Dear authors of the manuscript tc-2021-85,

In the manuscript a widely studied topic has been studied. The method is computationally quite heavy but the results in classification are good. The manuscript is quite thorough, the data set and the evaluation are quite comprehensive. There are still some aspects which need to be taken into account before publishing this manuscript. In the following are my comments.

We are grateful to the reviewer for the constructive comments on our manuscript (tc-2021-85) entitled “Sea ice and water classification on dual-polarized Sentinel-1 imagery during melting season”. We have addressed all the comments. Our point-by-point responses are attached below in blue, while the original Reviewers’ comments are in black.

Thank you again for valuable comments on our manuscript.

Sincerely,

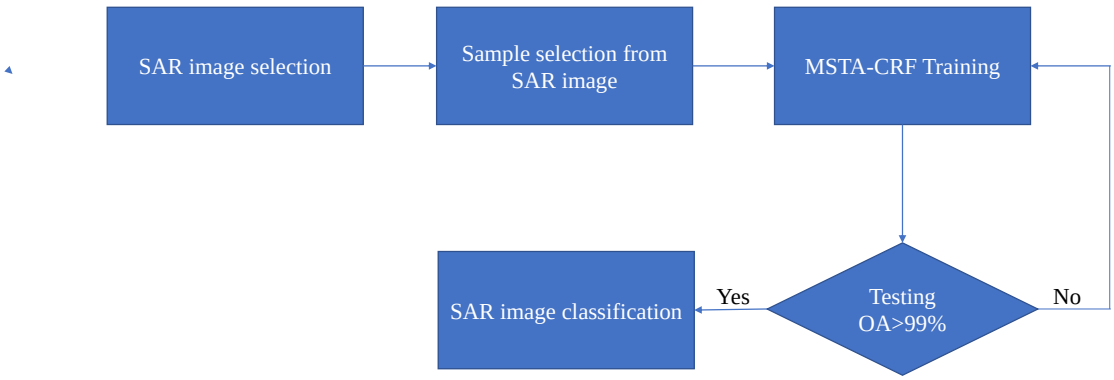
Test and training data set: It is not very clear how the data has been divided into independent training and test data sets. Evaluation should be performed using a test data set which is independent of the training data set, i.e. the training data set must be excluded from the test data set. Now division into these two independent data sets is not very clear to me. Please, in detail describe the division to independent training and test (evaluation) data sets to confirm the reader that they are independent.

Response: For MSTA-CRF model training, one Sentinel-1 SAR image on each day from June to Sept in 2015-2018 was selected. We first randomly select 10 samples of each category (sea ice and water) from the selected SAR images to construct the training and testing data set (9760 samples in total). During the training procedure, we randomly selected 100 samples for each category from the training and test dataset to train the model, and the rest of the samples is then used to verify its accuracy. If the accuracy is lower than 99%, 50 samples are added for each class to update the model, and these added training samples are removed from the test samples until the final classification accuracy on the test data is better than 99 %. We have repeated the training procedure ten times and found that when the number of training samples reaches 1000 the accuracy is over 99%. We have finally selected 1000 samples for training model (500 for each category), which accounted for 10.25% of the entire training and testing data set. The following table describes the procedure of MSTA-CRF model training. The table and the corresponding flowchart below will be included in the next version of the manuscript.

| | |
|---|---|
| Step 1 | SAR image selection<br><br>One SAR image on each day from June to Sept in 2015-2018 is randomly selected to construct the training data set, finally we get 488 images. |
| Step 2 | Training and testing data set construction: |

| | |
|---|---|
| | 10 patches (samples) for each category (ice and water) with the size of 64*64 pixels are randomly selected from the 488 SAR image using MET Norway ice charts, then we get 9760 patches for constructing the training and testing data set. |
| Step 3 | MSTA-CRF training:<br><br>100 patches for each category are selected for training the MSTA-CRF model, and the rest are used as testing samples to decide by the overall accuracy whether the training will be repeated. |
| Step 4 | Testing:<br><br>If the overall accuracy of the testing samples is larger than 99%, then we get the final MSTA-CRF model, otherwise 100 patches (50 for each category) will be added to retrain the MSTA-CRF model, and the newly selected 100 patches will be removed from the testing samples. |
| Step 5 | SAR image classification:<br><br>Repeat step 3 until we train a satisfied model, and the newly trained model will be used for sea ice and water classification on all the SAR images. |

We also give the flowchart of the training procedure in the following figure.




Introduction:
Also sea ice concentration (SIC) estimates can be and are derived based on the proposed SI/OW classification scheme. I recommend to include missing references to SAR-based SIC estimation, there are many papers on this published during the recent years, e.g.:

Wang, L., K. A. Scott, L. Xu, D. A. Clausi, Sea ice concentration estimation during melt from dual-pol SAR scenes using deep convolutional neural networks: A case study, IEEE Trans. Geosci. Remote Sens., vol. 54, no. 8, pp. 4524–4533, 2016.

Wang, Scott, Clausi, Sea Ice Concentration Estimation during Freeze-Up from SAR Imagery Using a Convolutional Neural Network Remote Sens. 2017, 9(5), 408; https://doi.org/10.3390/rs9050408

W. Aldenhoff, A. Berg and L. E. B. Eriksson, "Sea ice concentration estimation from Sentinel-1 Synthetic Aperture Radar images over the Fram Strait," 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 7675-7677, doi: 10.1109/IGARSS.2016.7731001.

Karvonen, Evaluation of the operational SAR based Baltic Sea ice concentration products, Advances in Space Research 56(1), 2015, DOI: 10.1016/j.asr.2015.03.039

And some references combining microwave radiometer and SAR for SIC estimation:

Karvonen, J., Baltic Sea Ice Concentration Estimation Using SENTINEL-1 SAR and AMSR2 Microwave Radiometer Data, IEEE Transactions on Geoscience and Remote Sensing (Volume: 55, Issue: 5, May 2017), pp. 2871-2883, 2017, DOI: 10.1109/TGRS.2017.2655567.

Malmgren-Hansen, D., Pedersen, L. T., Nielsen, A. A., Brandt Kreiner, M., Saldo, R., Skriver, H., Lavelle, J., Buus-Hinkler, J., Harnvig, K., A Convolutional Neural Network Architecture for Sentinel-1 and AMSR2 Data Fusion. IEEE Transactions on Geoscience and Remote Sensing, v. 59, n. 3, pp. 1890-1902. 2021, https://doi.org/10.1109/TGRS.2020.3004539

Especially convolutional neural networks in sea ice classification and parameter estimation have gained popularity during the recent years. These methods are computationally heavy but software for their parallel efficient execution on graphics adapters exist.

Response: Thanks for your comments. We will include some summary of the papers about sea ice concentration retrieval based on microwave data fusion, convolutional neural network based sea ice classification, as well as the above mentioned papers in the revised version.

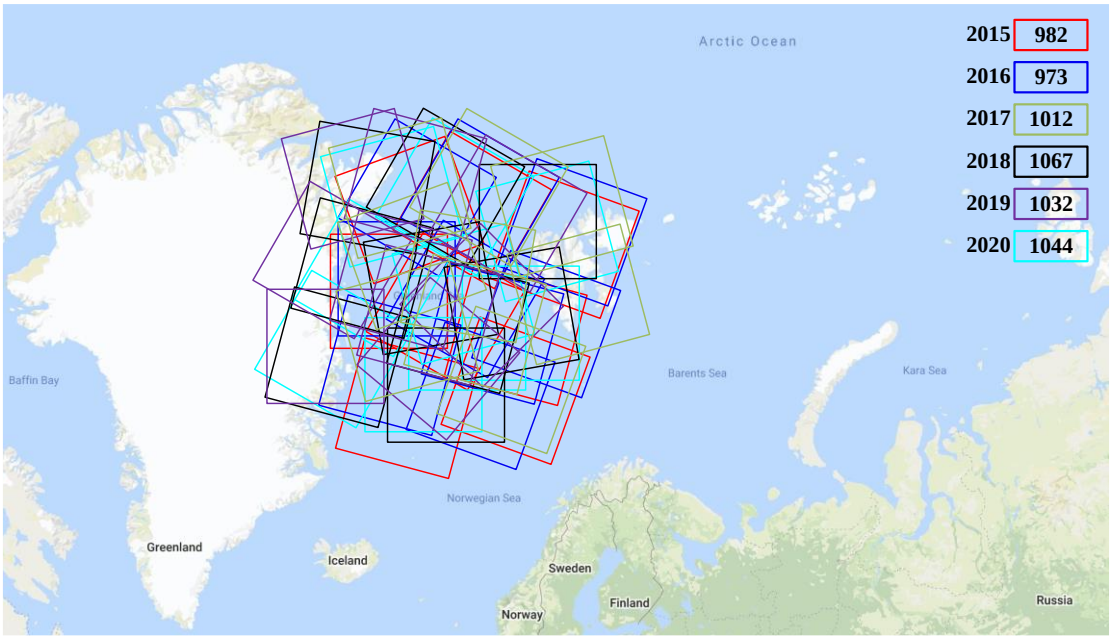
P4 2.1 Research area:

The sentence "To consider the spatial contextual information and preserve the spatial details of each pixel in SAR imagery, the energy function based maximum a posteriori (MAP) estimation in MSTA-CRF framework is proposed for operational ice water classification during melting seasons in Fram Strait." does not belong to this subsection, it could be in introduction or methodology section rather. Just start the section by "This study was performed in the area of Fram Strait during the melting season." or something similar.

Response: We agree with the comments, and we will rewrite the sentence in the manuscript.

P4 L21: "Figure 1 shows an overview of the research area and some satellite scenes used in this manuscript." and Figure 1 / Figure 1 caption.

Why just some scenes are shown? Could the figure for example show the total amount images at each location of the study area (by using some color coding), it would be much more informative.

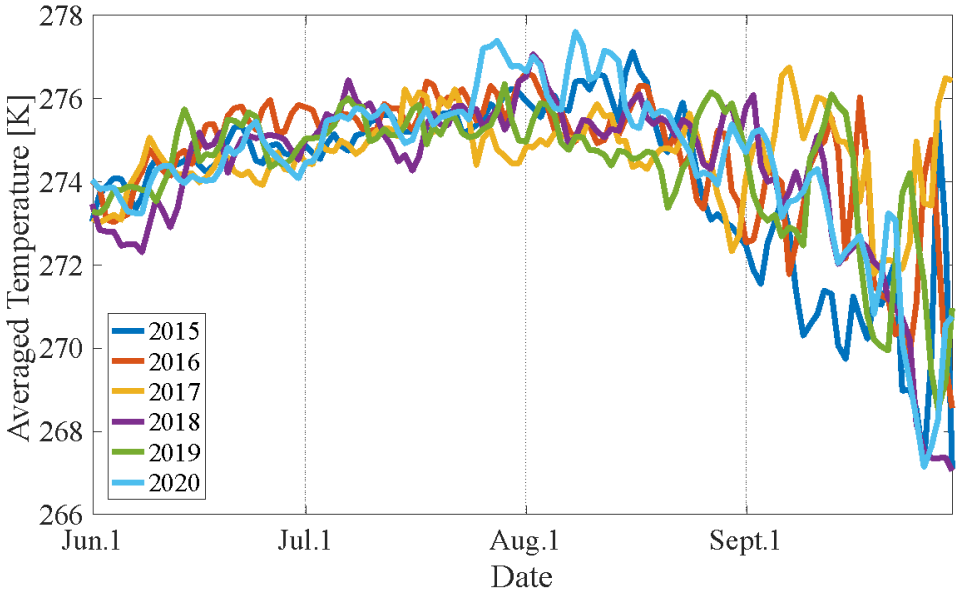



Response: We provide a new figure 1 in the revised manuscript. The different colors of the rectangles indicate the SAR images acquired in different years, and the number in the rectangle on the top right of the figure means the number of SAR images used for sea ice classification in the corresponding year.

P5 Sentinel-1 SAR Data:

L6: "... data during melting seasons from 2015 to 2020 are used.". Please, be more specific, give the periods. Is the melting period the same every winter? E.g. some kind of temperature statistics from nearby weather stations to confirm that the data represents melting period every winter would be useful here.

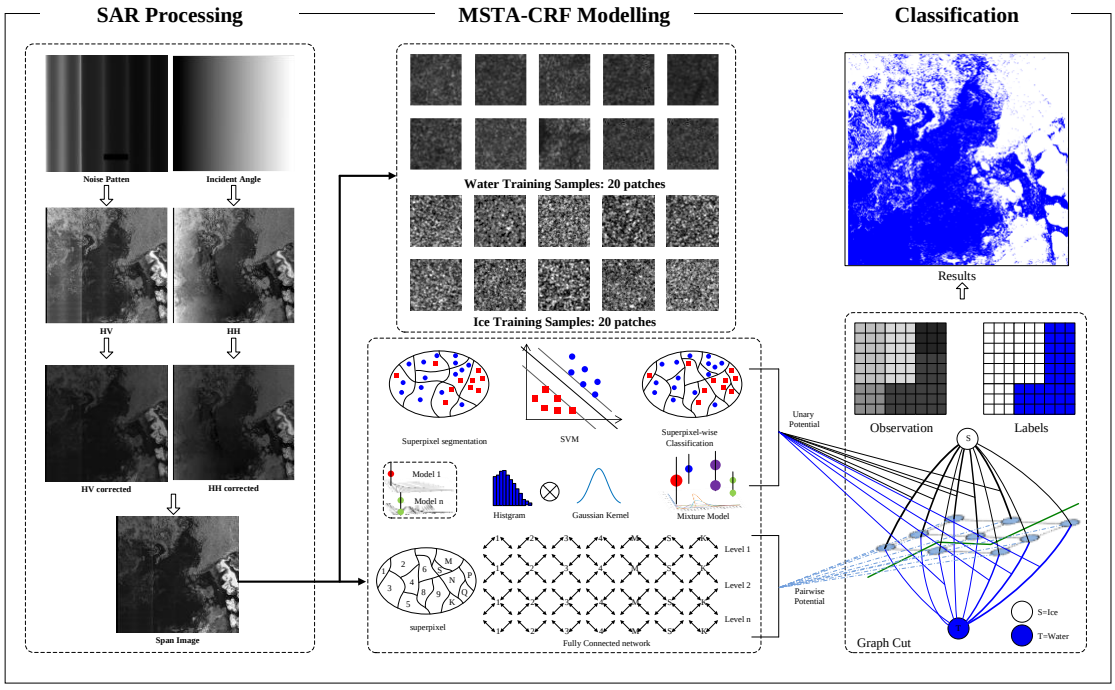
Response: It is not our objective to define the melting season, our purpose is to propose an algorithm for sea ice classification in melting seasons, and it can also be used for sea ice classification in other seasons. To analyze the melting condition in Fram Strait, we have downloaded the hourly averaged ERA5 2-meter temperature (spatial resolution of 0.25°) from the ECMWF website. We illustrate the daily averaged temperature in Fram Strait in Figure 2. It is clear to see that the temperature starts to increase from the beginning of June, and reaches top in the beginning of August, then it starts to decrease and finally drops below 268K at the end of September. Except for the year 2015, when the surface temperature falls below 273K at the beginning of September, in 2016-2020 the surface temperature falls below 273K in the middle of September. A surface temperature above 273K means that the sea ice is still in the melting condition. As a result, we have selected the Sentinel-1 SAR data from June to September each year and defined this time span as the melting period in Fram Strait.




P6 Methodology:

Figure 2. If I have understood correctly SPAN image is used as an input? Now in the figure there is an arrow from the leftmost SAR processing block to the MSTA-CRF block and it looks like the uppermost row SAR data were input to the MSTA-CRF, possibly the arrow could be started from the lower part of the block as is the second arrow. Assuming I have understood this correctly.

Response: You are right. We have revised Figure 2.



P5 L5: SPAN of the HH and HV channels (sqrt(HH^2 + HV^2)). Later SPAN is defined as square root of sigma0_HH^2 + sigma0_HV^2. Possibly the square root could be dropped from here and just say that SPAN represent the joint total power of the two SAR channels and leave the more precise definition later.

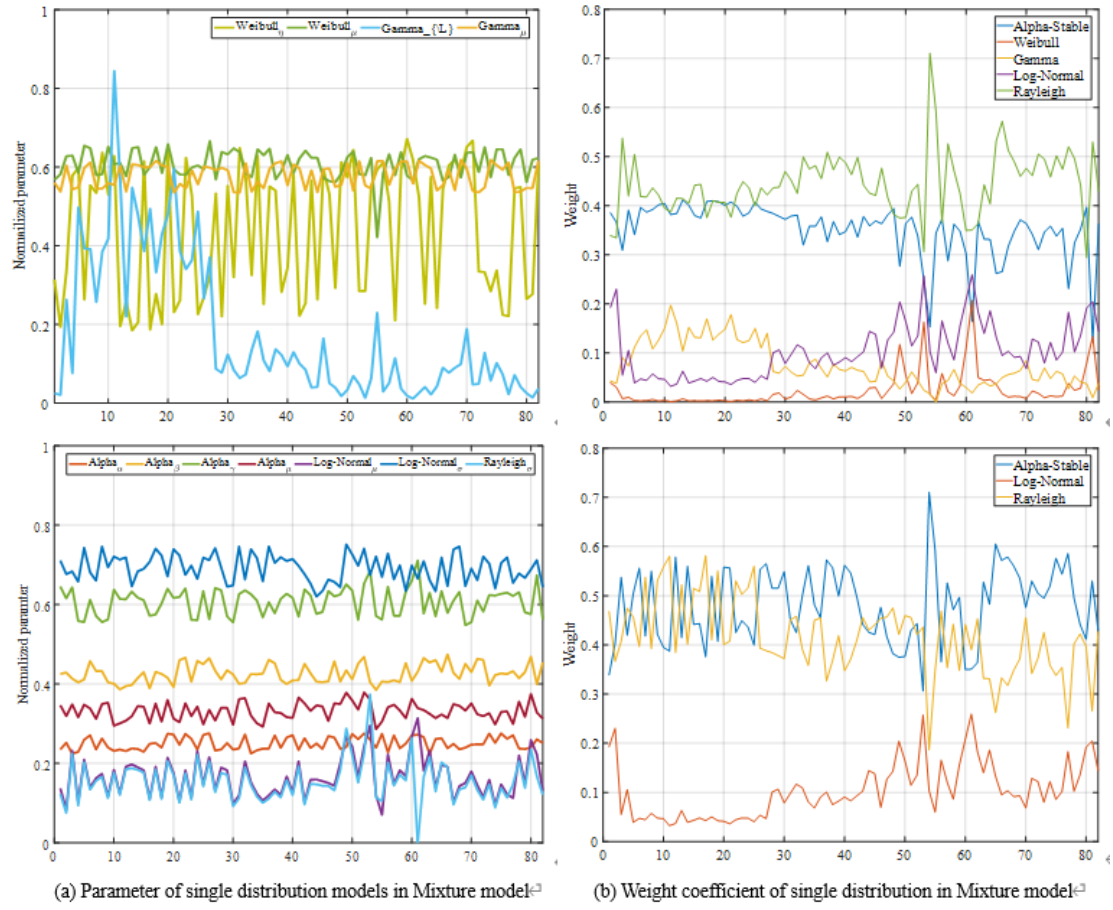
Response: You are right, it is sqrt(HH^2 + HV^2).

P9 L19: "using the MET Norway ice charts". Please, be more specific and describe exactly how the ice charts have been utilized.

Response: MET Norway ice charts provide manually classified sea ice categories daily. The reference (Zakhvatkina, 2017) also uses the MET Norway ice chart products for training and verification of sea ice classification. Therefore, we chose the MET Norway ice chart in the paper for sample selection and validation. In order to improve the accuracy of sample selection, we have also combined the visual inspection to improve the accuracy of sample selection. Besides, we have also analyzed in the paper that due to the difference between the SAR data acquisition time and the MET Norway ice chart acquisition time, the drift and freeze-thaw changes of the sea ice also affect the classification accuracy

P14 Fig. 6: Be more specific in Y-axis label, now there is just "normalized". "normalized" what? I guess "normalized parameter" would be better here. Fig. 6a is not very clear with so many curves in one figure. Would there be any alternatives to make a more clear image (or more than one image)?

Response: We have revised Figure 6(a), where in the top left we only give the normalized parameters of Weibull and Gamma (these two models are not used in MSTA-CRF modelling), and the remaining three models in the bottom left. The normalized parameter means that we have adjusted the value of the parameter to the range from 0 to 1.




(a) Parameter of single distribution models in Mixture model (b) Weight coefficient of single distribution in Mixture model

P13 23: "CV (coefficient of variance)"? Do You mean "coefficient of variation"? At least for me coefficient of variance is an unknown concept. If You use it, please, define it.

Response: You are right. CV in the manuscript means coefficient of variation, and it is defined as the ratio of the standard deviation and the mean accuracy, as can be seen in the following reference, and it will be added in the next version of the manuscript.

Keller M R, Gifford C M, Winstead N S, et al. Active/Passive Multiple Polarization Sea Ice Detection During Initial Freeze-Up[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020.

Minor:

P9 L13: "Figure 5 (e)", You probably mean Figure 4 (e)?

Response: You are right. It is Figure 4(e).

P9 L18 "for each category", as there are only two categories it would be better to say "for both categories".

Response: We have corrected it.

P10 L2: "each categories" -> "both categories"

Response: We have corrected it.

P10 L10: "...that when the training samples reaches..."? Do you mean "...that when the number of training samples reaches..."?

Response: We mean that when the number of training samples reaches 1000 (500 samples for each category) the overall accuracy is over 99%.

P10 L32: "...contains a great number of scatters of radiation..." -> "...contains a large number of scatterers of radiation...".

Response: We have corrected it.

P11 Table 3 caption: "...probability density function (pdf)..." -> "...probability density function (PDF)..." or rather even "...probability density function..." and give the acronym PDF in the text.

Response: We have corrected it by using probability density function (PDF).

P11 L12: "PDF" -> "probability density function (PDF)", PDF always with capital letters.

Response: We correct it by using probability density function (PDF).

P11 L22 and Eq. 8: Explain what is M (is it number of PDF's here?).

Response: M in the equation means the number of PDFs. In the manuscript, we finally use three different PDFs including Alpha-Stable distribution, Log-Normal distribution and Rayleigh distribution.

P12 L6: "...into several sub-superpixel using a random number..."? What does this mean? "...into several sub-superpixel using a random number of pixels..."?

Response: You are right, the superpixel will be segmented into several sub-superpixels using a more restrict mean-shift procedure, and the maximal number of sub-superpixel within a given superpixel is less than 50.

P12 Eq. 9: What are K and N in the equation?

Response: K is the number of kernels in the fully connected network and N is the number of pixels. We will add this information to the text.

P12 "...y_i and y_j are the SAR backscatter coefficients at a pair of subsuperpixel..."? Are these really SAR sigma0 values or SPAN values?

Response: it is sigma0.

P14 Fig. 6: Add x-axis labels ("model number" or something describing what is on the x-axis).

Response: We add it.

P15 L17-18: "PDF (probability density function)" This has already been opened on p. 11, so just write "PDF".

Response: we remove it.

There seem to be some sentences which are not very easy to understand. I am not a native English speaker and may not have noticed all of these sentences or possible grammar or typing errors. I recommend to let a native English speaker (Your co-author Nick Hughes) to check the sentences and language of the revised manuscript before submission.

Sincerely,