Dear Anonymous Reviewer #2,
We thank you for taking the time to provide comments, and have revised the manuscript according to them. Our response to each comment is included in red text after each comment below.

Specific comments and technical corrections:
P4541L22-25: A slightly contrasting claims in two subsequent sentences. First, "… C- and X-band ..
largely equivalent", then "X-band … add information when used .. with C-band".
C: It seems that this issue is still open. Reformulate the sentences.
We agree in this, and the sentences was reformulated:
"Results from the Baltic Sea suggest that the information content in C- and X-band are largely equivalent (Mäkynen and Hallikainen, 2004; Eriksson et al., 2010), while X-band was found to add information when used in combination with C-band in the Arctic Ocean (Brath et al., 2013)."

P4543L10: Individual
P4544 L8: opportunity
The misspellings were corrected.

P4545L12: The footprint of the EM-bird has a diameter of about 50 m.
General question (not related only to the EM measurements):
Please add in the text your opinion about the scales in which SAR image is geophysically reasonable to analyze. For which purposes, except for scientific research, sea ice type classifications in resolutions from 10 m to 100 m are needed? (Leads, heat budget, …).
We find the question two-folded.
1.The input scale or resolution of the segmented scene govern what kind of information you can get out. Information about structures like ridges, melt ponds and small leads all require a high resolution, while general sea ice type segmentation possibly could be performed on scenes with lower resolution.
2.When it comes to the required output resolution, this depends on the application. Climate scientist are often interested in large areas, while high resolution information is needed for instance for ship navigation or structures like oil platforms.

To comment this in the text, the following sentences were added to the new discussion part:
"The choice of window size will also determine which kind of information one can retrieve about the sea ice surface. If information about small-scale structure like ridges, melt ponds and small leads are important, this requires a small window. For more general information for instance about sea ice age, larger window sizes could be more appropriate. Choice of sensor would set restrictions on how high resolution it is possible to achieve, and high resolution is at the moment coupled to small swath width."

P4549L7: dual-polarimetric
The misspelling was corrected.

P4549L10: ‘G’
C: Did you mean B or where is G defined?
This was supposed to be B, and was corrected.

P4550L5: ‘supervised classifier’
C: An essential part of the supervised classification is the selection of the training areas. Here this question is ignored. How were the classwise training areas selected, how large were they? Were they selected just from one SAR scene (a good choice) or from several scenes? In the latter case the temporal evolution of the features has been implicitly included. Hence, the classification results are not as representative as in the former case.
The MAP supervised classifier was used scene-wise to evaluate the individual features discrimination capability, and was not a part of the segmentation algorithm (which is described in P4550L25-P4551L15). The training areas corresponds to the pixels in the ROIs, described in section 2.2.1. The evaluation was done scene-wise, using leave-one-out cross validation for training. To clarify this, the following sentence has been added (P4550L12):
“…in Bowman and Azzalini (1997). The pixels in the five ROIs were used as training areas, and each the satellite scenes were classified individually.”

The scene-wise procedure has also been stressed (P4550L17):
“The resulting classification accuracies obtained for each individual feature are used to evaluate the discrimination abilities of the features in each of the five scenes.”

P4550L14-16: A 7×7 pixels neighbourhood, L = 49, is used in the classification and a stepping window with steps of 5×5 pixels was employed to reduce neighbourhood overlap.
C: Do you have used the same resolution for RS-2 and TSX imagery? If not, then why not? The variation of ice surface roughness and other ice features on sea ice layer remain the same independent of the resolution of the sensor. So it would be logical to analyze them in the same resolution, especially if one wishes to compare two different frequencies.
In the manuscript, we chose to use the same neighbourhood/filter size for all scenes, even if the resolution of the scenes were different. This was done from a statistical point of view, we wanted equal sample sizes. Especially the statistical feature, RK, was expected to depend strongly on sample size. However, in the work with the manuscript we did try to regulate the filter size after the resolution of the scenes as the reviewer suggest, applying different filter sizes for each scene. We experienced that choosing the filter size the one way or the other did not affect the results to a noticeable degree.

To clarify our choice for the readers of the manuscript, the following sentence was added to the manuscript (P4551L7):
“…granular segmentations. The size of the neighbourhood does not take into account the difference in resolution between the scenes, but assure an equal sample size in the extraction of the features.”

The question is also discussed in the new-written discussion-section:
“To gain equal sample sizes in our study, the same neighbourhood size was used in filtering all scenes even if the scenes resolution differed. The scenes with highest resolution would therefore have smaller filter sizes in meters. This difference in scale possibly influences the signature of

physical properties of the surface, like surface roughness variation. We did, however, during our investigations, also try to use filter sizes adjusted to the resolution, but this made little difference to the results."

P4552L1 ' segmentation uses ..'
C: Do you utilize here a six-dimensional pdf or 6 independent one –dimensional pdfs? If the former option, how do form a 6-D nonparametric density function? Clarify text.
The Gaussian mixture modelling is parametric. It uses a 6-D multivariate Gaussian distribution, parameterised by a mean vector and a covariance matrix, per model. To clarify this in the text, multivariate was added to the sentence:
"The segmentation uses multivariate Gaussian mixture models to model the features' PDF, and employs an expectation-maximization algorithm."

P4551L7-9: A stepping window with steps of 5×5 pixels was used to reduce neighbourhood overlap, and an additional sub stepping window of 10× 10 pixels was used during the algorithm for computation efficiency.
C: This sentence is an unnecessary computational detail. Remove.
We agree in this, and the sentence was removed.

P4551L10-11: The algorithm was set to segment the scenes into six different classes, to allow for detection of the major sea ice types visually expected to be present in the scenes.
C: Write out the selected ice classes. Assign also to each color in Figs. 8 and 10 the corresponding ice class, like in Fig. 4 in Moen et al. (2013). It is difficult for the reader to assess the images when the color label explanations are imbedded in the text.
This question add to the classification/segmentation confusion discussed under point 3 in the reply to reviewer #1.

P4551L10-11: To specify which sea ice types we have in mind, the sentence was changed to:
"The algorithm was set to segment the scenes into six different segments. The number was chosen to allow for the five sea ice types described by the ROIs, in addition to one extra segment to allow for detection of other sea ice types and to assure some flexibility for the algorithm."

Fig. 4 in Moen et al. (2013) presents images classified by sea ice scientists, while Figs. 8 and 10 in our manuscript present segmentations, not classified images. With information from only one single flight-line through the scenes, we don't think it is appropriate to label all segments in segmented images. We have however tried to give similar positioned segments similar colors in all five scenes, to ease the access for the reader. In the revised results-section (see reply to reviewer #1), the parts describing the segmentation results have been rewritten for easier access.

P4552L8: 'received'
Misspelling corrected.

P4554L1: Differences in … resolution could also cause the lower accuracies in the X-band scenes.
C: Do you mean that a finer resolution is a disadvantage? You can always decrease the resolution.
Please clarify.

We agree that this part of the sentence is confusing. The resolution term is brought in due to the differences in scale, as already commented on by the reviewer. We have rephrased the sentence, removing the resolution part (P4553L1):
"…information. The lower incidence angles of the TS-X scenes could also contribute to the observed differences."
The issue of different scales is commented on in the new-written discussion part (see reply to reviewer #1).

P4556L1-15 and P4557L1-12: You can remove the references to the colors when you have added the color labels in Figs. 8 and 10 (see an earlier comment). This also improves the clarity of the text.
P4556-4557: I would prefer using the class labels in Section 3.2. instead of the ROI abbreviation. Then it would be easier for the reader to follow the text.
We refer to answer about class labeling in earlier comment.

P4558L12-20: '…The meteorological conditions could explain the poor segmentation of T1….'
C: It is worth noting that R2 was acquired just 24 hours later than T1. Around R/V Lance the air temperatures remained during this period close to zero degree Celsius. However, the results for R2 were good unlike for T1. In this case C-band SAR scene was more informative than X-band scene although the weather conditions were likely very similar for T1 and R2.
Add a sentence or two about this in the text.
We agree in this, and have reformulated the sentence, and the following paragraph (P4558L21-L26) to account for this. The text is included in the new-written discussion section:
"Both T1 and R2 were acquired during a period with air temperatures close to or above zero degrees Celsius, conditions which is on the limit of suitable for sea ice type discrimination by SAR. As reported by (Scharien et al., 2010), moisture in the upper sea ice layer could mask out volume scattering and hence lower the backscatter contrast between different sea ice types. The difficult conditions could explain the poor segmentation performance of T1. However, R2 was acquired during similar meteorological conditions with good segmentation results. Lower frequency, higher incidence angle and extra information contained in the cross-pol channel (lacking for T1) could all have contributed to a better segmentation of R2."

P4560L11-13: In the other scene the segmentation performed
poorly, probably due to air temperatures above freezing point and hence difficult conditions for sea ice characterization by SAR.
C: The classification results were poor for X-band SAR (T1) but not for C-band SAR (R2). Reformulate text.
The text was reformulated:
"… poorly. The poor performance might be a result of air temperatures above zero degrees Celsius combined with low incidence angle and polarimetric channel combination (HH-VV)."

P4569. Table 4 text: The best result for each ROI and the best overall accuracy for each scene are highlighted in bold.
C: Where is the best overall accuracy? I can not see that in Table 4.
This was a misprint. The sentence is corrected to:
"The best result for each ROI in each scene is highlighted in bold."