Post-Print Version:

Reliability Measures for Sea Ice Motion Retrieval From Synthetic Aperture Radar Images

Thomas Hollands, Stefanie Linow and Wolfgang Dierking

Reliability measures for sea ice motion retrieval from synthetic aperture radar images

Thomas Hollands, Stefanie Linow, and Wolfgang Dierking

Abstract—**Sea ice motion is triggered by wind and ocean currents. Its magnitude and direction can be automatically retrieved using pairs of satellite images acquired over the same area. However, external reference data for validation of drift retrievals, such as tracks from buoys, are sparse. Information about the reliability of the retrieved ice drift field is crucial for applications such as operational sea ice mapping or validation of computer models for simulations of sea ice dynamics. In this paper we introduce an intrinsic measure based on the properties of radar image pairs to assess the reliability of the retrieved ice drift vectors. The proposed method combines different parameters, e. g. correlation coefficient and two textural quantities, to provide information about the suitability of subimage regions for pattern matching. In this way, we generate a quality parameter (called confidence factor, CFA) for the calculated ice drift velocities. The CFA is compared to results obtained by 'backmatching'. The latter requires that the drift field is computed twice using the image pair, first in sequential and then in reversed order. For stable ice conditions, the results show that areas regarded as unreliable by the CFA compare well with the areas revealing larger differences from backmatching.**

I. INTRODUCTION

HIGH resolution ice drift fields can be retrieved automatically from time series of satellite radar imagery, for example by using algorithms based on the method of pattern matching, e. g. [1]. In practical applications, such as operational mapping of sea ice conditions, general process studies [2] or validation of models for simulations of atmosphere – sea ice – ocean interactions [3], [4], knowledge about the accuracy of the retrieved ice drift fields is required. In [5], the importance of uncertainty information for data assimilation is emphasized. Although most drift retrieval algorithms are tested against tracks from buoys and ice drift stations [1], [6] with good results concerning the overall performance, a method for evaluating the accuracy of each retrieved drift vector does not exist. In this paper, we suggest an approach to provide additional information on the reliability of each individual drift vector, whereby 'reliable' means that a certain level of accuracy is maintained. For this, we introduce two methods for an assessment of reliability. The first one, called backmatching, is a robust indicator for the reliability of sea ice motion, but doubles the computation time. Therefore we propose a second method by calculating a 'confidence factor' (CFA), which quantitatively characterizes certain image properties important for ice drift retrievals.

The automated retrieval of sea ice displacement vectors from two subsequently acquired, overlapping satellite images

T. Hollands, S. Linow and W. Dierking are with the Department of Climate Sciences, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, D-27570, Germany, E-mail: thomas.hollands@awi.de.

can be described as a problem of pattern matching, as mentioned above. In many cases, the algorithms employ a correlation measure to quantify the similarity of regions in two input images. Intuitively one might assume that high correlation values indicate highly reliable displacement vectors [7]. However, that is not necessarily the case since the correlation coefficient is influenced by various factors and can vary depending on the image texture characteristics [8]. We discuss this item in more detail in the following sections.

The paper is organized as follows: in the next section we describe our test data, the method used for calculation of the ice drift vectors, and our approaches to assess their reliability. In section III, we analyse and compare the results of backmatching and CFA calculations. Finally, a discussion and the conclusions are presented.

II. METHODS

In this section, we provide information about the used SAR images and the algorithm we employ for ice drift retrievals. We introduce the parameters that were selected to evaluate the CFA and explain the principle of backmatching.

A. Calculation of ice drift

As a test case we calculated drift fields using Envisat ASAR wide swath data acquired over the Ronne Polynia[1] in 2008, which are described in [3]. Image pairs are from the following days: February 18/19, 19/20, 22/23, May 30/31, and May 31/June 1. Figure 1 shows the location of our test site. The

Fig. 1. Location of the test site. The zoomed region contains the overlapping area of the image pair recorded on the 22. / 23. February 2008. The dashed line denotes the approximate position of the polynia.

retrieval algorithm determines the position change of a given area element on the ice between the acquisitions of image 1

[1]A polynia is a region of open water in the pack ice. It forms when katabatic winds or ocean currents push the sea ice away from the coastline, or when warm water is upwelling near the coast. Due to the lack of the insulating ice cover, polynias are regions of increased heat flux between ocean and atmosphere.

and 2. The displacement is then obtained as the connecting line between those two positions. Since the ice may move curvilinearly, the displacement represents only the shortest possible distance between the two reference points in image 1 and image 2 but not necessarily the true path of motion. The ice drift velocities are calculated from the ratio between displacement and time interval between the image acquisitions.

For drift detection we apply a cascaded multi-scale multi-resolution algorithm first described by [1] and modified by [9]. This is an area-based pattern-matching algorithm which employs a combination of phase correlation and normalised cross-correlation which are calculated over search windows ('templates') to determine the displacement of sea ice between the acquisition times of two satellite images. In order to increase the robustness and reduce the computational efforts, the drift algorithm is based on a resolution pyramid in which the ice displacement is first calculated at a coarse resolution. This information is then used to initialise the search for corresponding patterns at a higher resolution level. In this way, the displacement vector is refined at every level of the resolution pyramid until the original resolution of the input image is reached. The algorithm itself forms a cascade of multiple runs through the resolution pyramid (in the following denoted "CRP-scheme"[2]). In the cascade, the displacement field is calculated repeatedly. With each cascade step, the number of pixels (i. e. the size of the template) to be matched for a single displacement vector is reduced, leading to an increased spatial density of the displacement field. To compensate for the reduced robustness of the match due to the reduced number of pixels in the search template, the algorithm uses the coarser displacement field from the previous cascade as initial displacement information and refines it locally during each pass of the resolution pyramid. The similarity measure of the employed algorithm is a normalised cross-correlation (NCC) with a preceding phase correlation for candidate selection [1], [10]. The performance of phase correlation and NCC in the presence of noise was analysed in [11]. The phase correlation is more robust with respect to non-linear motion and computationally more efficient, while the NCC is more robust against noise [12], [13].

Due to the combination of both correlation measures, the employed algorithm is relatively susceptible to high frequency noise. However, there are other error sources which are directly linked to properties of the image data:

- In homogeneous regions without any textural patterns, the matching leads to spurious displacement vectors. This effect can occur in areas of new ice which do not show any distinct structures such as cracks, rafting zones or ridges.
- Periodic variations of brightness in the images can also lead to false matches. Such variations are observed, e.g., in polynias where often distinct belts of ice develop.
- Strong changes in ice structures between the acquisitions of the two images, caused by advancing deformation, rotation of ice floes, ice melt, or other processes are also problematic due to the reduced similarity between the

images.
- At the image borders, it may not be possible to match patterns when corresponding ice structures have left the monitored area or have just moved into it at the acquisition of the second image. The width of the zone where patterns are lost depends on the ice drift velocity.

Other error sources are directly related to the layout of the algorithm itself. Since results obtained at coarser spatial resolution are used to initialise calculations at higher resolution, the retrieval may fail if the field of displacement vectors is very heterogeneous in a given area. In this case, the initial displacement may cause the algorithm to search in the wrong direction. Another problem occurs when the overlap area between consecutive satellite images is irregularly shaped. In this case, it may not be possible to position a sufficiently large window for correlation calculations close to the image border. This can lead to a reduced number of iterations.

B. Concept of multiple reliability measures

In the following, we outline our approach to employ a combination of different parameters, derived from the respective image pair, for an assessment of the reliability of the final ice displacement vectors. The motivation is the observation that the correlation coefficient in itself is not an appropriate measure of the error of automatically retrieved displacements relative to reference data (in our case the latter were determined manually). Figure 2, for example, does not reveal a correlation between the error and the cross-correlation coefficient obtained in the final step of the CRP-algorithm described above. This is also valid for the phase correlation
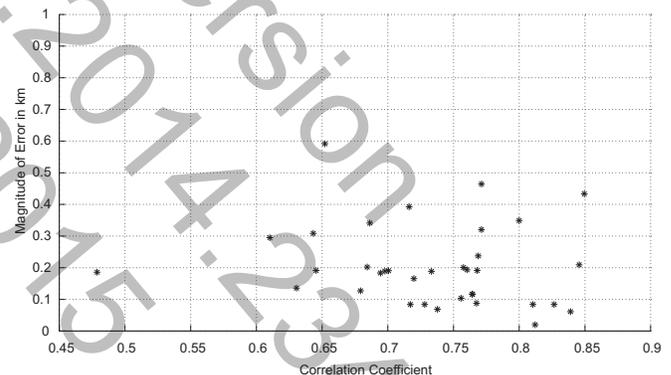

Fig. 2. Magnitude of error vs. cross-correlation coefficient for a selected set of drift vectors obtained using an image pair from the Ronne Polynia, Antarctica (22. / 23. 02. 2008)

[8]. We note that one in addition has to consider the error sources related to the image properties discussed in section A above. Therefore, we defined a 'confidence factor' (CFA) which combines the thresholds of different parameters to separate 'reliable' from 'unreliable' retrievals of displacement (the meaning of 'reliability' is discussed below). The selected parameters are useful for judging whether the properties of each image of an image pair and the comparison of both images fulfil the conditions for a successful pattern-matching. For operational use of the algorithm, the computational speed is important which means that the selected parameters must be

---

[2]Cascade-Resolution-Pyramid-scheme

relatively simple to evaluate. In the present form of the CFA, we use the following quantities:

- Mean Intensity Gradient (MIG)
- Mean Gradient Slope (MGS)
- Variance-to-squared-mean-ratio (VMR)
- Intensity threshold (IT)
- Correlation Coefficient
- Confidence Interval of Correlation Coefficient

One major premise of the CRP-algorithm is that ice structures can be recognized in both images of the image pair, i. e. parameters are needed that quantify the perceptibility of local image structures. The former two (MIG and MGS) are from the field of image texture analysis. SAR images are characterized by the presence of speckle and strong casual mirror reflections, which have also to be taken into account in the selection of parameters for calculating the CFA. The VMR is a measure for the effect of speckle on image intensity variations. By applying the IT it is possible to reduce the effect of mirror reflections (outliers of pixel brightness). These four parameters characterize the properties of each individual image but do not reflect any links between the two images forming a pair. This is achieved by employing the correlation coefficient and its confidence interval in the reliability assessment. Even if the correlation coefficient on its own does not provide a suitable measure for the reliability of a displacement vector, it is indisputably an indicator for similarities between intensity patterns in two different images. The single parameters are described in more detail in sections II C and D below.

For each parameter, we define a threshold to separate reliable from unreliable displacement vectors. Since the drift detection algorithm is based on pattern matching at multiple resolutions [9], it is necessary to consider the reliability at all levels of the CRP-scheme. This can be achieved in two ways: (1) One approach is to exclude individual displacement vectors at any step of the CRP-scheme if they exceed or fall below the threshold value of the CFA valid at this step. (2) Another approach is to combine the results of the different threshold checks and flag the individual displacement vectors accordingly. We found that method (1) reduces the reliability of the pattern matching if the total fraction of vectors discarded anywhere in the CRP-scheme exceeds 5 - 10 percent. Therefore we employ method (2) in our drift retrievals. Since the numerical values for the different parameters of which the CFA is comprised cover different ranges of magnitude, we introduced a counter and increment it by one for each threshold violation of any of the parameters at any stage in the CRP-scheme. In this manner, a map showing the number of threshold violations for each drift vector is generated. Consequently, a drift vector with a high number of threshold violations is less reliable than a vector with a lower number. The CFA is the sum of registered threshold violations normalised by the number of steps in CRP-scheme. Due to the complexity of the drift algorithm, it is impracticable to determine each threshold theoretically. Instead, we devised an empirical method. Our test site contains homogeneous (texture-free) areas on a stationary ice shelf (high intensity) as well as areas of dark-appearing level sea ice (low intensity close to the

noise level). First drift calculations showed spurious vectors in these regions. The thresholds were chosen in a way that drift vectors in such regions are clearly marked as unreliable. To this end, a combination of visual checks and histogram analyses for the individual steps of the drift algorithm were performed. The resulting empirical thresholds for each parameter contributing to the CFA are shown in Table I.

TABLE I
OVERVIEW OF THE PARAMETERS CONTRIBUTING TO THE CFA

Parameter		Threshold
Variance-to-squared-mean-ratio	$>$	0.5
Mean intensity gradient	$<$	1.7
Mean gradient slope	$<$	0.35
Intensity threshold	$>$	-3dB
Correlation coefficient r	$<$	0.3
Confidence interval for r	$<$	0.1

Finally we note that the crossing of each single threshold depends on the local image characteristics in each window used for calculating the drift field. A homogeneous ice area, for example, may reveal a certain amount of speckle, no indications of texture or strong reflections, but still a certain degree of correlation with a sufficiently low uncertainty. In such case, only MIG and MGS and possibly the VMR contribute to the CFA. In heavily ridged areas, the image texture is large and the relative contribution of speckle low, but stochastic mirror reflections from tilted ice fragments may bias the correlation (see below). This means that only IT will indicate a problem. However, if the texture patterns have changed in the time interval between the image data acquisitions, the correlation may be as well below the threshold. These examples demonstrate that the use of one single parameter is not sufficient for an adequate assessment of all possible ice conditions for a successful retrieval of the ice displacement. It does hence not make sense to quantify the individual effectiveness of each single parameter. The chosen combination of parameters is balanced to handle a large variety of ice conditions.

C. Parameters for single images (Texture Measures)

Textural parameters can be computed for regions in single images to quantify structure characteristics of image intensity patterns (e.g. high or low frequency grey-tone variations). Methods for describing the properties of image patterns can be divided into statistical, structural, and spectral approaches [14]. In our study we focus on statistical methods.

For the calculation of the mean intensity gradient (MIG) [15] we use the Sobel filter. If there are strong brightness differences, the mean intensity gradient is large, while it is close to zero for homogeneous regions. The mean gradient slope (MGS) is the second derivative of the intensity image, and hence, the slope of the gradient. The MGS as used here has been adapted from [16]. It is calculated by employing a Laplacian filter on the image. The MGS shows spatial changes of the gradient which can be an additional hint for the presence of structures. The benefit of its use is clear looking at the following example: In the case of a constant monotonic

increase of the image intensity inside a given window, the MIG reveals higher values, whereas the MGS is zero. The MGS is more closely related to higher-frequency intensity variation in the image.

The variance-to-squared mean ratio (VMR) is calculated for a given window of intensity values (as are MIG and MSG). It can be used to estimate the contribution of the SAR-specific granular multiplicative noise (speckle) to the image intensity variation if the number of looks of the SAR image is known [17]. Speckle results from the interference of signals from several randomly distributed scatters in on radar resolution cell. Since speckle is reduced by averaging neighbouring pixels, its contribution at the coarse-resolution levels of the drift retrieval can be neglected. However, it may be effective at higher spatial resolutions. Additionally, the VMR includes the contribution of real intensity variations (textural component). For sea ice motion tracking with SAR images, [6] employed the VMR to assess whether variations of the image brightness (radar intensity) in a given window are larger than the value expected due to speckle. We set the threshold of the VMR empirically in order to not only account for speckle but also for spatially dense brightness variation at the high resolution levels of the CRP-scheme which are not related to any ice structures. Its magnitude is large enough to exclude intensity variations caused by speckle at any resolution level.

The brightness threshold limits the maximum intensity within a given window. It takes into account that high intensity values within a pattern have a stronger influence on the resulting cross-correlation result than low intensity values. A single high-intensity peak as it is caused by a mirror (specular) reflection of the radar signal from the ice surface may dominate the resulting cross-correlation. Specular reflections occur, e.g., if larger fragments of ice (size of multiples of the radar wavelength) are oriented such that one side directly faces the radar. Hence, the occurrence of specular reflections is very sensitive to the imaging geometry, i.e. incidence and look angle of the radar, and orientation and shape of the reflecting ice structure. Mirror reflections observed in two consecutive images will hence occur from different locations. As a consequence, the correlation might link two randomly occurring specular events instead of estimating the correct underlying displacement of the sea ice. This limitation of the correlation approach has already been discussed in the context of Particle Image Velocimetry (PIV) [15] and stereo matching [18], [19]. The effect of peaks in radar intensity is reduced using the phase correlation for pre-selection of possible candidates for the pattern matching. However, in order to account for a possible effect of such peaks on the drift estimation we chose a threshold for the local backscatter intensity peak which is not exceeded in the absence of mirror reflections.

D. Parameters for image pairs (Cross-Correlation)

The cross-correlation can be used to assess the similarity of two given windows of an image pair. Its absolute value ranges from 0 to 1. It is 1 if the respective windows are identical and zero if they have nothing in common. Even if one single image contains detectable ice structures, the characteristics of corresponding intensity patterns can vary between two consecutive images, which reduces the chance of successful pattern matching. Therefore it is important to take not only into account the texture of the individual given window but also the correlation between both corresponding windows.

The statistical uncertainty of the correlation coefficient depends on the window size over which it is calculated. Hence, in the CFA we additionally include the confidence interval for the correlation coefficient as a function of the window size. For a small number of samples N in the window (i.e. the window is too small), the correlation coefficient r becomes unreliable. We assume a standardised normal distribution $f(z)$ for the error of the correlation coefficient and estimate the confidence interval Δr at a significance level of $\alpha = 99\%$:

$$\Delta r = \frac{z(\alpha) \cdot (1 - r^2)}{\sqrt{N - 1}} \quad (1)$$

For sea ice motion tracking, the usefulness of this parameter for an assessment of the statistical uncertainty of the correlation coefficient was already discussed in [20].

E. Backmatching

The concept of backmatching originates from the field of photogrammetry and is described by e.g. [21]. For the backmatching, we calculate the respective field of displacements twice, the first one by matching image 1 with image 2 and the second one by reversing the image order. For a direct comparison of both displacement fields, the result obtained from image pair 2-1 needs to be inverted and re-sampled to the grid used to calculate the displacement vectors for image pair 1-2. We emphasize the fact that the potential error sources discussed on page 2 that might inhibit a successful pattern matching apply in the same way for image pairs 1-2 and 2-1. In the ideal case, the resulting displacements are identical in magnitude and direction, and the Euclidean distance between both fields (referred to as backmatching distance) is zero. Under realistic conditions, this distance is larger than zero. It assumes low values for regions where the pattern matching works and larger values for regions where it fails. Contrary to the CFA, this procedure assesses the overall consistency of the retrieved displacements independently of the texture characteristics of the images.

Calculated displacement values are very different, ranging from 0 to tens of kilometres. Therefore it is more descriptive to normalise the backmatching distance with the absolute displacement values. For the drift velocity, the normalised back matching difference $\|\Delta \vec{v}\|$ is calculated from the velocity components u_f, v_f of the forward and u_b, v_b of the backward run:

$$\|\Delta \vec{v}\| = \frac{(u_b - u_f)^2 + (v_b - v_f)^2}{\sqrt{u_b^2 + v_b^2} \cdot \sqrt{u_f^2 + v_f^2}} \quad (2)$$

III. RESULTS

For calculating the accuracy of the retrieved ice drift, data of buoys or drift stations are rarely available, and they provide

only very few paths of motion for a given region. In most cases, one needs to determine spatially dense reference drift fields by visually tracking ice structures that can be identified in both images [9]. We did this to assess the performance of our algorithm. We found systematic errors (mean error) between -100 and 240 m/d, and random errors (standard deviations) between 13 and 430 m/d (Table II).

The reference drift velocities have magnitudes on the order of 9-30 km per day. For the automatically retrieved drift velocities, this translates to relative mean errors below 2% and standard deviations below 8%. An exception is the image pair from February 18/19. For this pair, a relative mean error of 11% and a standard deviation of 21% were determined. If

TABLE II
DRIFT FIELD STATISTICS FOR DIFFERENT SATELLITE IMAGE PAIRS FROM COMPARISON OF AUTOMATICALLY RETRIEVED AND REFERENCE DATA. ALL UNITS ARE (M/D).

Date	18./19.02. 2008	19./20.02. 2008	22./23.02. 2008	30./31.05. 2008	31.05./ 01.06. 2008
Zonal mean error	96	-135	3	26	25
Zonal standard deviation	314	367	139	242	223
Zonal mean velocity	8925	6225	1575	6750	-7800
Meridional mean error	215	311	-27	-96	-162
Meridional standard deviation	597	530	159	228	371
Meridional mean velocity	-2025	23494	24675	28875	12750

we take into account that the spatial resolution of the images is 150×150 m, the mean errors correspond to displacements smaller than 2 pixels. For the standard deviations, we found corresponding displacements of ± 0-3 pixels. The calculated error values are affected by the quality of the automatic drift retrieval and by the quality of the respective manually measured reference data. We estimate the uncertainty in the visual determination of the reference displacement field to be on the order of ± 1-2 pixels (± 150-300 m/d). Figure 3 shows the zonal and meridional components of the drift field for the forward (+) and backward ($\diamond$) run during the back-matching, compared to the respective vector components of the corresponding reference for the image pair 22./23.2.2008. The mean errors for x- and y-components are 0.003 and -0.027 km/d, the corresponding standard deviations 0.139 and 0.159 km/d, respectively. These numbers are valid in forward direction. In backward direction, we obtained -0.486 and -0.002 km/d for the mean error, and 0.542 and 0.408 km/d for the standard deviation, respectively. Since both images were acquired with a time difference of 23.5 hours, it is straightforward to convert the velocity components given in the figure to displacements. Both runs of the backmatching agree well with the reference data in most cases. Larger

deviations indicate that the algorithm did not succeed in finding corresponding ice structure, but in a few cases it was also difficult to identify matches visually. Nevertheless we kept those cases in the reference data since they represent situations in which the retrieval algorithm works less well because of the lack of clearly defined ice structures. In the interpretation of Figure 3 one has also to consider that the positions of the grid points used for the drift retrieval differ between the runs image 1 $\rightarrow$ image 2 and image 2 $\rightarrow$ image 1. For a direct comparison of the two results, the drift field of the second run is interpolated onto the grid of the first run.

Depicted in Figure 4 is an example of an automatically retrieved drift field together with reference vectors plotted on image 1 of the pair (22./23.2.2008) for visual reference. The output of the drift algorithm is represented by red arrows, and the manually collected reference vectors by yellow arrows. Both drift fields agree quite well. Figure 5 shows the respective relative backmatching difference $\|\Delta \vec{v}\|$ based on a forward and backward run of the algorithm. One can clearly identify regions in which the drift fields of image pairs 1-2 and 2-1 compare well, and regions where the search in pair 2-1 provides completely different results. Figure 6 shows the corresponding CFA. It takes high values on the ice shelf and at the margins of the image. Due to this increased number of mean threshold violations compared to the central part of the scene, these areas are potentially less reliable than areas with a low CFA value.

We find a strong consistency between Figure 5 and Figure 6 except for the area of the polynia. Here, the CFA reveals only a weak contrast relative to the adjacent pack ice, whereas the contrast of the normalized backmatching difference is significant. The ice cover in a polynia often reveals identifiable structures that have a strong impact on the CFA. However, ice conditions may change fast, dependent on wind conditions, and hence the algorithm might fail due to a lack of pattern matches even if the texture features of each separate input image are promising for a reliable match.

It is now necessary to define a threshold for the back-matching as well as for the CFA to separate regions where the algorithm works reliably from regions where it might fail. To this end, we analysed 2D histograms which relate the backmatching distances to the CFA for each image pair. An example of such a histogram is shown in Figure 7. This example is typical for all cases we analysed.

The histogram shows three main clusters which we can use to identify reliability thresholds. The CFA represents the mean number of threshold violations for the individual texture parameters over the entire CRP scheme and hence is a measure for the mean reliability of each drift vector. Based on a detailed analysis of the histograms for our time series, we define a vector to be reliable if the CFA value is less than 2. This corresponds to less than 33% threshold violations of all texture and correlation parameters included in the CFA over the entire CRP scheme. If the CFA value for a drift vector exceeds 2 we regard the corresponding drift vector as potentially unreliable. This threshold is supported by the observation that most of the manually collected reference vectors, acquired in regions with a sufficient amount of texture, can generally be linked to CFA
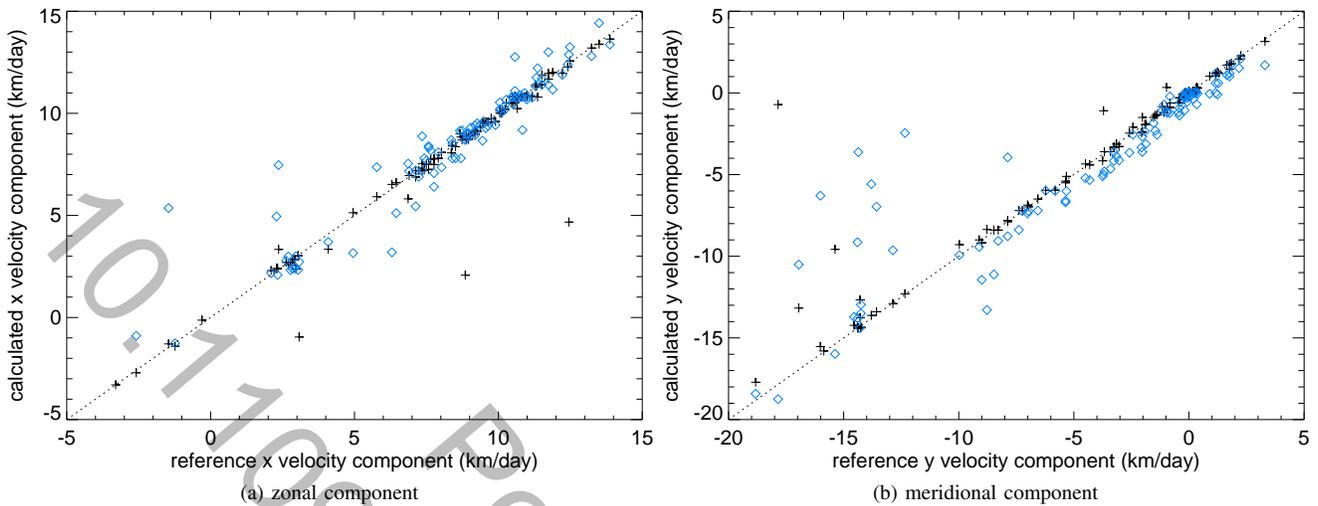
(a) zonal component

(b) meridional component

Fig. 3. Forward and backward calculation of the displacement field vs. manually measured displacements as reference for Image pair 22./23.2.2008 with + representing the velocity component calculated in forward direction and ◇ representing the velocity component calculated in backward direction
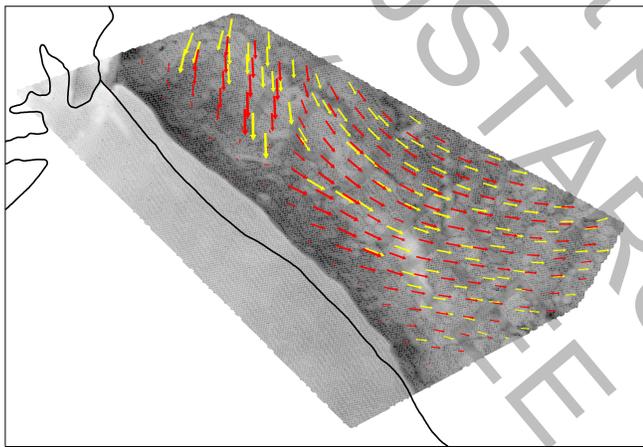

Fig. 4. Envisat WS SAR image taken on the 22.02.2008. The red arrows show the output of the sea ice drift algorithm, the yellow arrows are reference drift vectors. The approximate area of the polynia is indicated in Fig. 1.
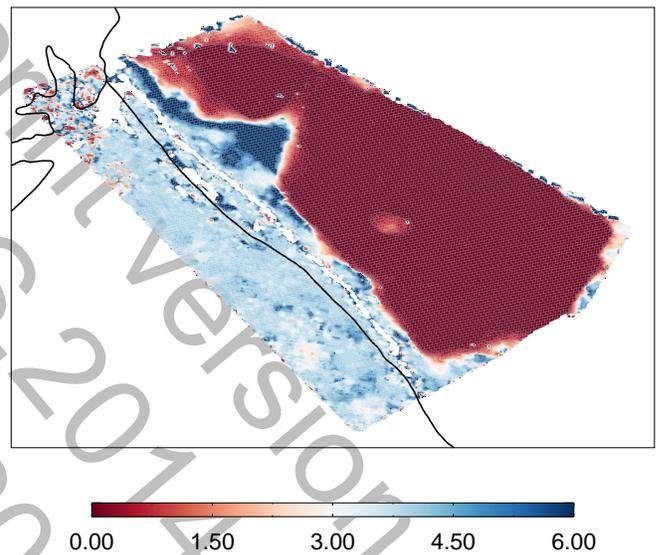

Fig. 5. Normalized backmatching difference according to Eq. 2. The approximate area of the polynia is indicated in Fig. 1.

values below 2. In the case of the normalised backmatching distance we chose a value of 3. The analysed image pairs are very suitable for this kind of analysis since they contain a number of characteristic patterns: the ice shelf with only small amount of texture, the polynia with a high amount of texture but low pattern stability and the pack ice region with many relatively stable textural features.

The chosen thresholds subdivide the histogram into four sectors:

1) CFA values lower than 2 and backmatching differences below 3 indicate that the region contains a distinct image texture that is sufficiently stable for reliable pattern matching. One example for such a region is pack ice, which usually appears highly structured in SAR images.

2) low CFA scores of less than 2 and large backmatching differences of more than 3 are linked with areas which reveal a sufficient texture for correlation but where the algorithm fails because of pattern instability or pattern

loss at image borders. One example for such conditions is thin ice located in a polynia, which is easily deformed and therefore reveals fast-changing ice structures.

3) higher CFA values (> 2) and large backmatching differences indicate the lack of texture combined with a relatively high backmatching difference. This cluster occurs mainly on the ice shelf and at the edges of the overlap area between the two images where the ice drifts into or out of the scene.

4) For CFA-values larger than 2 and backmatching differences below 3 we did not find a distinct cluster in the histogram depicted in Figure 7. However, the margins of clusters 1 and 3 extend into this zone. This class is mainly found at the edges where the ice drifts into or out of the overlap area, but also in the pack ice and and
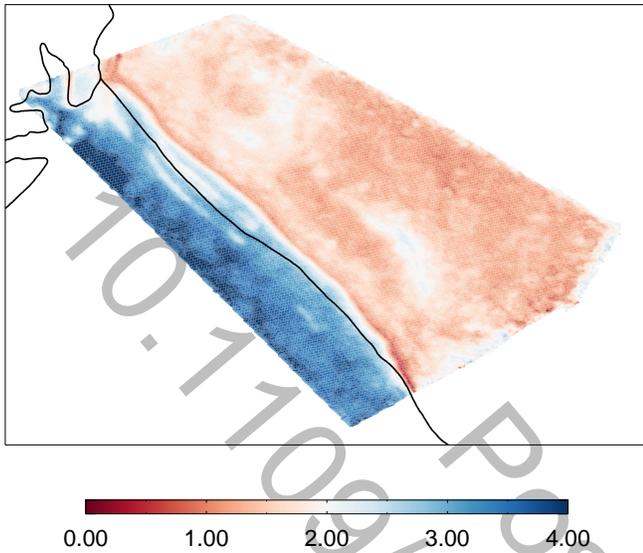
Fig. 6. Confidence factor CFA for the result shown in Figure 4. Larger values of the CFA indicate that drift vectors are less reliable. The approximate area of the polynia is indicated in Fig. 1.

Fig. 8. Map of regions with similar reliability for image pair 22./23.02.2008 based on the separation indicated in Fig. 6 and outlined in the text: • cluster 1 ○ cluster 2 • cluster 3 • cluster 4
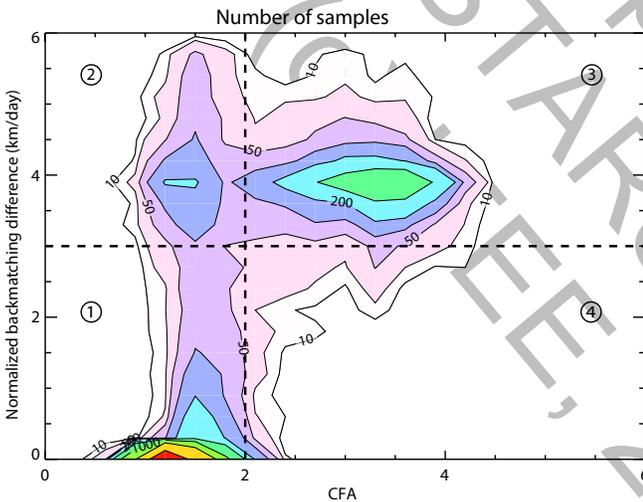

Fig. 7. 2D histogram showing the number of samples for the CFA values plotted over the corresponding normalised backmatching distances for image pair 22./23.2.2013.

on the ice shelf.

To study whether these four zones correspond to certain regions in the image, we employed the thresholds for the normalised backmatching distance and CFA indicated in Figure 7 and mapped them to the corresponding vector position in the SAR image. The result of this approach is visualised in Figure 8. It reveals that three main regions in the image pair correspond to one of the three local maxima and their neighbourhood that can be recognized in Figure 7. Those regions are the pack ice (1: orange), the polynia (2: white) and the ice shelf (3: violet). The fourth group (blue) is linked with a low amount of texture but a small relative backmatching distance. Visual analysis shows that the dark blue (cluster 4) patch in the middle of the orange area (cluster 1) in Figure 8, highlighted by a red circle, is related to a large homogeneous
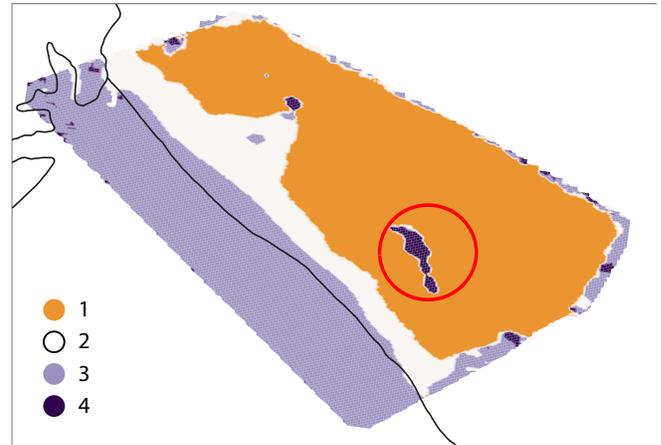
ice floe. Since it drifts together with the surrounding sea ice, it is consistent in the forward and backward run of the backmatching. This is a consequence of the cascaded pyramid approach: the coarse resolution drift constraints the drift at higher resolution. The CFA suggests low reliability for this floe since it reveals no suitable pattern for tracking. At the borders of the image overlap, where patterns enter or leave the overlapping area, we find a mixture of the four different classes due to the lack of actual correspondence in this area.

How can 'reliability' now be linked to accuracy? In our example, areas with normalized backmatching differences smaller than 3 (Figure 5) correspond to the pack ice area which dominates the calculation of the drift field accuracy (in Figure 4 the zone covered by the arrows). Hence in this case 'reliable' means that the expected uncertainty is between ± 0.14 and ± 0.54 km/d (corresponding to the standard deviation obtained in the sequential and reversed run of the drift retrieval algorithm). Except the bight-shaped area of large backmatching differences in the polynia (Figure 5), this is valid for a CFA < 2.0 (Figure 6). If computation time needs to be as short as possible, the CFA can be used as stand-alone without additional backmatching runs. In this case, however, it is difficult to separate strongly textured but unstable sea ice regions such as polynias or the transition from pack ice to the open ocean from stable structures, since such areas might be classified as weakly reliable (compare Figure 6 and Figure 5).

IV. DISCUSSION

The assessment of the retrieved drift fields as compared to manually retrieved drift vectors shows the good performance of the drift algorithm, also for difficult regions like polynias, where existing patterns change extremely fast. It has to be kept in mind, however, that reference vectors can be determined visually only in image regions with well-recognisable structures. The calculated error values listed in Table II are therefore based mainly on those regions where one expects a good performance of the algorithm anyway. Hence one can expect that the overall error over the entire image is larger. Nevertheless, the visually determined reference data are very

useful to prove the applicability of the presented methods. We found that the cross-correlation value itself is not sufficient to judge the reliability of a drift vector (see Fig. 2). One reason is that the result is affected by filtering and interpolation operations in the tracking algorithm. Another point to consider, e. g., is the casual occurrence of mirror reflections (i. e. high intensity peaks) which dominate the correlation (section II C).

The two methods for estimating ice drift reliability (back-matching and CFA) introduced here are fundamentally differ-ent but both are independent of external reference data. In general the reliability check by backmatching is more robust than with the CFA. An example are the already mentioned potentially erroneous drift vectors in the polynia area which are marked as reliable by the CFA-method. If the ice reveals fast-changing visible structures, the drift retrieval scheme can generate high "pseudo-"correlations between different areas. Such structural changes of thin ice areas and the ice in polynias may take place within hours. By taking into account information about wind and ice conditions, polynias and thin ice zones can be marked as unreliable, if the time interval between image acquisition is too large. The drawback of the backmatching technique is a significant increase in algorithm runtime. In operational mapping of sea ice conditions the temporal performance of the ice drift retrieval algorithm is crucial. Hence the CFA, which can be determined "on the fly" may be preferable to backmatching in some cases (section III).

In the study presented here, thresholds for the CFA were derived using Envisat WS data. In additional tests, it was found that they are valid for RADARSAT-2 imagery as well. Since the choice of parameters that comprise the CFA is largely independent of sensor characteristics, they are expected to work also for the upcoming Sentinel-1 mission [22]. Note that all these SAR missions are operated at C-band. We assume that images acquired at another radar frequency (e. g. L-band) require different thresholds, since the appearance of ice structures varies between frequency bands.

Another major issue in this context is the noise level of the image product used for the drift retrieval that might influence some of the employed texture measures (at least at the highest resolution level). The CFA suggested in this work may therefore be extended by introducing additional parameters to increase the contribution of similarity measures to the CFA. One option could e.g. to take into account some of the parameters suggested by [8], that characterize the peaks calculated by phase correlation and assess their appropriateness for a reliable pattern matching.

The major challenge of ice drift tracking algorithms is the difficulty to provide uncertainty values for every retrieved drift field. Alternatively to the methods for reliability checks proposed here, it might be suitable to implement some bundle adjustment techniques from the field of photogrammetry and surveying or approaches described by [23] and [24] which deal with the estimation of uncertainty in SSD (sum of squared differences) based feature tracking and Estimation-Theoretic frameworks for image flow computation. However, to adapt such methods to the cascaded multi-resolution approach is a challenging task.

V. CONCLUSION

In this paper we deal with the automated retrieval of sea ice drift vectors from a pair of SAR images. We found that it is possible to assess the reliability of the drift field from the input image pair. Here, 'reliable' means that the accuracy remains within a certain interval. One possibility to classify drift vectors as reliable is to run the retrieval using the two images of the input pair in sequential and reversed order (backmatching). This approach, however, doubles the computation time. In this article, we propose a 'Confidence Factor' (CFA) for the a-priori analysis of SAR images in order to test whether automatically retrieved drift vectors are reliable. The CFA combines different measures, namely four texture parameters and the correlation coefficient and its confidence interval, into a single score which provides a quantitative number for an assessment of the expected accuracy of the derived drift field. The advantage is that the extra time needed for computation is negligible. A disadvantage of the CFA compared to the backmatching approach is that it may fail to identify areas of fast changing ice conditions (e. g. in polynias and thin ice areas) if the temporal gap between image acquisitions is too large. In such cases, ice structures often exist that are easily detected by the textural analysis. Since those ice structure are temporally unstable, correlations between them may be erroneous. Such areas can be masked by using additional information (e. g. spatial distribution of ice types, wind conditions). We provide thresholds both for CFA and the normalized backmatching difference that separate reliable and unreliable areas. These thresholds are valid for C-band SAR images. Since, e. g. ice deformation structures can be more easily identified at lower radar frequencies, and the radar intensity contrast between old and young ice is larger at higher frequencies, we expect that the thresholds depend on the SAR system used for image acquisition.

REFERENCES

[1] M. V. Thomas, "Analysis of large magnitude discontinuous non-rigid motion," dissertation, University of Delaware, Delaware, USA, 2008. [Online]. Available: http://proquest.umi.com/pqdweb?did=1654490401&sid=5&Fmt=2&clientId=8331&RQT=309&VName=PQD

[2] R. Kwok, "Ross sea ice motion, area flux, and deformation," *Journal of Climate*, vol. 18, no. 18, pp. 3759–3776, September 2005.

[3] T. Hollands, V. Haid, W. Dierking, R. Timmermann, and L. Ebener, "Sea ice motion at the ronne polynia, antarctica: Sar observations vs. model results," *Journal of Geophysical Research*, vol. 118, no. 4, pp. 1940–1954, 4 2013.

[4] R. Kwok, "Observational assessment of arctic ocean sea ice motion, export, and thickness in cmip3 climate simulations," *Journal of Geo-physical Research*, vol. 116, no. C00D05, p. 8, 2011.

[5] F. Massonnet and A. Jahn, "Observational needs for sea ice models," CLiC Arctic Sea Ice Working Group, Short note, 2012. [Online]. Available: www.climate-cryosphere.org/activities/products/406-observational-needs-for-sea-ice-models

[6] R. Kwok, J. C. Curlander, R. McConnell, and S. S. Pang, "An ice-motion tracking system at the alaska sar facility," *IEEE Journal of Oceanic Engineering*, vol. 15, no. 1, pp. 44–54, 1 1990.

[7] A. Komarov and D. Barber, "Detection of sea ice motion from co- and cross-polarization radarsat-2 images," in *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, july 2012, pp. 3277 –3280.

[8] J. Karvonen, "Operational sar-based sea ice drift monitoring over the baltic sea," *Ocean Science*, vol. 8, no. 4, pp. 473–483, 2012.

[9] T. Hollands and W. Dierking, "Performance of a multiscale correlation algorithm for the estimation of sea ice drift from sar images: initial results," *Annals of Glaciology*, vol. 52, no. 57, pp. 311–317, 5 2011. [Online]. Available: http://www.igsoc.org/annals/v52/57/x57A037.pdf

[10] M. Thomas, C. Kambhamettu, and C. A. Geiger, "Motion tracking of discontinuous sea ice," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 5064–5079, 12 2011.

[11] R. Manduchi and G. A. Mian, "Accuracy analysis for correlation-based image registration algorithms," in *IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 5 1993, pp. 834–837.

[12] J. P. Lewis, "Fast normalized cross-correlation," in *Proceedings of Vision Interface 95*, Quebec City, Canada, 5 1995, pp. 120–123.

[13] A. Eckstein and P. P. Vlachos, "Digital particle image velocimetry (dpiv) robust phase correlation," *Measurement Science and Technology*, vol. 20, no. 5, p. 14, 4 2009.

[14] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed., M. J. Horton, Ed. Upper Saddle River, New Jersey: Pearson Education, Inc., 2008.

[15] B. Pan, Z. Lu, and H. Xie, "Mean intensity gradient: An effective global parameter for quality assessment of the speckle patterns used in digital image correlation," *Optics and Lasers in Engineering*, vol. 48, no. 4, pp. 469–477, 4 2010.

[16] L. Pearlstine, K. M. Portier, and S. E. Smith, "Textural discrimination of an invasive plant, schinus terebinthifolius, from low altitude aerial digital imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 3, pp. 289–298, 3 2005.

[17] F. T. Ulaby, F. Kouyate, B. Brisco, and T. H. L. Williams, "Textural information in sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, no. 2, pp. 235 – –245, 3 1986.

[18] M. R. Jenkin, A. D. Jepson, and J. K. Tsotsos, "Techniques for disparity measurement," *CVGIP: Image Understanding*, vol. 53, no. 1, pp. 14–30, 1 1991.

[19] B. K. P. Horn, "Non-correlation methods for stereo matching," *Photogrammetric Engineering & Remote Sensing*, vol. 49, no. 4, pp. 535–536, 4 1983.

[20] M. Fily and D. A. Rothrock, "Sea ice tracking by nested correlations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-25, no. 5, pp. 570–580, 9 1987.

[21] O. Schreer, *Stereoanalyse und Bildsynthese*. Springer, 2005.

[22] W. Dierking, "Mapping of different sea ice regimes using images from sentinel-1 and alos synthetic aperture radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1045–1058, 3 2010.

[23] K. Nickels and S. Hutchinson, "Estimating uncertainty in ssd-based feature tracking," *Image and Vision Computing*, vol. 20, no. 1, pp. 47–58, 1 2002.

[24] A. Singh, "An estimation-theoretic framework for image-flow computation," in *Proceedings of the third conference on computer vision*, Osaka, 1990, pp. 168–177.

Thomas Hollands received his Diploma degree in geoecology from the Technische Universität Bergakademie Freiberg, Freiberg, Germany in 2008 and a Ph.D. degree in geosciences from the University of Bremen, Bremen, Germany in 2012. In 2005 and 2007 he was with the Joint Research Centre of the European Commission, where he worked on 3-D change detection based on very high-resolution satellite imagery. From 2009 to 2012 he was as a doctoral researcher with the Earth Observing Systems group at the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine research (AWI), working on the tracking of sea ice motion from SAR satellite imagery. Since 2013 he works as Postdoc at the AWI and studies the potential of the new Sentinel satellite for Polynia research.

Stefanie Linow received her Diploma degree in geodesy from the Technical University of Berlin (Germany) in 2006. She received the Ph.D. degree in physics from the University of Bremen (Germany) in 2011. Since then, she is employed as a PostDoc at the Alfred Wegener Institute for Polar and Marine Research in Bremerhaven (Germany). Her research interests focus on microwave remote sensing. Current projects include radiative transfer modeling of microwave radiation in polar firn, as well as the analysis of sea ice motion and deformation from remote sensing data.

Wolfgang Dierking received the Diploma degree in geophysics from the University of Hamburg, Hamburg, Germany, and the Ph.D. degree in physics from the University of Bremen, Bremen, Germany, in 1985 and 1989, respectively. His Ph.D. project dealt with radar remote sensing of the ocean. Between 1990 and 1993, he held a post-doctoral position at the Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany, where he worked on microwave remote sensing of ice and snow and on airborne measurements of the sea ice surface topography by means of laser altimetry. From 1993 to 1995, he joined the Remote Sensing Group, Chalmers University of Technology, Gteborg, Sweden. During this period, he focused on ground-based scatterometry and microwave scatter modeling with applications to sea ice monitoring. From 1996 to 2002, he was an Associate Research Professor with the Department of Electromagnetic Systems, Technical University of Denmark. His work dealt with the utilization of polarimetric and interferometric SAR data in geoscientific research. Since March 2002, he has been a Senior Scientist with the Alfred Wegener Institute Foundation for Polar and Marine Research, working on the utilization of remote sensing techniques in climate studies. He participated in several land-based, ship-based, and airborne geoscientific field studies, among them a number of campaigns related to sea ice studies in the Arctic, Antarctic, and in the Baltic Sea.