Sea ice classification using Bayesian statistics

Version 1.0

J.A. Verspeek

Royal Netherlands Meteorological Institute (KNMI)
De Bilt, Netherlands

2006-02-06

Summary

Using the scatterometer data from the ERS-2 satellite, a new classification algorithm for sea/ice discrimination, based on Bayesian statistics is introduced. Tests for inner consistency are described and the results discussed. The influence of spatial and temporal averaging is examined. Also a comparison with ice maps obtained by other models and measurements is made. It is compared with IFREMER ice maps for ERS and Quikscat, and with surface temperature data retrieved from ECMWF. This algorithm may be applied to the ASCAT scatterometer on board of the MetOp satellites as well.

Contents

1 Introduction

Scatterometers are active remote sensing systems whose main application is to measure the wind velocity and direction above sea. The resulting wind fields can be used as stand-alone product or for assimilation in a numerical weather prediction. Quality control on the measurements is necessary to avoid assimilation of non-valid data. In most cases this quality control is straightforward, e.g. the data is missing, or the measurement is above land where no wind information can be retrieved. In other cases the quality control is not so simple, e.g. in the situation of ice contamination. Ice has very different backscatter characteristics from water. When a substantial part of the footprint is covered by ice and the other part is water, the measurement will not give any information that is useful for wind retrieval. A sea/ice discrimination algorithm can help in the quality control procedure to flag a measurement as non-valid (for wind retrieval). Apart from this, the output of the sea/ice discrimination algorithm can be seen as a product in itself. It can be used to monitor the development of the ice edge of the Arctic and Antarctic region on a time scale of days. Also it can be used to distinguish between different types of ice, of which the most prominent are first year ice and multiyear ice. A third application is the use of the algorithm for calibration/validation of the scatterometer instrument, because ice can be a static and stable source with isotropic backscatter properties.

The differences in backscatter properties between water and ice can be used to infer discrimination criteria. Firstly, ice has isotropic backscatter properties, it doesn't matter from which azimuth angle the scatterometer looks at the ice, the backscatter will be the same for a flat surface. Water on the other hand has anisotropic backscatter properties, which fact is used to retrieve wind direction. Secondly, ice has a characteristic function of backscatter versus incidence angle (σ^0 versus θ) which also differs from water. Thirdly, ice has specific backscatter properties regarding the polarisation of the incoming and reflected beam. See e.g. [Gohin, et al (1995)], (Ezraty and Cavanié (1997)].

Traditionally people have used simple parameters representing the geophysical properties, like an anisotropy coefficient or a slope parameter. Another approach is to define a geophysical model function (GMF) that describes the measurement backscatter data as function of one or more physical parameters. This method is used for wind retrieval but can be equally well applied for ice retrieval. It has the advantage that the full information content of the measurements is used. We applied this method to ERS [Haan de, Stoffelen (2001)] and it is planned to be used for the ASCAT scatterometer. Note: Polarisation properties cannot be used for ERS and ASCAT because of the fixed polarisation of the transmit/receive antennae.

For wind the backscatter geophysical model function naturally is dominated by two parameters, being the wind speed V and wind direction ϕ. Using this assumption the well-known cone shaped geophysical model function is empirically derived. For ice it is less obvious how many physical parameters define the backscatter. It is likely to be many, because many types of ice can be defined, each with its own backscatter properties. To give some examples, ice may contain a certain amount of salt water, it may be more or less porous, deformed, covered with snow, it can

be molten en refrozen giving a different ice structure, ice shelves can be broken into peaces (pan-cakes) which are frozen together later on. Under the assumption of isotropic backscattering all (σ^0-triplets lie in the plane $\sigma^0_{fore} = \sigma^0_{aft}$ in σ^0-space) plots of measured triplets in σ^0-space show that for each node all ice points lie on a straight line. This means that ice is described with one parameter (the coordinate along the ice line).

For a fixed node (fixed θ) the ice GMF is only dependent on the scalar *a*, defining the position on the ice curve. Along the ice curve a comoving coordinate system is defined, where the coordinates are denoted (*a*, *b*, *c*). *a* is the distance along the ice curve, *b* is the distance from the plane $\sigma^0_{fore} = \sigma^0_{aft}$ and *c* is the third coordinate perpendicular to the ice curve and lying in the plane $\sigma^0_{fore} = \sigma^0_{aft}$. In practice the ice curve can be accurately approximated by a line:

$$\vec{\sigma}^{\,0}(a) = O + a\vec{e}$$

Equation 1-1

The abscissa along the ice line *a* has a physical meaning representing the type of ice. The distance to the ice line d_{ice} is now defined as:

$$d_{ice} = \sqrt{b^2 + c^2}$$

Equation 1-2

We assume σ^0 in dB here, but this is not very relevant. Firstly the origin of the ice line is determined using the mean of ice characteristic backscatter function $f_{ice}(\theta)$ Then the slope is determined using interpolation on the measurement ensemble. As a result of this procedure the mean values of *a*, *b*, and *c* will be zero. Scaling factors are applied to match the standard deviation of each parameter, so that the value of the parameters will have the same physical meaning independent of the node number. Also a scaling factor for the distance to the ice line d_{ice} is applied so that the distance to the ice line is always measured in units of the standard deviation of *b* and *c*, and becomes independent of the node number. For a detailed description of the derivation of the ice model see [Haan de, Stoffelen (2001)]. In Figure 1 the wind cone and ice line are depicted in σ^0-space together with some measurements.

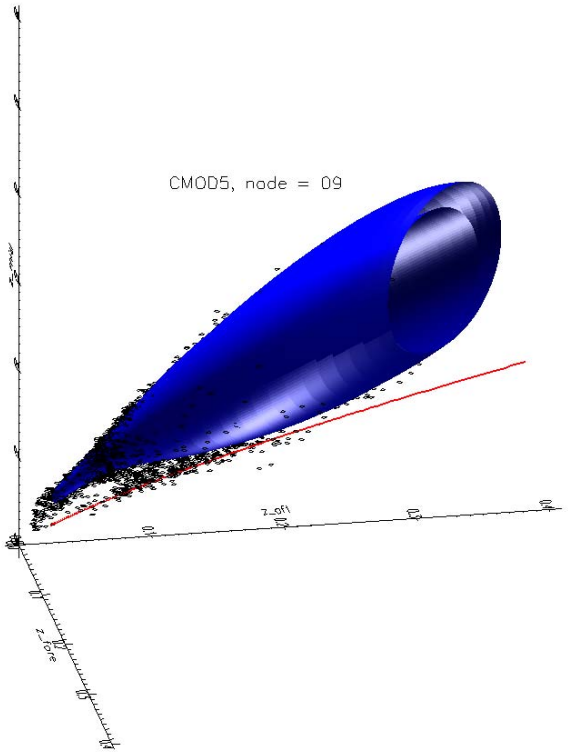

Figure 1 Wind cone (blue) and ice line (red) in z-space. Also some data points are shown (black)

2 Icemodel Description

2.1 Sea/ice classification

The sea ice classification as described in [Haan de, Stoffelen (2001)] has been extended and modified using Bayesian statistics. The original classification defines four classes for a σ^0-triplet measurement (Table 1).

Class	Description	Color
(a)	Water	Blue
(b)	Ice	Red
(c)	either Water or Ice	Orange
(d)	Neither Water nor Ice	Green

Table 1 Basic Ice/Sea lasses.

The colors in this table refer to the output bitmaps in which the classes are indicated by color. This classification is further refined using historical data and using spatial averaging, resulting in seven subclasses in total (Table 2).

Class	Description	Color
(a1)	probably Water	Purple
(a2)	Water	Blue
(b1)	probably Ice (not enough measurements)	Green
(b2)	probably Ice (SD too large)	Orange
(b3)	Ice	Grayscale
(c)	either Water or Ice	Red
(d)	neither Water nor Ice	Black

Table 2 Basic sea/ice subclasses

2.2 Bayesian Statistics

Among others, Bayesian statistics is used in decision theory, where you have a set of measurements and a number of "classes" into which the measurements can be categorised. When the a priori chance of each class is given, as well as an error model for each class, the a posteriori probability that a measurement belongs to a certain class can be calculated using Bayes' theorem. In our model we assume there are only two classes, *ice* and *water*, with a priori probabilities *P(ice)* and *P(water)*. Because there are only these two classes, the following holds true:

$$\begin{aligned} P(ice)+P(water) &= 1 \\ p(ice|\mathbf{x})+p(water|\mathbf{x}) &= 1 \end{aligned}$$

Equation 2-1

here **x** represents a number of measurements in space and time:

$$\mathbf{x} = x_0, \ldots, x_{N-1}$$

Equation 2-2

The a posteriori probability on ice $p(ice|\mathbf{x})$ is now given by Bayes' theorem [Breivik et al (2001)], [Bernardo (2003)]:

$$p(ice_0 \mid \mathbf{x}) = \frac{P(ice_0)\,p(\mathbf{x} \mid ice)}{P(ice_0)\,p(\mathbf{x} \mid ice_0) + P(water_0)\,p(\mathbf{x} \mid water_0)}$$

Equation 2-3

We assume the measurements x_i are independent of each other (naive Bayesian statistics):

$$p(\mathbf{x} \mid ice_0) = \prod_{i=0}^{N-1} p(x_i \mid ice_0)$$

Equation 2-4

Equation 2-3 now becomes:

$$p(ice_0 \mid \mathbf{x}) = \frac{P(ice_0)\prod_{i=0}^{N=1} p(x_i \mid ice_0)}{P(ice_0)\prod_{i=0}^{N=1} p(x_i \mid ice_0) + P(water_0)\prod_{i=0}^{N=1} p(x_i \mid water_0)}$$

Equation 2-5

The same equation can be written for $p(water_0|\mathbf{x})$:

$$p(water_0 \mid \mathbf{x}) = \frac{P(water_0)\prod_{i=0}^{N=1} p(x_i \mid water_0)}{P(ice_0)\prod_{i=0}^{N=1} p(x_i \mid ice_0) + P(water_0)\prod_{i=0}^{N=1} p(x_i \mid water_0)}$$

Equation 2-6

Combining the two gives:

$$\frac{p(ice_0 \mid \mathbf{x})}{p(water_0 \mid \mathbf{x})} = \frac{P(ice_0)}{P(water_0)} \prod_{i=0}^{N-1} \frac{p(x_i \mid ice_0)}{p(x_i \mid water_0)}$$

Equation 2-7

Using the logit function $Logit(p) = log(p/(1-p))$ this can be written as (ref [Wikipedia (2006)]):

$$Logit\big(p(ice_0 \mid \mathbf{x})\big) \;=\; Logit\big(P(ice_0)\big) + \sum_{i=0}^{N-1} \ln\left(\frac{p(x_i \mid ice_0)}{p(x_i \mid water_0)}\right)$$

Equation 2-8

Considering only one pixel $\boldsymbol{x} = x_0$, $N=1$:

$$Logit\big(p(ice_0 \mid x_0)\big) \;=\; Logit\big(P(ice_0)\big) + \ln\left(\frac{p(x_0 \mid ice_0)}{p(x_0 \mid water_0)}\right)$$

Equation 2-9

Afterwards a weighted spatial average $P(ice_o|x_o)$ of the a posteriori probabilities $p(ice_k|x_k)$ is taken:

$$P(ice_0 \mid x_0) \;=\; \sum_{k=0}^{N-1} p(ice_k \mid x_k)\, w_{k0}$$

$$\sum_{k=0}^{N-1} w_{k0} \;=\; 1$$

Equation 2-10

2.3 Classification Scheme

The new classification uses Bayesian statistics to calculate a certain ice probability for a point, using temporal and spatial averaging. Only valid measurements are used in the averaging. The subdivision into classes as given in Table 1 reduces to valid measurements (a+b+c) and outliers (d):

Class	Description
(a)	either Water or Ice (valid measurements)
(b)	neither Water nor Ice (outliers)

Table 3 Ice/Sea Classes

The outliers are rejected, because they are lying far from both the wind cone and the ice line and do not provide information on either sea or sea ice condition. All valid points are given a certain weight factor and used in the further calculation accordingly. This results in a (a posteriori) ice probability $p(ice|\boldsymbol{x})$ for all useful points. The six remaining subclasses of Table 2 are reduced to four subclasses:

Class	Description	Color
(α)	Water (P(ice) < 0.50)	Blue
(β_1)	Ice (P(ice) >= 0.50)	Grayscale
(β_2)	Ice but SD too large	Orange
(γ)	Not enough measurements	Green

Table 4 Ice/Sea Subclasses

2.4 Space-time averaging

It takes some days before the ERS-2 scatterometer has covered the whole of the Arctic region and Antarctica. Also in the center nodes of the swath, the discriminative power of the measurements is low because the ice line and wind cone coincide there. Therefore in order to calculate a complete ice map it is necessary to introduce spatial and temporal averaging of the measurements. This has the advantage that a complete coverage of the region is achieved, and the ice map is updated whenever the ERS satellite is passing by. A disadvantage is that the information is smeared out over space and time so a trade off has to be made in choosing the right parameters for reducing spatial and temporal resolution.

2.4.1 Spatial averaging

Spatial averaging is done using a 5x5 weight matrix. The measurement under consideration is the central pixel of the matrix. The weights are determined using a decay length constant L in pixel distance units.

$$\begin{aligned} W_s(r) &= \exp(-r/L) \quad ,0 \le r \le B \\ W_s(r) &= 0 \qquad\qquad r > B \\ r &= \sqrt{x^2 + y^2} \end{aligned}$$

Equation 2-11

For the spatial averaging a 5x5 matrix is used giving the weighting factors for the pixels surrounding the "central" pixel, the pixel under consideration. In the original sea ice classification only a 0 or 1 could be used for an element in the matrix, in the new sea ice classification any number may be used. The matrix can be either completely specified by the user, or it's elements can be calculated from a spatial decay length parameter which is also specified by the user. The weight is then an exponential decaying function from the distance to the location of the measurement.

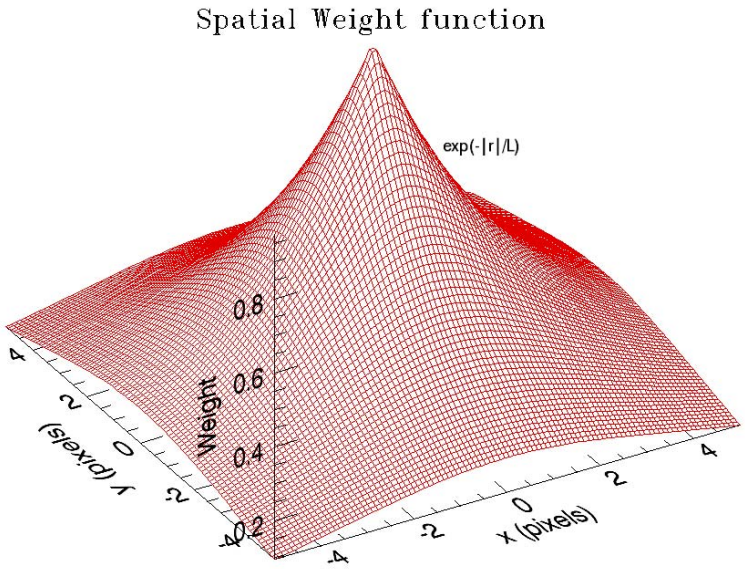

Figure 2 Spatial weight function for a decay length L=3 pixels.

In figure 2 the spatial weight function is plotted for a decay length of 3 pixels. Below the corresponding spatial weight matrix:

$$W_s(x,y) = \begin{bmatrix} 0.39 & 0.47 & 0.51 & 0.47 & 0.39 \\ 0.47 & 0.62 & 0.72 & 0.62 & 0.47 \\ 0.51 & 0.72 & 1.00 & 0.72 & 0.51 \\ 0.47 & 0.62 & 0.72 & 0.62 & 0.47 \\ 0.39 & 0.47 & 0.51 & 0.47 & 0.39 \end{bmatrix}$$

2.4.2 Temporal averaging

In the original classification for each point the last ten historical measurements were stored and used in the calculation of the subclass. There was no limit for the age of the measurements and all measurements had equal weight. In the new model a time averaging function is introduced. An exponential decay function with a sharp cutoff time is chosen:

$$\begin{aligned} W_t(\Delta t) &= 0 && ,\Delta t < 0 \\ W_t(\Delta t) &= \exp\left(-\frac{\Delta t}{A}\right) && ,0 \le \Delta t \le B \\ W_t(\Delta t) &= 0 && ,\Delta t > B \end{aligned}$$

Equation 2-12

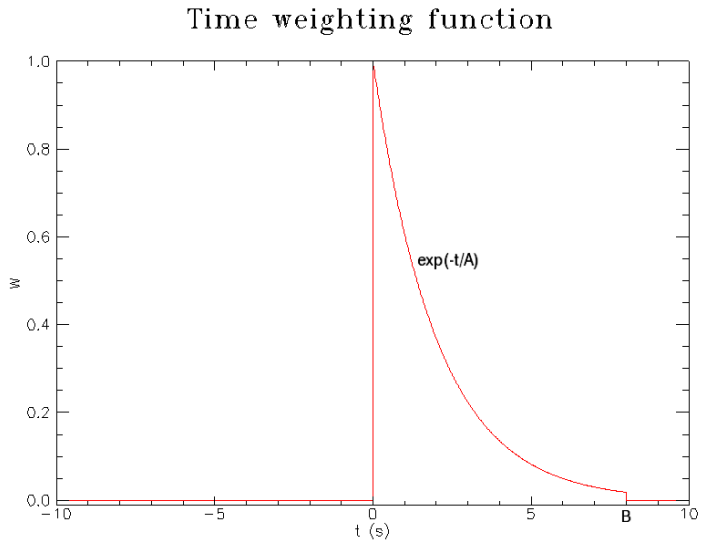

Figure 3 Time weighting function

where $\Delta t = t_o - t$ is the time difference between the time of observation t of a certain pixel (x, y) and the time of the latest observation to of the central pixel (x_o, y_o). Future observations ($\Delta t < 0$) get zero weight and are thus not taken into account. A is the user input time decay constant, and B the user input cutoff time. Input of $A = 0$ yields a flat function. Observations that are older than B will not be taken into account. Input of $B = -1$ means there is effectively no cutoff time.

2.4.3 Combined Spatio-temporal averaging

All measurements are used with a weight factor $W_{st}(x, y, t)$ that is the product of the spatial weight factor $W_s(x, y)$ and the temporal weight factor $W_t(t)$. The space-time weight factor $W_{st}(x, y, t)$ is normalised to give the resulting weight factor $w(x, y, t)$.

$$W_{st}(x,y,t) = W_s(x,y)W_t(t)$$

$$w(x,y,t) = \frac{W_{st}(x,y,t)}{\sum_i W_{st}(x_i,y_i,t_i)}$$

Equation 2-13

For calculating average values of ice-dependent parameters, a final weighting factor $W_f(x, y, t)$ including the a posteriori probability $p(ice|x, y, t)$ is introduced, as well as its normalised equivalent $w_f(x, y, t)$:

$$W_f(x,y,t) = W_{st}(x,y,t)\,p(ice \mid x,y,t)$$

$$w_f(x,y,t) = \frac{W_f(x,y,t)}{\sum_i W_f(x_i,y_i,t_i)}$$

Equation 2-14

Now the average value for any ice-dependent variable, e.g. ice parameter *a* may be calculated accordingly:

$$\tilde{a}(x,y,t) = \sum_i w_f(x_i,y_i,t_i)\,a(x_i,y_i,t_i)$$

Equation 2-15

2.5 The a priori and a posteriori ice probability

The a posteriori probability is given by Equation 2-9

$$Logit\big(p(ice \mid x)\big) = Logit\big(P(ice)\big) + \ln\left(\frac{p(x \mid ice)}{p(x \mid water)}\right)$$

Equation 2-16

where the a priori probability *P(ice)* still has to be chosen. When using Bayesian statistics, it is not allowed to use the same information (measurement) in both the a priori and a posteriori probability. In our model, all historical measurements at time t_o up to time t_{n-1}, which are older than the measurement under consideration x_n at time t_n, are used in the calculation of the a priori probability $P_n(ice)$. The new measurements at time t_n , also those of the surrounding pixels, are used for the calculation of the a posteriori probability $p(ice|\mathbf{x})$. This new a posteriori probability $p(ice|x)$ is stored and used for the calculation of the a priori probability P_{n+1} at a future time t_{n+1}.

We use Equation 2-16 to recursively calculate the a posteriori probabilities. The index *n* refers to measurement *n* at position $x = x_n$ and time $t = t_n$.

$$t = t_n: \quad Logit\big(p(ice \mid \mathbf{x_n})\big) = Logit\big(P_n(ice)\big) + \ln\left(\frac{p(\mathbf{x_n} \mid ice)}{p(\mathbf{x_n} \mid water)}\right)$$

Equation 2-17

To make a start the a priori probability $P_0(ice)$ is set to a climatological value $P_{cl}(ice)$ for which the value of $P_{cl}(ice) = 0.5$ can be chosen if no information is available. Later on the a priori probability will be a function of previous measurements as well. In our model the a priori probability $P_n(ice)$ is chosen to be a weighted average of a climatological value $P_{cl}(ice)$ and the a posteriori probability from the previous measurement $p(ice|\mathbf{x}_{n-1})$. The weighting factor $w_n(t)$ is a decreasing exponential function of time (see paragraph 2.4.2), so that the weight from a past measurement $\mathbf{x}_{n-1}$ is gradually reduced over time:

$$P_n(ice) = P_{cl}(ice)^{[1-w_n(t_{n-1})]}\, p(ice \mid \mathbf{x_{n-1}})^{w_n(t_{n-1})}$$

$$t = t_n: \quad Logit\big(P_n(ice)\big) = [1 - w_n(t_{n-1})]Logit\big(P_{cl}(ice)\big) + [w_n(t_{n-1})]Logit\big(p(ice \mid \mathbf{x_{n-1}})\big),$$

$$w_n(t_i) = \exp\left(-\frac{(t_n - t_i)}{A}\right)$$

Equation 2-18

If the time between two measurements t_n and t_{n-1} is very large then $w_n(t_{n-1}) \rightarrow 0$ and the a priori probability $P_n\ (ice) \rightarrow P_{cl}\ (ice)$. The information of the previous measurement is lost and the climatological value is used as the a priori probability.

If the time between two measurements t_n and t_{n-1} is very short then $w_n(t_{n-1}) \rightarrow 1$ and the a priori probability $P_n\ (ice) \rightarrow p(ice|\mathbf{x_{n-1}})$. The information of the previous measurement is still valid and it is used as the a priori probability.

For the a priori probability $P_n\ (ice)$ at time $t=t_n$ one gets (see Appendix A):

$$t = t_n: Logit\big(P_n(ice)\big) = Logit\big(P_{n-1}(ice)\big) + w_n(t_{n-1}) \ln\left(\frac{p(\mathbf{x_{n-1}} \mid ice)}{p(\mathbf{x_{n-1}} \mid water)}\right)$$

Equation 2-19

Together with Equation 2-17 this equation is efficient from a computational and memory management point of view, because what need to be stored at time t_{n-1} is only the a priori probability $P_{n-1}(ice)$, the time of measurement t_{n-1} and the measurement $\mathbf{x_{n-1}}$ itself. Using this information and the new measurement $p(\mathbf{x_n}\ |ice)$ at time t_n the new a posteriori probability $p(ice|\mathbf{x_n})$ can be calculated.

2.6 Distribution of σ^0-triplets

In order to be able to calculate the a posteriori probability a geophysical model and error model for both ice and wind is necessary. The model for wind and ice are the wind cone and ice line respectively, which both have a dependency on the incidence angle θ only. The error model is described below.

2.6.1 Error model for sea points

The distribution of sea points around the wind cone in the measurement space (σ^0 in dB) is assumed to be a standard normal distribution:

$$P_{normal}(r) = \frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{r^2}{2s^2}\right)$$

Equation 2-20

Here r is the distance to the wind cone, and s is the standard deviation of this distance. For the normalised distance to the wind cone $d_{wind} = r/s$, the value of $d_{wind} = 1$ corresponds to the

standard deviation.

2.6.2 Error model for ice points

The distribution of the *b* and *c* ice parameter is assumed to be normal. For the distance to the ice line in the measurement space this assumption results in a Rayleigh-distribution (Weisstein (2006)]:

$$P_{rayleigh}(r) = \frac{r}{s^2}\exp\left(-\frac{r^2}{2s^2}\right), r = \sqrt{\left(b^2 + c^2\right)}$$

Equation 2-21

Here *r* is the distance to the ice line and *s* corresponds to the standard deviation of the (normal) distribution when regarding the *b* or *c* parameter. The average value of this distribution is

$$\tilde{P}_{rayleigh} = s\sqrt{\frac{\pi}{2}}$$

Equation 2-22

and the maximum value occurs for $r = s$. In the new sea/ice classification the normalised distance to the ice line $d_{ice} = r/s$ is used. The top of the distribution will then occur at the value of $d_{ice} = 1$.

3 Icemodel Verification

3.1 Introduction

This section gives a description and a discussion of some test cases that have been run to verify the ice model. All test cases are run for the Antarctic region. For the wind computation CMOD5 has been used with inversion and 2D-var ambiguity removal. For the tests ice model version 5.0 has been used. ERS BUFR files have been reprocessed for producing the ice maps. The reprocessing was necessary because normally points below a temperature of 273 K (SST limit) are flagged out and not processed at all. This means no wind solution is present and no distance to cone is calculated. The SST mask is an area larger than that enclosed by the ice edge. For wind calculations it is a conservative mask but for producing ice maps all points need to be processed.

The collocated input files are taken from the MOS-archive:

/fa/ao/sat/ers/scat/reprEra40/

The reprocessed output files are written to the MOS again:

/ua/verspeek/sat/ers/scat/reprIce/

3.2 Typical run

Output of the ice model run is mainly in the form of graphical files in the format S<yyyymmdd>*.ppm or N<yyyymmdd>*.ppm, where S/N represents Southpole/Northpole, and <yyyymmdd> represents the date. These bitmap files can be combined into an animated gif movie and viewed with e.g. a web browser. Files S<yyyymmdd>.ppm show the ice subclass (Table 2, Table 4), and for class = ice (α_2, or β_1) the average value of the a ice parameter.

Files S<yyyymmdd>t.ppm show the basic ice class (Table 1)

Files S<yyyymmdd>postprob.ppm show the a posteriori ice probability $p(ice|\boldsymbol{x})$.

In the middle of the swath (node 10), the ice line is closely lying to the wind cone and the difference between ice and wind is hard to make. This can be seen in the left figure and the middle figure where the middle of the satellite track can be seen as orange, respectively gray area. In the right figure this effect is averaged out.

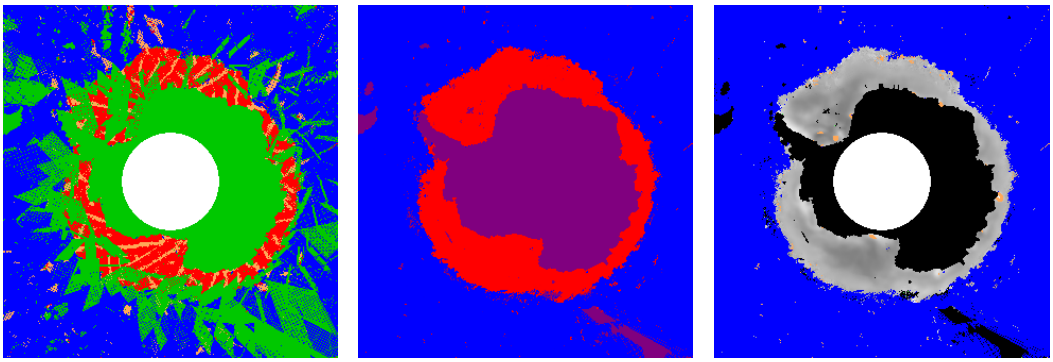

Figure 4 Graphical output of typical run (2000-05-24). The left picture shows the basic sea/ice classes (Table 1), the middle picture the a posteriori ice probability on a gray scale, and the right picture shows the sea/ice subclasses (Table 4) and the ice parameter a on a gray scale.

3.3 Parameter sensitivity

In the following tests the influence of several input parameters is examined, especially the parameters that deal with the spatial and temporal averaging.

3.4 Spatial averaging

Different levels of spatial averaging can be introduced by specifying different values for the decay length L in paragraph 2.4.1 with the "--decaylength" command line option. The decay length is given in units of pixel length (25 km).

For the default decaylength of $L = 3$ the spatial weighting matrix reads:

$$W_s(x,y) = \begin{bmatrix} 0.390 & 0.475 & 0.513 & 0.475 & 0.390 \\ 0.475 & 0.624 & 0.717 & 0.624 & 0.475 \\ 0.513 & 0.717 & 1.000 & 0.717 & 0.513 \\ 0.475 & 0.624 & 0.717 & 0.624 & 0.475 \\ 0.390 & 0.475 & 0.513 & 0.475 & 0.390 \end{bmatrix}$$

In figure 5 the icemodel is run with a number of different values of the decay length. In all cases the default value for the decay time is used. Green points, representing points with not enough measurements, are only present for L = 0 where effectively no spatial averaging is applied. The level of detail does not seem to decrease much with increasing decay length. This is probably caused by the fact that the radar footprint is 50x50 km^2 compared to the pixel size of 25x25 km^2. Also the time averaging already adds to the smearing out of minor details. In the last case where a flat spatial function is taken, the averaging area is limited by the implementation that uses a 5x5 matrix around the central pixel for spatial averaging (an “all ones” matrix in this case).

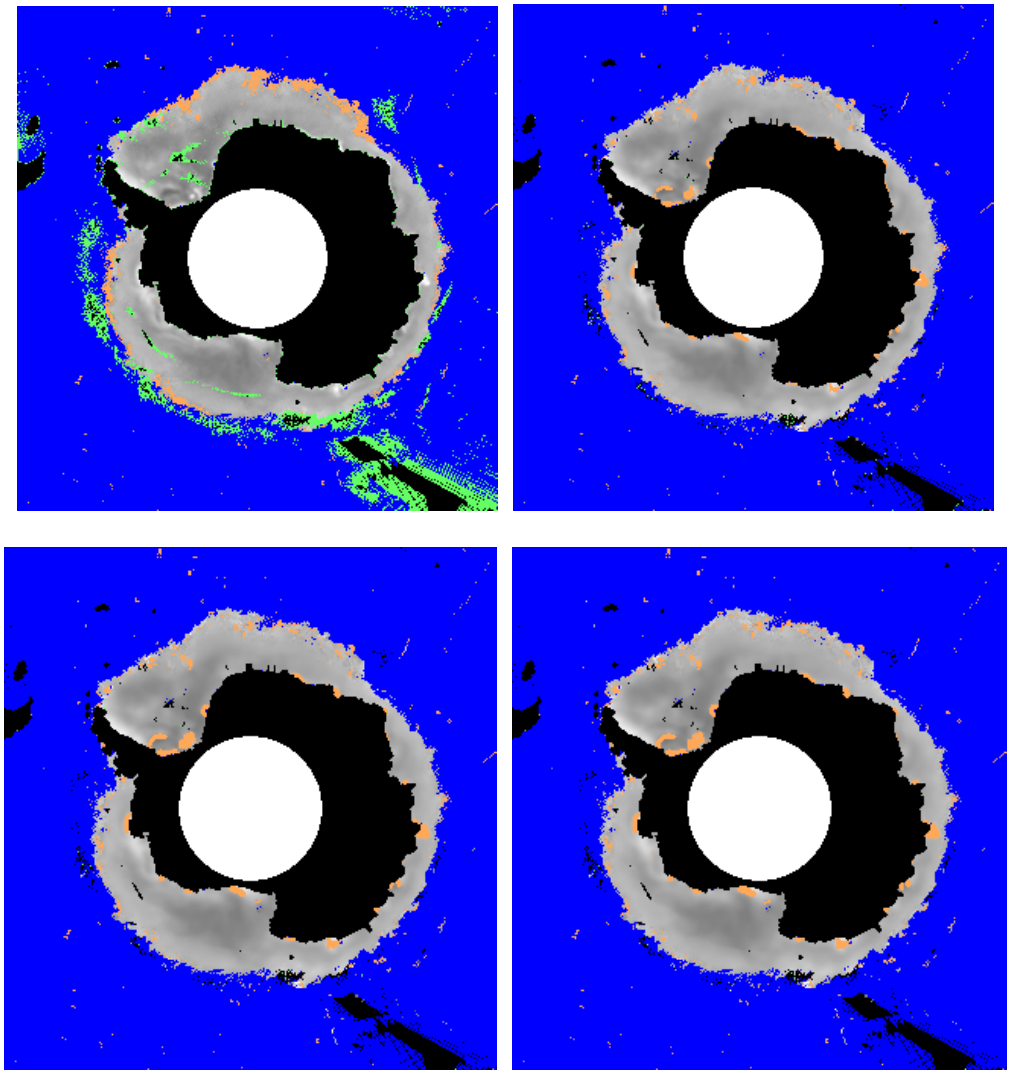

Figure 5 Spatial averaging in order of increasing averaging area, respectively decay length L= 0, L=1, L=3, and L=infinity (flat function).

3.5 Temporal averaging

In the original model the time averaging function was flat without a cutoff time, with a maximum of 10 historical measurements that were taken into account. In the new model the time decay constant *A* and the cutoff time *B* determine the time averaging behaviour (see paragraph 2.4.2). A number of test runs with increasing influence of old measurements have been performed. The default decay length has been applied for spatial averaging in these runs.

As can be expected, the number of green points representing points that have not enough measurements for calculating the ice probability, is reduced when going to a larger time averaging interval. The number of orange points representing points where the ice-parameter standard deviation is too large, is increasing. The number of falsely classified points, water where it should be ice, or ice where it should be water, is decreasing also. For the case of a decaytime of A = 1 hour, where effectively only the latest measurement is taken into account, spurious water points can be seen in the middle of the ERS track. For these middle wind vector cells (WVCs), the iceline is lying closely to the windcone and the classification is the most difficult to make. For the last figure with a decaytime of 192 hour some of the details seem to be smeared out.

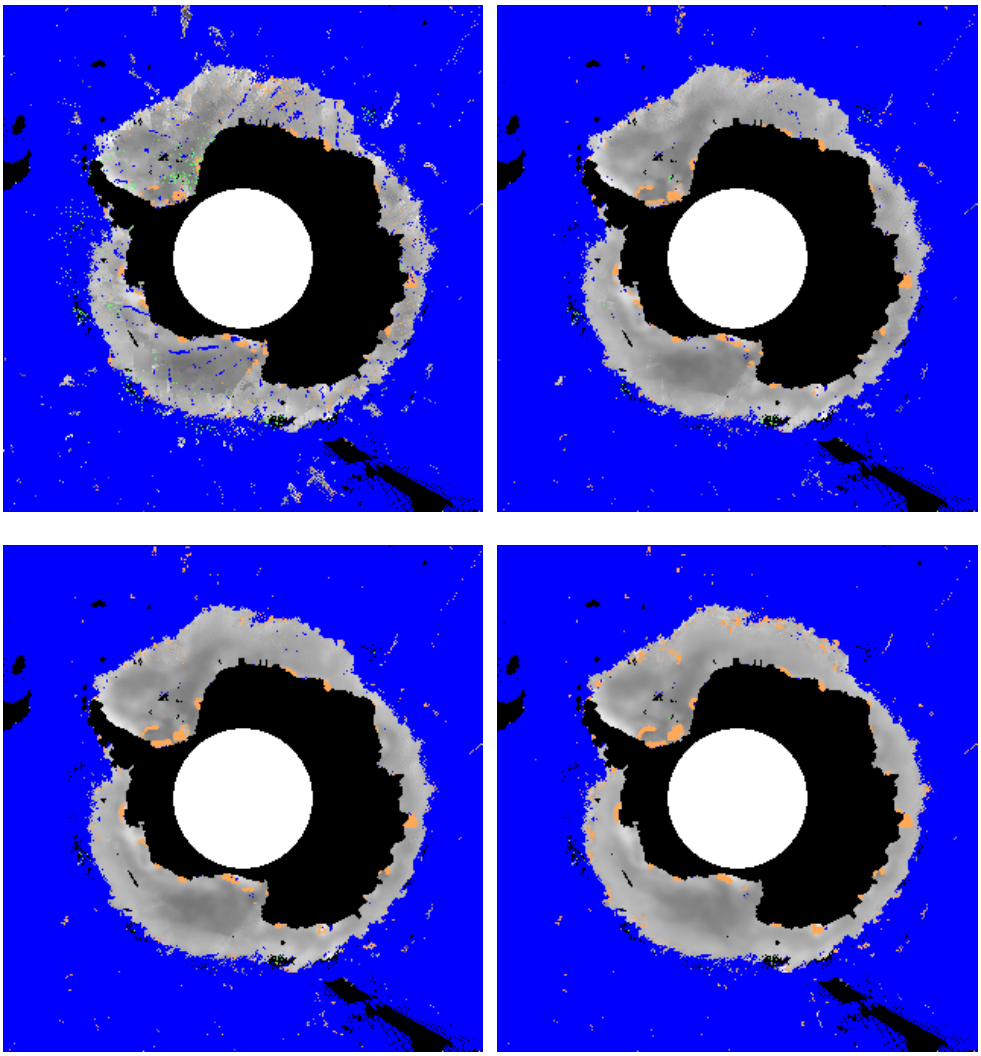

Figure 6 Time averaging in order of increasing averaging time. Corresponding to a decay time of *A*=1 *h*, *A*=12 *h*, *A*=24 *h* and *A*=192 *h*.

3.6 Wind information

The wind information (distance to the wind cone) may be omitted with the command line option "-dontuse windinfo". This mode is useful for testing unprocessed BUFR files or the ice model itself. It can also be useful when an ice mask is used that is retrieved by other means, or when a model is added to examine tropical rain forests or other vegetation. The calculated ice parameter *a* is the same in areas which are truly ice, but in the water areas lots of spurious ice areas occur.

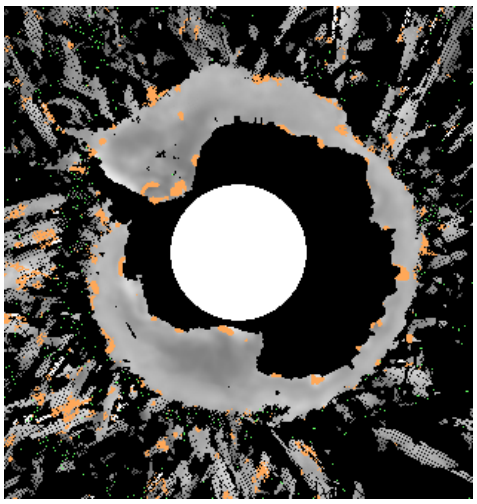

Figure 7 No wind information used

3.7 Land points

Land points may be included with the command line option "-use landpoints". The check on the land mask bit in the ESA quality flag is then omitted. Also the check on the value of the standard deviation for the ice parameter *a* has been omitted (no orange points). The land point near the Antarctic "coast" have a very high value for the *a* parameter, indicating high reflectivity. The color in the bitmap is clipping to white, indicating a value of *a* >= 15. The high brightness is caused by snow crust and refrozen ice in the snow cover of the glacial ice sheets.

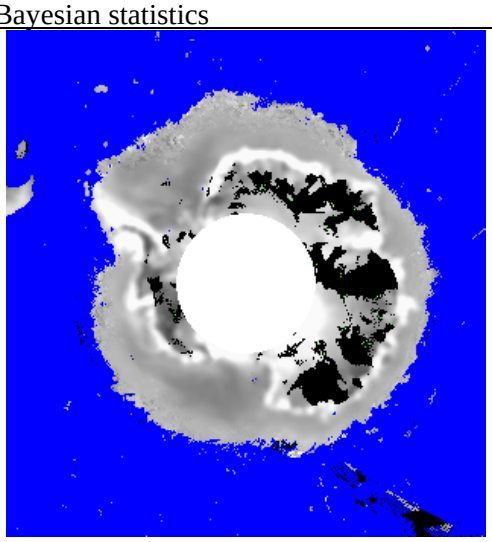

Figure 8 Including land points, the limit on the standard deviation for subclass b2 (orange) has been relaxed.

3.8 Comparison with SST

The output of the ice model is compared with an ice map produced with Surface Temperature (SST) data from the ECMWF Era40 database. In parameter STL1 above sea the Sea Surface Temperature (SST) and above land the Soil Surface Temperature is stored in this parameter. This parameters is read in from a file with daily data on a predefined regular grid and collocated with the ERS BUFR data using interpolation. By specifying a small temperature interval for the SST contour plot from -4°C to +4°C in steps of 1°C, it can be seen that a SST temperature of -2 °C is conservative, i.e. is enclosing the ice edge.

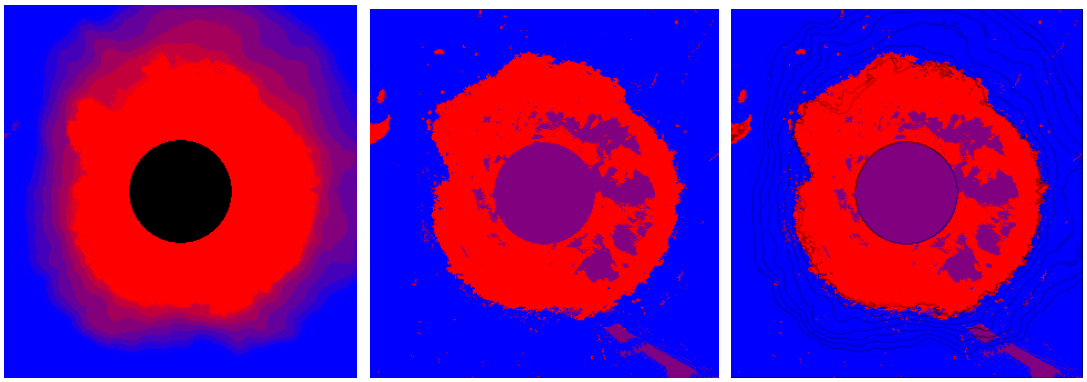

Figure 9 SST ice map (left), icemodel a posteriori ice probability (middle) and the two pictures overlayed on 2000-05-24

3.9 Comparison with other ice models

The ice maps from the KNMI ice model are compared with the ice maps as produced by IFREMER. IFREMER uses the anisotropy and derivative parameters to visualise the ice map. In Figure 10 the three figures represent the KNMI ice map, the IFREMER ice map from ERS-2

data, and the IFREMER ice map ice map from Quikscat. The IFREMER maps represent the backscatter at 40°. The ice edge and major ice parameter structures are consistent among the three cases.

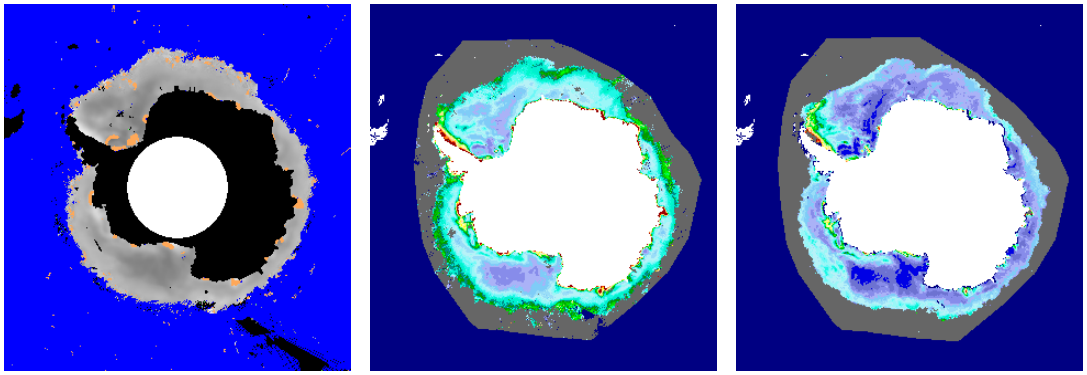

Figure 10 Comparison of ice maps. On the left the KNMI ice map on 2000-05-24, in the middle the IFREMER weekly ice map (2000-05-22/2000-05/28) showing the σ^0 at 40° from ERS-2 data, on the right the IFREMER daily ice map from Quikscat data on 2000-05-24.

3.10 Error model verification

In paragraph 2.6.1 and 2.6.2 the error models for wind and ice triplets have been described. In order to see if the assumption of these distributions are correct it can be compared with the distribution from actual measurements. Figure 11 shows the distance to wind cone for one day of data. Negative cone distance corresponds to a measurement point lying inside the cone, positive distances correspond to points lying outside the wind cone. Only points that passed the quality control have been used. The shape of the measured distribution is more sharply peaked than the theoretical Gaussian distribution.

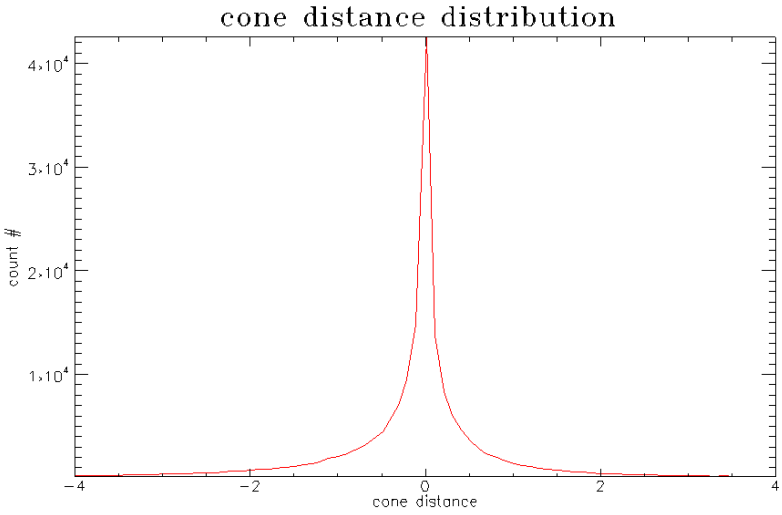

Figure 11 Distribution of the distance to wind cone. A negative value corresponds to a triplet lying inside the cone, a positive value corresponds to a triplet outside the cone.

In Figure 12 the distribution of the ice parameters *a*, *b* and *c* is plotted. Here only points belonging to ice class (b) ("Ice" – see Table 1) have been used. Parameter *a*, the along-line coordinate, looks like a superposition of several Gaussian distributions which could correspond to different types of ice. The coordinates perpendicular to the ice line, *b* and *c*, together determine the distance to ice line. The *c*-distribution is skew and the extra points are lying on the side of the wind cone which suggests that these measurements are actually wind points. This could be caused by measurements for which the distance to wind is large but not correct. These points may be flagged but the flag is not checked in this case, only the distance to cone is considered. Apart from this skewness the *b* and *c* distributions look Gaussian in accordance with the assumption.

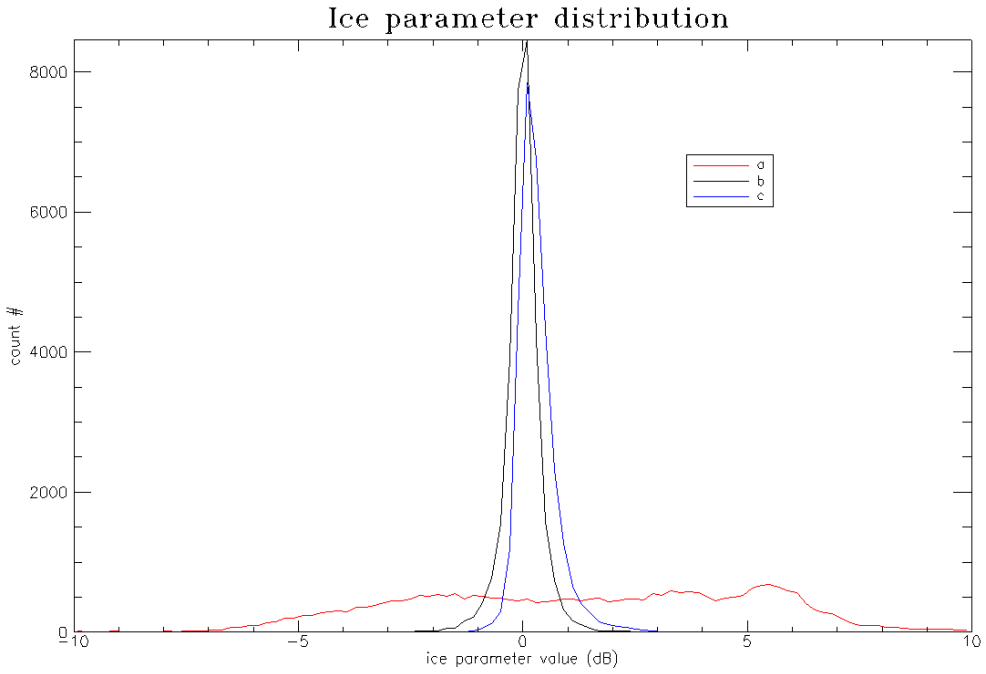

Figure 12 Distribution of the a, b, and c ice parameters.

4 Conclusions

The ice model proves a useful tool for sea/ice discrimination and could help in the quality control for the wind retrieval procedure. The calculated ice map is a product on its own based on the scatterometer measurements. The generalised spatial and time averaging gives the user more control and produces pictures of the ice map without artificial boundaries. Default values for the adjustable parameters are built in to give meaningful results, based on the outcome of this study. For the time averaging a value of A = 192 h is taken as a default, and for the spatial averaging a decay length of L=3. The threshold for the sea/ice discrimination of $p(ice|x)$.= 50 % seems good. The ice edge is not very sensitive to this parameter. For conservative calculations a threshold of $p(ice|x)$. = 10 % could be taken as well. The limit for enough measurements (total weight W_f >= 5) could be lowered or made dependent on the averaging constants. The limit for the standard deviation of the ice parameter *a* could be made dependent on the absolute value of *a*.

5 Expectations for ASCAT

The MetOp satellites will each carry an ASCAT scatterometer that has about the same antenna geometry as the ERS scatterometer, but will duplicate the swath on both sides of the satellite nadir track. The same ice model can be used for ASCAT, with the ice line origin and slope, and the scaling factors for the ice parameters recalculated using a new measurement set. ASCAT has higher incidence angles than the ERS scatterometer, which is beneficial for the sea/ice discrimination. The larger coverage, especially when more satellites will be operational, will allow for higher spatial and time resolution. The ERS scatterometer is always looking away from the South Pole, causing a large "blind spot" above Antarctic land. For ASCAT it will be much smaller because ASCAT looks in both directions. A region of ice with stable backscatter properties over a longer period can be used for calibration of the ASCAT scatterometers.

Appendix A – Derivation of the a posteriori probability

The a posteriori probability is given by Equation 2-9

$$Logit\big(p(ice_0 \mid x_0)\big) \;=\; Logit\big(P(ice_0)\big) + \ln\left(\frac{p(x_0 \mid ice_0)}{p(x_0 \mid water_0)}\right)$$

Equation 5-1

We use Equation 2-9 to recursively calculate the a posteriori probabilities. The index i refers to the i^{th} measurement at time $t = t_i$ and position $x = x_i$. Here the probabilities $p(\mathbf{x}_i|\text{ice})$ and $p(\mathbf{x}_i||\text{water})$ may be space-averaged values of measurements at time $t = t_i$. For simplicity this is not denoted in the formulas:

$$t = t_0: \quad Logit\big(p(ice \mid \mathbf{x_0})\big) \;=\; Logit\big(P_{cl}(ice)\big)$$

$$t = t_n: \quad Logit\big(p(ice \mid \mathbf{x_n})\big) \;=\; Logit\big(P_n(ice)\big) + \ln\left(\frac{p(\mathbf{x_n} \mid ice)}{p(\mathbf{x_n} \mid water)}\right)$$

Equation 5-2

In our model the a priori probability $P_n(ice)$ is chosen to be a weighted average of a climatological value $P_{cl}\ (ice)$ and the a posteriori probability from the previous measurement $p(ice|\mathbf{x_{n-1}})$ The weighting factor $w_n(t)$ is a decreasing exponential function of time, so that the weight from a past measurement $\mathbf{x_{n-1}}$ is gradually reduced over time (Equation 2-18):

$$P_n(ice) = P_{cl}(ice)^{[1-w_n(t_{n-1})]}\, p(ice \mid \mathbf{x_{n-1}})^{w_n(t_{n-1})}$$

$$t = t_n: \quad Logit\big(P_n(ice)\big) \;=\; [1 - w_n(t_{n-1})]Logit\big(P_{cl}(ice)\big) + [w_n(t_{n-1})]Logit\big(p(ice \mid \mathbf{x_{n-1}})\big),$$

$$w_n(t_{ni}) = \exp\left(-\frac{(t_n - t_i)}{A}\right)$$

Equation 5-3

By using Equation 5-3:

$$Logit\big(P_n(ice)\big) \;=\; [1 - w_n(t_{n-1})]Logit\big(P_{cl}(ice)\big) + [w_n(t_{n-1})]Logit\big(p(ice \mid \mathbf{x_{n-1}})\big),$$

and Equation 5-2:

$$Logit\big(p(ice \mid \mathbf{x_n})\big) \;=\; Logit\big(P_n(ice)\big) + \ln\left(\frac{p(\mathbf{x_n} \mid ice)}{p(\mathbf{x_n} \mid water)}\right)$$

We can derive a formula for the a priori probability $P(ice)$ at each time step:

For $t=t_0$ (no measurements yet):

$$Logit\left(P_0(ice)\right) = Logit\left(P_{cl}(ice)\right)$$

For $t=t_1$:

$$Logit\left(P_1(ice)\right) = [1-w_1(t_0)]Logit\left(P_{cl}(ice)\right)+w_1(t_0)Logit\left(p(ice\mid \mathbf{x_0})\right)=Logit\left(P_{cl}(ice)\right)$$

For $t=t_2$:

$$Logit\left(P_2(ice)\right) = [1-w_2(t_1)]Logit\left(P_{cl}(ice)\right)+w_2(t_1)Logit\left(p(ice\mid \mathbf{x_1})\right)=$$

$$=Logit\left(P_{cl}(ice)\right)+w_2(t_1)[Logit[\,p(ice\mid \mathbf{x_1})-\log it\left(P_{cl}(ice)\right)]=Logit\left(P_{cl}(ice)\right)+w_2(t_1)\ln\left(\frac{p(\mathbf{x_1}\mid ice)}{p(\mathbf{x_1}\mid water)}\right)$$

For $t=t_3$ (using $w_n(t_k)w_k(t_s) = w_n(t_s)$):

$$Logit\left(P_3(ice)\right) = [1-w_3(t_2)]Logit\left(P_{cl}(ice)\right)+w_3(t_2)Logit\left(p(ice\mid \mathbf{x_2})\right)=$$

$$Logit\left(P_3(ice)\right) = [1-w_3(t_2)]Logit\left(P_{cl}(ice)\right)+w_3(t_2)Logit\left(P_2(ice)\right)+w_3(t_2)\ln\left(\frac{p(\mathbf{x_2}\mid ice)}{p(\mathbf{x_2}\mid water)}\right)=$$

$$Logit\left(P_3(ice)\right) = Logit\left(P_{cl}(ice)\right)+w_3(t_2)w_2(t_1)\ln\left(\frac{p(\mathbf{x_1}\mid ice)}{p(\mathbf{x_1}\mid water)}\right)+w_3(t_2)\ln\left(\frac{p(\mathbf{x_2}\mid ice)}{p(\mathbf{x_2}\mid water)}\right)=$$

$$=Logit\left(P_{cl}(ice)\right)+w_3(t_1)\ln\left(\frac{p(\mathbf{x_1}\mid ice)}{p(\mathbf{x_1}\mid water)}\right)+w_3(t_2)\ln\left(\frac{p(\mathbf{x_2}\mid ice)}{p(\mathbf{x_2}\mid water)}\right)$$

For $t = t_n$.:

$$Logit\left(P_n(ice)\right) = Logit\left(P_{cl}(ice)\right)+\sum_{i=1}^{n-1}w_n(t_i)\ln\left(\frac{p(\mathbf{x_i}\mid ice)}{p(\mathbf{x_i}\mid water)}\right)$$

Equation 5-4

From this equation using recursion follows Equation 2-19:

$$Logit\left(P_n(ice)\right) = Logit\left(P_{n-1}(ice)\right)+w_n(t_{n-1})\ln\left(\frac{p(\mathbf{x_{n-1}}\mid ice)}{p(\mathbf{x_{n-1}}\mid water)}\right)$$

Equation 5-5

The a posteriori probability for $t=t_n$ is (Equation 5-2):

$$Logit\left(p(ice \mid \mathbf{x_n})\right) = Logit\left(P_n(ice)\right) + \ln\left(\frac{p(\mathbf{x_n} \mid ice)}{p(\mathbf{x_n} \mid water)}\right)$$

or

$$Logit\left(p(ice \mid \mathbf{x_n})\right) = Logit\left(P_{cl}(ice)\right) + \sum_{i=1}^{n} w_n(t_i) \ln\left(\frac{p(\mathbf{x_i} \mid ice)}{p(\mathbf{x_i} \mid water)}\right)$$

Equation 5-6

Appendix B - Command line options

The ice model has been compiled with several Fortran-90 compilers and run on Linux and Unix systems. A stand-alone of the ice model exists and it is also incorporated in the ERS Data Processor (ESDP). It makes use of several genscat modules, e.g. for BUFR handling. A makefile is available with the code. The ice model code resides in CVS on the bcp1 machine. The latest version can be retrieved with the command:

>vbcvs checkout icemodel

or

>cvs -d :ext:'logname'@bclap1:/data/cvs/nwpsaf checkout icemodel

The icemodel program has a lot of optional parameters which may be specified. When parameters are omitted they will be given their default value.

`icemodel.x [-shortoption|--longoption] <optionValue>`

with <> indicating non-obligatory input, [] indicating obligatory input, and | indicating alternatives. The following command line options are available:

`[-i|--in] <metainfile>`	Specify meta inputfile containing the list of BUFR files to be processed.
`[-if] <infile>`	Specify one BUFR inputfile. Overrules <metainfile>.
`[-S|--Southpole]`	Icemap for Southpole region
`[-N|--Northpole]`	Icemap for Northpole region
`[-dl|--decaylength] <decayLength>`	Decay length. If decaylength=0 only the central pixel is used. If decaylength=-1 a flat function is used (all ones matix)

[-dt|--decaytime] <decayTime>
Decay time. If dt=0 decay time is infinite (flat function)

[-ct|--cutofftime] <cutoffTime>
Cutoff time. If ct=-1 cutoff time is infinite (no cutoff)

[-outputLevel] <outputLeveL>
Output level for graphical files. =0 means no graphical output files.

[-dontwriterst] Don't write a restart file when the processing is finished.

[-rst|--restartfile] <restartFileName>
Rstart file to be read in to start the processing.

[-dontuse windinfo] Don't use wind information

[-use landpoints] Use the land points

[-o|--out] <outputfile> Specify ASCII outputfile

[-n|--Nodes] <nodeNrs> Nodes to be used for generating ASCII output file, e.g. –n "[3,4,5]"(don.t forget the hooks and quotes)

[-m|--mask] <MaskNr> Mask number to be used for generating the ASCII output file.

[-h|--helpall] Shows command line options.

[-v|--version] Shows version number

[-d|--debug] <debugLevel> Debug level. =0 means no debug output.

Acronyms and abbreviations

Name	Description
ASCAT	Advanced scatterometer
CMOD	C-band geophysical model function used for ERS and ASCAT
ECMWF	European Centre for Medium-Range Weather Forecasts
ERS	European Remote sensing Satellite
ESDP	ERS Data Processor
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
GMF	Geophysical model function
KNMI	Koninklijk Nederlands Meteorologisch Instituut (Royal Netherlands Meteorological Institute)
MetOp	Meteorological Operational satellite
MLE	Maximum Likelihood Estimator
NWP	Numerical Weather Prediction
SST	Sea Surface Temperature
WVC	Wind vector cell, also called node or cell

References

[Breivik et al (2001)] Breivik, L.A., EastWood, S., Schyberg, H., Andersen, S., Tonboe, R. Sea ice products for EUMETSAT Satellite Application Facility, *Canadian Journal of Remote Sensing*, 2001, page 403-410

[Bernardo (2003)] Jose M. Bernardo, "Bayesian Statistics", *Encyclopedia of Life Support Systems, Probability and Statistics*, Oxford, UK, 2003

[Ezraty and Cavanié (1997)] Ezraty, Robert and Cavanié, Alain, Development and evaluation of an NSCAT 25 km resolution sea ice product, *IMSI report no 4*, IFREMER Technical report DRO/OS No.97/03, October 1997

[Gohin, et al (1995)] Gohin, F., Some active and passive microwave signatures of Antarctic sea ice from mid-winter to spring 1991, *Int. J. Remote Sensing,* vol **16,** No. 11, 2031-2054, 1995

[Haan de, Stoffelen (2001)] de Haan, Siebren and Stoffelen, Ad, "Ice Discrimination using ERS scatterometer", *KNMI, OSISAF report*, de Bilt, 18 September 2001

[Weisstein (2006)] Eric W. Weisstein, "Rayleigh distribution ", From MathWorld, A Wolfram Web Resource, http://mathworld.wolfram.com/RayleighDistribution.html

[Wikipedia (2006)] Naive Bayesian classification, http://en.wikipedia.org/wiki/Naive_Bayesian_classification