

**DANISH METEOROLOGICAL INSTITUTE**  
**SCIENTIFIC REPORT**

**02 - 08**

**SAR surface cover classification using distribution  
matching**

**July 2002**

**Rashpal S. Gill**



**Copenhagen 2002**

**ISSN 0905-3263**  
**ISBN 87-7478-460-9**

# PREFACE

Discrimination between open water and sea ice in SAR imagery can still pose a problem to the ice analysts in their daily task of charting the sea ice for safe navigation. To help them in this task, new algorithms have been developed and tested. The algorithms that are described in this report rely on a user first manually identifying a particular region in a SAR image (e.g., open water area or sea ice of particular concentration or ice type) then the algorithm(s) will automatically determine similar regions in the remainder of an image. These algorithms are based on a common principal of matching the statistics of the known and unknown regions using (a) Kolmogorov-Smirnov (KS), and (b) Chi-Square (CS) distribution matching tests. The main advantage in using these distribution matching tests is that the probability distribution function (pdf) of the known region does not need to be known. Both KS and CS tests determine whether the two data sets belong to the same or different, yet undetermined, distributions. The main difference between KS and CS tests is that they are valid for un-binned and binned data respectively.

In this report the relative performance of the KS and CS tests is presented. The tests were carried out using the amplitude SAR image and the image products: (a) Power-to-Mean Ratio (PMR), and (b) Gamma-pdf which are computed from it. Both PMR and Gamma-pdf are useful tools for discriminating between open water and sea ice type in SAR images and have been reported in the open literature by the author. The results presented in this report shows that the KS test is very efficient (both reliable and computationally fast) at identifying similar surface types. It performed best with the amplitude data and Gamma-pdf while results using the Gamma-pdf and PMR images were prone to ambiguities. CS test did not perform as well as the KS test. This is because the data first has to be arbitrarily binned which results in some information being inevitably lost. It was also found to be many times slower to run on the computer. For these reasons it was decided not to use the CS test for matching known and unknown regions in a SAR image.

The information obtained using the KS tests can be considered as the ‘best statistical guess’ during situations when the ice analysts have difficulty in interpreting parts of a SAR image.

**Keyword:** Sea ice, RADARSAT, image interpretation, distribution matching, Kolmogorov-Smirnov test, Chi-Square test, Greenland.

## List of contents

1.	INTRODUCTION .....	3
2.	DISTRIBUTION MATCHING TESTS .....	6
2.1	Kolmogorov – Smirnov test .....	6
2.2	Chi-Square test.....	7
3.	EVALUATION METHOD .....	9
4.	RESULTS .....	11
4.1	RADARSAT image from West Coast of Greenland .....	11
4.2	RADARSAT image from East Coast of Greenland.....	13
5.	DISCUSSION and CONCLUSIONS .....	14
6.	REFERENCES .....	15
7.	FIGURES .....	16
8.	DANISH METEOROLOGICAL INSTITUTE Scientific Reports.....	32

# 1. INTRODUCTION

The Danish Meteorological Institute (DMI) is responsible for the operational charting of the sea ice in the waters around Greenland for the safety of ship navigation. Between 1996 - 1998 extensive validation campaigns were undertaken which included numerous aerial underflights to determine the quality of the RADARSAT ScanSAR Wide data for ice mapping in different ice and weather conditions (Gill et. al., 2000). Based on these evaluations operational ice mapping using primarily RADARSAT ScanSAR Wide data was started in autumn 1998. Today DMI is the third largest commercial user of RADARSAT data in the world after the Canadian Ice Services and the National Ice Centre.

The interpretation of these data is made difficult by the fact that the backscatter signatures from the open water and ice regions are not unique (Sandven et. a., 1994). The backscatter from the water regions are dominated by the local wind conditions and those from the ice regions are dependent on the ice type, ice concentration, surface roughness of the individual floes and level of surface melting during the summer months (Gill and Valeur, 1999). The backscatter signals are also critically dependent on the radar incidence angles (far - and near - range effects). For example, it is quite common during manual interpretation that the belts of ice appear nearly white (on a grey tone scale) in the far- and nearly black in the near-range against the background sea clutter. Manual interpretation of these data is particularly difficult in the navigationally most important Cape Farewell waters (the southern most tip of Greenland) which are characterised by strong winds (wind speeds  $\approx$  20 m/s - 30 m/s are common ) and scattered sea ice of low concentration ( $\approx$  1/10 - 3/10). This sea ice is mainly of arctic origin of thickness between 2 m - 5m and floe sizes typically  $<$  50 m mixed in with the locally formed ice and icebergs from the east coast of Greenland (Gill et. al., 2000). Detecting regions of ice in these data can also be difficult close to coasts where wind patterns are often complex or the relevant region is totally devoid of winds (lee areas) making image interpretation extremely difficult. It requires ice analysts which have high skills at interpreting SAR images and, in addition, have a knowledge of the local ice regime.

To provide fast and easy to interpret additional tools/products to the ice analysts that would help them to analysis SAR images during routine operations, a number of parameters based on the first and second order statistics, probability distributions, wavelet transform and constant false alarm rate have been developed at DMI in the last 6 - 7 years. A suitable method to estimate the performance of these products was found to be to display their values on a computer screen (after appropriate scaling) and use manual interpretations. This method was preferred as it was found to be more accurate (reflecting even the spread in variations of the parameter over the same region type within the different parts of the same image) than the traditional methods which usually involve making graphical or scatter plots of the parameter for different ice types or concentration. Further, using this approach it was possible to make the use of ice analyst's image interpretation skills and it enabled their involvement from a very early stage during validations. By displaying these grey tone 'images' of the parameters along side the original image it was possible for the ice analysts to evaluate the performance (and the limits) of the various parameters over different regions of sea ice and open water. The data set used for evaluation were the RADARSAT ScanSAR Wide images (from an archive pool of over 700) processed at Gatineau and West Freugh from 1996 - 2000. In particular, these images were of diverse complexity reflecting the different sea ice and weather conditions characteristic for these waters.

Results from the investigations on the first order statistics have already been reported (Gill and Valeur, 1999). These investigations showed that all first order parameters were ambiguous for really 'noisy' images (usually those containing very complex surface wind patterns). Nevertheless, it was found that the grey tone 'images' of the normalised second moment of the probability distribution, the Power-to-

Mean-Ratio (PMR), given by the simple expression

$$PMR = \frac{\langle I^2 \rangle}{\langle I \rangle^2} \quad (1)$$

where  $I$  is the intensity value, were in most cases very useful to discriminate between the different regions of ice and water and for determining the positions of possible icebergs. The higher order moments (skewness, kurtosis) were also very useful but generally did not add to the information already available from PMR. As a result of these findings when operational ice monitoring based on RADARSAT first started at DMI it was decided to produce PMR image product(s) as part of the overall productions in its image analysis and ice charting system. Currently, PMR grey images are routinely used by the ice analysts during ice charting operations as supplement to the original contrast enhanced RADARSAT images.

The grey tone ‘images’ of the Gamma probability distribution for open water - sea ice discrimination was also investigated by the author (Gill, 2001). It was found that these grey tone images also contain useful information for sea ice and open water discrimination and supplement the PMR ‘images’. Gamma probability distribution was found to be especially useful for detecting calm water regions.

Briefly, the Gamma distribution is valid for a homogeneous region and describes the statistics of the background speckle when there is no texture variations. Earlier studies by the author (Gill and Valeur, 1999) was found that the first order moments averaged over large regions of open water, fast ice and ice of concentration between 4/10 - 9/10 matched well with the theoretical moments given by the K-pdf. Further the values for open water, followed by fast ice approached the theoretical values given by the Gamma ( $\gamma$ ) pdf. Regions of low concentration of sea ice and very turbulent open water areas, both of which are very heterogeneous (large PMR values) could not be represented by any of the pdf tested in the study.

Gamma probability distribution,  $P_g(I)$ , for single pixel intensity is given by the following mathematical expression

$$P_g(I) = \left( \frac{L}{m_b} \right)^L \frac{I^{L-1}}{\Gamma(L)} \exp \frac{-LI}{m_b} \quad (2)$$

where  $L$  is the number of looks ( $L=7$  for ScanSAR Wide image product),  $\Gamma$  is the Gamma function and  $m_b$  is the mean intensity of the background and is given by

$$m_b = \langle I \rangle \quad (3)$$

The Gamma probability distribution for the average intensity,  $\bar{I}$  can be derived by performing incoherent averaging over  $m$  pixels and is given by the following equation (Oliver and Quegan, 1998):

$$P_g(\bar{I}) = \left(\frac{mL}{\mathbf{m}_B}\right)^{mL} \frac{\bar{I}^{(mL-1)}}{\Gamma(mL)} \exp\left(-\frac{mL\bar{I}}{\mathbf{m}_B}\right) \quad (4)$$

This article essentially consists of three sections. In the next section (2) a brief description of the Kolmogorov-Smirnov (KS) and Chi-Square (CS) tests is given. The method used to evaluate the two tests is presented in section (3). The results from the evaluations of the two tests are presented in section 4. Finally, section 5 concerns with discussion and conclusions.

## 2. DISTRIBUTION MATCHING TESTS

The algorithms reported in here are given in the “Numerical Recipes in C by Press et. al., Second Edition, University of Cambridge Press, pages 620 – 627. A brief description of the two tests; (a) Kolmogorov-Smirnov and (b) Chi-Square, is given below. For detail description the reader should refer to the above reference. As mentioned above the main difference between KS and CS tests is that they are valid for binned and un-binned data, respectively. What this means is that in the KS case there is no restriction or demand put on the data sets to be compared; they can be integers or floats and in any order (the ordering is carried out within the algorithm). However this is not the case with CS test, here data must be binned which means some sort of arbitrary histograms of the data must first be created. This inevitably results in loss of information. The binning of the test data sets is critically important in the evaluations that are reported in the following sections. In particular, for the amplitude averaged products this is not important as the amplitude values are naturally binned and are in the range 0.0 – 255.0 floating points as they are computed from the original 8-bit raw RADARSAT ScanSAR Wide data by averaging pixels. However, in the case of PMR and Gamma-pdf it is not so simple. PMR values are typically in the range 1.0 – 3.0, while Gamma-pdf probabilities lies between 0.0 – 1.0. There is clearly not adequate dynamic range in PMR and Gamma-pdf values that reflect the small but very significant changes in the local texture in a given SAR image. To overcome this problem what is done is not to ‘work’ with integer or 8-bit approximations but instead to use the full floating point values of these parameters. In particular, the following amplifications to the PMR and Gamma-pdf (GAM) values are made when binning:

$$PMR_{\text{binned}} = 10^6 (PMR_{\text{original}} - 1.0) \quad (5)$$

$$GAM_{\text{binned}} = 10^5 (GAM_{\text{original}}) \quad (6)$$

From the above two expressions it is clear that PMR and GAM values are binned into  $10^6$  and  $10^5$  different bins, respectively. These multiplications factors were obtained after extensive trials.

### 2.1 Kolmogorov – Smirnov test

The main assumption in applying the Kolmogorov – Smirnov test is that the underlying unknown probability distributions that describe the two data sets that are to be compared are a function of a single variable. What is done in this test is that the cumulative distribution of the two data sets are computed. Then the maximum of the absolute difference between the two cumulative distributions,  $D$ , is computed. The smaller  $D$  is of course the more likely that the two data sets belong to the same distribution. However, the significant level of  $D$  i.e., the value that disapprove the null hypothesis that the two data sets are from the same distribution is given by the following expression which are given in the above reference

$$P(D) = Q_{KS}(N_{\text{constant}} D) \quad (7)$$

where

$$N_{constant} = \sqrt{N_e} + 0.12 + 0.11/\sqrt{N_e} \quad (8)$$

and

$$N_e = \frac{N_1 N_2}{N_1 + N_2} \quad (9)$$

is the effective number of data points.  $N_1$  and  $N_2$  are the number of data points in data sets 1 and 2, respectively.

Q is a monotonic function and is given by the following series

$$Q_{ks}(I) = 2 \sum_{j=1}^{\infty} -1^{j-1} e^{-2j^2 I^2} \quad (10)$$

The above expressions are most sensitive near the median value and least sensitive towards the tails of the distributions. Other expressions for P(D) are also discussed in the above reference. These were also tested by the author but were not used in the routine tests as they did not show significant improvements to the overall results reported below.

## 2.2 Chi-Square test

The Chi-Square probability function is an incomplete Gamma function,  $Q(\chi^2 | \nu)$ , the expression for which can be found in any standard textbook on probabilities. For example “Numerical Recipes in C”, second edition page 221 give the following expression

$$Q(\mathbf{c}^2 | \mathbf{n}) = Q(\mathbf{n} / 2, \mathbf{c}^2 / 2) = \text{gammq}(\mathbf{n} / 2, \mathbf{c}^2 / 2) \quad (11)$$

where ‘gammq’ is a standard library function in the C library. In the above expression  $\nu$  is the number of degrees of freedom, while  $\chi^2$  is given by the following equation

$$c^2 = \sum_i \frac{(A_i \sqrt{B/A} - B_i \sqrt{A/B})}{A_i + B_i} \quad (12)$$

where

$$A = \sum_i A_i \quad \text{and} \quad B = \sum_i B_i \quad (13)$$

are the number of points in data set A and B, respectively. These are equations 14.3.3 and 14.3.4 in the above reference.

### 3. EVALUATION METHOD

To determine the usefulness of the KS and CS tests at matching known and unknown regions of SAR AMPLITUDE image and GAMMA and PMR products the following method was used.

1. From the original ScanSAR Wide 8-bit, 100 m pixel size, 500 km swath width wide, approximately 100 Mbytes amplitude data files the following three **float** products are generated:
  - AMPLITUDE product which is obtained by averaging window of size  $4 \times 4$  pixels, and fixing the distance between two of these consecutive windows at 4 pixels in both directions
  - GAMMA product obtained by using expression (4), above, the size of computation window and inter window spacing was same as for the AMPLITUDE product, and finally
  - PMR product obtained by using expression (1) , for window size of  $20 \times 20$  pixels, with inter spacing of two consecutive windows again fixed at 4 pixels in both directions.

These averaging of windows is necessary to compute the parameters used in the expressions in equation (1) and (4), to reduce the background speckle noise by smoothing the data. Further this has the advantage of reducing the data volume of each image product used in the evaluation and hence shorten the computation times.

1. Training areas of different surface types (e.g., calm and turbulent water, sea ice of low and high concentration, multi- and first- year sea ice, and if appropriate sea ice in different stages of development, etc.) are generated by displaying the AMPLITUDE product on a computer terminal (SILICON GRAPHICS ERSDAS IMAGINE work station). In particular, regions of size  $\sim 50 \times 50$  pixels of different surface types are manually identified in the AMPLITUDE image. This creates the mask file, example of which is shown in figure 1 which shows the mask file superimposed on top of the AMPLITUDE image. In the figure only two regions have been manually identified: open water (green box) and sea ice (red box). In practice there is no limit to the number of different regions of open water or sea ice of different concentration or type that can be manually identified and hence used for matching tests.
2. The mask file generated above is then used to identify and store in a buffer the data points pertaining to different surface types from each of the three products: AMPLITUDE, GAMMA and PMR. For example, these are the points within the colour boxes shown in figs. 1 and 3 and 4.
3. Then KS and CS tests are used to match the data points in each of the three products to their respective data points pertaining to a different surface type to determine the probabilities of a match (probabilities =1.0) and no match (probabilities =0.0). To compute these probabilities a test window of size  $4 \times 4$  pixels, was determined by carrying out trials with windows of different sizes and was found to be both computationally optimal and statistically sufficient,, is slid across the three products. The pixels in these  $4 \times 4$  test windows are compared with the manually identified surface classes ( $\sim$  window size  $50 \times 50$  pixels) using the CS or KS

tests to determine whether the two data sets belong to the same distribution. These results are given as probabilities which are then displayed on a computer graphical screen, after appropriate scaling, for manual interpretation. It should be noted that if matching between the pre-chosen surface class and the rest of the image is carried out using the KS test, then no further special consideration of the data are necessary. However, if CS test is used then the AMPLITUDE, PMR and GAMMA float data products must first be binned into integer bins using expressions (5) and (6), respectively.

4. The probabilities computed above for each of the three products using either of the two tests are then linearly amplified so that they could be displayed on a 8-bit computer graphical. The grey tone values of these probabilities were then manually interpreted and are presented as results in the next section.

## 4. RESULTS

This consists of, for the purpose of illustrating the method, a classification of two RADARSAT ScanSAR Wide images shown as figures 1 and 2, respectively, and given at the back of this report. The first image is from Disco Bay region of West Greenland from 2000-05-24. This image has been chosen because it is relatively easy to interpret manually and thus is adequate for illustrating the results. The second SAR image is much more difficult to interpret. It contains many different surface types: sea ice of both very low and very high concentration, calm and turbulent open water regions and first and multi-year sea ice floes. This second image is of Scoresby Sound from the East coast of Greenland from 2000-07-22. In both of these figures typical regions of different surface types of interest for this report have been indicated by different colour boxes. For example, fig. 1 contains just two boxes, red and blue boxes for sea ice and open water, respectively. Similarly, fig. 2 contain 4 boxes, two each for sea ice (low and high ice concentration) and open water (calm and turbulent water). The data points from three different products (AMPLITUDE, GAMMA and PMR) within these (and only these !) boxes are stored in a buffer and are used in the distribution matching tests as described in the last section. The results from these two distribution matching tests are presented below.

### 4.1 RADARSAT image from West Coast of Greenland

Figures 3 and 4 show the GAMMA and PMR ‘image’ products, respectively, on a grey tone scale, after appropriate scaling, corresponding to the AMPLITUDE image product shown in figure 1. Both of these figures shows the usefulness of these products for interpreting the original SAR image (fig. 1). As can be seen from fig. 3, the GAMMA product is very efficient at identifying regions of calm water, including those within the ice pack, and fast ice (Gill, 2001). Similarly, the PMR product is good at detecting sea ice boundaries, icebergs and also at delineating the structure, such as ice ridges, within ice floes.

Figure 5 shows the histogram plots of the AMPLITUDE (top row; figs. 5a and 5b), GAMMA (middle row; figs. 5c and 5d) and PMR (bottom row; figs. 5e and 5f) data points for sea ice and open water regions found within the red and blue square boxes in figs. 1, 3 and 4. To make these plots the GAMMA and PMR float data pixels were binned as described in section 2 above (equations (5) and (6)). As can be seen from each pair of these curves, they appear distinctively different for sea ice and open water. The KS and CS techniques are thus exploiting this information to locate regions with similar data distributions.

Figure 6 shows a pair of images showing the probabilities of a distribution match computed using the KS test from the AMPLITUDE float product shown in fig. 1 for sea ice (fig. 6a) and open water (fig. 6b), respectively. In the figures the probabilities have been linearly scaled by a factor of  $10^9$ . From fig. 6a it can be seen from the high grey values that the amplitude pixels from the ice regions matches well with the manually located (those within the red box), thus indicating that they belong to the same distribution. On the other hand, low grey values in the figure indicates that the amplitude pixels from these regions (mainly in the open water and land regions) does not belong to the same distribution as sea ice pixels found in the red box. Similarly, fig. 6b shows the results from the KS test for determining how well the amplitude image pixels matches to those from the water region indicated by the blue box. Here high probabilities of a distribution match are obtained for part of the water region. Low values are obtained for not only the sea ice and land regions in the image, but more significantly, also for some part of

the water region which have been indicated on this figure with the text 'Also water'. From this it can be concluded that the water pixels from the latter regions does not belong to those water pixels within the blue box. One possible interpretation of this could be that the water in these regions have a different 'backscatter state' which physically could be turbulent water compared to the relatively calm water found within the blue box.

Figure 7 is same as fig. 6 except in this case the probabilities are computed using the CS test. In the figures the probabilities have been linearly scaled by a factor of  $10^7$ . As can be seen, figs. 7a and 7b are very similar to figs. 6a and 6b, respectively. However, a closer examination reveals that the CS test results presented in fig. 7a shows that some of the open water regions in the Disco Bay, indicated on the figure with the text 'Also water', are misclassified as sea ice. Further, the ice edge in fig. 7a appears more diffused than it in fig. 6a. Similarly, comparing figs. 7b with 6b it can be seen from the grey values that there is less details in the sea ice regions in fig. 7b than there is in fig. 6b. The reason why CS test appears to perform not as well as the KS test is obvious; in the CS test the data are arbitrarily binned which result in some information being inevitably lost while in the KS test the original floating point data are used.

Figure 8 (8a and 8b) and 9 (9a and 9b) shows the KS and CS probabilities computed from the GAMMA product shown in fig. 3. In the figures the KS and CS probabilities have been linearly scaled by a factor of  $10^5$  and 2.5, respectively. These figs. 8 and 9 correspond to figs. 6 and 7 which were computed from the amplitude SAR image shown in fig. 1. A comparison of figs. 6 with 8 and 7 with 9 shows that they are very similar to each other with the exception of figure 9b which shows the CS probabilities computed from the GAMMA product for water. In this latter figure most of the sea ice region, indicated on the figure with the word 'misclassified', also has very high CS probabilities which clearly is incorrect. From this it appears that CS distribution matching test using the GAMMA product for the water pixels performs poorly and should be used with care.

Figure 10 (10a and 10b) and 11 (11a and 11b) shows the KS and CS probabilities computed from the PMR values shown in fig. 4. In the figures the KS and CS probabilities have been linearly scaled by a factor of  $10^{10}$  and 10.0, respectively. These figs. 10 and 11 correspond to the figs. pairs 6 and 7 and 8 and 9 discussed above. Again a comparison between figs. 10 and 11 with figs. 6 and 7 shows that they are very similar with the exception of fig 11a which shows the CS probabilities computed using the PMR product for sea ice matching. In the latter figure there is much noise and the sea ice and open water regions are poorly separated. A open water region in the figure, indicated on the figure with the word 'misclassified', is clearly misclassified as sea ice. Thus again it appears that CS distribution matching test performs rather poorly.

From the example presented in figures 6 – 11 it is clear that KS distribution matching test is much more successful at locating similar surface types than the CS test. Even in the cases when the CS test performs well it, however, does not appear to contain any more information that already is not available from the KS test. This is true for all 3 image products used in the tests. The reason for the poor performance of the CS test are clear and have been outlined above. Furthermore, the CS algorithm was found to be 4 - 5 times slower than the KS algorithm, which is a very important parameter in any operational environment such as the national sea ice charting services. For these reasons it was decided not to use the CS test in any further evaluation of this technique.

## 4.2 RADARSAT image from East Coast of Greenland

Figure 12 and 13 shows the GAMMA probabilities and PMR values on a grey scale. These figures correspond to the AMPLITUDE image from the East coast of Greenland shown in fig. 2.

Figure 14 and 15 shows the histogram plots of the binned floating point values of the AMPLITUDE, GAMMA and PMR products for sea ice and open water found within the coloured boxes indicated in the figs. 2, 12 and 13. In particular, the curves shown in the left columns of figures 14 and 15 are of sea ice of high and low concentration, while those in the right columns are for turbulent and calm water, respectively. As mentioned above these histograms represent the data points within the different colour boxes shown in figs. 2, 12 and 13.

Figure 16 (16a and 16b) shows the KS probabilities computed for sea of high (fig. 16a) and low (fig. 16b) concentration. As can be seen from fig. 16a that KS distribution matching is successful at locating the high ice concentration region (white regions in the figure off the coast). Similarly, it can be seen from fig. 16b that it is also reasonably good at matching the region consisting of sea ice of low concentration. However, in this case it is also misclassifying a region of open water as sea ice of low concentration. This region of open water is the edge of a water front separating regions of calm and turbulent waters and is marked on fig. 16b with the text 'Not sea ice'. As mentioned above this image is relatively complex and it is encouraging that the KS test has been able to locate most of the sea ice regions.

Figure 17 (17a and 17b) shows the KS probabilities matched for calm and turbulent water using the GAMMA probability product. Here again most of the regions containing calm and turbulent water have been successfully located. However, again there are few regions, indicated in the figures as 'misclassified' are clearly not correctly classified.

Finally figure 18 (18a and 18b) shows the KS probabilities computed using the PMR float values for calm and turbulent water, respectively. As can be seen from these two figures they appear almost identical. Thus the KS test with the PMR data does not appear to perform too well in discriminating between calm and turbulent water regions. One possible explanation for this is that the PMR float product used in the test is computed for a window of size  $20 \times 20$  pixels while to compute the GAMMA and AMPLITUDE windows of size only  $4 \times 4$  pixels are used. Large windows are required for the PMR product to reduce the background speckle noise. However, these large windows for PMR results in the smoothing of the data which in turn could result in the probability distributions for the different surface types being very similar. Nevertheless KS test with the PMR data is still reasonably successful at locating the water and sea ice regions and the information in these figures supplement the information in figures 16 and 17, from the AMPLITUDE and GAMMA products.

Thus given the texture complexity of this SAR image, the results presented in figs. 16 – 18 are encouraging. Overall, the KS distribution matching test has been judged to be reasonably good at locating regions of similar surface types. This conclusion is based on the trials carried out with this algorithm with a much larger RADARSAT data set including the images from the Cape Farewell region.

## 5. DISCUSSION and CONCLUSIONS

In this report the technique of matching the data from an unknown region with that of previously identified areas were explored to determine if useful information could be obtained about the former. Two type of data matching tests were examined; (a) Kolmogorov-Smirnov and (b) Chi-Square. The first test, KS, works on the original floating point data while the CS test is used with binned data. To illustrate the methods two RADARSAT images, one from the West coast and the other from the East coast of Greenland were used. The data sets used in the tests were the floating point AMPLITUDE, GAMMA and PMR products. These 3 data products were derived from the original raw 8-bit amplitude RADARSAT data by averaging windows of size  $4 \times 4$  pixels in the case of AMPLITUDE and GAMMA probabilities and windows of size  $20 \times 20$  pixels for the PMR product. The full details are given in section 2 above.

The results of the KS and CS tests were presented as ‘images’ which are nothing more than the respective probabilities on a grey tone scale. This method of presenting the results was chosen as it lends easily to manual interpretation. The evaluation of the results shows that both KS and CS distribution matching techniques contain some very useful information at matching the unknown surface regions to known surface types, for example sea ice and open water. In particular, this type of information is very useful during operational ice charting where even the most experienced ice analyst can sometime be in doubt in interpreting part of a SAR image. The information available from the KS and CS tests can thus be considered as a ‘best statistical guess’.

Concerning the relative performance of the KS and CS tests, it was found that KS test was far more superior than the CS test at locating surfaces of same types. In terms of computer execution times KS test was found to be 4 - 5 times faster than the CS test. Further for the cases where the CS and KS tests performance was similar, it was found that the CS test did not appear to contain any ‘new’ information that was not already available from the KS test. For these reasons it was concluded not to use the CS test for further evaluations.

However, it must be pointed out that the Kolmogorov-Smirnov test has its own limitations. These are listed below.

1. KS test is most sensitive around the median of the cumulative distribution function,  $P(x)$ , it is less sensitive at the tail end of the distribution where  $P(x)$  is approaching 0 or 1. In the C-recipe book, modified version of the KS test are proposed, these were tested by author and results were not too different from the original when displayed on a grey tone scale.
2. KS test cannot discriminate between all types of distributions, such as a distribution with 2 maximums.

In the subsequent reports it will be reported how the information obtained using the KS test on the AMPLITUDE, GAMMA and PMR products is used in the fuzzy logic rules called Multi Experts – Multi Criteria Decision Making (ME-MCDK) to classify a SAR image both semi-automatically and fully automatically i.e., without the need for any operator supervision or input.

## 6. REFERENCES

Gill R. S., (2001). "Sea Ice Edge and Icebergs Detection using routine operations". Canadian J. REMOTE SENSING, special issue on Sea Ice and Icebergs, ivol. 27, no. 5, pp 411 – 432.

Gill R. S., Rosengreen M. K: and Valeur H. H. (2000). "Operational Ice Mapping with RADARSAT for ship navigation in the Greenlandic Waters", Canadian Journal of Remote Sensing, vol. 26, No. 2, pp. 121 - 132.

Gill R. S. and Valeur H. H. (1999). "Ice cover discrimination in the Greenland waters using first-order texture parameters of ERS SAR images", Int. J. Remote Sensing, vol. 20, no. 2, pp. 373 - 385.

Oliver C. J. and S. Quegan. (1998). "Understanding Synthetic Aperture Radar Images", book publ. Artech House.

Sandven S., O. M. Johannessen, K. Kloster and M. Miles. (1994) "SIZE'92 ERS-1 SAR validation experiment", EARSeL Advances in Remote Sensing, vol., 3, No. 2, pp 50 -56.

## **7. FIGURES**

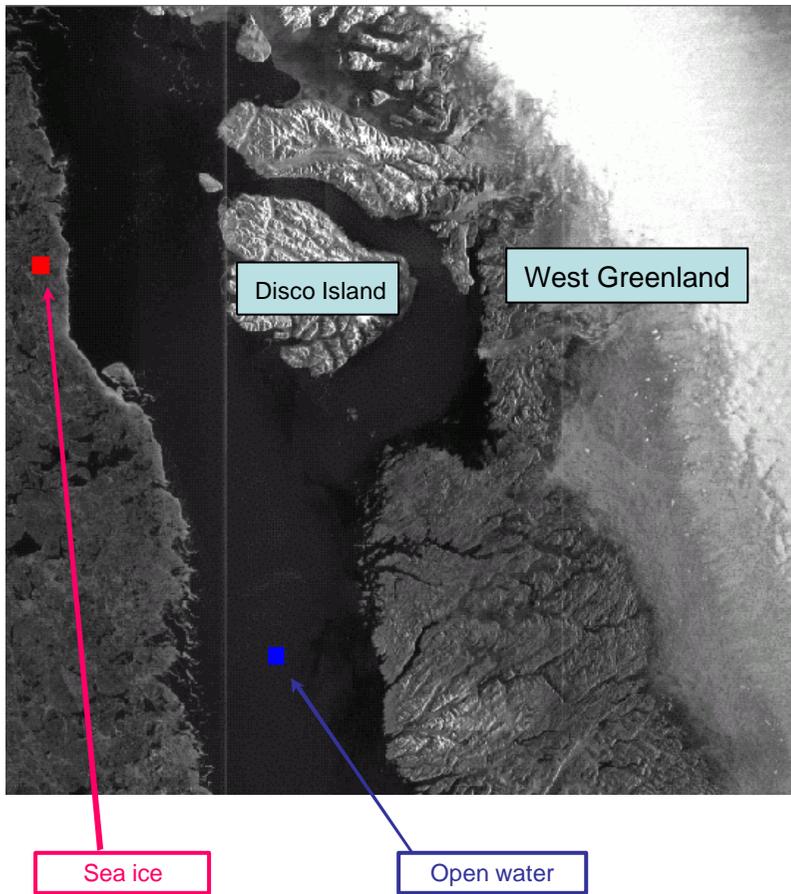


Figure 1. 4\*4 pixel averaged RADARSAT AMPLITUDE image from West Greenland from 2000-05-24.

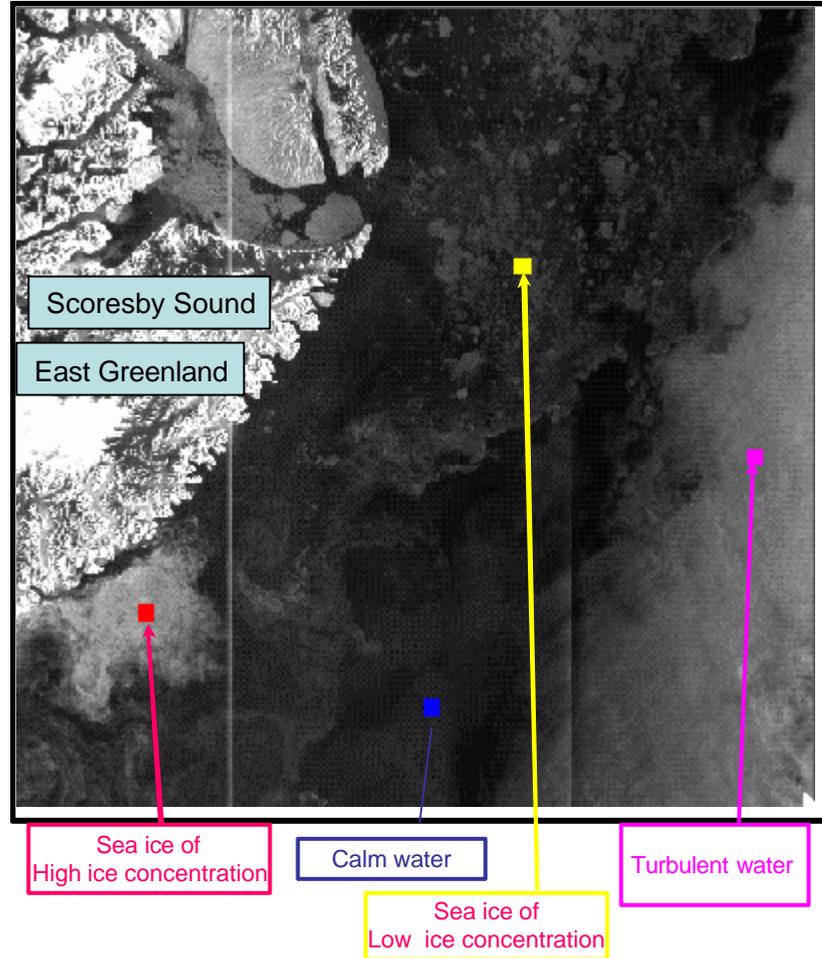


Figure 2. 4\*4 pixel averaged RADARSAT AMPLITUDE image from East Greenland from 2000-07-22.

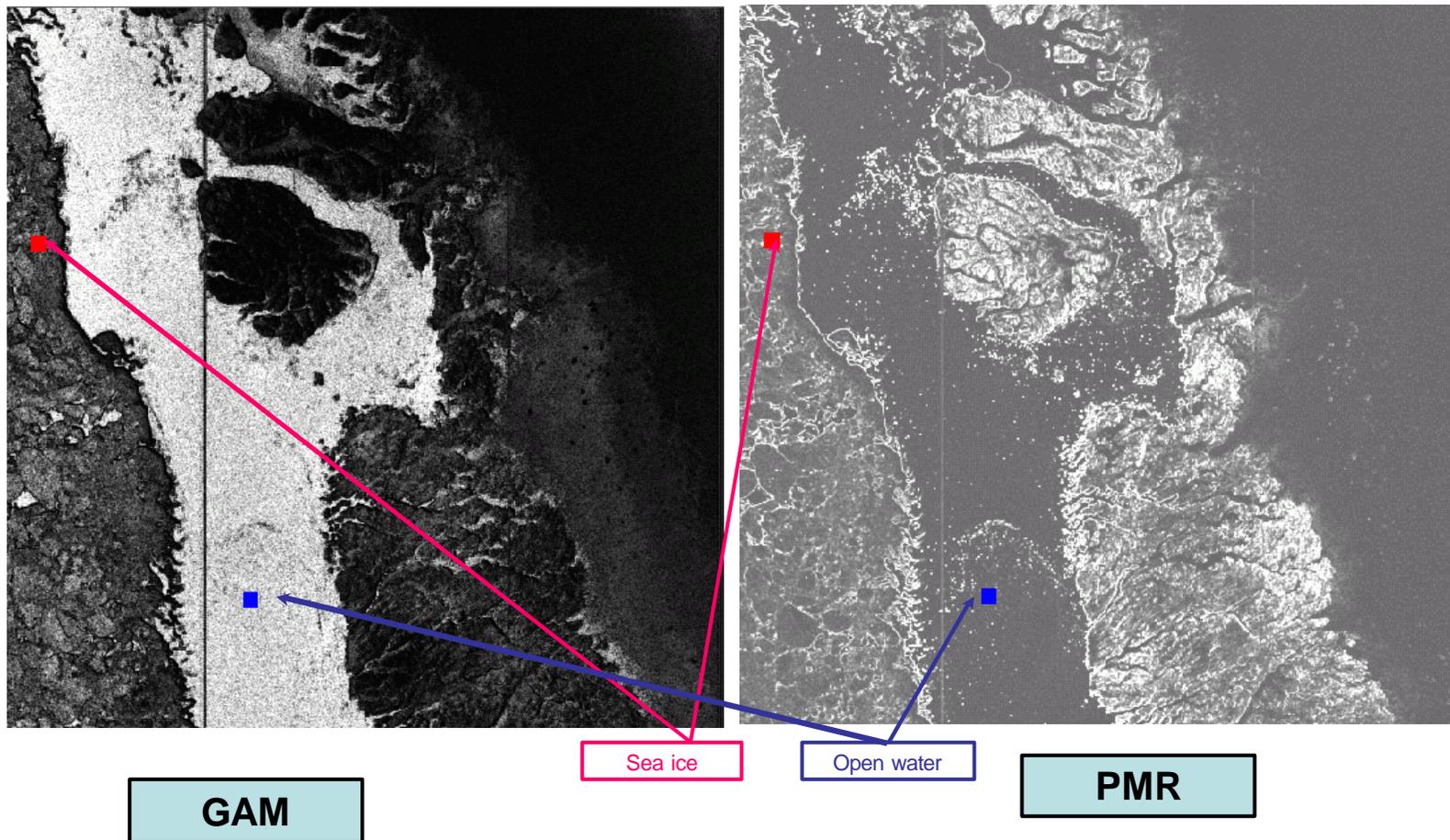
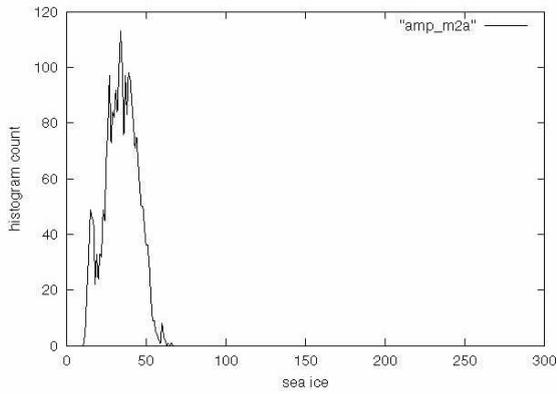
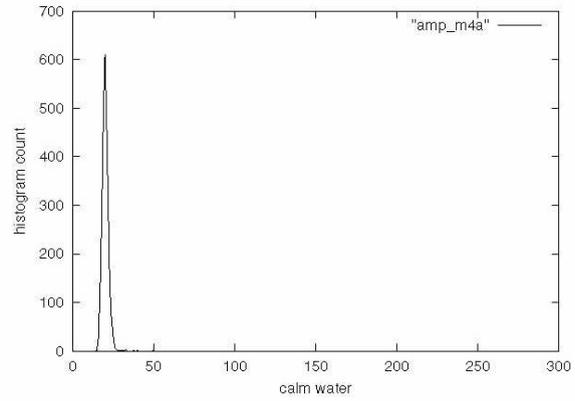


Figure 3. GAMMA product corresponding to fig.1 above shown on a grey level scale. Gamma probabilities have been multiplied by a factor of  $10^5$ .

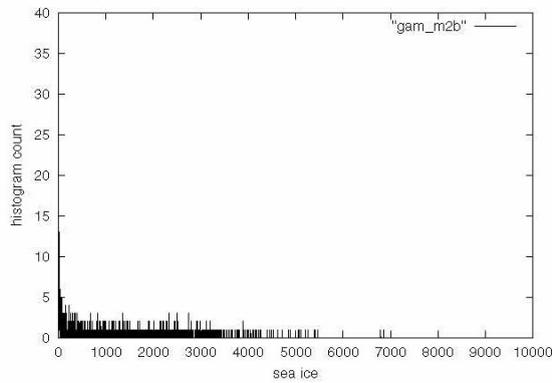
Figure 4. PMR product corresponding to fig.1 above shown on a grey level scale. PMR values have been multiplied by a factor of 100.



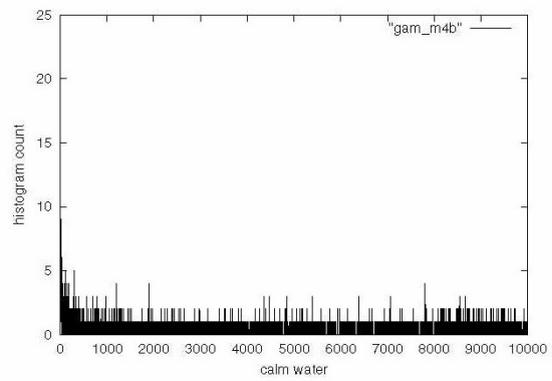
**Fig. 5a. AMPLITUDE – SEA ICE**



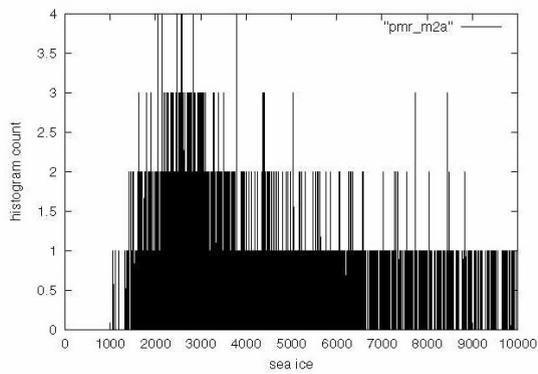
**Fig. 5b. AMPLITUDE – WATER**



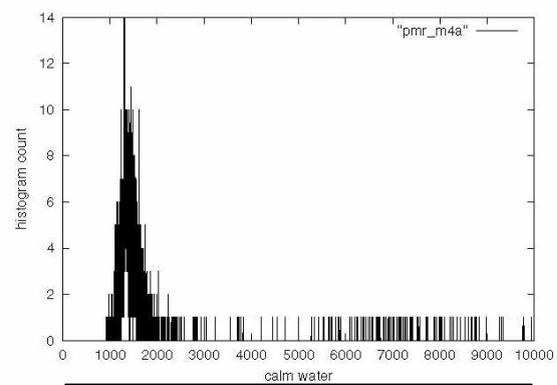
**Fig. 5c. GAMMA – SEA ICE**



**Fig. 5d. GAMMA – WATER**



**Fig. 5e. PMR – SEA ICE**



**Fig. 5f. PMR – WATER**

**Figure 5. The curves above show the binned data points of sea ice (left column) and water (right column) of the AMPLITUDE (top row), GAMMA (middle row) and PMR (bottom row) products indicated by red and blue boxes in figs. 1, 3 and 4, respectively, from 2000-05-24.**

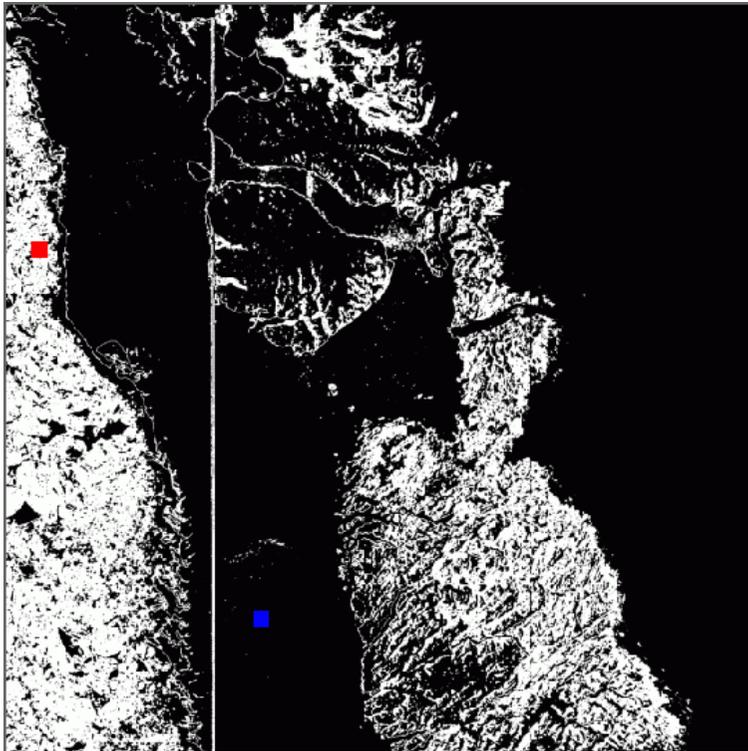


Fig. 6a. AMPLITUDE: KS probs. for sea ice

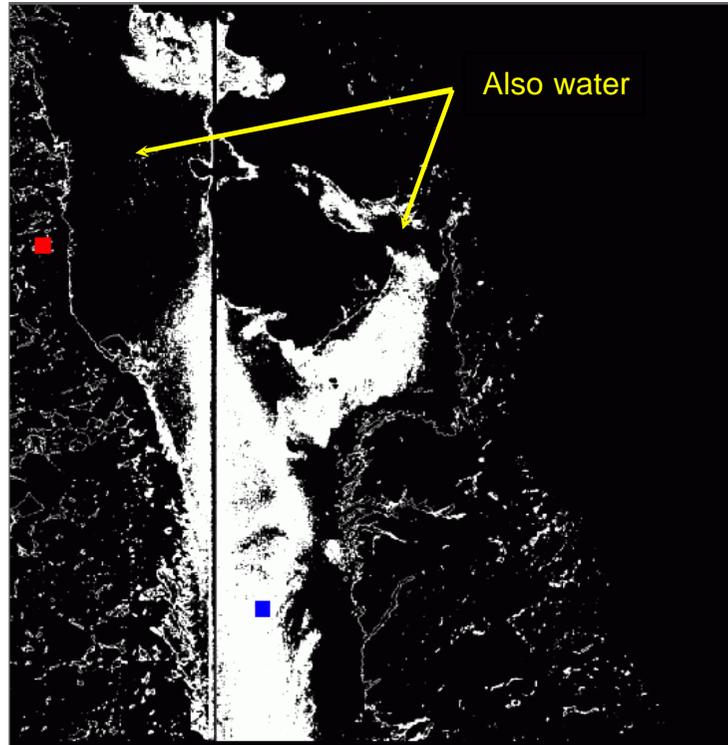


Fig. 6b. AMPLITUDE: KS probs. for water

Figure 6. The Kolmogorov-Smirnov probabilities computed by matching the AMPLITUDE pixels for sea ice in the red box (left – fig. 6a) and for open water in the blue box (right- fig. 6b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa.

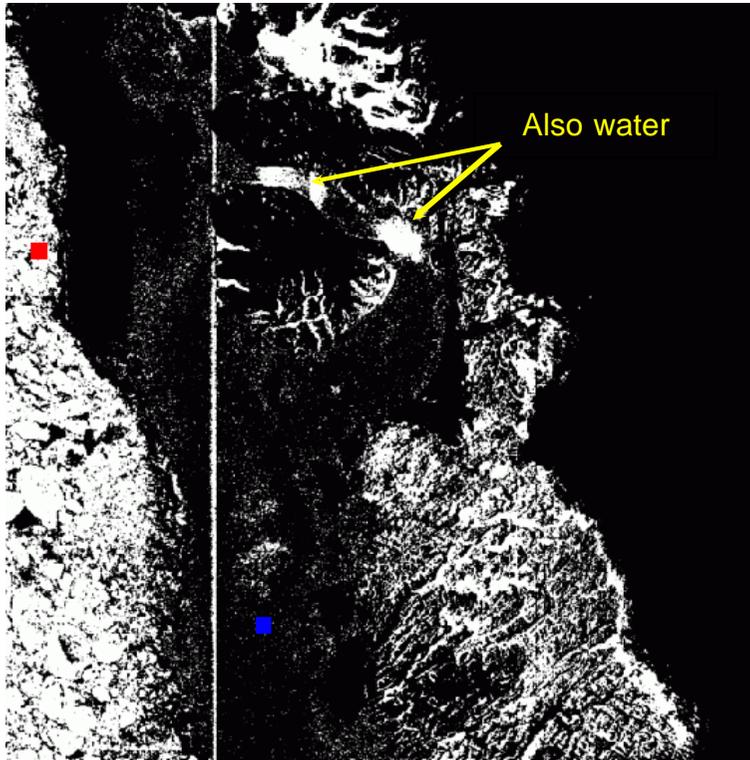


Fig. 7a. AMPLITUDE: Chi-Square probs. for sea ice

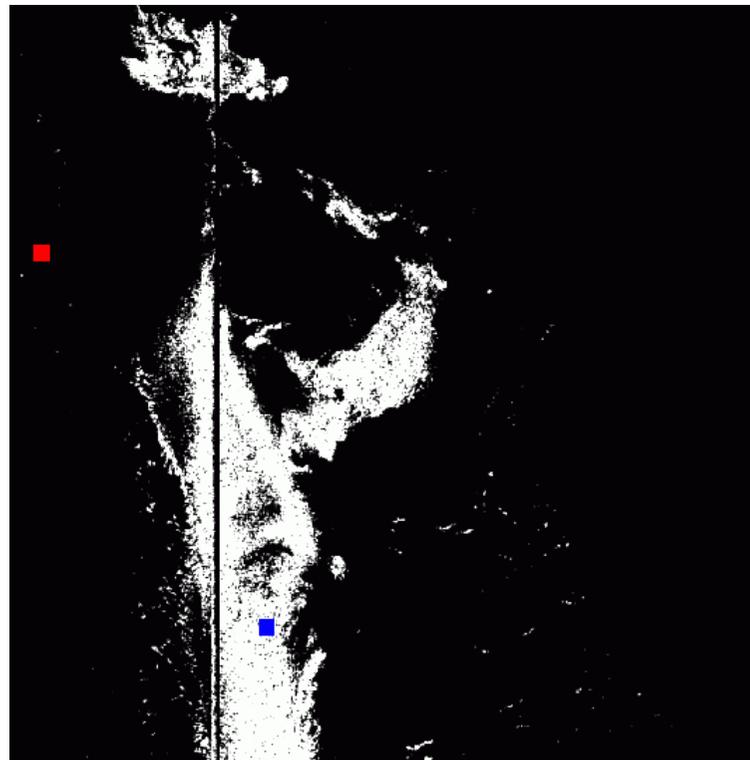
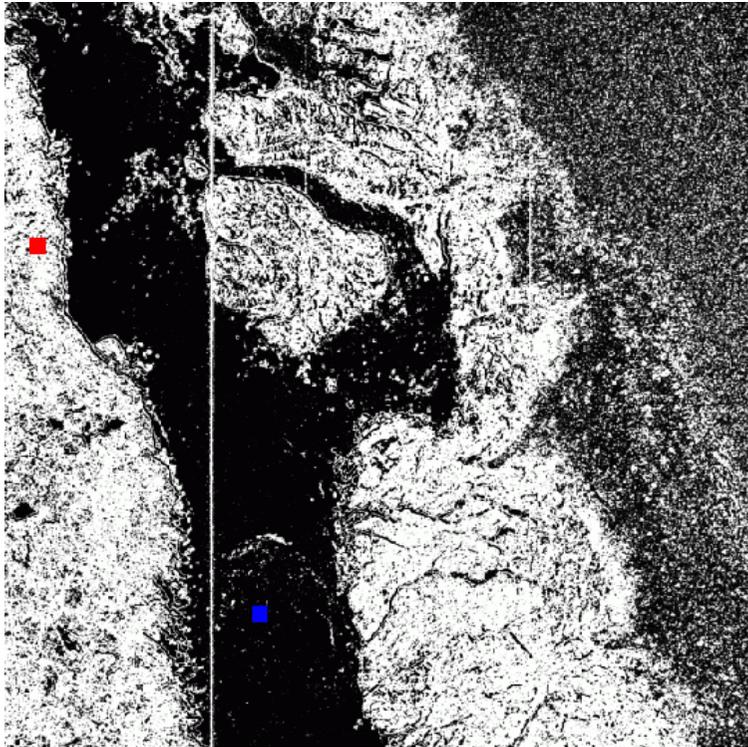
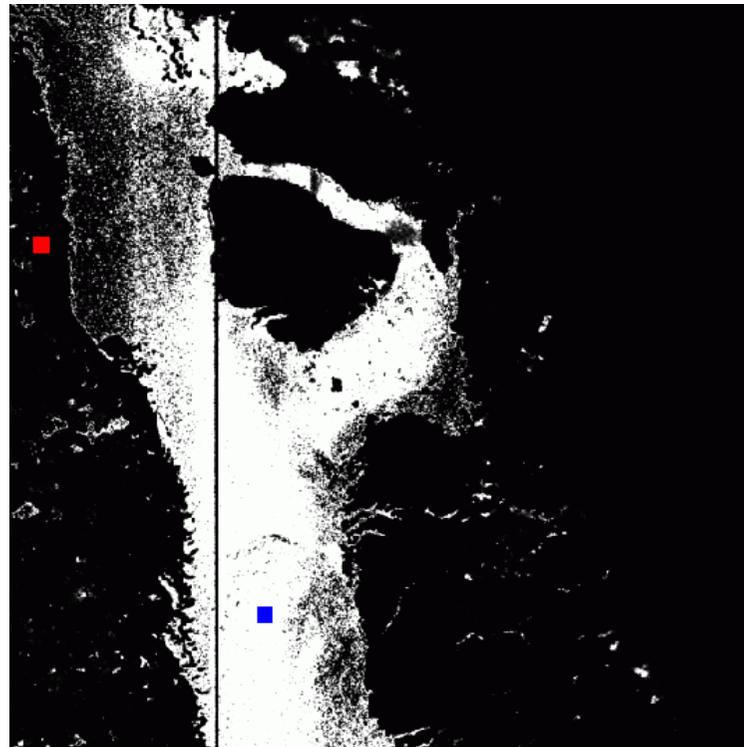


Fig. 7b. AMPLITUDE: Chi-Square probs. for water

Figure 7. The Chi-Square probabilities computed by matching the AMPLITUDE pixels for sea ice in the red box (left – fig. 7a) and for open water in the blue box (right- fig. 7b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa.



**Fig. 8a. GAMMA: KS probs. for sea ice**



**Fig. 8b. GAMMA: KS probs. for water**

**Figure 8. The Kolmogorov-Smirnov probabilities computed by matching the GAMMA pixels for sea ice in the red box (left – fig. 8a) and for open water in the blue box (right- fig. 8b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa.**

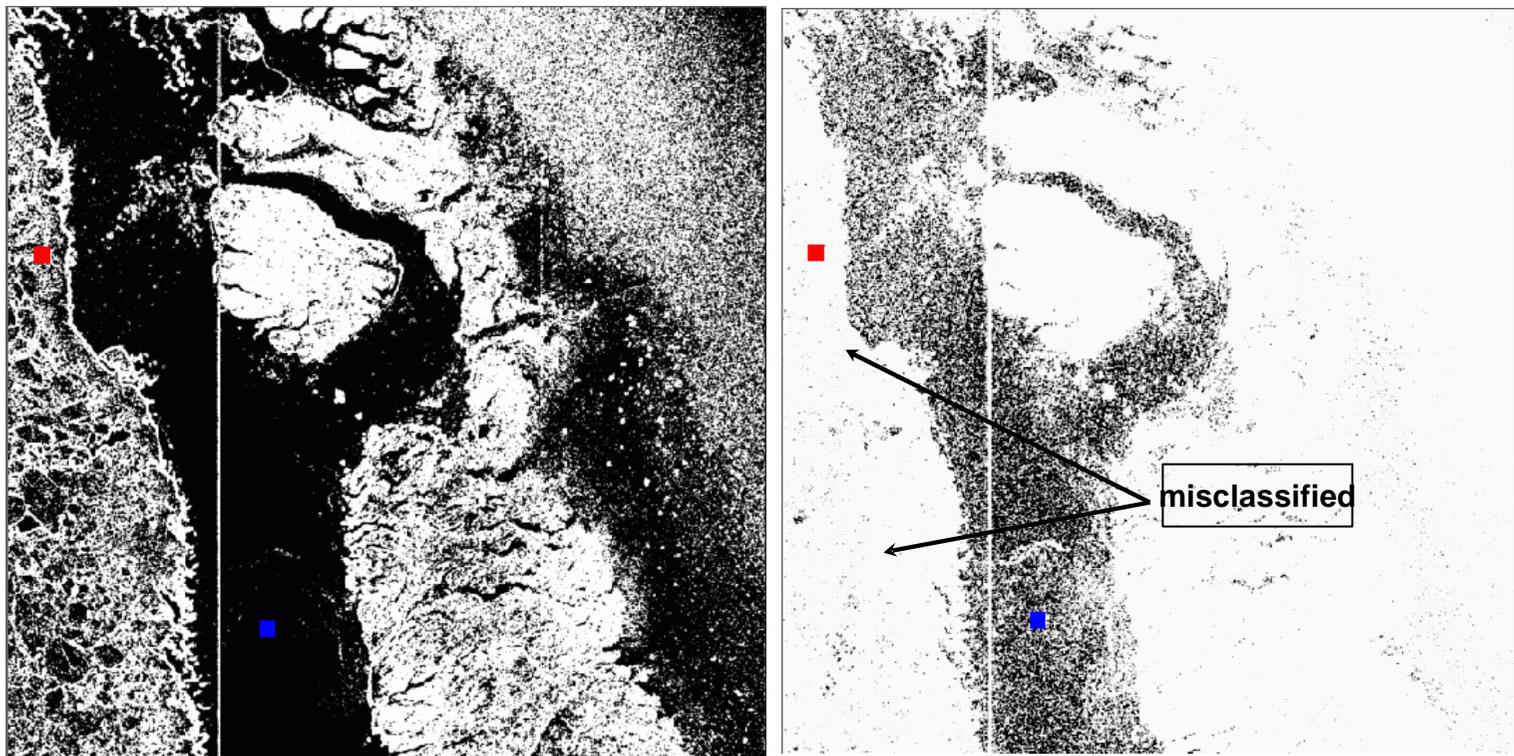


Fig. 9a. GAMMA: Chi-Square probs. for sea ice

Fig. 9b. GAMMA: Chi-Square probs. for water

Figure 9. The Chi-Square probabilities computed by matching the GAMMA pixels for sea ice in the red box (left – fig. 9a) and for open water in the blue box (right- fig. 9b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa.

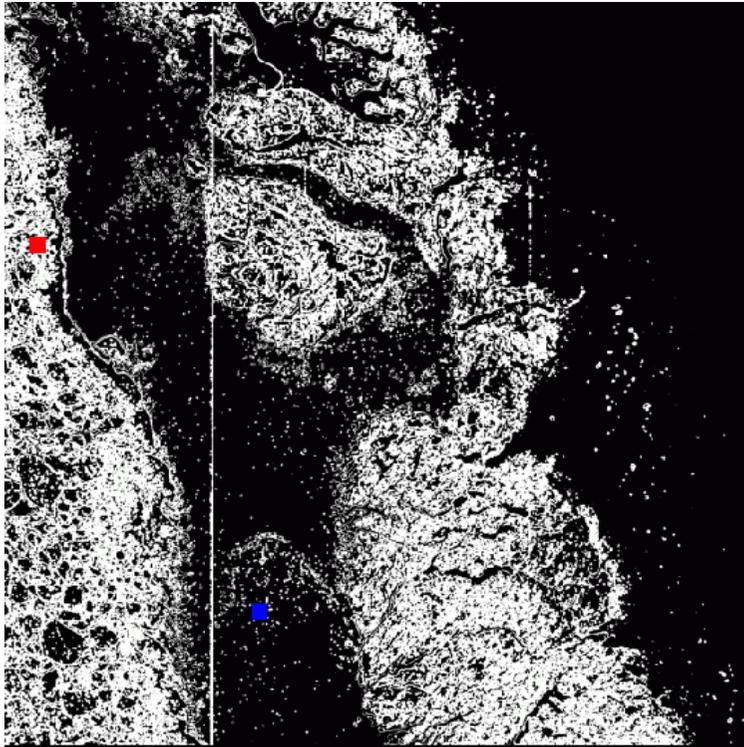


Fig. 10a. PMR: KS probs. for sea ice

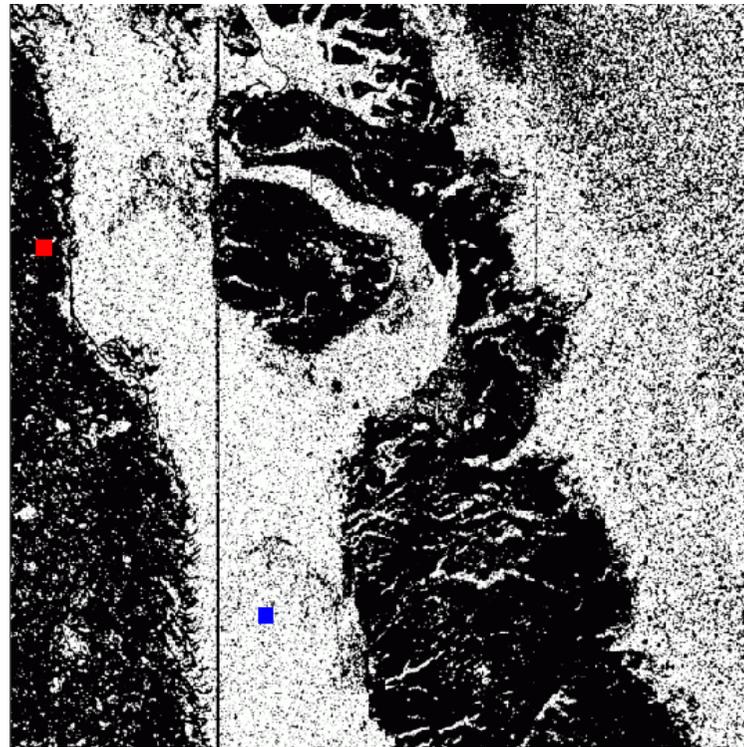
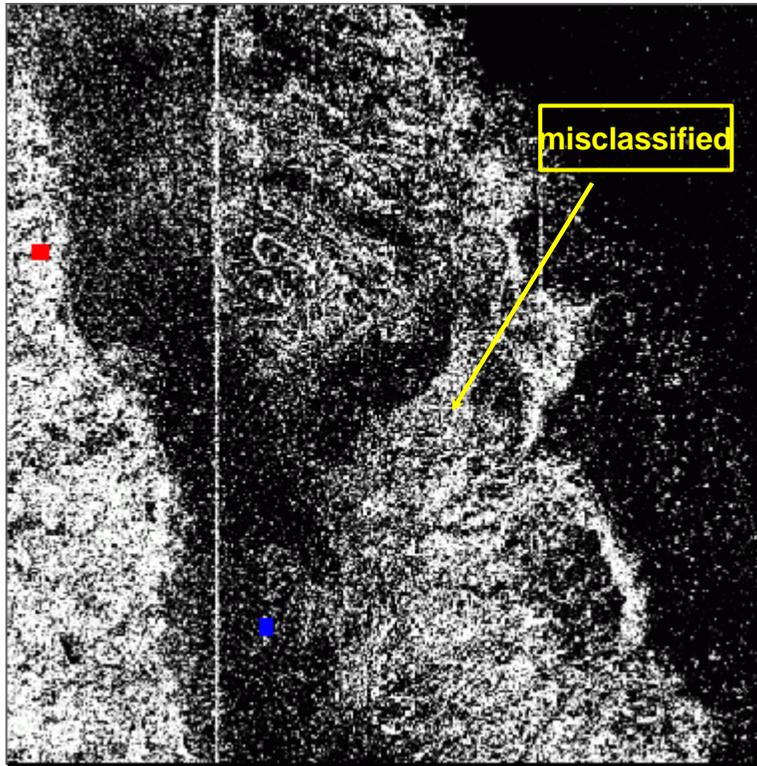
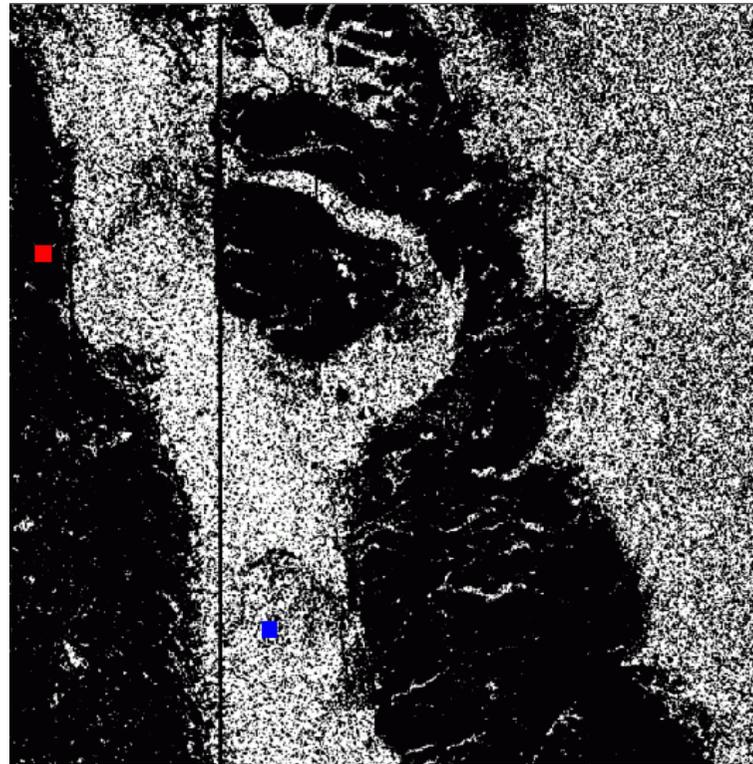


Fig. 10b. PMR: KS probs. for water

Figure 10. The Kolmogorov-Smirnov probabilities computed by matching the PMR pixels for sea ice in the red box (left – fig. 10a) and for open water in the blue box (right- fig. 10b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa.



**Fig. 11a. PMR: Chi-Square probs. for sea ice**



**Fig. 11b. PMR: Chi-Square probs. for water**

**Figure 11. The Chi-Square probabilities computed by matching the PMR pixels for sea ice in the red box (left – fig. 11a) and for open water in the blue box (right- fig. 11b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa.**

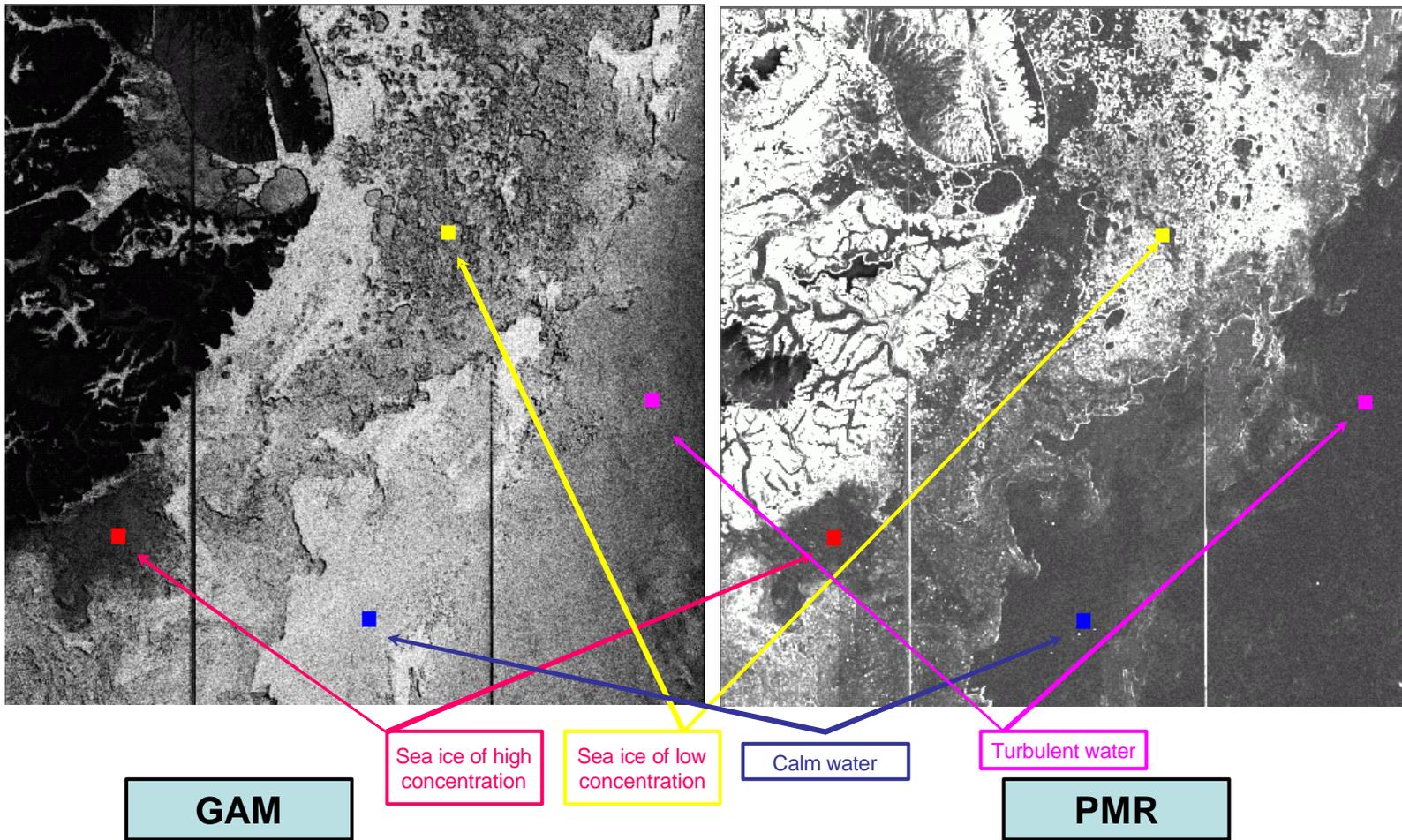
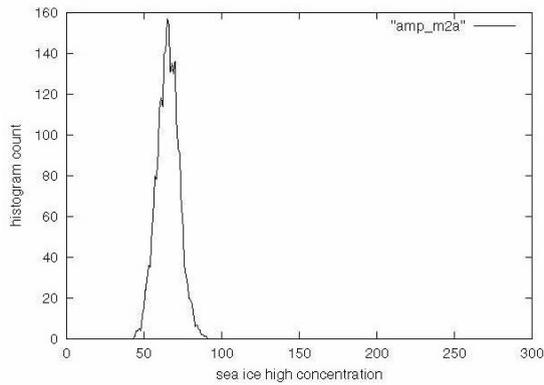
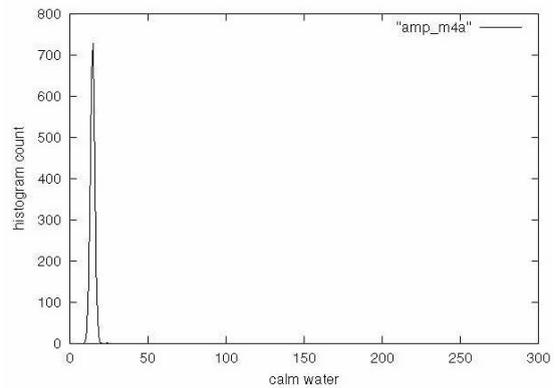


Figure 12. GAMMA product corresponding to fig.2 above shown on a grey level scale. Gamma probabilities have been multiplied by a factor of  $10^5$ .

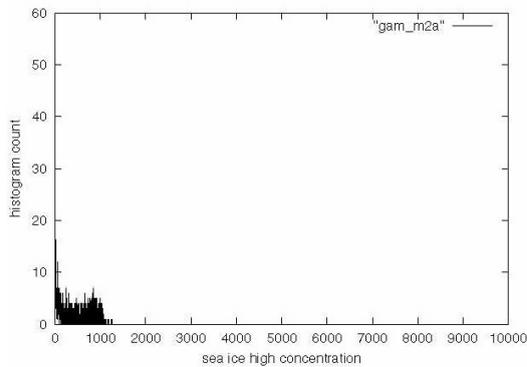
Figure 13. PMR product corresponding to fig.2 above shown on a grey level scale. PMR values have been multiplied by a factor of 100.



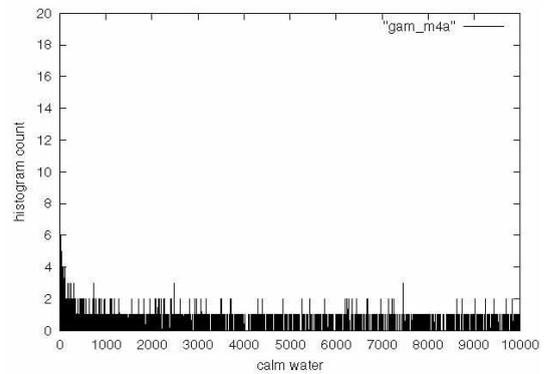
**Fig. 14a. AMPLITUDE – high conc. SEA ICE**



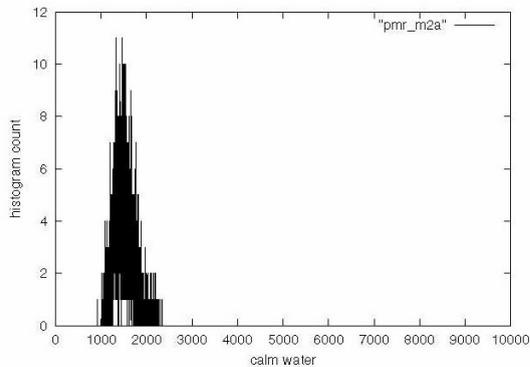
**Fig. 14b. AMPLITUDE – calm WATER**



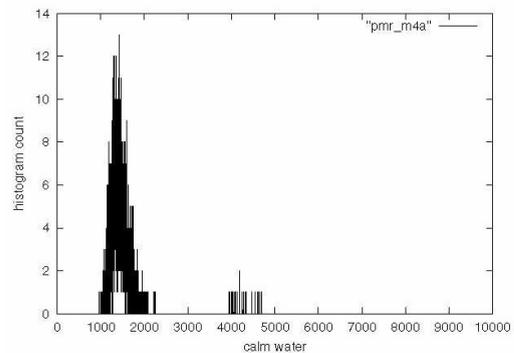
**Fig. 14c. GAMMA – high conc. SEA ICE**



**Fig. 14d. GAMMA – calm WATER**

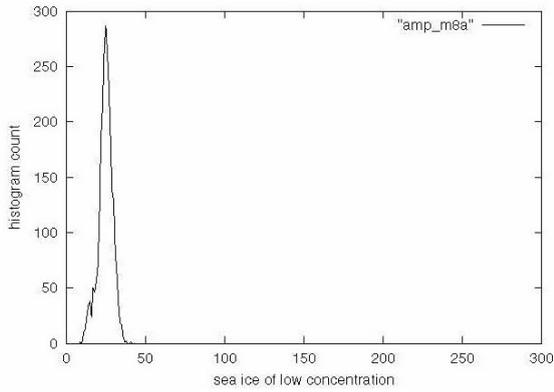


**Fig. 14e. PMR – high conc. SEA ICE**

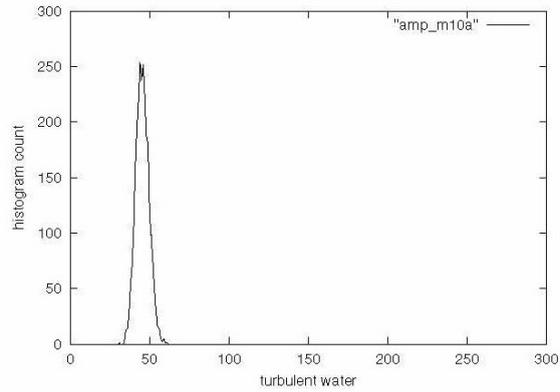


**Fig. 14f. PMR – calm WATER**

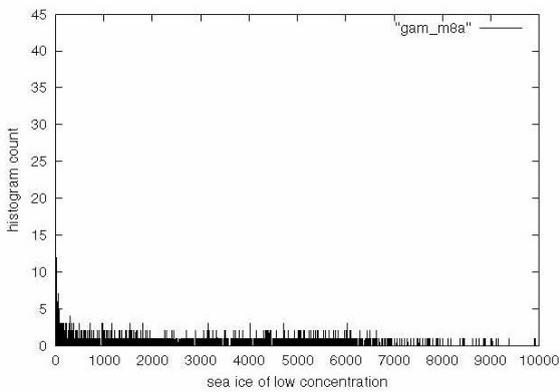
**Figure 14. The curves above show the binned data points of sea ice of high concentration (left column) and calm water (right column) of the AMPLITUDE (top row), GAMMA (middle row) and PMR (bottom row) products indicated by red and blue boxes in figs. 2, 12 and 13, respectively, from 2000-07-22.**



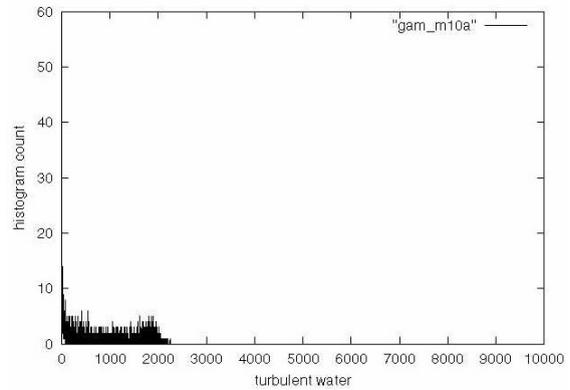
**Fig. 15a. AMPLITUDE – low conc. SEA ICE**



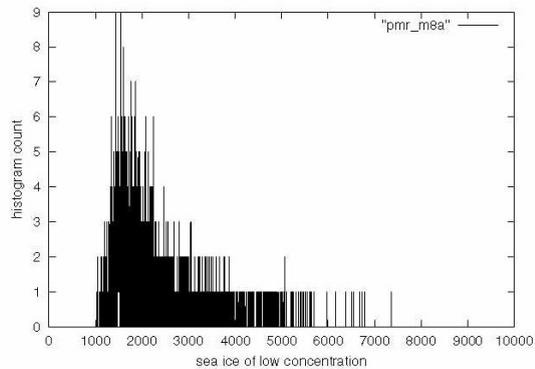
**Fig. 15b. AMPLITUDE – turbulent WATER**



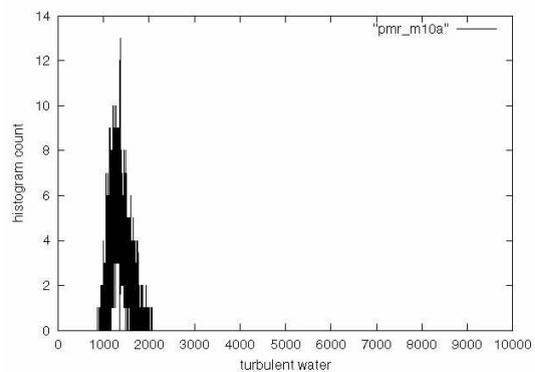
**Fig. 15c. GAMMA – low conc. SEA ICE**



**Fig. 15d. GAMMA – turbulent WATER**



**Fig. 15e. PMR – low conc. SEA ICE**



**Fig. 15f. PMR – turbulent WATER**

**Figure 15. The curves above show the binned data points of sea ice of low concentration (left column) and turbulent water (right column) of the AMPLITUDE (top row), GAMMA (middle row) and PMR (bottom row) products indicated by yellow and magenta boxes in figs. 2, 12 and 13, respectively, from 2000-07-22.**

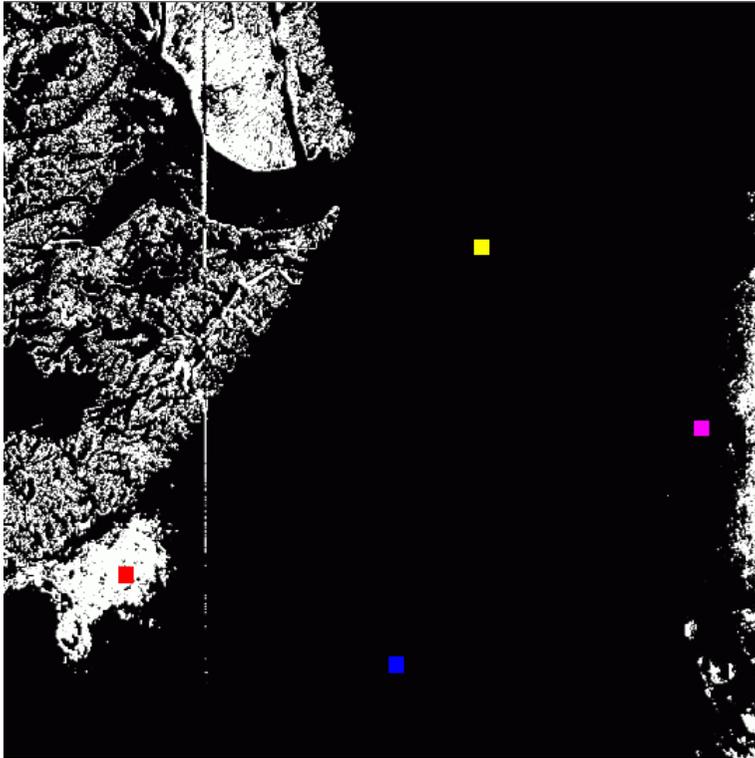


Fig. 16a. AMPLITUDE: KS probs. for sea ice of high concentration.

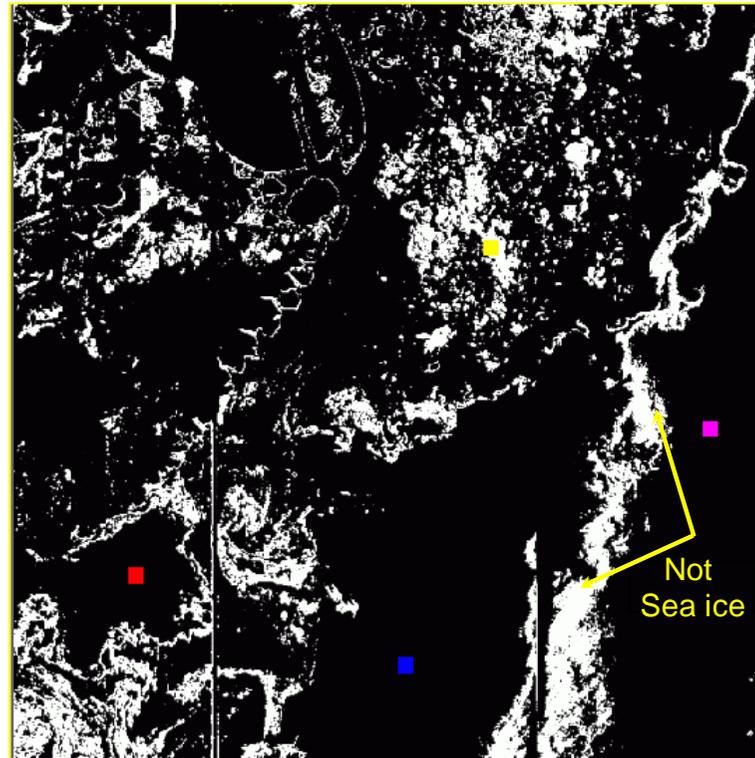


Fig. 16b. AMPLITUDE: KS probs. for sea ice of low concentration.

Figure 16. The Kolmogorov-Smirnov probabilities computed by matching the AMPLITUDE pixels for sea ice of high concentration in the red box (left – fig. 16a) and of low concentration in the yellow box (right- fig. 16b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa. The corresponding AMPLITUDE image is given in fig. 2.

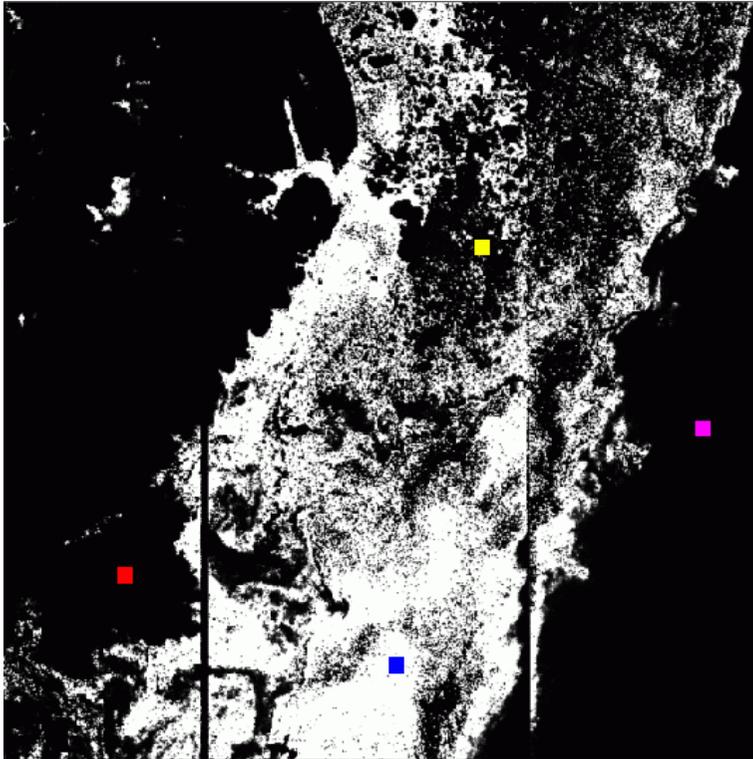


Fig. 17a. GAMMA: KS probs. for calm water.

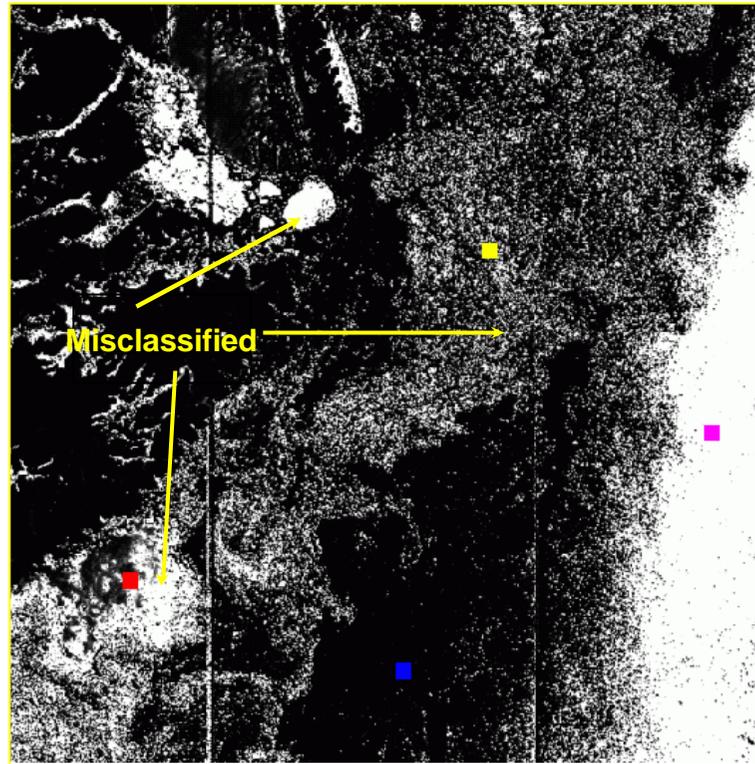


Fig. 17b. GAMMA: KS probs. for turbulent water.

Figure 17. The Kolmogorov-Smirnov probabilities computed by matching the GAMMA pixels for calm water in the blue box (left – fig. 17a) and turbulent water in the magenta box (right- fig. 17b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa. The corresponding AMPLITUDE image is given in fig. 2.

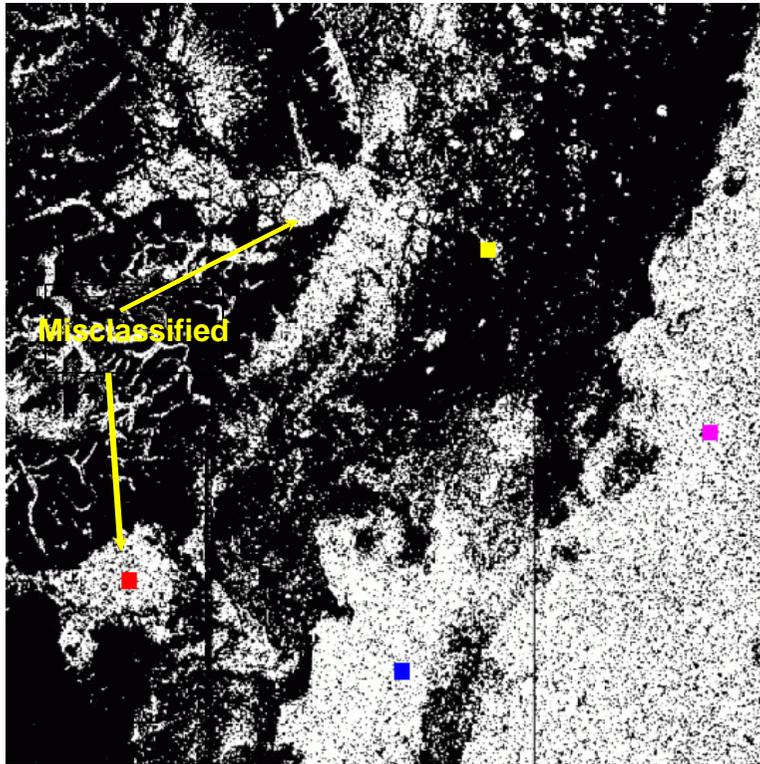


Fig. 18a. PMR: KS probs. for calm water.

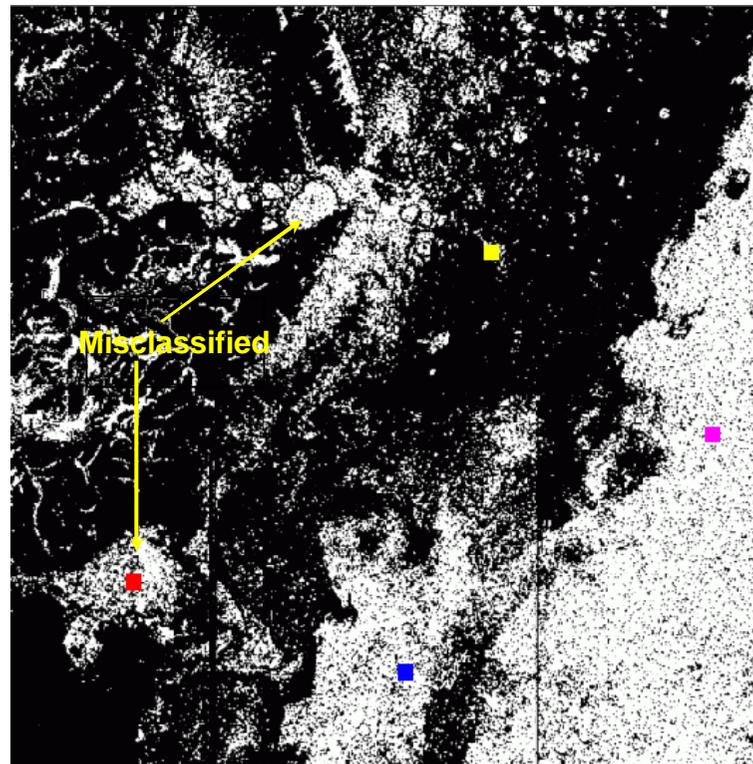


Fig. 18b. PMR: KS probs. for turbulent water.

Figure 18. The Kolmogorov-Smirnov probabilities computed by matching the PMR pixels for calm water in the blue box (left – fig. 18a) and turbulent water in the magenta box (right- fig. 18b) with the rest of image pixels shown on a grey tone scale. High grey tone values indicate high probabilities and vice-versa. The corresponding AMPLITUDE image is given in fig. 2.

## 8. DANISH METEOROLOGICAL INSTITUTE

### Scientific Reports

Scientific reports from the Danish Meteorological Institute cover a variety of geophysical fields, i.e. meteorology (including climatology), oceanography, subjects on air and sea pollution, geomagnetism, solar-terrestrial physics, and physics of the middle and upper atmosphere.

Reports in the series within the last five years:

No. 97-1

**E. Friis Christensen og C. Skøtt:** Contributions from the International Science Team. The Ørsted Mission - a pre-launch compendium.

No. 97-2

**Alix Rasmussen, Sissi Kiilsholm, Jens Havskov Sørensen, Ib Steen Mikkelsen:** Analysis of tropospheric ozone measurements in Greenland: Contract No. EV5V-CT93-0318 (DG 12 DTEE): DMI's contribution to CEC Final Report Arctic Tropospheric Ozone Chemistry ARCTOC.

No. 97-3

**Peter Thejll:** A search for effects of external events on terrestrial atmospheric pressure: cosmic rays.

No. 97-4

**Peter Thejll:** A search for effects of external events on terrestrial atmospheric pressure: sector boundary crossings.

No. 97-5

**Knud Lassen:** Twentieth century retreat of sea-ice in the Greenland Sea

No. 98-1

**Niels Woetman Nielsen, Bjarne Amstrup, Jess U. Jørgensen:** HIRLAM 2.5 parallel tests at DMI: sensitivity to type of schemes for turbulence, moist processes and advection.

No. 98-2

**Per Høeg, Georg Bergeton Larsen, Hans-Henrik Benzon, Stig Syndergaard, Mette Dahl Mortensen:** The GPSOS project Algorithm functional design and analysis of ionosphere, stratosphere and troposphere observations.

No. 98-3

**Mette Dahl Mortensen, Per Høeg:** Satellite atmosphere profiling retrieval in a nonlinear troposphere Previously entitled: Limitations induced by Multipath.

No. 98-4

**Mette Dahl Mortensen, Per Høeg:** Resolution properties in atmospheric profiling with GPS.

No. 98-5

**R.S. Gill and M. K. Rosengren:** Evaluation of the Radarsat imagery for the operational mapping of sea ice around Greenland in 1997.

No. 98-6

**R.S. Gill, H.H. Valeur, P. Nielsen and K.Q. Hansen:** Using ERS SAR images in the operational mapping of sea ice in the Greenland waters: final report for ESA-ESRIN's: pilot projekt no. PP2.PP2.DK2 and 2<sup>d</sup> announcement of opportunity for the exploitation of ERS data projekt No. AO2..DK 102.

No. 98-7

**Per Høeg et al.:** GPS Atmosphere profiling methods and error assessments.

No. 98-8

**H. Svensmark, N. Woetmann Nielsen and A.M. Sempreviva:** Large scale soft and hard turbulent states of the atmosphere.

No. 98-9

**Philippe Lopez, Eigil Kaas and Annette Guldborg:** The full particle-in-cell advection scheme in spherical geometry.

No. 98-10

**H. Svensmark:** Influence of cosmic rays on earth's climate.

No. 98-11

**Peter Thejll and Henrik Svensmark:** Notes on the method of normalized multivariate regression.

No. 98-12

**K. Lassen:** Extent of sea ice in the Greenland Sea 1877-1997: an extension of DMI Scientific Report 97-5.

No. 98-13

**Niels Larsen, Alberto Adriani and Guido DiDonfrancesco:** Microphysical analysis of polar stratospheric clouds observed by lidar at McMurdo, Antarctica.

No. 98-14

**Mette Dahl Mortensen:** The back-propagation method for inversion of radio occultation data.

No. 98-15

**Xiang-Yu Huang:** Variational analysis using spatial filters.

No. 99-1

**Henrik Feddersen:** Project on prediction of climate variations on seasonal to interannual time-scales (PROVOST) EU contract ENVA4-CT95-0109: DMI contribution to the final report: Statistical analysis and post-processing of uncoupled PROVOST simulations.

No. 99-2

**Wilhelm May:** A time-slice experiment with the ECHAM4 A-GCM at high resolution: the experimental design and the assessment of climate change as compared to a greenhouse gas experiment with ECHAM4/OPYC at low resolution.

No. 99-3

**Niels Larsen et al.:** European stratospheric monitoring stations in the Arctic II: CEC Environment and Climate Programme Contract ENV4-CT95-0136. DMI Contributions to the project.

No. 99-4

**Alexander Baklanov:** Parameterisation of the deposition processes and radioactive decay: a review and some preliminary results with the DERMA model.

No. 99-5

**Mette Dahl Mortensen:** Non-linear high resolution inversion of radio occultation data.

No. 99-6

**Stig Syndergaard:** Retrieval analysis and methodologies in atmospheric limb sounding using the GNSS radio occultation technique.

No. 99-7

**Jun She, Jacob Woge Nielsen:** Operational wave forecasts over the Baltic and North Sea.

No. 99-8

**Henrik Feddersen:** Monthly temperature forecasts for Denmark - statistical or dynamical?

No. 99-9

**P. Thejll, K. Lassen:** Solar forcing of the Northern hemisphere air temperature: new data.

No. 99-10

**Torben Stockflet Jørgensen, Aksel Walløe Hansen:** Comment on "Variation of cosmic ray flux and global coverage - a missing link in solar-climate relationships" by Henrik Svensmark and Eigil Friis-Christensen.

No. 99-11

**Mette Dahl Meincke:** Inversion methods for atmospheric profiling with GPS occultations.

No. 99-12

**Hans-Henrik Benzon; Laust Olsen; Per Høeg:** Simulations of current density measurements with a Faraday Current Meter and a magnetometer.

No. 00-01

**Per Høeg; G. Leppelmeier:** ACE - Atmosphere Climate Experiment.

No. 00-02

**Per Høeg:** FACE-IT: Field-Aligned Current Experiment in the Ionosphere and Thermosphere.

No. 00-03

**Allan Gross:** Surface ozone and tropospheric chemistry with applications to regional air quality modelling. PhD thesis.

No. 00-04

**Henrik Vedel:** Conversion of WGS84 geometric heights to NWP model HIRLAM geopotential heights.

No. 00-05

**Jérôme Chenevez:** Advection experiments with DMI-Hirlam-Tracer.

No. 00-06

**Niels Larsen:** Polar stratospheric clouds micro-physical and optical models.

No. 00-07

**Alix Rasmussen:** "Uncertainty of meteorological parameters from DMI-HIRLAM".

No. 00-08

**A.L. Morozova:** Solar activity and Earth's weather. Effect of the forced atmospheric transparency changes on the troposphere temperature profile studied with atmospheric models.

No. 00-09

**Niels Larsen, Bjørn M. Knudsen, Michael Gauss, Giovanni Pitari:** Effects from high-speed civil traffic aircraft emissions on polar stratospheric clouds.

No. 00-10

**Søren Andersen:** Evaluation of SSM/I sea ice algorithms for use in the SAF on ocean and sea ice, July 2000.

No. 00-11

**Claus Petersen, Niels Woetmann Nielsen:** Diagnosis of visibility in DMI-HIRLAM.

No. 00-12

**Erik Buch:** A monograph on the physical oceanography of the Greenland waters.

No. 00-13

**M. Steffensen:** Stability indices as indicators of lightning and thunder.

No. 00-14

**Bjarne Amstrup, Kristian S. Mogensen, Xiang-Yu Huang:** Use of GPS observations in an optimum interpolation based data assimilation system.

No. 00-15

**Mads Hvid Nielsen:** Dynamisk beskrivelse og hydrografisk klassifikation af den jyske kyststrøm.

No. 00-16

**Kristian S. Mogensen, Jess U. Jørgensen, Bjarne Amstrup, Xiaohua Yang and Xiang-Yu Huang:** Towards an operational implementation of HIRLAM 3D-VAR at DMI.

No. 00-17

**Sattler, Kai; Huang, Xiang-Yu:** Structure function characteristics for 2 meter temperature and relative humidity in different horizontal resolutions.

No. 00-18

**Niels Larsen, Ib Steen Mikkelsen, Bjørn M. Knudsen m.fl.:** In-situ analysis of aerosols and gases in the polar stratosphere. A contribution to THESEO. Environment and climate research programme. Contract no. ENV4-CT97-0523. Final report.

No. 00-19

**Amstrup, Bjarne:** EUCOS observing system experiments with the DMI HIRLAM optimum interpolation analysis and forecasting system.

No. 01-01

**V.O. Papitashvili, L.I. Gromova, V.A. Popov and O. Rasmussen:** Northern polar cap magnetic activity index PCN: Effective area, universal time, seasonal, and solar cycle variations.

No. 01-02

**M.E. Gorbunov:** Radioholographic methods for processing radio occultation data in multipath regions.

No. 01-03

**Niels Woetmann Nielsen; Claus Petersen:** Calculation of wind gusts in DMI-HIRLAM

No. 01-04

**Vladimir Penenko; Alexander Baklanov:** Methods of sensitivity theory and inverse modeling for estimation of source parameter and risk/vulnerability areas.

No. 01-05

**Sergej Zilitinkevich; Alexander Baklanov; Jutta Rost; Ann-Sofi Smedman, Vasiliy Lykosov and Pierluigi Calanca:** Diagnostic and prognostic equations for the depth of the stably stratified Ekman boundary layer.

No. 01-06

**Bjarne Amstrup:** Impact of ATOVS AMSU-A radiance data in the DMI-HIRLAM 3D-Var analysis and forecasting system.

No. 01-07

**Sergej Zilitinkevich; Alexander Baklanov:** Calculation of the height of stable boundary layers in operational models.

No. 01-08

**Vibeke Huess:** Sea level variations in the North Sea – from tide gauges, altimetry and modelling.

No. 01-09

**Alexander Baklanov and Alexander Mähura:** Atmospheric transport pathways, vulnerability and possible accidental consequences from nuclear risk sites: methodology for probabilistic atmospheric studies.

No. 02-01

**Bent Hansen Sass and Claus Petersen:** Short range atmospheric forecasts using a nudging procedure to combine analyses of cloud and precipitation with a numerical forecast model.

No. 02-02

**Erik Buch:** Present oceanographic conditions in Greenland waters.

No. 02-03

**Bjørn M. Knudsen, Signe B. Andersen and Allan Gross:** Contribution of the Danish Meteorological Institute to the final report of SAMMOA. CEC contract EVK2-1999-00315: Spring-to-autumn measurements and modelling of ozone and active species.

No. 02-04

**Nicolai Kliem:** Numerical ocean and sea ice modelling: the area around Cape Farewell (Ph.D. thesis).

No. 02-05

**Niels Woetmann Nielsen:** The structure and dynamics of the atmospheric boundary layer.

No. 02-06

**Arne Skov Jensen, Hans-Henrik Benzon and Martin S. Lohmann:** A new high resolution method for processing radio occultation data.

No. 02-07

**Per Høeg and Gottfried Kirchengast:** ACE+: Atmosphere and Climate Explorer.