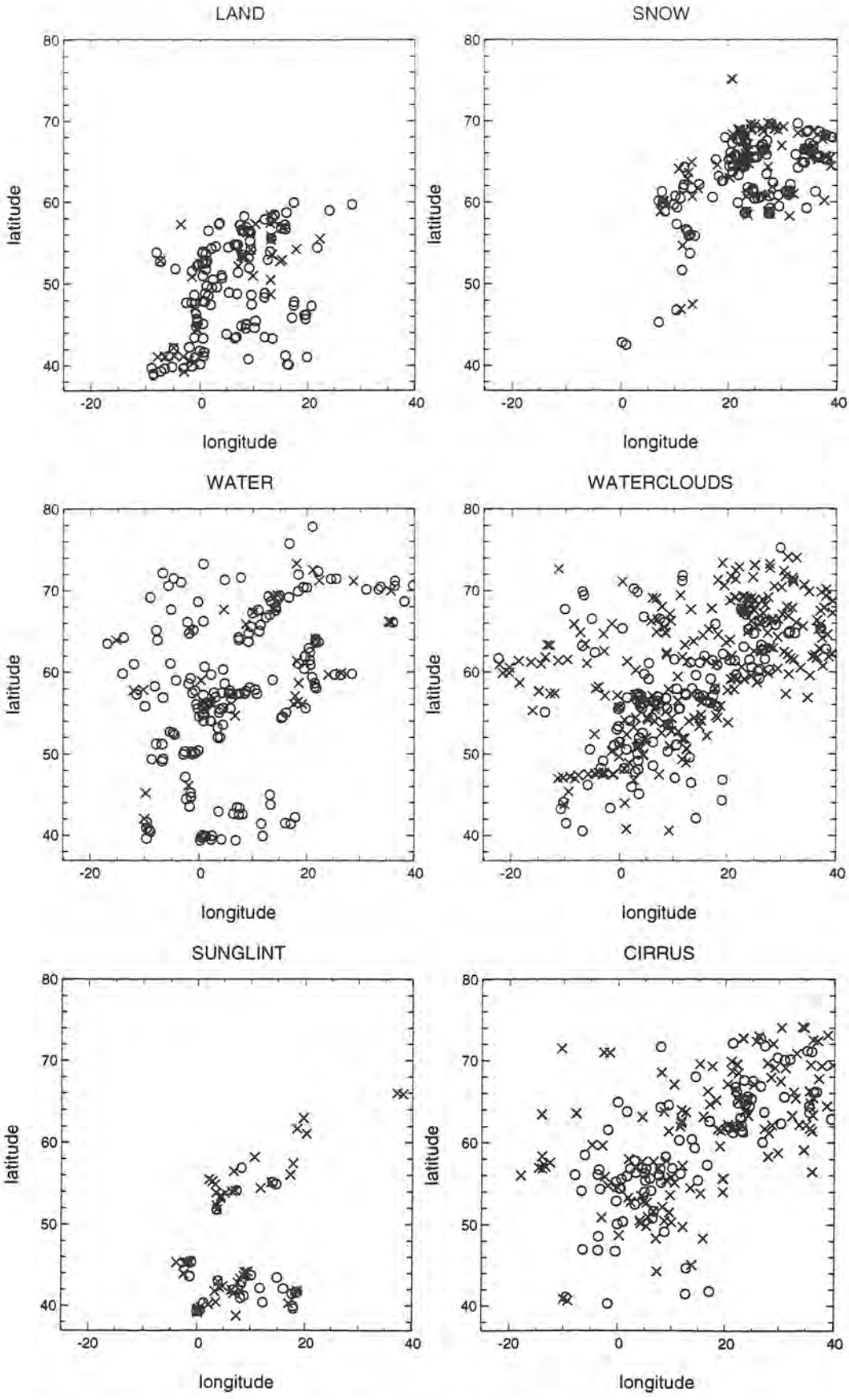


Figure 3: The spatial distribution of training targets by classes. "x" OH and "o" GO

3 The reflectance behaviour in channel 3A

3.1 Separability using channel 3A exclusively

In Figures 4 and 5 normalized histograms are presented showing channel 3A reflectances for different classes and Gaussian distribution curves calculated from them. The normalized histogram has the same property as the Gaussian distribution curve; it integrates to one. The plots were made using all targets. Reflectance values were calculated by dividing calibrated pixel values by the cosine of the solar zenith angle. This should reduce any variability caused by the varying solar zenith angle. The dataset naturally contain a great range of different satellite viewing angles. Thus systematic dependencies on the satellite zenith angle and in particularly the difference between the satellite zenith angle and the solar zenith angle may be possible to reveal. As the dataset is rather limited as concerns the available solar zenith and azimuth angles, their full impact on the results is harder to assess.

When comparing the surface targets to the water clouds (Figure 4) it can be seen, how snow and water surfaces are rather well discriminated using only channel 3A data, but discriminating land is more problematic. Notice however the very wide distribution of water cloud reflectances meaning that some dark (presumably thin) water clouds may still be confused with snow and water surfaces.

Discriminating ice clouds using only channel 3A is more difficult (Figure 5). The histogram for Cirrus clouds over surface is very similar to the histogram for snow and the histogram for cirrus clouds over lower (water) clouds resembles very much the histogram for water clouds. On the other hand, thin cirrus can be separated rather well from water clouds.

Sunglint and water clouds are very difficult to discriminate using only channel 3A as shown in Figure 5. This can be partly explained by a large range of variability for sunglint. This variability also explains why the sunglint targets deviate significantly from the Gaussian approximation. More training targets are needed to get a fully representable distribution of reflectances. The change from water to sunglint is in addition a very gradual process. Therefore the process of classifying targets into either the *sunglint* or *water* classes is bound to be a matter of subjective definition, much more so than for most other pair of classes.

The Gaussian approximation suits rather well for most of the classes. Noticeable exceptions are sunglint and cirrus over lower clouds. This can be at least partly explained by the relatively small number of available targets. Histograms of land and water decrease slower on the right side of the maximum, so the maximums of histograms and Gaussian curves are not at the same place.

3.2 Separability using channels 1 and 3A

For optimal cloud classification with AVHRR/3 data one must of course take advantage of the information available in the other spectral channels. Here we will restrict our discussion to the effect of adding channel 1 at 0.6μm to the 1.6μm data concerning the separability of cloud free surfaces and clouds. Some discussion on the influence of the remaining AVHRR channels can be found below in section 5.1.

Scatter plots using channel 1 and channel 3A reflectances are shown in Figures 6 and 8 using the same data partitioning for the OH and GO datasets, respectively. We can also use the quota between channels 3A and 1 (Figures 7 and 9).

Classes are better separated with these two channels than with 3A channel alone. Snow seems to be well separated from all classes except thin cirrus clouds. The most serious problems are seen for the separation of sunglint from waterclasses and the separation of cirrus clouds from snow. Even though strong sunglint can be easily discriminated from the water clouds, there is an overlap with water clouds and weak sunglint. Likewise, cirrus over lower clouds is still overlapping with the water clouds.

The use of the quota makes it easier to separate land and water classes and most of the separating thresholds can be drawn parallel to the axes, but otherwise it does not seem to add new information.

There are some noticable differences between the two datasets. As the GO dataset have more surface targets, those classes have a greater range of values than corresponding classes of the OH dataset. However, for cloud targets differences are quite small.

In conclusion, the combined use of AVHRR channels 1 and 3A seems to further enchance the separability of snow surfaces, water clouds and sunglint. Furthermore, the use of the reflectance quota contributes to a better discrimination of cloud free land and water surfaces. However, separability problems still remain for a fully succesful discrimination of water clouds, sunglint and thin cirrus. More features seem to be needed and this is further explored in the section 5.1.

3.3 Exploring the degree of anisotropy for surfaces and clouds

Figure 10a for the channel 3A reflectance shows how "sunglint"-effects can be seen in other classes than the true sunglint class. All the targets in this figure belong to the forward scattering mode and have the sun zenith angle between 55 and 65 degrees. Significant sunglint targets are identified when the difference between sun and satellite zenith angles falls below 20 degrees. Water and snow targets also show an increase in reflectance with decreasing difference between the satellite and sun zenith angles. But for clouds this feature is more pronounced, as the values have almost doubled compared to the values away from the "specular point".

It is interesting to compare this to the cases where the targets belong to the backward scattering mode and where the sun zenith angle varies between 75 and 85 degrees (Figure 10b). There are no sunglint targets of course because the satellite is viewing away from the sun. The increase of the reflectance when the difference of the sun and satellite zenith angles decreases can be understood by an enchanced backscattering due to the cloud geometry (illuminated cloud sides) and the character of Mie-scattering for cloud droplets at near-infrared wavelengths. In the latter case, scattering is more systematically occuring in the forward and backward scattering direction since diffraction effects become more important than pure reflection.

Possible additional effects like the increase of the size of the AVHRR footprint with increasing viewing angle is harder to assess, but is anyway thought to be of less importance here.

In Figure 11 the same plots as in Figure 10 are shown for the channel 1 reflectance. The same "sunglint"-effects as in Figure 10 can be seen. The most noticable difference is how the value of the reflectance is dependent of the satellite zenith angle for water targets. This is because at this shorter wavelength, contribution from atmospheric aerosols has a greater effect on channel 1 than for channel 3A reflectance.

The effect of the anisotropic scattering is further explored in Figure 12. It shows histograms for all classes for the two different clusters of the azimuth angle differences. All data are used, so the backward scattering mode includes also higher sun zenith angle targets compared to the data shown in Figures 10 and 11.

In the land targets there are large differences between the two different azimuth angle sets for channel 3A. This indicates a large backscattering effect for the land targets. This difference mostly disappears when the quota between channels 1 and 3A is used. The origin of this enchanced backscattering for land targets is not completely understood since this behaviour is not known from earlier experience of using AVHRR/2 data. However, some mis-classifications confusing low clouds with cloudfree land surfaces at the same viewing conditions have earlier been noticed for SMHI SCANDIA model [2]. This occurred also in early spring and it is possible that land areas at this particular time have specific behaviour (e.g. caused by the large fraction of non-vegetated farm land areas with wet and bare soil) causing anisotropic reflection in AVHRR channels 3A and 3B.

For snow and water cloud targets there are some differences between the two different azimuth angle sets. These can also be largely removed using the quota between channels 1 and 3A. Pronounced differences for the quota are found for the water targets. These come most probably from sunglint contamination. Because of the continuous nature of the sunglint, it may have made some influence on water targets that the observer did not notice.

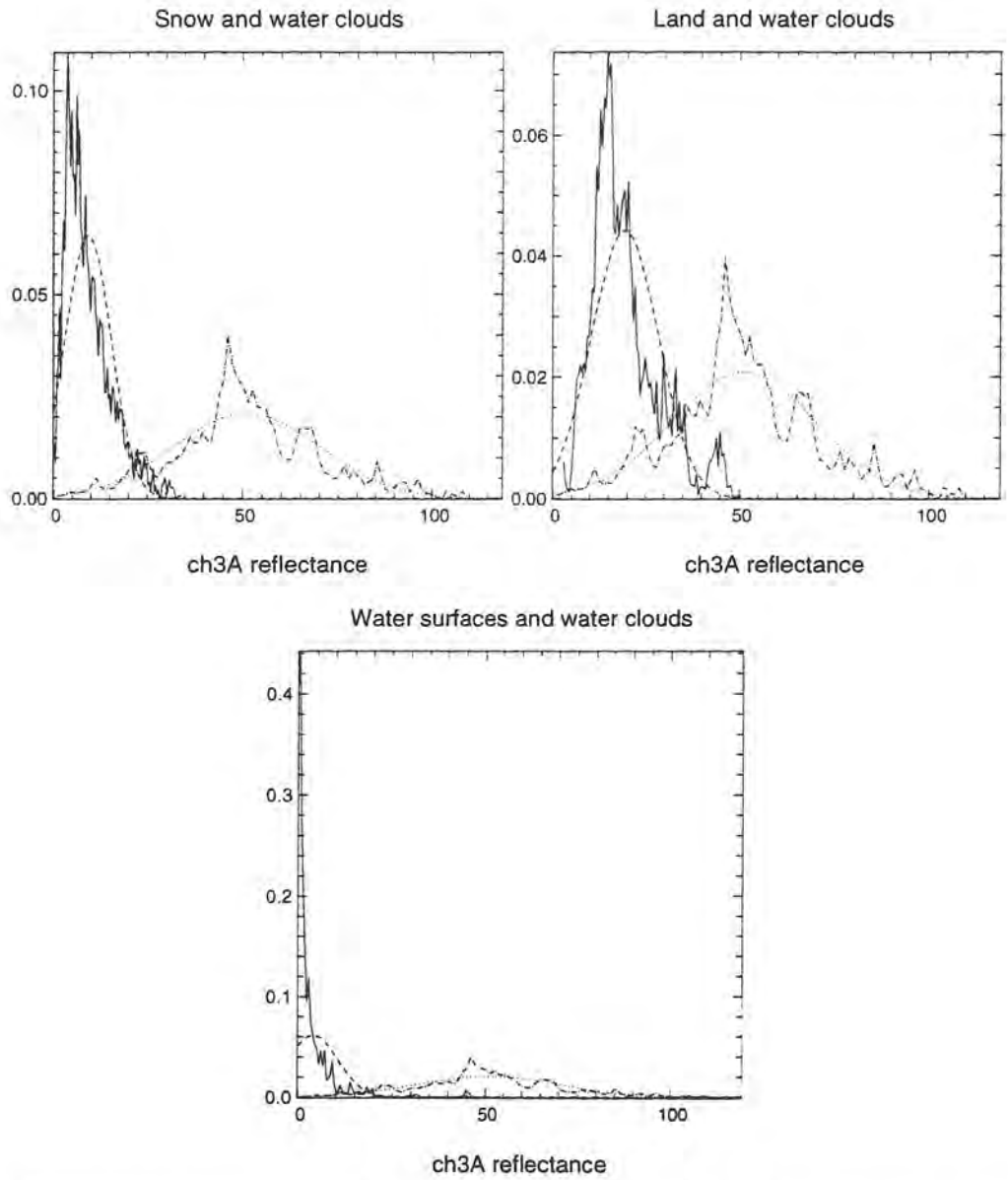


Figure 4: Normalized histograms and Gaussian distribution curves for different classes as a function of channel 3A reflectances. In every figure the histogram for water clouds is given by the dashed-dotted curve and the Gaussian distribution curve is displayed dotted. The histogram of the other class is given the solid line and the Gaussian distribution curve is shown as the dashed line. These other classes are snow, land and water.

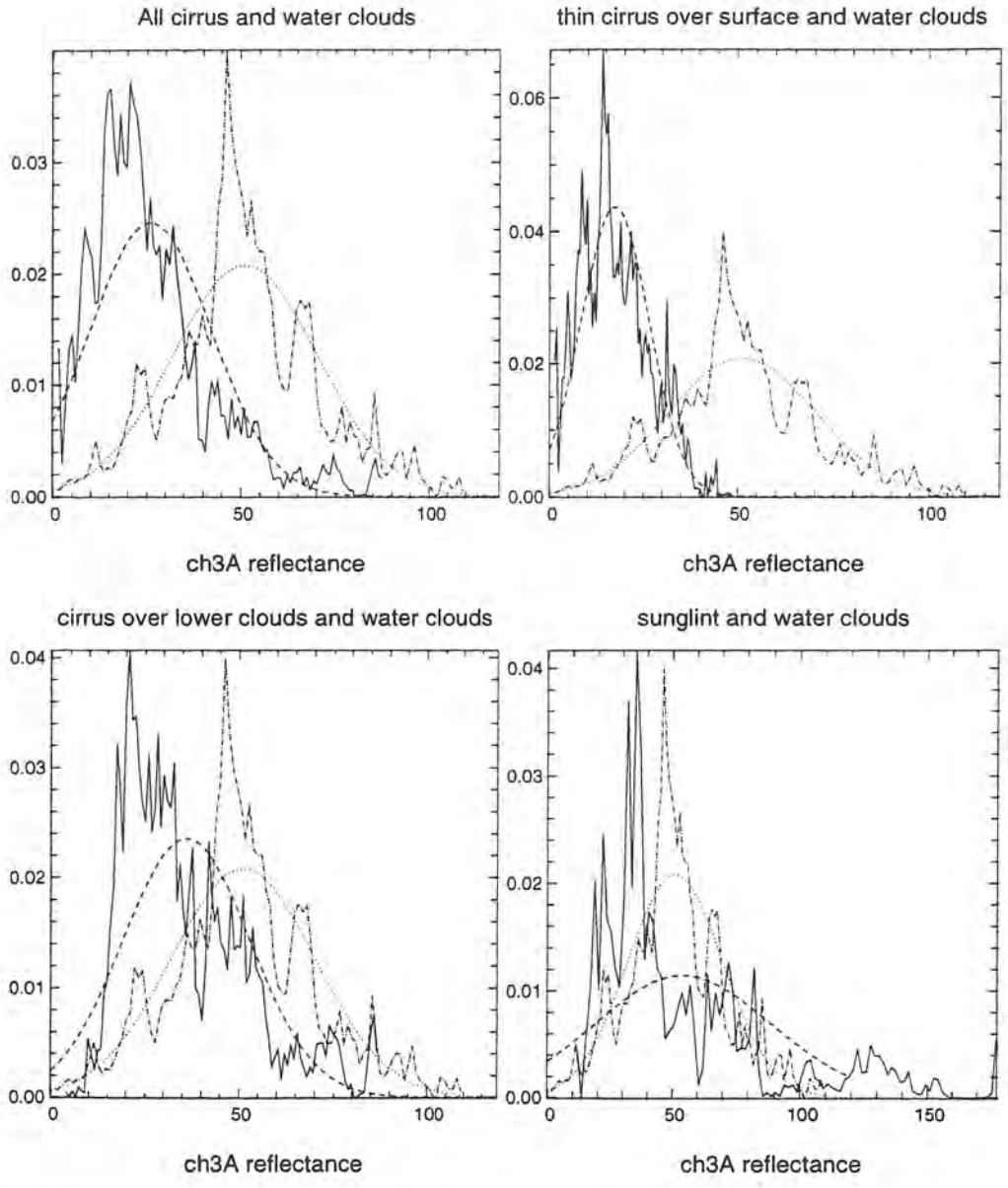


Figure 5: Normalized histograms and Gaussian distribution curves for different classes as a function of channel 3A reflectances. In every figure the histogram for water clouds is given by the dashed-dotted curve and the Gaussian distribution curve is displayed dotted. The histogram of the other class is given the solid line and the Gaussian distribution curve is shown as the dashed line. These other classes are all cirrus, thin cirrus, cirrus over lower clouds and sunglint.

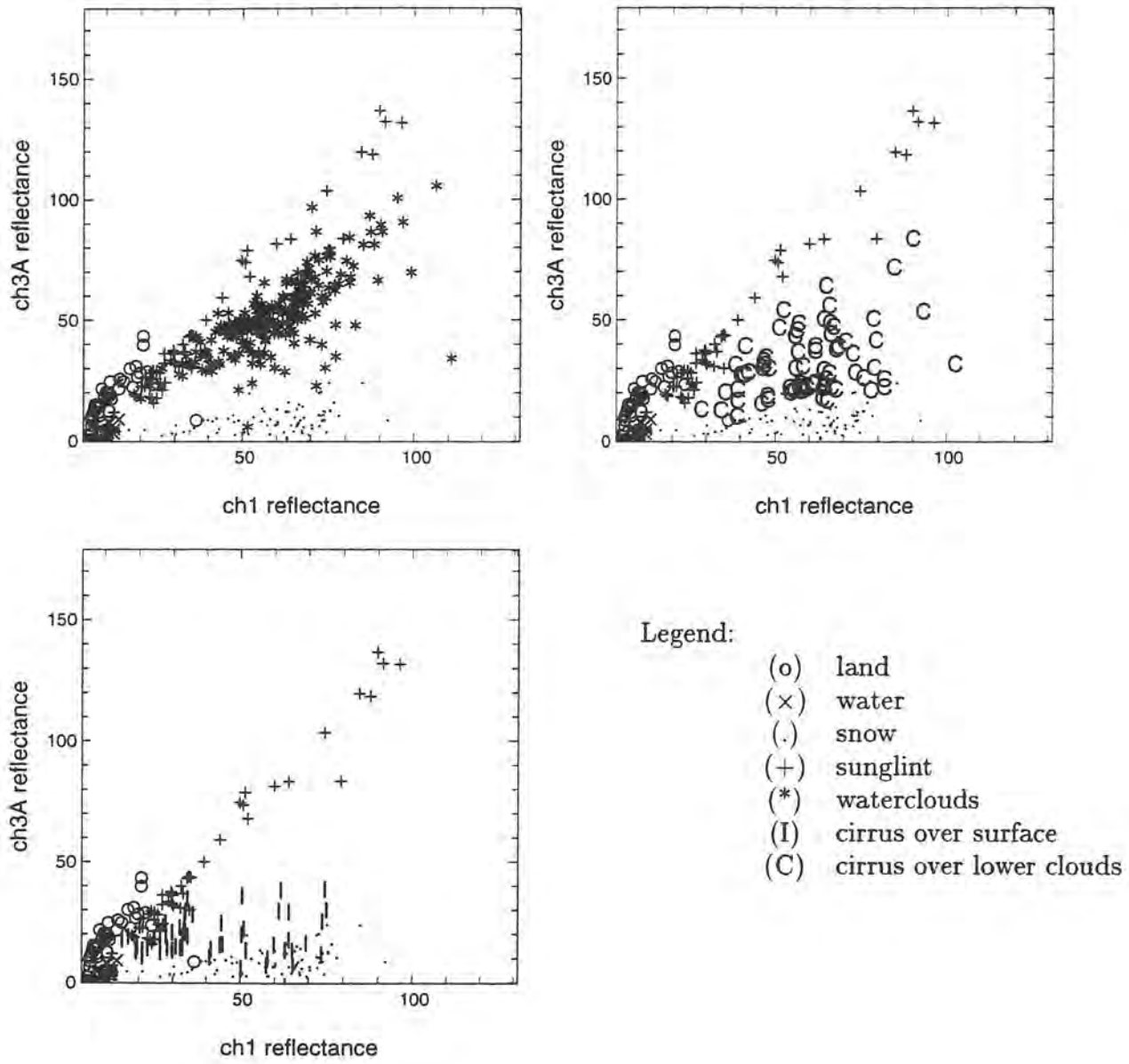


Figure 6: The 3A reflectances as a function of channel 1 for land, water, snow, sunglint and different cloud classes, for the water clouds in the upper left figure, for cirrus over surface in the upper right figure and for cirrus over lower clouds in the bottom figure. (data from OH dataset)

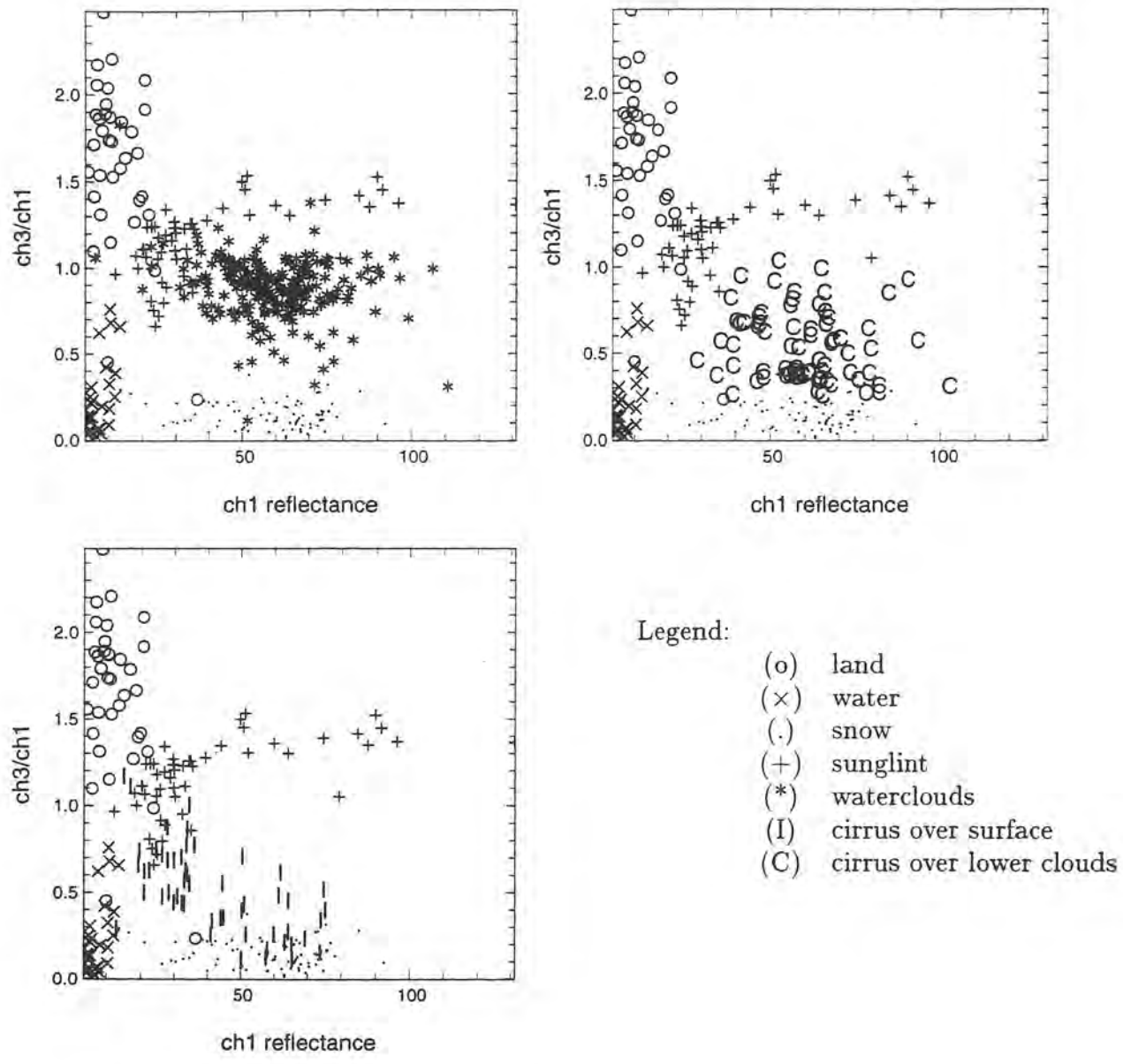


Figure 7: The quota of channel 3A and 1 reflectances as a function of channel 1 for land, water, snow, sunglint and different cloud classes, for the water clouds in the upper left figure, for cirrus over surface in the upper right figure and for cirrus over lower clouds in the bottom figure. (data from OH dataset)

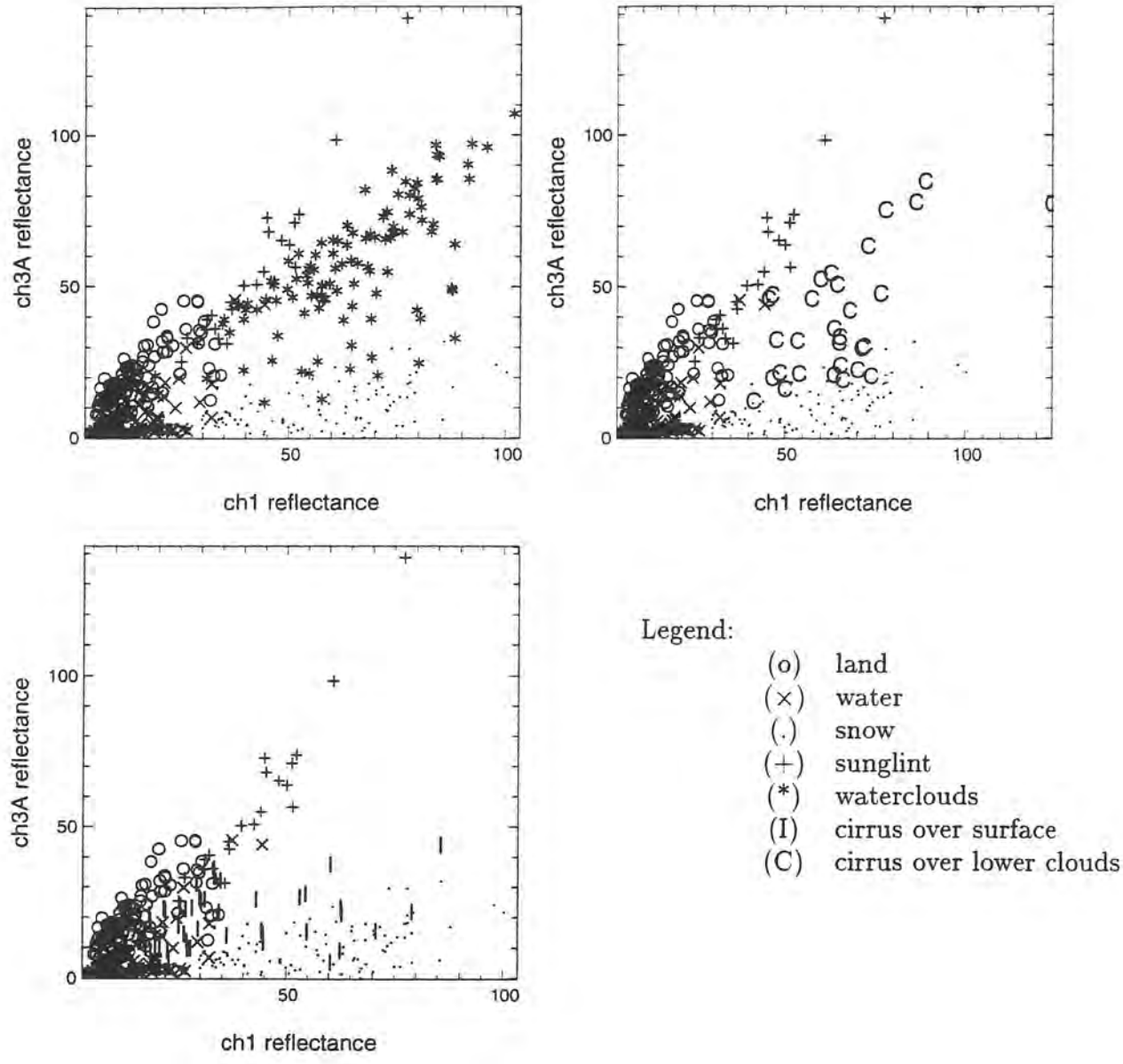


Figure 8: The 3A reflectances as a function of channel 1 for land, water, snow, sunglint and different cloud classes, for the water clouds in the upper left figure, for cirrus over surface in the upper right figure and for cirrus over lower clouds in the bottom figure. (data from GO dataset)

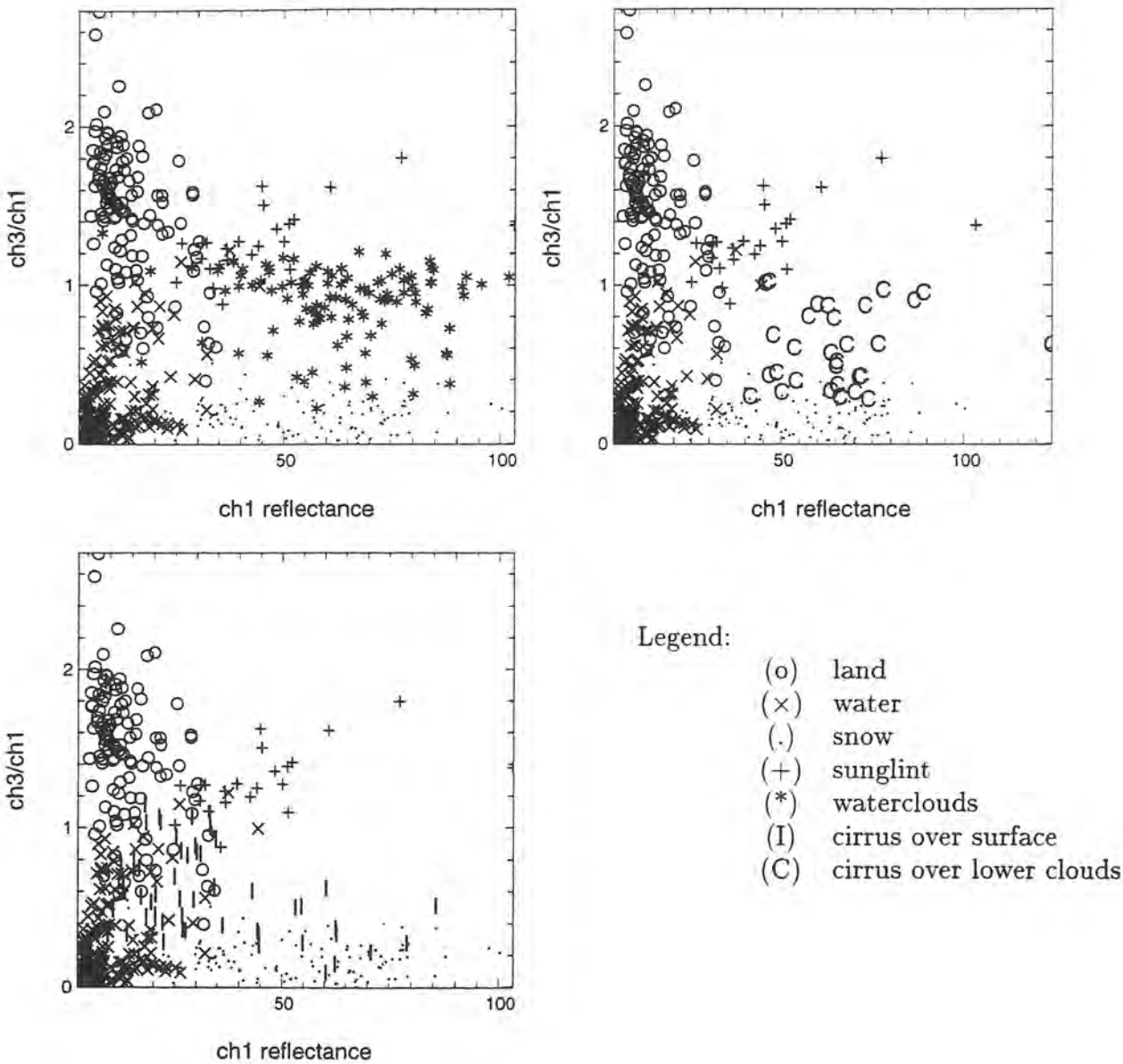


Figure 9: The quota of channel 3A and 1 reflectances as a function of channel 1 for land, water, snow, sunglint and different cloud classes, for the water clouds in the upper left figure, for cirrus over surface in the upper right figure and for cirrus over lower clouds in the bottom figure. (data from GO dataset)

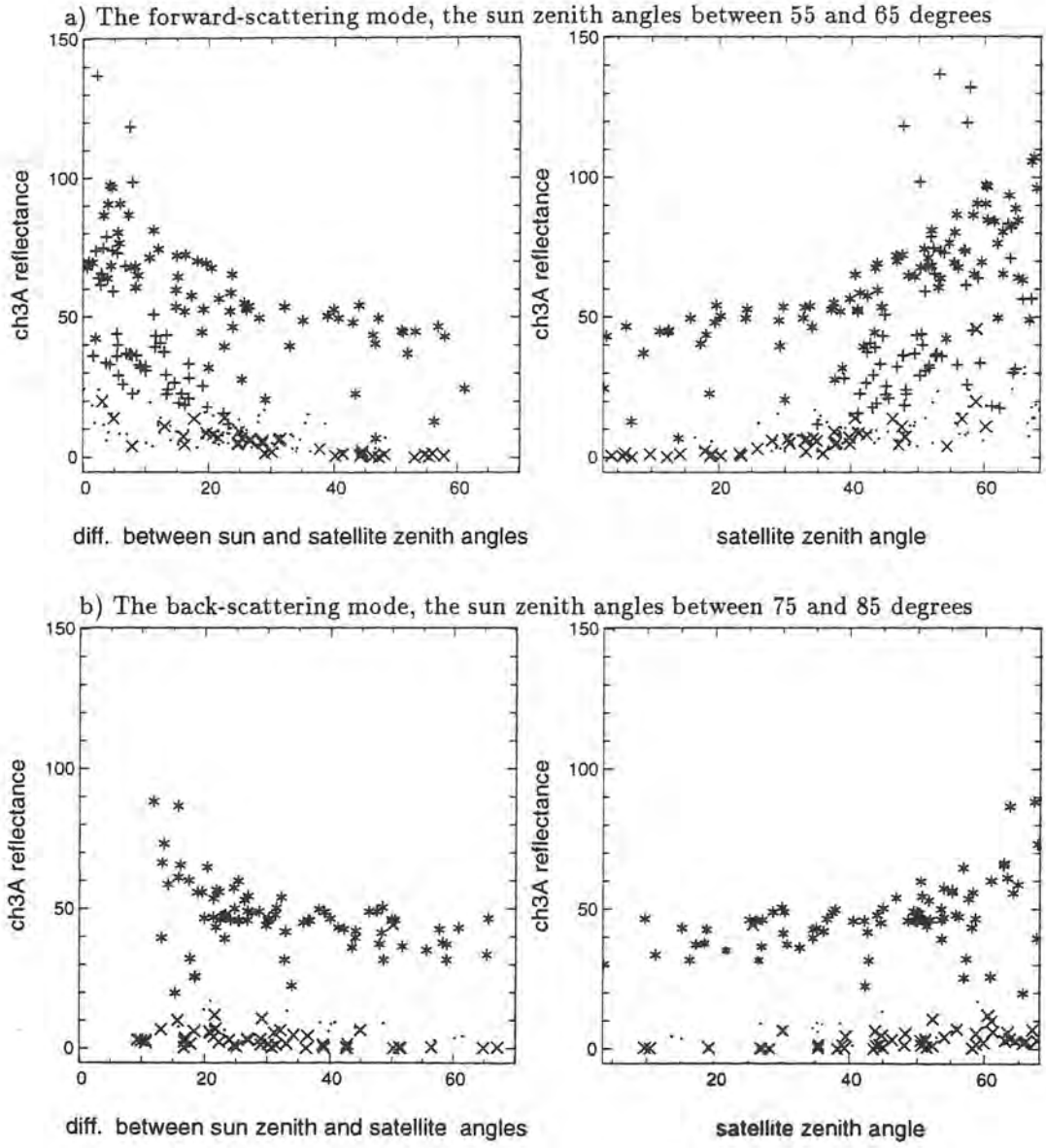


Figure 10: Behaviour of different classes for the channel 3A (a) in the forward-scattering mode where conditions for sunglint are present and (b) in the back-scattering mode where the conditions for sunglint are not present. Plots for water (x), snow (.),water clouds (*) and sunglint (+) (Data from OH+GO)

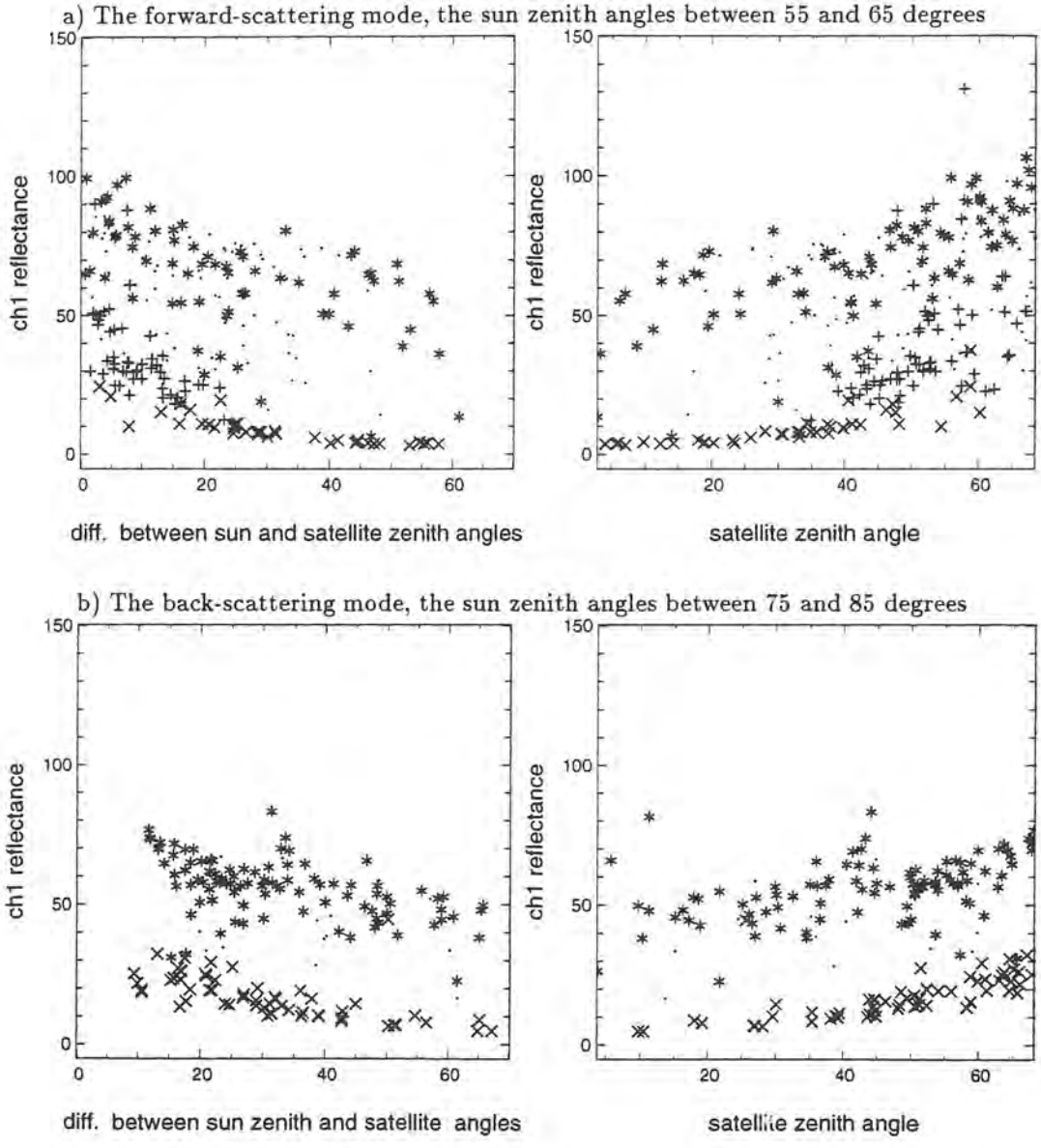


Figure 11: Behaviour of different classes for the channel 1 (a) in the forward-scattering mode where conditions for sunglint are present and (b) in the back-scattering mode where the conditions for sunglint are not present. Plots for water (x), snow (.),water clouds (*) and sunglint (+) (Data from OH+GO)

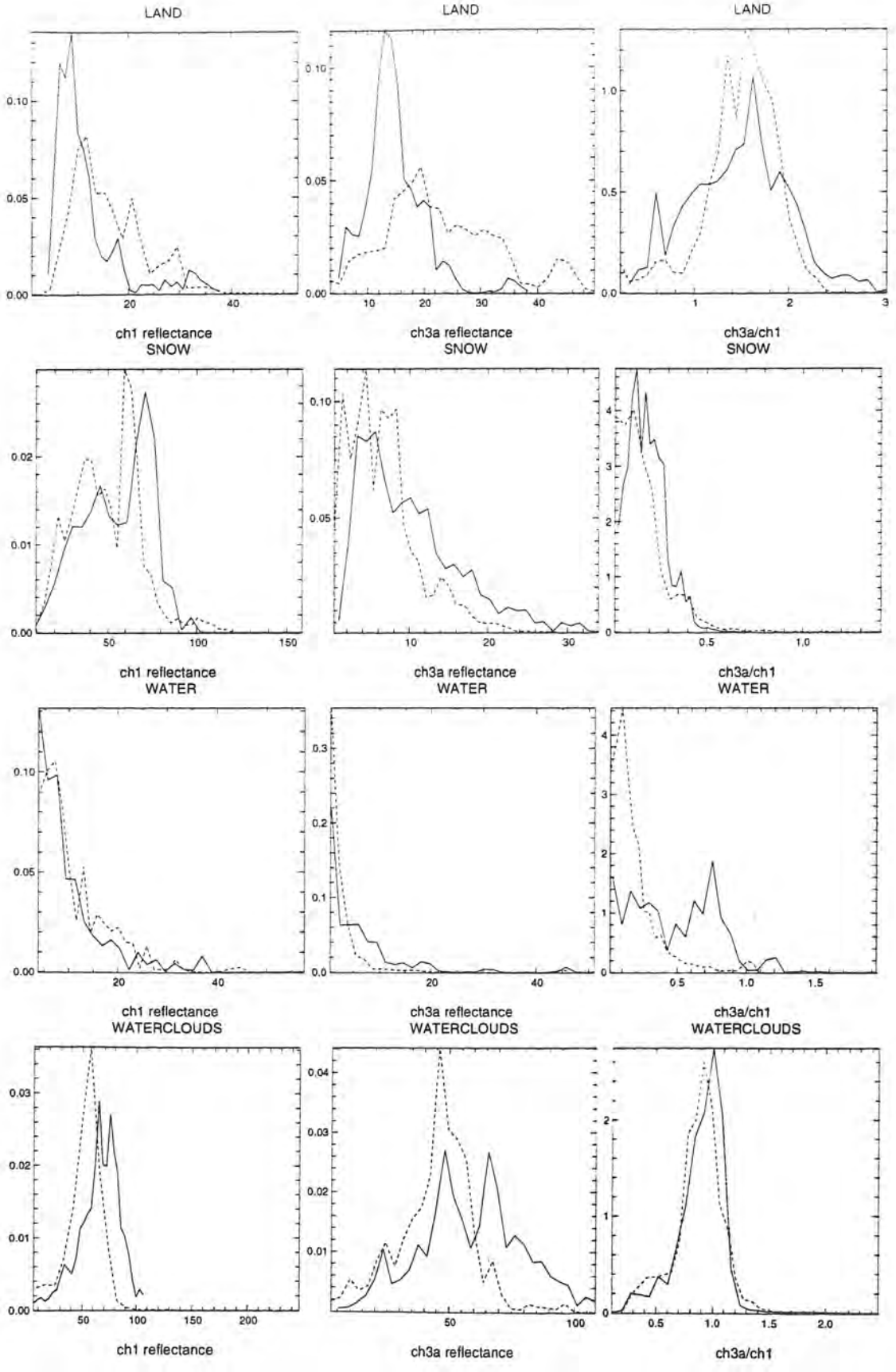


Figure 12: Histograms of the channel 1 and 3A reflectances and their quota for different classes, the forward-scattering mode (solid line) and the backward-scattering mode (dashed line) (Data from OH+GO)

4 Differences between 3A and 3B reflectances

To have a qualitative understanding of the differences in separability using channels 3A and 3B, the approximation of the 3B reflectance was calculated according to Appendix A.

Targets used for these tests were collected in the autumn 1998 and during the 3A test period. There were approximately equal number of targets from both periods, except for snow, that were mostly from the test period. Figure 13 show scatterplots for the separability between sunglint and water clouds and the separability between snow and water clouds. The only constraint for these targets was that the sun zenith angle must be less than 80 degrees. Note that some targets have reflectance values exceeding 100%. For channel 3B these values take place when the channel 3B is saturating (the brightness temperature about 320K) and sun zenith angle is rather high (over 70 degrees) thus indicating even higher reflectances than shown here. Figure 13 can be compared with Figures 6 and 8 showing scatterplots for the reflectances of channel 1 and 3A.

The separability of sunglint and waterclouds in 3B seems to be significantly better than for 3A. For both channels strong sunglint is well separated, but for 3B weaker sunglint is well separated too. For the further illustration of this separability a straight line is drawn in Figure 13 representing the threshold of one of the tests used in the Pathfinder project to mask clouds over sunglint regions: a pixel is classified as light to moderate sunglint if the quota between channel 3B and 1 reflectances is greater than 0.7 and the channel 1 reflectance is greater than 10%.[1] As the figure shows almost all of the sunglint targets can be separated from waterclouds using only this test. But, there are some sunglint targets for which the quota is much less than 0.7. Further examination of these targets did not reveal any explanation of this behaviour. One possibility could be that the targets did contain also very thin water clouds, thus causing a mixed spectral signature.

On the other hand, the snow separability with 3A seems to be much better than with 3B. Both channels have a very low reflectance for snow, but the 3B reflectance can be low for water clouds too. Some snow targets in Figure 13 have rather high 3B reflectance values; over 10%. For almost all of these targets the conditions for direct specular reflection was prevailing, so high values can at least partly be explained by "snowglint".

Notice that the variability of the azimuth angle difference for these targets was very large as opposed to the channel 3A data set, so more tests with more complete data sets are probably needed to confirm these conclusions.

5 Classification

The last task of the study was to propose a preliminary cloud discriminator based on channel 3A data. Two different kind of methods were investigated, a maximum likelihood method and a thresholding method. In addition, some indications on possible improvements of separability by introducing a few other channels and features were investigated. The choice of an additional maximum likelihood classifier is explained by the fact that such a method can give some additional indication of the quality of classification.

To assist in choosing the appropriate image features and to get some quantitative results for the separability of classes, the Jeffries-Matusita distance separability index was used.

5.1 The Jeffries-Matusita Distance Separability Index

We assume that classes follow the normal probability density distribution $p(X_k|i)$ as given by

$$p(X_k|i) = \frac{1}{(2\pi)^{n/2}|V_i|^{1/2}} \exp[-\frac{1}{2}(X_k - U_k)^T V_i^{-1}(X_k - U_k)] \quad (1)$$

where U_i is the mean vector, V_i is the covariance matrix,$|U_i|$ denote the determinant of a matrix U_i, U_i^{-1} denotes the inverse of matrix U_i and $|U_i|^T$ denotes the transpose of the matrix U_i. Now we can compute

[1] For a description of the Pathfinder project in general and cloud discrimination in sunglint regions in particularly check `http://asd-www.larc.nasa.gov/~baum/Pathfinder/sunglint.html`

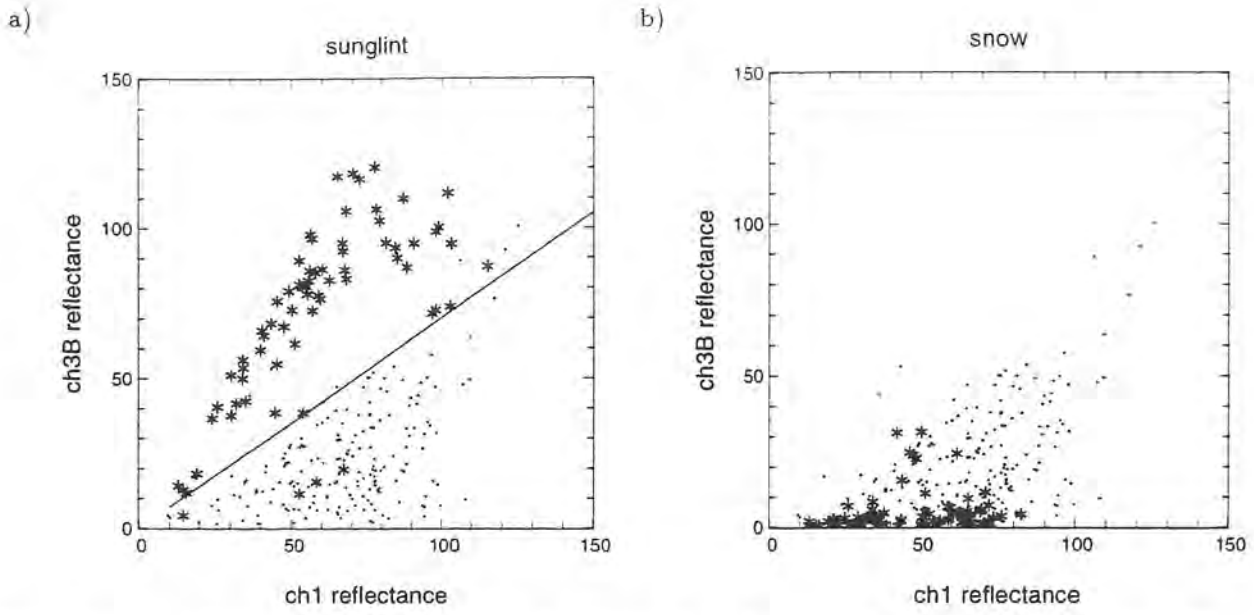


Figure 13: The reflectance of channel 3B as a function of the reflectance of channel 1 for a) sunglint "*" and waterclouds "." and b) snow "*" and waterclouds "." from AVHRR/2 targets. For the sunglint/water versus cloud separability, the threshold used in the Pathfinder project to mask clouds over sunglint regions is displayed by the straight line in a).

the separability of different classes using a separability index. For this study, the Jeffries-Matusita Distance Separability Index (hereafter denoted J-M Index) was chosen. It is given by [3] as

$$J - M_{ij} = [2(1 - \exp(-a_{ij}))]^{1/2} \tag{2}$$

where

$$a_{ij} = \frac{1}{8}(U_i - U_j)^T(\frac{V_i + V_j}{2})^{-1}(U_i - U_j) + \frac{1}{2}\ln(|(V_i + V_j)/2|)/(|V_i||V_j|)^{1/2}) \tag{3}$$

The J-M Index is a saturating formula meaning that as the separability increases and approaches infinity, the value of the J-M Index gets closer to $\sqrt{2}$. It tends to over-emphazise results for small interclass separations and under-emphasize results for the greater separations. The J-M Index can also only be calculated if the inverse of the covariance matrix can be produced, meaning that no feature based on a linear combination of any pair of the other used features is allowed.

The features used were visible channel reflectances (ch1, ch2, ch3A), various differences and quotas between them (ch3A - ch1, ch3A - ch2, ch2 - ch1, ch3A / ch1, ch2 / ch1, ch3A / ch2) and the brightness temperature differences between infrared channels and HIRLAM-surface temperature (ch4 - Tsurf, ch5 - Tsurf, ch4 - ch5). The separability indices were calculated using the combined OH and GO data sets.

The results for each class against all other classes using the best combination of one, two and three features are given in the tables 2-9. Note that separabilities for ALL CIRRUS/OTHER CIRRUS and ALL CIRRUS/THIN CIRRUS are not shown because it does not make sense to calculate separability between a main class and its subclasses.

These separability results suggest the following: When using only one feature, only water clouds and thin cirrus are best separated from each other using just the channel 3A. This is consistent with our previous results from the histograms.

Snow is best separated using the channel 3A combined with other channels. Snow and water clouds can be separated using the quota between channels 3A and 2, snow and water using the difference between channels 3A and 2, snow and land using the difference between channels 3A and 1. Cirrus classes are the most difficult classes to separate from the snow.

Table 2: Maximum separability of SNOW and other classes for one, two and three chosen features.

Compared class	One feature (separability)	Two features (separability)	Three features (separability)
LAND	ch3A - ch1 (1.313)	ch3A / ch1 ch3A / ch2 (1.387)	ch3A / ch1 ch3A / ch2 ch3A - ch1 (1.406)
ALL CIRRUS	ch5 - Tsurf (1.098)	ch5 - Tsurf ch3A / ch2 (1.307)	ch5 - Tsurf ch3A / ch2 ch3A (1.344)
OTHER CIRRUS	ch5 - Tsurf (1.250)	ch4 - Tsurf ch3A / ch1 (1.400)	ch4 - Tsurf ch3A / ch1 ch2 / ch1 (1.404)
THIN CIRRUS	ch3A / ch2 (1.019)	ch3A / ch2 ch5 - Tsurf (1.219)	ch3A / ch2 ch5 - Tsurf ch3A / ch1 (1.303)
WATERCLOUDS	ch3A / ch2 (1.349)	ch3A / ch2 ch4 - Tsurf (1.408)	ch3A / ch2 ch4 - Tsurf ch3A / ch1 (1.410)
SUNGLINT	ch3A / ch2 (1.414)	ch3A / ch2 ch3A / ch1 (1.414)	ch3A / ch2 ch3A / ch1 ch1 (1.414)
WATER	ch3A - ch2 (1.329)	ch3A / ch1 ch2 / ch1 (1.402)	ch3A / ch1 ch2 / ch1 ch1 (1.413)

Table 3: Maximum separability of OTHER CIRRUS and other classes for one, two and three chosen features.

Compared class	One feature (separability)	Two features (separability)	Three features (separability)
LAND	ch5 - Tsurf (1.304)	ch1 ch2 / ch1 (1.399)	ch1 ch2 / ch1 ch5 - Tsurf (1.410)
SNOW	ch5 - Tsurf (1.250)	ch4 - Tsurf ch3A / ch1 (1.400)	ch4 - Tsurf ch3A / ch1 ch2 / ch1 (1.404)
SUNGLINT	ch4 - Tsurf (1.273)	ch3A / ch1 ch3A / ch2 (1.360)	ch3A / ch1 ch3A / ch2 ch4 - Tsurf (1.390)
THIN CIRRUS	ch2 / ch1 (0.911)	ch2 / ch1 ch2 - ch1 (1.131)	ch2 / ch1 ch2 ch1 (1.248)
WATER	ch2 (1.396)	ch2 ch2 / ch1 (1.410)	ch3A ch2 / ch1 ch3A / ch1 (1.414)
WATERCLOUDS	ch5 - Tsurf (0.752)	ch2 / ch1 ch2 - ch1 (1.006)	ch2 / ch1 ch2 - ch1 ch5 - Tsurf (1.135)

Table 4: Maximum separability of LAND and other classes for one, two and three chosen features.

Compared class	One feature (separability)	Two features (separability)	Three features (separability)
ALL CIRRUS	ch5 - Tsurf (1.178)	ch5 - Tsurf ch2 / ch1 (1.314)	ch5 - Tsurf ch2 / ch1 ch2 - ch1 (1.361)
OTHER CIRRUS	ch5 - Tsurf (1.304)	ch1 ch2 / ch1 (1.399)	ch1 ch2 / ch1 ch5 - Tsurf (1.410)
SNOW	ch3A - ch1 (1.313)	ch3A / ch1 ch3A / ch2 (1.387)	ch3A / ch1 ch3A / ch2 ch3A - ch1 (1.406)
THIN CIRRUS	ch5 - Tsurf (1.111)	ch5 - Tsurf ch2 / ch1 (1.269)	ch5 - Tsurf ch2 / ch1 ch2 - ch1 (1.336)
WATERCLOUDS	ch2 / ch1 (1.112)	ch2 / ch1 ch1 (1.348)	ch2 / ch1 ch1 ch4 - Tsurf (1.375)
SUNGLINT	ch2 / ch1 (1.205)	ch2 / ch1 ch2 - ch1 (1.333)	ch2 / ch1 ch3A / ch1 ch3A / ch2 (1.400)
WATER	ch2 / ch1 (1.115)	ch2 / ch1 ch3A (1.272)	ch2 / ch1 ch3A / ch1 ch2 - ch1 (1.354)

Table 5: Maximum separability of THIN CIRRUS and other classes for one, two and three chosen features.

Compared class	One feature (separability)	Two features (separability)	Three features (separability)
LAND	ch5 - Tsurf (1.111)	ch5 - Tsurf ch2 / ch1 (1.269)	ch5 - Tsurf ch2 / ch1 ch2 - ch1 (1.336)
SNOW	ch3A / ch2 (1.019)	ch3A / ch2 ch5 - Tsurf (1.219)	ch3A / ch2 ch5 - Tsurf ch3A / ch1 (1.303)
SUNGLINT	ch3A / ch2 (1.278)	ch3A / ch2 ch5 - Tsurf (1.357)	ch3A / ch2 ch5 - Tsurf ch2 / ch1 (1.383)
OTHER CIRRUS	ch2 / ch1 (0.911)	ch2 / ch1 ch2 - ch1 (1.131)	ch2 / ch1 ch2 ch1 (1.248)
WATER	ch2 (1.045)	ch2 / ch1 ch3A - ch2 (1.222)	ch2 / ch1 ch3A - ch2 ch5 - Tsurf (1.305)
WATERCLOUDS	ch3A (1.028)	ch3A ch2 / ch1 (1.092)	ch1 ch3A / ch2 ch3A - ch2 (1.169)

Table 6: Maximum separability of WATERCLOUDS and other classes for one, two and three chosen features.

Compared class	One feature (separability)	Two features (separability)	Three features (separability)
LAND	ch2 / ch1 (1.112)	ch2 / ch1 ch1 (1.348)	ch2 / ch1 ch1 ch4 - Tsurf (1.375)
SNOW	ch3A / ch2 (1.349)	ch3A / ch2 ch4 - Tsurf (1.408)	ch3A / ch2 ch4 - Tsurf ch3A / ch1 (1.410)
SUNGLINT	ch4 - Tsurf (0.983)	ch4 - Tsurf ch3A / ch2 (1.151)	ch4 - Tsurf ch3A / ch1 ch2 / ch1 (1.234)
WATER	ch2 (1.329)	ch2 ch2 / ch1 (1.383)	ch2 ch2 / ch1 ch3A (1.396)
ALL CIRRUS	ch3A / ch2 (0.695)	ch3A ch5 - Tsurf (0.832)	ch3A ch2 ch3A / ch2 (0.936)
OTHER CIRRUS	ch5 - Tsurf (0.752)	ch2 / ch1 ch2 - ch1 (1.006)	ch2 / ch1 ch2 - ch1 ch5 - Tsurf (1.135)
THIN CIRRUS	ch3A (1.028)	ch3A ch2 / ch1 (1.092)	ch1 ch3A / ch2 ch3A - ch2 (1.169)

Table 7: Maximum separability of SUNGLINT and other classes for one, two and three chosen features.

Compared class	One feature (separability)	Two features (separability)	Three features (separability)
ALL CIRRUS	ch3A / ch2 (1.278)	ch3A / ch2 ch4 - Tsurf (1.358)	ch3A / ch2 ch4 - Tsurf ch3A - ch1 (1.383)
LAND	ch2 / ch1 (1.205)	ch2 / ch1 ch2 - ch1 (1.333)	ch2 / ch1 ch3A / ch1 ch3A / ch2 (1.400)
OTHER CIRRUS	ch4 - Tsurf (1.273)	ch3A / ch1 ch3A / ch2 (1.360)	ch3A / ch1 ch3A / ch2 ch4 - Tsurf (1.390)
SNOW	ch3A / ch2 (1.414)	ch3A / ch2 ch3A / ch1 (1.414)	ch3A / ch2 ch3A / ch1 ch1 (1.414)
THIN CIRRUS	ch3A / ch2 (1.278)	ch3A / ch2 ch5 - Tsurf (1.357)	ch3A / ch2 ch5 - Tsurf ch2 / ch1 (1.383)
WATERCLOUDS	ch4 - Tsurf (0.983)	ch4 - Tsurf ch3A / ch2 (1.151)	ch4 - Tsurf ch3A / ch1 ch2 / ch1 (1.234)
WATER	ch3A / ch1 (1.274)	ch3A / ch1 ch3A (1.351)	ch3A / ch1 ch1 ch3A - ch2 (1.376)

Table 8: Maximum separability of ALL CIRRUS and other classes for one, two and three chosen features.

Compared class	One feature (separability)	Two features (separability)	Three features (separability)
LAND	ch5 - Tsurf (1.178)	ch5 - Tsurf ch2 / ch1 (1.314)	ch5 - Tsurf ch2 / ch1 ch2 - ch1 (1.361)
SNOW	ch5 - Tsurf (1.098)	ch5 - Tsurf ch3A / ch2 (1.307)	ch5 - Tsurf ch3A / ch2 ch3A (1.344)
SUNGLINT	ch3A / ch2 (1.278)	ch3A / ch2 ch4 - Tsurf (1.358)	ch3A / ch2 ch4 - Tsurf ch3A - ch1 (1.383)
WATER	ch2 (1.165)	ch2 ch2 / ch1 (1.291)	ch5 - Tsurf ch2 / ch1 ch3A - ch2 (1.344)
WATERCLOUDS	ch3A / ch2 (0.695)	ch3A ch5 - Tsurf (0.832)	ch3A ch2 ch3A / ch2 (0.936)

Table 9: Maximum separability of WATER and other classes for one, two and three chosen features.

Compared class	One feature (separability)	Two features (separability)	Three features (separability)
ALL CIRRUS	ch2 (1.165)	ch2 ch2 / ch1 (1.291)	ch5 - Tsurf ch2 / ch1 ch3A - ch2 (1.344)
LAND	ch2 / ch1 (1.115)	ch2 / ch1 ch3A (1.272)	ch2 / ch1 ch3A / ch1 ch2 - ch1 (1.354)
OTHER CIRRUS	ch2 (1.396)	ch2 ch2 / ch1 (1.410)	ch3A ch2 / ch1 ch3A / ch1 (1.414)
SNOW	ch3A - ch2 (1.329)	ch3A / ch1 ch2 / ch1 (1.402)	ch3A / ch1 ch2 / ch1 ch1 (1.413)
SUNGLINT	ch3A / ch1 (1.274)	ch3A / ch1 ch3A (1.351)	ch3A / ch1 ch1 ch3A - ch2 (1.376)
THIN CIRRUS	ch2 (1.045)	ch2 / ch1 ch3A - ch2 (1.222)	ch2 / ch1 ch3A - ch2 ch5 - Tsurf (1.305)
WATERCLOUDS	ch2 (1.329)	ch2 ch2 / ch1 (1.383)	ch2 ch2 / ch1 ch3A (1.396)

Separating land and water from each other doesn't need channel 3A, use of other visible channels is enough. This does change a little when more features are added.

Results of separating sunglint from other classes may be a bit more suspect than others, because sunglint is not well approximated with the Gaussian curve. Sunglint behaves mostly like water clouds, so it is well separated from other classes like water clouds but separation from water clouds is difficult. For the separation from water clouds, the surface temperature differences are used. When using only one feature, the J-M Index has one of its lowest values. According to other tests, discrimination of sunglint and water clouds can be improved using the channel 4 and channel 5 temperature difference, though the J-M Index does not indicate it.

Different kinds of cirrus are mostly separated by the temperature difference, but some classes can be separated using channel 3A (water clouds / thin cirrus), channel 2 (water / cirrus) or a combination of these two (snow / thin cirrus). Separation of all cirrus and water clouds is very problematic and the J-M Index has its lowest values here. Thin cirrus can be somewhat better separated than other cirrus from water clouds, but the separability is still low.

Notice that these are only results for the class separability and not results from any classification. In the classification there is usually an optimum number of features and the quality of classification decreases, if more features are added. The separability index never decreases when new features are added. This makes the comparison of the separability with different number of features problematic. In the end, it is the classification results that decide the best choice of features.

5.2 The thresholding classifier

The results using the J-M Index were used as a first guide to choose features for a classifier. The finally chosen features are described in section 5.2.1 and results are given in section 5.2.2.

5.2.1 Tests using chosen features

The flowchart for the classification is shown in Figure 14. Features used are the channel 1 reflectance (ch1), the quota between the channels 3A and 1 (ch3A/ch1), the quota between the channels 2 and 1 (ch2/ch1), the difference between the channel 3A and 2 reflectances (ch3A- ch2) and the difference between the channel 4 temperature and the surface temperature (ch4-Tsurf).

Cirrus is detected using the difference between channel 4 and the Hirlam surface temperature. If the difference is large enough, the pixel is labeled as cirrus. This test is a rather crude but effective way to discriminate cirrus and snow. As the channel 3A appearance of cirrus clouds resembles very much that of snow, one possibility to improve snow/cirrus separability would be to utilise prior knowledge of snow coverage. However, this isn't much of a help during winter in Northern Europe due to the limited snow cover information available. Cirrus tests need obviously more work.

Sun glint is only allowed if the geometrical conditions for sunglint exist. These conditions are said to be present if the difference between the sun and satellite zenith angles is less than 25 degrees and the pixels belongs to the forward scattering area. Strong sunglint is discriminated from the water clouds using high values of ch3A/ch1 and the channel 1 reflectance. Weak sunglint is also discriminated with the help of same features: pixels otherwise resembling waterclouds are labeled as sunglint if they are very warm. Notice that sunglint cannot be well discriminated from the water clouds using just these tests and better test are thus needed.

All land targets and most of water targets can be discriminated from other classes using the channel 1 reflectance. If it is low enough, the pixel is either land or water. These two classes are discriminated using the quota between the channels 2 and 1. If it is larger than one, the pixel is land.

Both water clouds and snow have high values of the channel 1 reflectance, but water clouds have high values of the quota between the channels 3A and 1 and snow has low values of the quota between the channels 3A and 1. There is some overlap with water and snow, so they are discriminated with the difference between the channel 3A and 2 reflectances, snow having lower values.

Notice that this scheme does not use the difference between channels 4 and 5 (a commonly used feature for cirrus detection) at all. For the current test dataset, the separability study suggested somewhat

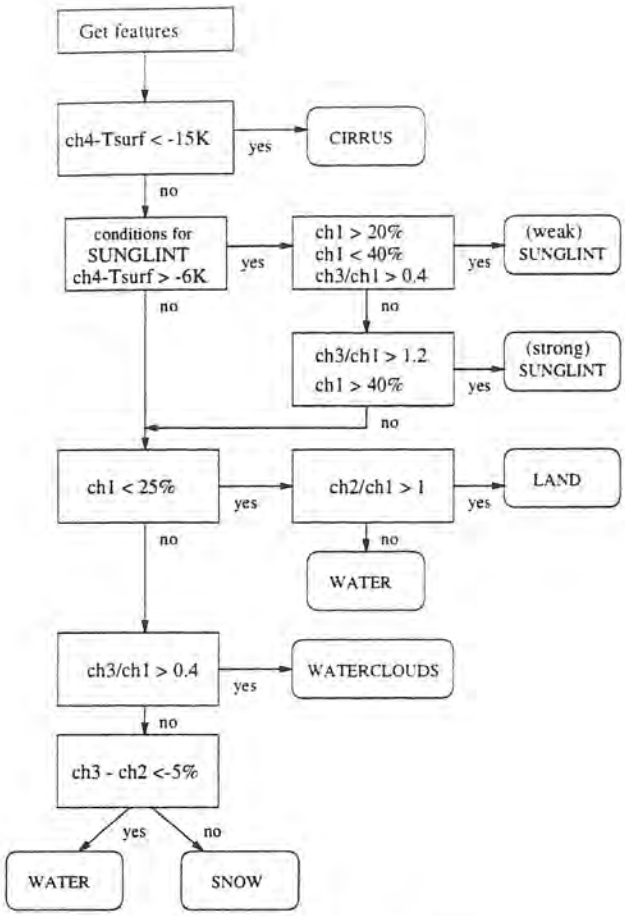


Figure 14: The flowchart for the thresholding classifier. The conditions for sunglint are present if the difference between the sun and satellite zenith angles is less than 25 degrees and the pixels belongs to the forward scattering area.

surprisingly that the feature ch4-ch5 would have only a marginal influence on class separability. Even though the separability index studies reported earlier in section 5.1 do not select this feature specifically, it is evident that it has a separability potential (shown in Figure 15, left part). However, the same kind of information is probably available also in other features (e.g. as seen in Figure 15, right part). To get additional information, other features were apparently chosen.

One explanation for the above documented relative unimportance of feature ch4-ch5 in class separability, compared to what is usually expected of this feature, may lie in the spectral response functions for channel 4 and 5. Channel 4 and 5 on NOAA 15 are significantly closer together compared to earlier satellites, and thereby the channel difference from NOAA 15 data should be less sensitive to the presence of clouds.

Another explanation may be found in the character of the data for this study and that the importance of the ch4-ch5 feature to cloud detection may vary with season. E.g. it is likely that it would be given higher weight in the summer time when surface temperatures and moisture contents in the lower troposphere are higher.

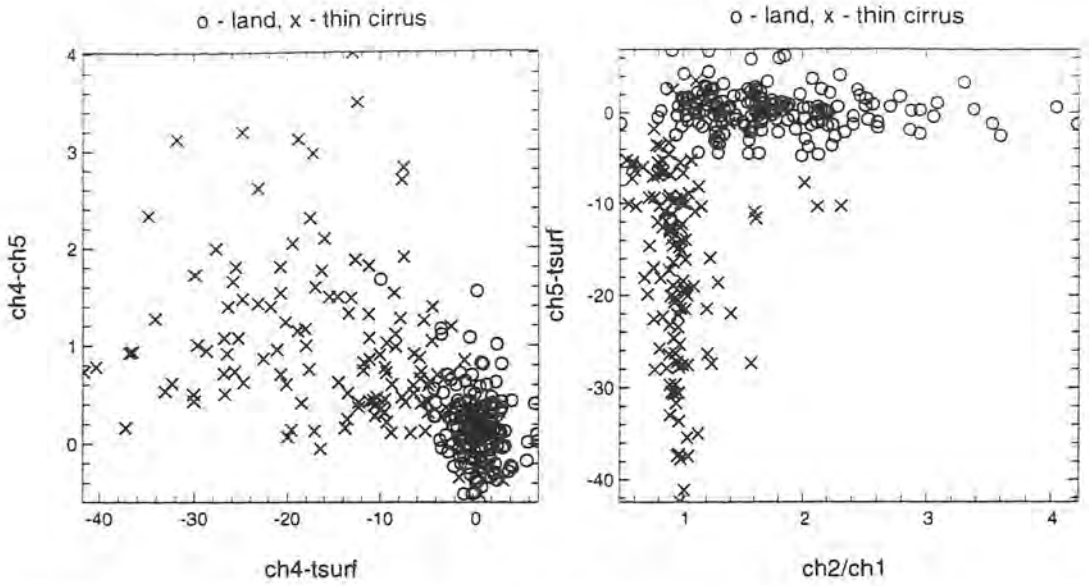


Figure 15: The separability of land and thin cirrus using features ch4-tsurf and ch4-ch5 (left side) and using features ch2/ch1 and ch5-tsurf (right side).

5.2.2 Results

Two kinds of results are given here, full results using each class and combined results for only surface/cloud discrimination. The results are all computed for combined OH+GO-targets.

The results are shown in the confusion matrix in Table 10. The rows tell how well the known targets were classified. For example, 175 water targets or 96.2% of them where correctly classified as water, 4 targets where wrongly classified as sunglint and 3 as waterclouds. The columns tell how well the classifier managed to classify targets. For example, 175, or 82.9% of them, of all targets classified as water were really water targets. Misclassified targets include 3 sunglint targets and 16 cirrus targets.

The surface classes are mostly well separated between themselves and between cloud classes. The classifier has considerable problems in separating cirrus from waterclouds. There are also problems separating cirrus from water and snow surfaces. The test for cirrus is very crude, so this is not suprising. These classification problems for the full classification can also seen in the cloud mask results. Most of the cloudy pixels erronously classified as cloud-free are cirrus pixels classified as water and snow. On the other hand, because of the rather cold temperature threshold, no cirrus targets are classified as surface. In the group cloud-free targets classified as cloudy, sunglint targets seem to be the most difficult to classify accurately.

The results are encouraging, but results can surely be improved as tests are still rather crude.

5.3 The maximum likelihood classifier

Let us again assume that classes follow the multivariate normal probability density distribution as given by equation 1. Now we can calculate the posterior probability that a target is in class i using the Bayes' theorem

$$p(i|X_k) = \frac{p(X_k|i)p(i)}{p(X_k)} \tag{4}$$

The target is then classified to the class with the highest posterior probability.

The mean vectors and covariances were computed using one half of the data set and the results are given for the other half of the data set, which then could be considered as independent. The prior

Table 10: The confusion matrices for a) the complete classification and b) the cloud mask using the thresholding classifier.

a)

		classifier						
		WATER	SUNGLINT	LAND	SNOW	WATERCL	CIRRUS	
	WATER	175	4	0	0	3	0	96.2%
	SUNGLINT	3	64	0	0	7	0	86.5%
targets	LAND	5	3	127	2	10	0	86.4%
	SNOW	6	0	4	182	3	0	93.3%
	WATERCLOUDS	6	0	3	1	227	104	66.6%
	CIRRUS	16	2	7	19	19	147	70.0%
		82.9%	87.7%	90.1%	89.2%	84.4%	58.6%	

Correctly classified 80.2% targets

b)

		classifier		
		cloud-free	cloudy	
targets	cloud-free	575	23	96.2%
	cloudy	54	497	90.2%
		91.4%	95.6%	

Correctly classified 93.3% targets

probability $p(i)$ is supposed to be the same for all classes, except that sunglint is not allowed outside the sun glint area. This area is defined the same way as for the thresholding classifier.

The same features as for the separability index study were defined. Classification results for all feature combinations having more than two but less than eight features were computed. The classification results were computed for all used classes and for the cloud mask by combining the results of all classed into two classes, "cloudy" and "cloud-free".

5.3.1 Results

The best classification results for the best set of features are given for the cloud mask in table 11 and for the complete classification in tables 12.

Note that the maximum number of features considered was eight, but only results for seven could be calculated in practice. This is because the inverse of the covariance matrix could not be calculated for eight features since any eighth feature would be possible to construct by a linear combination of some of the other seven features. In tables 13 and 14 are given the confusion matrices for the best cloud mask and the best complete classification, respectively.

At first, the quality of the classification results increases when more features are added, but then the results peak and the quality start to decrease. This peak happens for the cloud mask at 4 features and for the complete classification at 5 features.

The best features selected for the cloud mask don't differ much for the two different datasets. Only when the classification is made using two features, the other feature can be either ch3A/ch2 or ch3A/ch1. The optimum number of features is four, but even for three features the results are quite good. These three features are channel 1 and 3A reflectances and the difference between the HIRLAM surface temperature and channel 4 temperature. The efficiency of the first two of these features could be seen in the section 3.2. The adding of the fourth feature ch3A/ch1 doesn't increase the quality of results much.

For the complete classification the best features differ a lot more between the two different datasets. But it can be seen that for the best results, data from all channels must be used. The channel 5 and the channel 2 data are introduced earlier than for the cloud mask, so the information from the channels 2 and 5 increases the accuracy of the classification inside the "cloudy" and "cloud-free" metaclasses. The "pure" feature of the temperature difference between channels 4 and 5 is not present at any feature selections,

Table 11: The best classification results for the cloud mask with the maximum likelihood classifier (M-L).

M-L results for GO data trained on OH data

Features selected	Accuracy, %
ch4 - Tsurf, ch3A / ch2	91.7
ch 1, ch 3A, ch4 - Tsurf	93.4
ch 1, ch 3A, ch4 - Tsurf, ch3A / ch1	93.4
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch3A / ch1	92.7
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch5 - Tsurf, ch3A / ch1	92.5
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch5 - Tsurf, ch3A / ch1, ch3A / ch2	90.1

M-L results for OH data trained on GO data

Features selected	Accuracy, %
ch4 - Tsurf, ch3A / ch1	91.7
ch 1, ch 3A, ch4 - Tsurf	96.5
ch 1, ch 3A, ch4 - Tsurf, ch3A / ch1	96.7
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch3A / ch1	96.1
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch5 - Tsurf, ch3A / ch1	95.4
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch5 - Tsurf, ch3A / ch1, ch3A / ch2	95.0

but it is likely, that the same kind of information is implicitly used when using the surface temperature differences for both channels 4 and 5.

The confusion matrices of results are shown in Table 14. The separation of waterclouds and cirrus clouds is the most problematic for the classifier. Compared to the results of the thresholding classifier in table 10, the maximum likelihood classifier classifies sunglint better. The cirrus is also better separated from the water.

There are also noticable differences between results for the different datasets, for example how well cirrus and water are classified. This can be explained by the differences between datasets. For example, as can be seen from Figures 7 and 9, the variability of water class is greater in GO data than in OH data, so it is not suprising if classification of water is worse when the covariances are computed using OH data and used on GO data.

The results of the maximum likelihood classifier indicate that the accuracy of the thresholding classifier can be increased in the further studies. But the accuracy doesn't differ very much between the thresholding and the maximum likelihood classifiers, so the improvement cannot be changed remarkably.

In addition to the tests on individual targets also "full" classification were carried out on a majority of all available scenes for the test period. One example of such a result based on the maximum likelihood classifier is shown on the front cover of this report. Generally, results looked quite reasonable but they contained also naturally all of the defects which were noticed in test on individual targets.

References

[1] Adam Dybbroe. *The interactive training manager - User's Manual.* SMHI, 1998.

[2] Karl-Göran Karlsson. *Cloud classification with the SCANDIA model.* Number 67. SMHI Reports Meteorology and Climatology, 1996.

[3] I. L. Thomas, N.P. Ching, V. M. Benning, and J.A. D'Aguanno. A review of multi-channel indices of class separability. *International Journal of Remote Sensing*, 8(3):331–350, 1987.

Table 12: The best classification results for the complete classification with the maximum likelihood classifier (M-L).

M-L results for GO data trained on OH data

Features selected	Accuracy, %
ch 1, ch3A / ch1	75.0
ch4 - Tsurf, ch3A - ch2, ch3A / ch1	78.8
ch4 - Tsurf, ch5 - Tsurf, ch3A - ch2, ch3A / ch1	81.1
ch4 - Tsurf, ch5 - Tsurf, ch3A - ch2, ch3A / ch1, ch3A / ch2	81.6
ch4 - Tsurf, ch5 - Tsurf, ch3A - ch1, ch3A - ch2, ch3A / ch1, ch3A / ch2	81.3
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch5 - Tsurf, ch3A / ch1, ch3A / ch2	80.5

M-L results for OH data trained on GO data

Features selected	Accuracy, %
ch3A - ch1, ch3A / ch1	79.4
ch 1, ch5 - Tsurf, ch3A / ch1	85.7
ch 2, ch4 - Tsurf, ch5 - Tsurf, ch3A / ch1	86.4
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch5 - Tsurf	86.8
ch 3A, ch4 - Tsurf, ch5 - Tsurf, ch2 - ch1, ch3A / ch1, ch3A / ch2	85.3
ch 1, ch 2, ch 3A, ch4 - Tsurf, ch5 - Tsurf, ch3A / ch1, ch3A / ch2	83.7

Table 13: The confusion matrices for the cloud mask with the maximum likelihood classifier (M-L) using features ch 1, ch 3A, ch4 - Tsurf, ch3A / ch1 .

M-L results for GO data trained on OH data

		classifier		
		cloud-free	cloudy	
targets	cloud-free	384	19	95.3%
	cloudy	21	180	89.6%
		94.8%	90.5%	

Correctly classified 93.4% targets

M-L results for OH data trained on GO data

		classifier		
		cloud-free	cloudy	
targets	cloud-free	191	4	97.9%
	cloudy	14	336	96.0%
		93.2%	98.8%	

Correctly classified 96.7% targets

Table 14: The confusion matrices for the complete classification with the maximum likelihood classifier (M-L).

M-L results for GO data trained on OH data using features ch4 - Tsurf, ch5 - Tsurf, ch3A - ch2, ch3A / ch1, ch3A / ch2

		classifier						
		WATER	SUNGLINT	LAND	SNOW	WATERCL	CIRRUS	
	WATER	126	6	6	2	6	6	82.9%
	SUNGLINT	0	21	0	0	1	1	91.3%
targets	LAND	1	0	104	0	2	5	92.9%
	SNOW	1	0	2	101	0	12	87.1%
	WATERCLOUDS	1	2	2	0	92	18	80.0%
	CIRRUS	7	0	2	4	24	49	57.0%
		92.6%	72.4%	89.7%	94.4%	73.6%	53.8%	

Correctly classified 81.6% targets

M-L results for OH data trained on GO data using features ch 1, ch 2, ch 3A, ch4 - Tsurf, ch5 - Tsurf

		classifier						
		WATER	SUNGLINT	LAND	SNOW	WATERCL	CIRRUS	
	WATER	30	0	0	0	0	0	100.0%
	SUNGLINT	5	40	1	0	5	0	78.4%
targets	LAND	0	0	33	1	1	0	94.3%
	SNOW	0	0	0	79	0	0	100.0%
	WATERCLOUDS	2	0	3	1	194	26	85.8%
	CIRRUS	1	0	0	9	17	97	78.2%
		78.9%	100.0%	89.2%	87.8%	89.4%	78.9%	

Correctly classified 86.8% targets

A Defining the Channel 3B reflectance

The monochromatic reflectivity (or reflectance) ρ_λ is the ratio of the reflected (backscattered) radiance to the incident radiance. In the case of solar reflection one can write:

$$\rho_\lambda = \frac{L_\lambda}{\mu_0 L_{\lambda 0}}$$

where L_λ is the measured radiance, $L_{\lambda 0}$ is the incoming solar radiance, and μ_0 is the cosine of the solar zenith angle θ_0.

Assuming the solar radiance is independent of direction, the equation for the reflectance can be written in terms of the solar flux $F_{\lambda 0}$:

$$\rho_\lambda = \frac{L_\lambda}{\frac{1}{\pi}\mu_0 F_{\lambda 0}}$$

For channel 3B the outgoing radiance is due to solar reflection and thermal emission. Thus in order to determine a channel 3B reflectance it is necessary to substract the thermal part from the satellite signal. To do this the temperature of the observed object is needed. The best candidate available is the channel 4 brightness temperature, since most objects behave approximately as blackbodies in this spectral interval.

The channel 3 reflectance, ρ_3 for simplicity, can thus be written as

$$\rho_3 = \frac{L_3 - \varepsilon_3 \int_0^\infty \Phi_3(\lambda) B_\lambda(Tb_4) d\lambda}{\frac{1}{\pi}\mu_0 F_{30}} \tag{5}$$

where L_3 is the measured radiance in channel 3B, $\Phi_3(\lambda)$ is the channel 3 spectral response function, B_λ is the Planck function, and Tb_4 is the channel 4 brightness temperature.

If the observed object is optically thick (transmittance = 0)

$$\varepsilon_3 = 1 - \rho_3$$

and using this in equation 5 one gets:

$$\rho_3 = \frac{L_3 - \int_0^\infty \Phi_3(\lambda) B_\lambda(Tb_4) d\lambda}{\frac{1}{\pi}\mu_0 F_{30} - \int_0^\infty \Phi_3(\lambda) B_\lambda(Tb_4) d\lambda} \tag{6}$$

If the satellite observation is given in brightness temperaure, it is of course possible to derive the corresponding radiance by folding the spectral response with the Planck function:

$$L_3 = \int_0^\infty \Phi_3(\lambda) B_\lambda(Tb_3) d\lambda$$

Determination of the solar flux

Values for the channel 3B solar flux were derived from tabulated values of "reflected solar radiance by a perfect diffusive reflector" published on the Internet by Bryan A. Baum (`http://asd-www.larc.nasa.gov/~baum/`)

For NOAA 12 the derivation looks like this:

$$F_{30} = \pi w_3 \; 5.213 \frac{mW}{m^2 \; sr \; cm}$$

where w_3 is the equivalent width of the spectral response function. Since the width is needed in wavenumbers ($k = 1/\lambda$) we have

$$w_3 = \int_0^\infty \Phi_3(k) dk = \int_0^\infty \frac{\Phi_3(\lambda)}{\lambda^2} d\lambda$$

Thus for NOAA 12 $F_{30} = 4.4303 \; W/m^2$.