J.Karvonen
pdftitle=A sea ice concentration estimation algorithm
pdfauthor=J. Karvonen
title Manuscript prepared for The Cryosphere Discuss.
with version 4.1 of the LaTeX class copernicus_discussions.cls.
Date: 13 June 2014

A sea ice concentration estimation algorithm utilizing radiometer and SAR data

Finnish Meteorological Institute (FMI), Helsinki, PB 503, 00101, Finland

Correspondence to: J. Karvonen (juha.karvonen@fmi.fi)

1 Response for the reviewrs 1 comments

Anonymous Referee 1 Received and published: 26 May 2014 In this paper an algorithm is presented to combine segmented SAR imagery with polarization and gradient ratio data from AMSR-2 in order to map an ice concentration value to each SAR segment. Overall the paper is well-organized and well-written. However, there are a few issues that require attention before publication.

Dear reviewer of my manuscript,

Thank You for the good and constructive comments. I have tried to take the Your comments into account in my revised manuscript. In the following are my reponses for the reviewer comments.

p. 2217 lines 11-21 - The authors state that the incidence angle correction does not work over water, but this does not affect the results as long as the open water and ice segments are separated. For the results presented here, they may have been separated, but how robust is this result? Could the authors state under what conditions this might be a problem. For example, what was the range of wind speeds for the image areas and times over which the algorithm was tested?

The segmentation algorithm has been tested in operational use during two winters in the Baltic now. It seems to work well. In vary rare cases they may be problems to distinguish between ice and open water areas into different segments. This of course edpends on the segmentation parameters, and we have selected the parameters experimentally such that distinguishing between water and sea ice is good. If there is a wide range of image intensities (pixel values) in the open water area due to the waves and different incidence angles, we get many open water segments, but still the ice segments are separate from the open water segments. Included some text on this topic.

In addition, how necessary the incidence angle correction? On p. 2222 lines the authors state that ”in the ice covered areas the SAR frame boundaries are not visible indicating that the incidence angle correction for sea ice has been successful” - but these image boundaries were not very visible over the ice in the original mosaics (or perhaps the authors could indicate which

Figure they can see these boundaries in and link this with the incidence angle discussion). Due to the fact that water and ice have different dependencies on incidence angle, one might think retaining this dependence might help the classification.

All the mosaics are made of imagery with the incidence angle correction applied, so the single image boundaries are typically visible only in the open water areas, for example in Fig 3a. Have included a reference to Fig 3a here. Also the ice edge ha been drawn by hand in the image 3a.

p. 2217 lines 26-27 Were any atmospheric corrections applied to the brightness temperatures? All channels listed would have contamination due to atmospheric effects (e.g. water vapour, cloud water, and windspeed). For example, the overestimation by the algorithm seen in Figure 4(a) (blue area in the Gulf of Bothnia) could be due to weather effects - which might be why it shows up when the ice concentration is compared with the ice charts, but not when it is compared to other passive microwave products (ASI ice concentration - which is nearly, but not completely, weather independent).

We have not explicitly applied any weather correction filters. Typically the weather filtering is based on gradient rations of the 36 and 18 and 23 and 18 GHz channels (NASA team, bootstrap and ASI). Our algorithm in based on a neural network and these gradien ratiosn are included in the inputs, so we can assume that the system will also learn tyhe weather filtering, assuming correct reference data in the training and a representative training data set. Our training data set in this experiment was probably not representative with respect to the weather effects and some affects of the weather may occur (also possibly the IC overestimation over open water in the Arctic example of Fig. 7b. Included some text on this topic.

p.2218 lines 16-22 Passive microwave brightness temperatures next to or near a land boundary will contain a signature from both the land and the water or ice (see for example Improving passive microwave sea ice concentration algorithms for coastal areas: applications to the Baltic Sea by Maab and Kaleschke, Tellus (2010), 62A, 393-410.) Unless the contribution from the land can be identified, the brightness temperatures located a specific distance from the land (e.g., the radius of the footprint for the given channel) are usually discarded. It's not clear from the method described on p.2218 if this is taken into account. Could this be contributing to the problem on p. 2226 lines 10-11 - 'in the case of a narrow ice zone near the coast the ice con-

centration is not estimated correctly'? Could the authors state in what way it is not estimated correctly (over estimated or under estimated)? A reference to one of the figures indicating where this problem is noted would be helpful to give the reader an example.

It is underestimated near the coast of Gulf of Finland (e.g. in Fig. 3d). This is probably due to the fact that mixed coastal pixels are left out of our algorithm, and the narrow ice zone in the coastal area can not be seen by the algorithm. The ASI algorithm with a better radiometer resolution sees this ice, for this reason we are also interested in applying our algorithm to data in the full resolution such that the mixed-pixel areas are reduced. Tried to improve this section. also included the reference.

Updated the text in the conclusions section.

p. 2218 line 25 Why was the mode the chosen metric instead of the mean? For example, if there were outliers in the passive microwave data, due to weather effects or variability in surface conditions, choosing the mode may reduce the impact of these outliers on the results.

Mode is describing the typical value within a segment, and the idea is just to exclude outliers. Mean is a measure typically used in Gaussian statistics, if the distribution is Gaussian, mean, mode and median are the same value. Have added a sentence on this. I usually avoid using mean and rather use mode or median which are more robust with respect to outliers.

p. 2221 line 5 and p. 2236 Figure 3 - Could the authors please indicate the resolution of the AMSR-2 bootstrap ice concentration, and the channels used. While the reader can look up the reference, a sentence would be helpful.

Added some information on the bootstap algorithm.

p. 2222 lines1-2 - The reference data were interpolated to the SAR mosaics (resolution 500m) in order to calculate the error statistics. However, this implies that the reference data can be represented at a resolution of 500m, for example a small-scale detail in the SAR mosaic that is not represented in the passive microwave data would show up as an error, but in fact may be correct. Have the authors checked the error statistics carrying out the interpolation the other way around? From this point of view the reference data is left at its original resolution, and the question then is how does the ice concentration from the proposed method compare to this estimate.

Made a comparison in the radiometer (ASI) resolution (resampled our result) also, the numbers are given in the text. This comparison is not very useful from our point of view because the idea of the algorithm is just to improve the precision of the boundaries of different concentration areas using the SAR segmentation.

Comparison in the FMI ice chart resolution (1km) does not differ significantly from the results presented, because the resolution differences are not so large, so they were not included.

p.2224 lines 21 and 22 - The authors state that the set of 10 mosaics is not sufficient to capture the range of brightness temperatures, but seems to be sufficient when using polarization ratio and gradient ratios. Could the authors comment on how they know that their current set 'seems' to be sufficient (maybe they can link this part of the discussion back to that on pg. 2220).

We just first trained the algorithm and then tested it with the training data set. The results using brightness temperatures directly as inputs were even visually poor, but using the ratios instead, the results for the training data set were reasonable with the same training. Some sentences about this added.

p. 2236 Figure 3 - Zooming in on this figure indicates the ice concentration from the proposed method contains more fine scale detail of the ice edge than what is present in the AMSR-2 or ASI ice concentration. An additional figure zooming in on part of Figure 3 would help emphasize this result, and would be complementary to the discussion on p. 2226.

Included a zoomed figure with a detail of the concentration images.

p. 2216 line 10 - It would be helpful to add a sentence or two describing the nature of the ice cover on the Baltic Sea at this time of year. Is it completely ice covered? How thick is the ice? How deformed is the ice? Would wet snow or water be expected on the ice at this time of year (this is regarding the comment on p. 2226 l.15)?

Included some sentences of the Baltic Sea ice characteristics.

p. 2214 lines 20-25 - An advantage of a passive microwave radiometer is that the low frequency channels are not affected by atmospheric conditions.The fact that AMSR-2 has daily coverage over most ice-covered areas is not an advantage of a radiometer, but of the orbit of the satellite carrying the radiometer.

Rephrased this. I think, it is also an advantage of the radiometer because it has such a wide swath.

Minor revisions Acronyms should not be used in the abstract without explanation

Corrected.

p. 2238 In the figure there are 4 panels, but the caption only contains (a), (b) and (c).

Corrected.

p. 2216 Replace 'SAR segmentation' with 'SAR image segmentation'

Corrected.

p. 2218 line 25 upsampled MODIS polarization ratios? should this be up sampled AMSR2 polarization ratios?

Yes, I have been working with MODIS and AMSR2 sinumtaniously and seem to mix things. Corrected.

p. 2215 line 19 - Leigh et al. 2014 used by HH and HV

Included "dual-band" in the sentence.

J.Karvonen
pdftitle=A sea ice concentration estimation algorithm
pdfauthor=J. Karvonen
title Manuscript prepared for The Cryosphere Discuss.
with version 4.1 of the LaTeX class copernicus_discussions.cls.
Date: 13 June 2014

A sea ice concentration estimation algorithm utilizing radiometer and SAR data

Finnish Meteorological Institute (FMI), Helsinki, PB 503, 00101, Finland

Correspondence to: J. Karvonen (juha.karvonen@fmi.fi)

1 Response for the reviewrs 2 comments

Anonymous Referee 2 Received and published: 31 May 2014 Kavonen (2014) present an improved SAR sea ice concentration algorithm that utilizes radiometer data for improved segmentation. The paper covers scientific questions that are relevant for TC. There are a few revisions needed before publication.

Dear reviewer of my manuscript,

Thank You for the good and constructive comments. I have tried to take the Your comments into account in my revised manuscript. In the following are my reponses for the reviewer (2) comments.

p. 2216 line 13. Why is only the HH channel utilized? This is unclear specially since it is also outlined that the AMSR-2 channels have both H and V polarizations.

In this study only SAR segmentation is utilized in addition to the radiometer data and we have good results of SAR segmentation using only one channel (HH). But classification of the resulting segments to open water and ice based on HH channel only is in many cases very difficult or even impossible even when using texture measures in addition to backscattering. However, in this application we use segment-wise radiometer data and we do not have this problem. The segmentation result can slightly be improved in some cases by using both the SAR channels, e.g. we have studied segmentation of a Principal Component (PC) image consisting of the first PC of the two channels. Bu this does not impprove the results significantly, typically only in some minor details. Have added some text about this, in the section on SAR preprocessing.

p. 2216 lines 15-20. Did the same kind of temperature and wind conditions prevail during the two time periods?

No. The weather conditions during the training period were varying, there was a short cold period (1-2 days), then a warmer period and then colder again. During the test period, there were two warmer periods and one colder period (not as cold as the one short training period cold period). Included this information in the text. In Sections "Study area, time and data".

p. 2220 line 14. How many hidden-layer neurons were normally used? What is rather fast?

An optimal number of neurons or units is dependent on the data typically chosen experimantally. We have made selection by increasing the number of neurons and performed training until the training error does not decrease notably any more. Then we have a suitable number of units. Included a sentence on this.

Rather fast means that the classification is available in a reasonable time. In practice this means some seconds on a desktop computer. I have tried to express this more exactly in the revised text.

p. 2221 line 10. How comparable are the sea ice conditions in the Baltic Sea and the Barents and Kara Sea? A description of the study areas "normal" sea ice conditions could perhaps be included in chapter 2?

They differe quite much. Have included some text on these.

p. 2221 line 10. What is a good agreement?

Based on visual inspection the areas of high and low concentration correspond to each other. Edited the sentence.

p. 2222 line 9-18. How do the error and standard deviation estimates presented here compare to previous studies of sea ice concentration? It would be nice to see an expanded discussion on the results from the comparison with the FMI ice charts for the two algorithms (yours and ASI).

There are some comparisons to other EO products, like to concentration from Landsat imagery (Cavalieri et al. 2006), AMSR-E inter-algorithm comparison (Spreen, 2008), and to some visual interpretations from ship (Spreen 2008). Very difficult to make an objective comparison between the results because different things are compared. I included this information in the text.

The new algorithm results are slightly better than the ASI results. This is expectable because we have used FMI ice charts in the training. However, they are quite close to each other. This is also an expected result, because the resolutions are already so close to each other. Have added some sentences on this.

p. 2222 lines 18-21. If possible include a line in the figures indicating where the sea ice edge is so that it is easy for the reader to identify the areas affected by wave conditions.

Ice egde (sketched by hand) added in the SAR mosaic and segmentation images (red line).

p. 2223 line 7-8. Here it is stated that the ASI overestimate the sea ice concentration in the coastal zone. How does this compare to the presented new algorithm results?

Added a sentence on this.

p. 2223 line 11-13. Which Arctic areas? Is it possible to provide an indication in the Figure as to the affected areas? Is this overestimation done by the presented new algorithm or by the ASI concentration estimates?

Tried to describe the location of the area by words.

p. 2223 line 15. Why wasn't a training data set compiled for the Arctic study area as well?

This test was made for the Baltic Sea data, and we had the FMI Baltic Sea high resolution ice charts as our training data set. The arctic test was just made of curiosity on how this method would work in the Arctic with the Baltic Sea training data, and it seemed to work suprisingly well.

p. 2223 line 16. Why were only a visual evaluation made? What does "quite well" imply?

We just wanted to know do the two products look similar (kind of general overview). The idea of the study was to test the new algorithm and to compare to some reference data and we selected the FMI ice charts and ASI radiometer data. "Quite well" means that their generally look the same, but there are different details, because our results are given in a higher resolution with more details (more precise boundaries). Tried to improve text here.

p. 2223 line 23-25. The author indicates that the difference in resolution would affect error estimates. Why wasn't the lower resolution used for such calculations in order to compare the results?

This has also been done now, numbers are given in the text (in section "Evaluation"). However, the basic idea is to improve the location of the area boundaries using the SAR segmentation and SAR resolution. In the radiometer resolution the more precise area boundaries produced by SAR are lost again, so this kind of comparison is not very useful in this case in my opinion.

p. 2223 line 27-29 - p. 2224 line 1. In your opinion what would be the best way to improve the performance, more training data or use the same amount of training data from the Arctic?

More training data both from the Arctic and the Baltic to include different weather and ice conditions, a really replresentative training data set.

p. 2224 line 6. How large were the standard deviations?

There were only two images, very little data also, these were inlcuded in the table now, but I also included a word of warning that the data set was small.

p. 2224 line 11. How does the thickness estimates contribute to the sea ice concentration estimates? This is not clear at present. Changed the text, hopefully it is more clear now.

This just refers to the methodology used for ice thickness, a similar approach could be used for ice concentration.

p. 2225 line 11-15. Would it have been possible to use another data source for the comparison than the FMI sea ice charts as they are based on SAR data?

They are not based solely on SAR data, they are based on other sources also, such as MODIS imagery, and also ice models are used. I think they are the best available information source in the Baltic in a good resolution. We could possibly have used for example the ASI concentrations in the training phase, but at this point we do not have resources to perform this work again with other data sets. Probably we are going to use e.g. radiometer data for training in the Arctic in the future, but we first need resources for this work. With MODIS there is the problem of clouds and cloud masking.

p. 2226 line 11-12. Which areas?

Included one reference to one of the images as an example.

The evaluation would benefit from an indication if certain ranges in sea ice concentration are easier to identify with the different sea ice concentration algorithms. E.g. is the presented new algorithm better in distinguishing high concentration sea ice than low concentration sea ice? Or similar?

The performance is quite similar, no special concentration ranges where the new algorithm can be said to perform better with this data set. The main advantage of the new algorithm is its ability to locate the boundaries of different ice concentration areas more precisely, because the result is in a high resolution defined by ther SAR mosaic. The new algorithm should sharpen the areal boundaries of different concentration areas. This lead to that the concentration distributions of our algorithm have less mid-range concentrations than the other algorithms (especially bootstrap with the lowest resolution), because the blurring at the edges is reduced by our high-

reolution algorithm. In the Arctic test area our algorithm overestimated the concentration in some open water areas compared to the other algorithms. This was probably due to weather and can be improved by training the MLP with a representative training data set (including Arctic data of different weather conditions).

Figures. The figures are generally quite small and it is difficult to see all the details described in the text. Maybe some kind of indication in the figures as to what the author is talking about could be included, maybe a reference to a position in Lat/Long or similar.

I have tried to describe the locations in the figures better in the text. Because this will be a digital publication, the reader can zoom in the figures to see details.

Generally some subjective words such as; little, worse and some could be changed into more precise indications.

Have tried to improve these.

Minor revision; p. 2219 line 4-5. "The boundaries of different..." This is not clear.

Just meaning that the new algorithm will produce sharper boundaries based on the SAR segmentation. Tried to improve this sentence.

p. 2219 line 9. The abbreviation MLP is explained already on page 2216 line 5.

Removed the second explanation.

p. 2219 line 11. What is the resolution of the FMI gridded ice charts?

The nominal resolution is about 1000m. Included.

p. 2221 line 6. What is the resolution of the AMSR-2 bootstrap algorithm?

The resolution is about 10 km. Included.

p. 2221 line 6. Bootsrap needs to be changed to bootstrap.

Changed.

p. 2221 line 12. Figure 5 is referred to before Figure 3 and 4.

Moved the reference later.

p. 2221 line 12. Should the reference to Figure 5 be changed into Figure 7?

Yes, it should be Figure 7.

p. 2222 line 18. Figure 5 is again reference before Figure 4. Order of Figures needs to be changed.

Order of figures changed.

p. 2222 line 23. ”In these gures...” Which gures, gure 3 and 5?

Yes, 3 and 5, in the updated version 3 and 4.

p. 2222 line 28. Does ”...the bootstrap algorithm result in both cases.” Refer to the Gulf or Riga and the Gulf of Finland?

Refering to both the images and Gulf of Riga. Changed ”both cases” to ”both cases of February 2, and February 11”.

p. 2222 last sentence - p. 2223 first sentence. This is not clear.

Rephrased this.

p. 2224 line 16-18. ”...the MLP convergence was slower and estimation results worse...” How much slower and how much worse?

We did not get reasonable results for our test data set with this training data set. The L1 errors were 30-40 percentage points. Also the algorithm did not converge to a a reasonable error level with even 100000 epochs. Added some text about this.

p. 2225 line 19. How much is a ”little multi-year”?

Less than 1 percent of the whole area. Small portion in the northeastern part of the image. Included this rough number in the manuscript.

Tables. The tables would benefit from a more comprehensive table captions. E.g. indication to test areas, datum for the different data sets, and an indication if the results are presented for the algorithm results compared to the FMI ice charts etc.

Have improved the captions, more detailed information on the projections have been included in the section describing the data.

Figure 1 and 3-8. Inclusion of the coordinates would be beneficial, e.g. lat long or similar on the axis.

Have included scales in the images and their explanations. Unfortunately I do not have good tools to include numbers in a suitable scale into the imagery automatically, so the scales are described in the image captions.

Figure 7. Which areas are not covered by the SAR mosaic?

The black areas are either land areas or not covered by the SAR mosaic.

Figure 8. Here the proposed algorithm is referred to as the FMI algorithm. This is the only instance where that happens.

Changed ”FMI” to ”our”.