Anonymous Referee #3

First of all, we would like to thank the reviewer for the providing important comments. We hope you reconsider your decision based on our responses and modifications made to the manuscript. Please find below our answers (in green) and modifications (deleted in red and added in blue) to your comments/suggestions/questions.

General comments

The manuscript by Park et al presents Random Forest based classifier (Python Scikit-Learn) for sea ice type classification from dual pol (HH-HV) Sentinel-1 images which were collected during the winter period only. Training dataset were collected from the National Ice Center weekly ice chart and the classification algorithm exploits standard GLCM features along with some additional features. Since the launch of Sentinel-1 SAR sensors (a+b), it continuously monitoring Arctic Sea ice with high spatial and temporal resolution and an automated sea ice type classification product from high resolution SAR is highly desirable and relevant for the sea ice monitoring community. Having said that, there are several limiting factors which are preventing researchers to come up with a robust SAR based sea ice type classification scheme, (1) Backscatter variation due to varying incidence angle, along with sensor specific noise related issues (2) Backscatter variation due to seasonal changes (winter–melt- early summer - summer). In this manuscript authors performed a denoising technique which was developed by authors previously (Park et al., 2018), and a standard linear incidence angle correction (Section 2.2.3). It is important to note here that different sea ice types have different incidence angle dependency. Moreover, if the incidence angle correction was applied over all classes (i.e. including open water), it will most likely not contribute to the robustness of the classifier. Backscatter variation due to seasonal changes is completely ignored in the presented manuscript. As authors aimed to develop an operational system, in my opinion authors cannot ignore this major issue completely.

First of all, we would like to thank the reviewer for the providing important comments. Please find below our answers (in green) and modifications (deleted in red and added in blue) to your comments and suggestions.

- The use of Sentinel-1 in ice type classification is highly demanded for sea ice monitoring community but has not going well because of the two reasons that the reviewer pointed. We partly solved the problem of radar backscatter variation due to sensor noise issue, and this is one of the main significances of this study.

- The fact that different sea ice types have different incidence angle dependency is well known, but it is not true that applying the incidence angle correction over all classes will not contribute to the robustness of the classifier. Typically, the slope for open water is higher than that for sea ice, thus the correction works in a way reducing the backscatter variation for open water as well. Moreover, ice type-specific correction prior to ice type classification is controversial.

- As the reviewer pointed, backscatter variation due to seasonal changes is important for operational ice charting. Since the developed algorithm was tested for winter season only, we changed the title as "Classification of **<u>Winter</u>** Sea Ice Types in Sentinel-1 SAR images" and the limitation related to the variations in radar backscattering and its corresponding image textures due to seasonal changes is added to the end of Section 3 in the revised manuscript.

- The only variable that we introduced to capture the seasonality, day of year, might not correspond to the same temperature, fluxes and weather regimes. Weather variant input parameters may be more suitable. However, as sea ice drifts continuously, weather variant information for each of the ice floes in the SAR image at the image acquisition time needs to be calculated and joined, and this may require complex and rigorous works. In order to simplify the problem, we used day of year only, but certainly other parameters need to be tested in the follow up paper. We are very interested in combining sea ice drift with ice type-specific texture changes in our future study.

Specific comments

The proposed classification scheme is based on Python Scikit-Learn library and GLCM features, this kind of classification scheme is well known and published several times for different frequency bands. Therefore the current manuscript is very limited in terms of innovation. What I find slightly different is use of weekly ice chart for training data generation.

Our manuscript includes several significance and innovations:

i) This is the first study to examine the classification of Sentinel-1 SAR images for determination of sea ice types.

ii) This study demonstrates the ability to classify ice types from the noisy Sentinel-1 image by adopting our previous development, textural denoising method, and in the revised manuscript, we show how the denoising methods they have developed lead to improved ice type classification.

[Section 3]

…

To see how the denoising step in Section 2.2.2 led to improvements in the classification accuracies, the same training and evaluation was conducted for the same dataset without applying the textural noise correction, and Table 4 shows the results. In both FC1 and FC2, the improvements in accuracies for young ice (+8.2-9.8%) and first-year ice (+9.2-11.6%) were most pronounced compared to those for open water (+1.7%) and old ice (+1.2-1.7%). On the contrary, a small decrease was observed for new ice (-2.8-4.7%). Nevertheless, the improvement in kappa (+0.05) demonstrates clear improvement in the overall classification result.

Table 4: Changes in classification accuracies before and after applying textural denoising

class	case								
	FC1			FC2			FC3		
	Thermal denoising only	Textural denoising applied	difference	Thermal denoising only	Textural denoising applied	difference	Thermal denoising only	Textural denoising applied	difference
OW	88.4	90.1	+1.7	88.9	90.6	+1.7	88.0	85.4	-2.6
NI	30.2	28.0	-2.8	27.7	23.0	-4.7	31.8	23.9	-7.9
YI	34.9	44.7	+9.8	36.2	44.6	+8.2	43.4	51.5	+8.1
FYI	29.3	38.9	+9.6	30.4	42.0	+11.6	38.0	47.0	+9.0
OI	91.5	92.7	+1.2	90.3	91.7	+1.4	75.2	66.3	-8.9
kappa	0.62	0.67	+0.05	0.62	0.67	+0.05	0.54	0.49	-0.05

iii) the use of public ice chart for training is new, and the difference from the conventional approaches were discussed to the introduction section.

[Section 1]

…

In ~~most~~many of the previous works on ice-water and/or sea ice classification (Soh and Tsatsoulis, 1999; Zakhvatkina et al., 2013; Leigh et al., 2014; Liu et al., 2015; Ressel et al., 2015; Zakhvatkina et al., 2017; Aldenhoff et al., 2018), the training and validation were done using manually produced ice maps. Although the authors claimed that the manual ice maps were drawn by ice experts, the selection of SAR scenes and interpretation can be ~~subjective~~inconsistent, and the number of samples were not enough to generalize the results because of the laborious manual work. Therefore, increasing objectivity is crucial, and automating the classification process is encouraged. The idea of training using SAR images and accompanying image analysis charts, which is raw interpretation of SAR images by trained ice analysts working at operational ice services, were tested for sea ice concentration estimation by Wang et al. (2017); however, such image analysis charts are not accessible to public.

However I am also concerned about the automated training data generation as it is not clear which images were used to generate the ice chart by NIC and there is a high probability that that the ice chart polygons will not match that Sentinel-1 mosaics ice types. Hence there is a high risk that the classifiers were trained with wrong training data. Authors mentioned that 57 images were selected manually where only 'ice egdes' match well with the ice chart. In my opinion, this is also a manual selection of training data which authors criticized in the introduction section.

We criticized conventional approach for two reasons: selecting training/testing data by anonymous human ice expert can be subjective, and preparing a large volume of such dataset is laborious. If public ice charts are used for training and validation, these two issues are partly resolved. Although the public ice chart is also made by human expert thus it cannot be completely free from subjective decisions, training using a large volume of these charts would reduce operator-to-operator bias. Another advantage of the use of public ice chart is that the best knowledge of certified ice analysts rather than anonymous expert is to be mimicked through the machine learning process. Regarding the manual selection of training data, the visual inspection along the ice edges overlaid on SAR image is also not automated and this is why the proposed method is "semi-" automated algorithm, but it requires much less effort and expertise compared to drawing manual ice chart from the SAR image itself.

[Sections 2.2.6]

…

~~To automate image selection for training, a good ice/water classifier for SAR image is needed.~~ In order to automate image selection, the ice edges in SAR images needs to be identified first. Since ~~even such a simple binary~~ an ice/water classifier has not been well developed yet for Sentinel-1, the image selection procedure has to be done manually in the beginning. However once a classifier is generated with high accuracy, it can be used to automate the procedure, then the whole process in the proposed scheme will be fully automated. This is why the proposed algorithm is named "semi-" automated for now. Nevertheless, the manual selection is done by visual inspection of ice-water boundaries overlaid on SAR images. The ice-water boundary can be extracted easily from the reprojected ice chart by selecting the pixel borders of open water class. Then the SAR backscattering image contrasts across the ice-water boundaries are examined both in HH- and HV-polarization because the image contrast between ice-water is larger in HV but smooth level ice is better recognizable in HH.

The selection (and definitions) of ice types for SAR based sea ice classification scheme is crucial. The 5 class classifier has some classes which might be very close to each other in terms of backscatter and texture. This might be the main reason for significantly low classification accuracy. I would recommend the authors to restrict the classifier for 4 classes (Open Water, Young Ice, FYI and old ice).

As the reviewer pointed, the five classes do not fit very well to the characteristic radar signatures from sea ice. We have tested for the four classes and the results are as below.

		Predicted											
		OW (open water)			NYI (New + Young ice)			FYI (First-year ice)			OI (old ice)		
	Case	FC1	FC2	FC3	FC1	FC2	FC3	FC1	FC2	FC3	FC1	FC2	FC3
Actual	OW	**91.9**	**92.1**	**87.3**	2.0	2.3	6.0	6.1	5.7	6.7	0.0	0.0	0.0
	NYI	11.2	10.2	10.3	**47.7**	**49.1**	**54.8**	32.7	34.2	30.2	8.4	6.6	4.6
	FYI	5.5	4.6	27.8	26.8	27.4	27.8	**41.0**	**43.5**	**48.1**	26.7	24.5	19.1
	OI	0.6	0.5	8.1	3.5	3.8	8.1	3.1	4.0	25.9	**92.8**	**91.6**	**65.5**

Compared to the results of five-class scheme in Table 3, the accuracy increases for merged young ice (new ice and young ice) were 3.0-4.5%. However, the misclassification among young ice and first-year ice was still significant.

A seasonal assessment of the classification scheme is missing. It the most important issue to address and without this assessment it would not be reasonable to claim the scheme to be either operational or innovative.

As the reviewer pointed, we did not conduct seasonal assessment. Since the developed algorithm was tested for winter season only, we changed the title as "Classification of Winter Sea Ice Types in Sentinel-1 SAR images" and the limitation related to the variations in radar backscattering and its corresponding image textures due to seasonal changes is added to the end of Section 3 in the revised manuscript.

Due to above mentioned major issues I didn’t listed any technical corrections. I would kindly invite the authors to address this issues in future. Due to the lack of innovation and failed to address the basic issues, at current stage I can only recommend the manuscript to be rejected despite it is well within the scope of The Cryosphere.

We kindly ask the reviewer to reconsider his/her decision based on our responses and modifications made to the manuscript.