Remote Sensing of Environment 297 (2023) 113780

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Unsupervised machine learning detection of iceberg populations within sea ice from dual-polarisation SAR imagery

Ben Evans [a,*], Anita Faul [a], Andrew Fleming [a], David G. Vaughan [a,†], J. Scott Hosking [b]

[a] British Antarctic Survey, High Cross, Madingley Road, Cambridge, CB3 0ET, Cambridgeshire, United Kingdom
[b] Alan Turing Institute, The British Library, London, NW1 2DB, United Kingdom

ARTICLE INFO

Edited by Menghua Wang

Keywords:
Iceberg
Machine learning
Automated
Bayesian
Classification
Radar

ABSTRACT

Accurate quantification of iceberg populations is essential to inform estimates of Southern Ocean freshwater and heat balances as well as shipping hazards. The automated operational monitoring of icebergs remains challenging, largely due to a lack of generality in existing approaches. Previous efforts to map icebergs have often exploited synthetic aperture radar (SAR) data but the majority are designed for open water situations, require significant operator input, and are susceptible to the substantial spatial and temporal variability in backscatter that characterises SAR time-series. We propose an adaptive unsupervised classification procedure based on Sentinel 1 SAR data and a recursive Dirichlet Process implementation of Bayesian Gaussian Mixture Model. The approach is robust to inter-scene variability and can identify icebergs even within complex environments containing mixtures of open water, sea ice and icebergs of various sizes. For the study area in the Amundsen Sea Embayment, close to the calving front of Thwaites Glacier, our classifier achieved a mean pixel-wise F1 score against manual iceberg delineations from the SAR scenes of 0.960 ± 0.018 with a corresponding object-level F1 score of 0.729 ± 0.086. The method provides an excellent basis for estimation of total near-shore iceberg populations and has inherent potential for scalability that other approaches lack.

1. Introduction

1.1. Context

The calving of icebergs to the Southern Ocean is a principal mode through which mass is lost from the Antarctic ice sheet (Depoorter et al., 2013; Liu et al., 2015). An increase in iceberg calving may herald an accelerating contribution to global sea level rise (Slater et al., 2020; Bamber et al., 2009). The calving, fragmentation and melting of icebergs contribute freshwater and nutrients directly to the ocean (Silva et al., 2006) with resulting changes to primary productivity (Duprat et al., 2016). They can alter the local ocean circulation (Robinson and Williams, 2012) by influencing temperature and salinity distributions, influence the formation and breakup of sea ice with consequences for phytoplankton and wildlife (Arrigo et al., 2002), and directly scour benthic communities (Barnes, 2017). Icebergs have also been shown to be vectors for nutrients and thereby to influence spatial distributions of, for example, primary productivity (Biddle et al., 2015; Lancelot et al., 2009; Raiswell et al., 2008; Smith et al., 2007), supporting localised ecosystems and nutrient cycles (Vernet et al., 2012). They therefore represent a key component of the bio-physical coupling between continental and ocean systems that must be better resolved in order to meet the nascent ambitions for digital twinning of the natural environment (Blair, 2021; Bauer et al., 2021). In addition, including more realistic iceberg distributions, trajectories, and volumes will improve climate models and understanding of Antarctic iceberg impact on the future Southern Hemisphere climate (Schloesser et al., 2019; Stern et al., 2016).

Icebergs also represent a hazard to commercial and research vessels (Bigg et al., 2018). Accurate and timely monitoring of iceberg populations, size distribution, and projected trajectories can therefore contribute to numerous scientific and operational objectives. Such monitoring remains challenging because of a lack of fully automated and scalable methods for iceberg identification and tracking (Barbat et al., 2021; Koo et al., 2021).

SAR is an active sensing technique which offers many opportunities for operational monitoring of icebergs, although low resolution radars such as satellite scatterometers have also been effective in tracking large Antarctic icebergs (Budge and Long, 2018). Icebergs are composed of fresh water and, in frozen conditions, are often covered in snow or firn. These attributes mean that both surface and volume

* Corresponding author.
E-mail address: benevans@bas.ac.uk (B. Evans).
† Deceased.

https://doi.org/10.1016/j.rse.2023.113780
Received 9 January 2023; Received in revised form 2 August 2023; Accepted 18 August 2023
Available online 1 September 2023
0034-4257/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

scattering coefficients tend to be higher for icebergs than for saline sea ice or open water, and consequently they usually present as relatively bright objects within C-band SAR imagery, particularly at high latitudes (Wesche and Dierking, 2012). SAR is capable of penetrating cloud cover, which is common over the Southern Ocean, and operates effectively throughout the year even during the polar night. It has long been recognised as a promising data source (Parashar et al., 1980; Young et al., 1998; Power et al., 2001) and many approaches have been deployed in attempts to exploit it. The majority have been based on variations of the Constant False Alarm Rate (CFAR) algorithm (Marino et al., 2016; Gill, 2001; Song et al., 2021; Karvonen et al., 2021), an approach that seeks to detect local anomalies based on contrast between intensities of pixels in a target window and a larger background window by applying a threshold for the acceptable probability of false alarm based on the background intensity distribution. This method is therefore best suited to sparse icebergs in open water settings where contrast between icebergs and background is high. If we are to understand the provenance of icebergs we need methods capable of identifying them at or near the point of calving. The CFAR approach has fundamental limitations for use-cases within sea ice, close to the land margin and where iceberg density is high. It is these environments that form the focus of this study. Soldal et al. (2019) demonstrated very poor performance of a modified CFAR algorithm within arctic sea ice, with true positive rates between 8 and 22 percent and over 90 percent of 'detections' being false. Sea ice normally has higher backscatter intensities than open water and therefore provides lower contrast to icebergs, reducing the effectiveness of the CFAR approach. This typically results in low sensitivity of the classifier or in high false alarm rates (Soldal et al., 2019). Likewise, a complex scene with densely packed icebergs will result in high background intensities, thereby precluding detection. When in proximity to land the CFAR windows may be incomplete due to masking of the land and their statistics unreliable. Furthermore, the window sizes used place a limit on the size of the icebergs that are detectable and larger windows become computationally expensive. Icebergs around Antarctica have been shown to vary in size by six orders of magnitude (Wesche and Dierking, 2015). This would imply the use of prohibitively large windows to detect larger icebergs, both in terms of computation but also the increased backscatter variability incorporated within the background window which would compromise detection of smaller icebergs. The thresholds used also require careful determination on a per-scene basis in order to account for the particular backscatter characteristics of that scene that arise from variable conditions and geometries. This latter obstacle was recently addressed by Karvonen et al. (2021), who proposed a non-parametric implementation which demonstrated advantages over existing operational methods for open water arctic settings, reporting a precision (true positives/prediction positives) of 0.72 and recall (true positives/manually-derived positives) of 0.92. Both Soldal et al. (2019) and Karvonen et al. (2021) recognised the potential for using the additional information available within dual-polarised SAR images to improve detection. Soldal et al. (2019) used the iDPolRad filter (Marino et al., 2016) and Karvonen et al. (2021) used channel cross-correlation to integrate information from both polarisations. Both studies showed improved detection when cross-polarised data were used.

There have recently been a number of alternative approaches proposed to the problem of iceberg detection in the Southern Ocean and, in particular, within sea ice. Mazur et al. (2017) used an object-based segmentation algorithm followed by a manually calibrated thresholding of object statistics to produce average object recall rates of 96.2%, equating to 93.2% of iceberg area for icebergs exceeding 0.625 km^2. This compares favourably to Williams et al. (1999) who detected 84% with a modal area of 0.15 km^2 using 100 m pixels, Silva and Bigg (2005) (70% and 70% of area for objects over 0.015 km^2) and Wesche and Dierking (2015) (46% of objects between 0.3 km^2 and 5 km^2 and 86.7% of the area of those detected objects). These performance statistics were all computed against icebergs derived manually from SAR data rather than complementary datasets. The method of Mazur et al. (2017), however, required operator input for each scene in terms of threshold determination and, given the variable nature of SAR imagery it would not be suitable for automated deployment. To contextualise amount of imagery available and the need for methods that do not require operator input, Liang et al. (2021) collated 69586 coastal scenes across Antarctica between 2014 and 2020.

Supervised classification approaches do not attract much research interest for the purposes of iceberg detection since they would probably be unsuitable because of inter-scene variability and the difficulty of establishing a sufficient labelled dataset for training purposes.

Barbat et al. (2019) also identified the need for iceberg detection methods that do not rely on significant operator input if they are to be widely applied either in time or space. They therefore proposed an adaptive approach using a random forests committee which seeks generality by integrating predictions across heterogeneous models. Their method required small numbers of training labels and limited operator input during training as a consequence of their use of incremental learning (Impoco and Tuminello, 2015). In adopting an adaptive and largely unsupervised approach they provided a potential basis for a scalable, automated monitoring system and they reported excellent performance with an average classification accuracy of $97.5 \pm 0.06\%$. They did, however, base their performance statistics on a post-hoc validation using an unspecified visual evaluation method for false positive predictions and obvious misses, which, based on application to our own classifications, could potentially have inflated their model's apparent capabilities. Further, their initial superpixel segmentation required optimal parameter determination and may limit the generality of the methodology. When applied to our test scenes using their parameters the method failed to identify superpixels associated with some of the larger icebergs, including the largest one central to the scene (B-22A), instead conflating these with the background object. Experiments with the scale and sigma parameters also showed potential for superpixels to have poorly defined boundaries relative to the true iceberg objects. Mazur et al. (2017) also recognised this limitation with their superpixel method. The determination of the optimal parameters therefore represents a requirement for scene-specific knowledge or tuning which we seek to avoid. We therefore question the robustness and generality of superpixel methods in congested near-calving-front environments that form the focus of this study.

1.2. Adaptive non-parametric clustering

The objective of this study is to provide a viable basis of ongoing near-real-time monitoring and tracking of iceberg populations in the Amundsen Sea Embayment to support efforts to better understand the dynamics and future, in particular, of the Thwaites Glacier. Understanding how the West Antarctic Ice sheet will change, in particular the Thwaites Glacier and the adjacent Amundsen Sea, is of high priority (Scambos et al., 2017) and requires better understanding of iceberg processes and observations to improve regional ice-ocean models. Furthermore, grounded icebergs in the study area play an important role in sea ice and ocean dynamics by forming the eastern boundary of the Amundsen Sea Polynya (Fig. 1) and influence its dynamics by becoming grounded within the polynya (Stammerjohn et al., 2015; Macdonald et al., 2023). A secondary objective is to develop an algorithm with the potential to be applied at scale for monitoring iceberg populations across the Southern Ocean. In order to achieve this we propose a recursive implementation of a pixel-based Dirichlet Process mixture model (DPMM) classifier (Faul, 2019) which removes any need for knowledge of scene backscatter characteristics or composition and eliminates requirements for fixed thresholds. The method does not require *a priori* determination of the classes to be identified, nor a labelled training dataset. It does not involve any kernel operations during clustering, thereby overcoming many of the fundamental limitations identified for the CFAR approach in proximity to land. The pixel-based method,

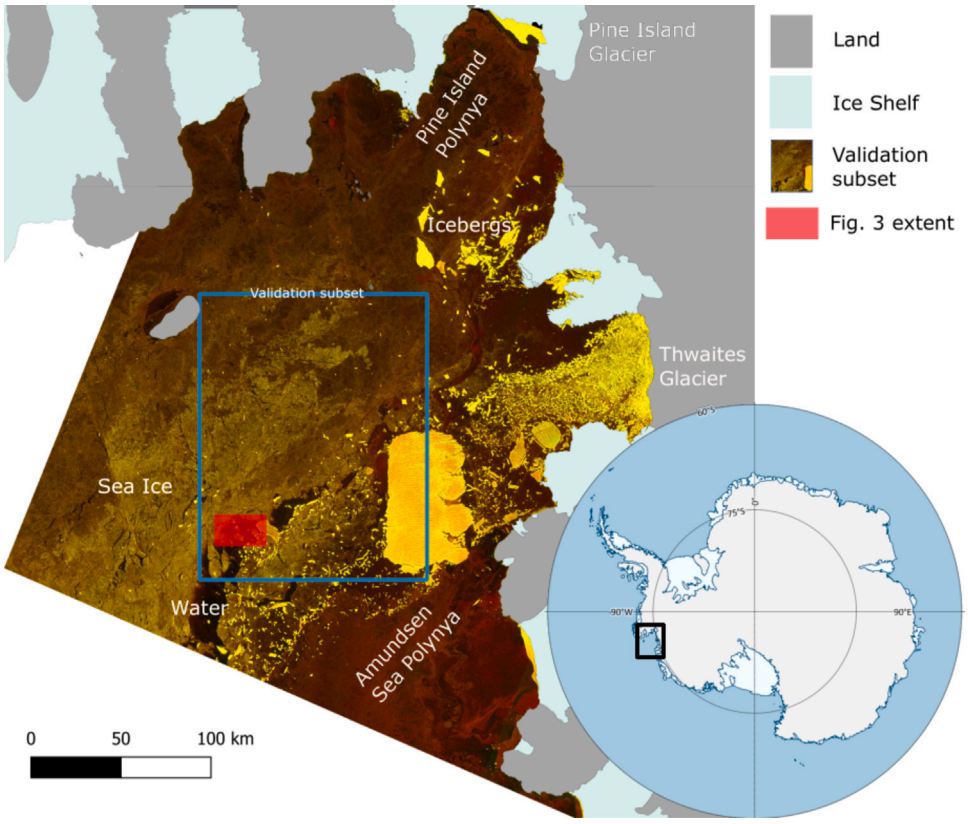

Fig. 1. Example of cross-polarised Sentinel 1 scene from 3rd October 2019 (Spring) for Amundsen Sea embayment (black outline on inset map). HH intensity rendered in red, HV intensity rendered in green (blue channel = 0). Central subset used for independent validation outlined in blue and various surface types labelled Land extent data from the SCAR Antarctic Digital Database, accessed 2021. Inset generated using Quantarctica (Matsuoka et al., 2021). Contains Copernicus Sentinel data 2019. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

unlike semantic segmentation approaches, is also insensitive to the high variability in the size of target iceberg features.

DPMMs initially assume the data are explained by the simplest possible model, a single gaussian distribution, but incorporate the option to increase the complexity by creating a new cluster if the data is not adequately explained by the model so far. While in a model with a fixed number of clusters the total probability mass is distributed between clusters, DPMMs hold back a proportion of the probability mass for the possibility of encountering data which cannot be explained with the current model. DPMMS are known as non-parametric, since the number of clusters is no longer a parameter to the model. In contrast to DPMM methods, traditional finite mixture models would require model averaging or selection procedures to determine the optimal number of components to model. DPMMs have unbounded complexity and therefore guard against underfitting while the estimation of full posteriors over the parameters mitigates overfitting. In brief, DPMMs consider each observation in turn in a random order in several epochs. First the observation is removed from the data in order to treat it like previously unseen data. If it is the only one in a cluster, that cluster gets removed. The probability distributions for all clusters are updated as well as the probability of unseen data. This means that their mean and variation are adjusted. Next it is assessed via the likelihoods which cluster the observation under consideration should be assigned to if at all. If none of the current clusters explains the observation well a new cluster is created.

DPMMs have been applied to image classification and segmentation tasks, although more commonly in medical rather than environmental imaging contexts (Castro and Glocker, 2016; Ferreira da Silva, 2007, 2009; Derraz et al., 2010; Cheng et al., 2016; Orbanz and Buhmann, 2006). Given the potential advantages that the DPMM method offers over parametric models when considering remotely-sensed data sources with high spatio-temporal variability, it is perhaps surprising that the approach has not attracted more attention. A few applications to SAR imagery have been attempted, though none have focused on iceberg detection, and they focus on the general methodological aspects rather than application to specific scientific domains or questions (Sun et al., 2009; Song et al., 2017). A recent attempt has also been made to apply DPMMs to classification of hyperspectral satellite imagery (Mantripragada and Qureshi, 2022), in which the authors observed robustness of the DPMM to noisy data that may otherwise have required pre-processing for noise reduction, another advantageous attribute for our proposed application to SAR imagery.

2. Data and methods

2.1. Data

The method presented here uses C-band SAR data from the Copernicus Sentinel 1 mission. This data source was selected because of its high temporal availability and cross-polarisation capability, which should allow for high frequency monitoring and additional detection capability compared to single polarisation platforms. Sentinel 1A launched on 3rd April 2014 and is still operational, providing approximately weekly overpass. Sentinel 1B launched on 25th April 2016, increasing overpass frequency to approximately 3 days, but failed on 23rd December 2021. Sentinel 1 data are freely available and therefore appropriate to support an open-source operational method. Extra Wide Swath (EW) mode data were used, downloaded in dual-polarisation Ground Range Detected (GRD) format from. This provides 40 m pixels across a swath width of 410 km acquired as five sub-swaths and at incidence angles ranging between 18.9 and 47.0 degrees. Pre-processing was conducted in Python using the 'snappy' module to interface with the European

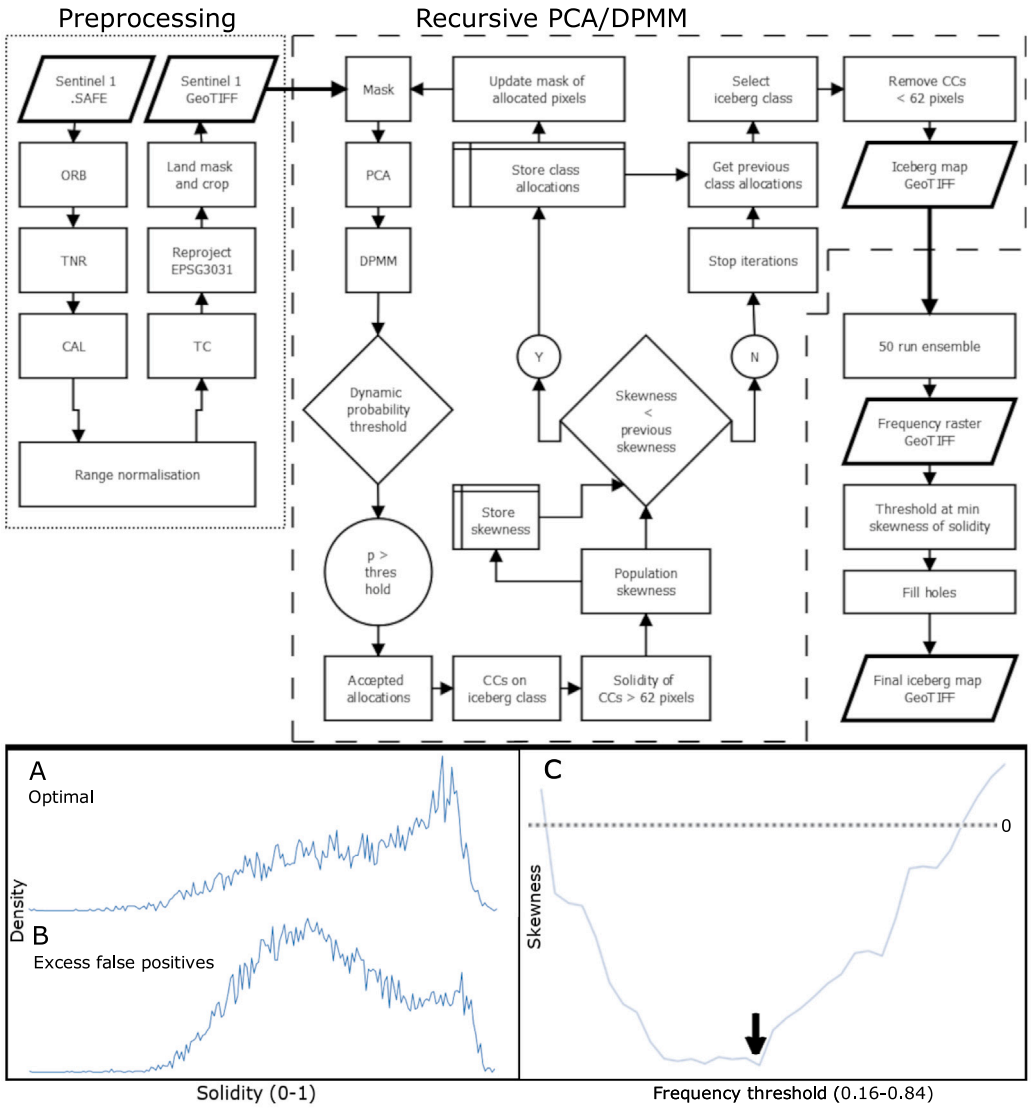

Fig. 2. Schematic of workflow separated into preprocessing, recursive DPMM and ensembling stages. Inset A shows an example histogram of object solidity with strongly negative skewness, representing the optimal stopping point, while inset B shows the subsequent iteration with higher skewness representing proliferation of false positives. Bin width 0.01. Inset C shows example progression of mean object skewness during threshold determination for frequency raster. Bin width 0.02.

Space Agency's (ESA) Sentinel Applications (SNAP) Toolbox. In this study we focus on the area around Thwaites and Pine Island glaciers (red box on inset to Fig. 1) in order to demonstrate the applicability of the proposed method within a mixture of sea ice and open water, and in settings with high iceberg density and a large range of iceberg sizes. This is an area of recent rapid glacial change, and has been said to be close to a transition in its calving regime.

A standard pre-processing workflow was used to apply orbit files (Fig. 2, ORB), conduct thermal noise removal (TNR) and calibration (CAL) using the SNAP tools provided for these purposes. Supplementary material S1 provides details of the parameters used. Subsequent to calibration, a range-normalisation process was conducted to account for the cross-scene variability in backscatter intensity that arises from the sideways look-angle of the satellite and resulting differences in path distance and incidence angle between near-scene and far-scene locations. This tends to lead to higher intensities closer to the sensor and lower intensities further away (e.g., Topouzelis and Singha, 2016). Range-normalisation was conducted by fitting a linear breakpoint algorithm to the median pixel values in individual columns across the scene, with one breakpoint allowing for some degree of nonlinearity to be accounted for. The coefficients of this fit were inverted and used to adjust all pixel values in each column in order to remove the cross-scene trends. The process was conducted separately for HH and HV bands of each scene. This process helped to reduce confusion between icebergs and sea ice arising from their position within the scene and its effect on their apparent relative backscatter intensities.

Subsequent to the range-normalisation process, scenes were further processed using the SNAP range-doppler terrain correction (TC) algorithm, parameters for which are detailed in S1. Scenes were reprojected to the WGS84/Antarctic Polar Stereographic Projection (EPSG:3031) and a land mask (Gerrish et al., 2021) applied to exclude grounded ice areas. Areas outside the minimum bounding box of the remaining valid pixels were cropped to minimise file sizes. The resulting images contained σ^0 bands for both HH and HV, with NoData being given a value of zero and masked out during subsequent analysis.

2.2. Iceberg detection

HH and HV bands of the scene were converted from linear to decibel scales and scaled to the range 0-1 prior to principal components analysis. The sign of components was checked for consistency and inverted, if necessary, such that higher values expressed in the first

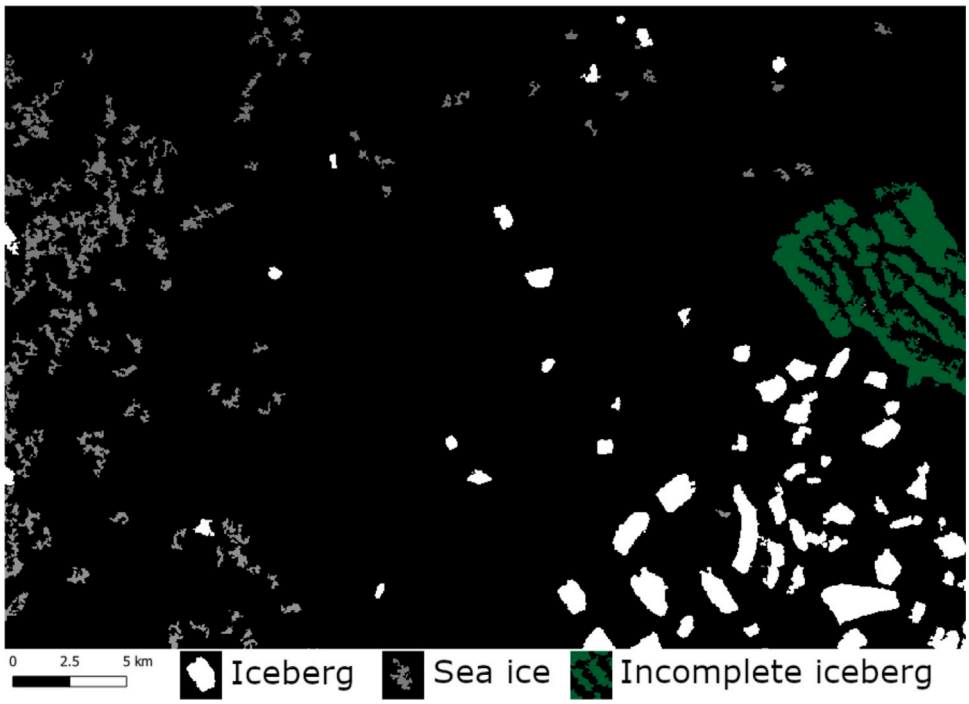

Fig. 3. Example of connected components identified from SAR DP clustering, with holes up to 500 pixels filled. Contrasting geometries of icebergs (white), sea ice (grey) and incomplete icebergs (green) are highlighted. Extent of figure marked in red on Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

principal component reflected higher σ^0 intensities. The first principal component was then used as the basis for univariate Dirichlet Process clustering (selected as preferable to multivariate clustering for computational reasons). Training was conducted on an approximate 0.1% of the pixels, achieved by regular sampling over the flattened array at a step of 1024 pixels. This sampling was found to be sufficient to represent the structure within the overall pixel population while minimising computation time and resulted in training sets in the initial iteration in the order of 115,000 pixels for a scene containing no land mask. DP clustering was implemented using the scikit-learn (Pedregosa et al., 2011) Bayesian Gaussian Mixture algorithm 'bgm' with a dirichlet process weight concentration prior type. Given the problem of inter-scene variability, priors were as weak as possible. Mean and covariance priors were determined by the mean and variance of the sample data respectively. This reflects the initial belief that the data can be explained as being generated by one stochastic process. Weight concentration and mean precision priors were both set to 1. Initially the probability of another cluster existing is $1/(1+concentration)$. Initialising the concentration to 1 means that we have equal probabilities of 0.5 whether data belongs to the first and only cluster or a new cluster which is uninformed. The mean precision, the inverse of the mean fraction, controls the spacing between clusters. A value of 1 allows overlap which is the initial uninformed state. These priors are appropriate because we have no *a priori* knowledge of scene composition. There was only one dimension so the degrees of freedom prior was also set to 1. We tested various limits to the number of clusters. Fewer clusters lead to larger class memberships, more accepted pixels in each iteration, and a less precise stopping point. Conversely, more clusters lead to the algorithm running for more iterations, which is resource intensive. A maximum permitted number of clusters of 5 was found to be an appropriate compromise.

The trained model was used to predict class allocations to all pixels in the scene and those with assignment probabilities exceeding the final peak in the histogram (of assignment probabilities) were accepted as valid cluster assignments. These pixels were labelled and stored. The class with the highest mean intensity values in PCA1, which implies the highest backscatter intensity, was assumed to represent icebergs. Subsequent classes were considered to be sea ice, water or ambiguous (e.g., shadows). For the purposes of this study the class label of these other classes is unimportant except in that they do not constitute icebergs. If there are no icebergs present in the scene then the algorithm mis-classifies areas of sea ice as icebergs.

All remaining pixels were passed back to the principal components analysis stage and the subsampling and clustering repeated on the reduced set. This iterative process had the effect of removing those pixels that clearly belong to distinct real-world classes early on, allowing for greater discrimination in later iterations where class separation was more challenging. The iterative approach further ensures that icebergs are identified even in scenes where they are very rare, since the scene pixel count declines to a point at which even scarce iceberg pixels represent a significant cluster within an iterative subsample. The method does, however, make the tacit assumption that at least one iceberg is present within the scene. Furthermore, the recursive PCA dimensionality reduction stage means that, since the dataset changes between iterations as a result of pixel cluster acceptance, the orientation of the principal component axes also changes, providing a new 'perspective' on the data in each subsequent iteration.

Icebergs, when correctly identified, are relatively geometric shapes with distinct edges and high solidity (defined as the proportion of pixels within the convex hull of the shape that are also within the shape itself). In contrast, those sea ice features with the highest backscatter intensities that initially become allocated to the iceberg class once the recursive algorithm approaches the interface between classes produce strikingly different geometries. This contrast in the morphology of classified features arises because the structure and spatial variability in backscatter of sea ice does not tend to produce distinct edges around the 'brightest' areas as it does for icebergs. Clusters of sea ice pixels therefore tend to have more tortuous perimeters, and consequently more background space within the convex hull, and lower solidity than true icebergs. Poorly resolved or incomplete icebergs, which are common in early iterations, also have low solidity (Fig. 3). Mazur et al. (2017) also observe and exploit this geometric contrast between

icebergs and background objects, although they employ the shape parameters of area, border index, compactness and roundness.

Subsequent to each clustering iteration, connected components (CC) analysis (Shapiro, 1996), using 8-connectivity (whereby iceberg pixels are considered part of the same object if they are adjacent either at their edge or diagonally at a vertex), was conducted on all pixels accepted to the iceberg class in all preceding iterations. A subsequent hole-filling operation removed holes within CCs that were less than 500 pixels in size. A minimum area filter was applied to remove CCs with a total area less than or equal to 62 pixels (0.0992 km^2), effectively eliminating the A0 (<0.1 km^2) class proposed by Wesche and Dierking (2015). This smallest class was excluded because it is difficult for a human expert to reliably separate icebergs from noise when they become small. This, coupled with increasing iceberg numbers as size decreases, would have made human labelling of the scenes for validation purposes impractical and unreliable. The solidity of the remaining connected components was computed as the ratio of the pixels in the component to the pixels in the convex hull of the component.

Iterations were repeated until iceberg false positives began to proliferate as a consequence of inclusion of sea ice features in the iceberg class. This stopping point was identified by an increase in the skewness of the distribution of connected component solidity values (Insets A and B, Fig. 2). Once such an increase was observed, the results of the preceding iteration were taken to be the optimal clustering, having the most negatively skewed (i.e., most 'solid' population of geometries). A minimum of the skewness indicates that features with high solidity dominate the population. Alternative shape metrics (Euler number, Hu moments, perimeter-area ratio) were trialled, as were alternative stopping criteria (maximum mean solidity, maximum median solidity, maximum solidity). The skewness of the solidity distribution was found to be the most robust indicator of the optimal stopping point. The shape metrics trialled are, however, sensitive to the representation of objects as connected pixels. This approach has the potential to represent adjacent icebergs as a single object or single icebergs as multiple objects. If the pixel-level classification is good, these situations are rare, becoming more common as pixel-level classification quality declines. Consideration of the distribution of the metric and its evolution through algorithm iterations rather than as individual or absolute values mitigates this issue somewhat. It is possible that using multiple geometric descriptors beyond just solidity could reduce the sensitivity of the method to connected component object representations but would require more complex heuristics around deciding which combination of those descriptors represents the optimal solution and is beyond the scope of the current study.

The algorithm is stochastic, so initialisations and data subsamples vary randomly between runs and iterations within runs, leading to non-deterministic results. This attribute was exploited through deployment as an ensemble of runs. This further increased the number of viewpoints on the dataset, with the commonalities between runs indicating where the algorithm is most confident in its identification of iceberg pixels. The algorithm was run 50 times (potentially parallelised) for each scene and the proportion of runs in which each pixel was allocated to the iceberg class computed, resulting in a map of classification frequency (Fig. 4)

In order to convert these frequencies into a binary classification of icebergs a similar logic to that used in determining the optimal stopping criterion was applied. The frequencies were thresholded at 0.02 increments between 0.16 and 0.84 and all pixels exceeding the threshold were retained. Connected components were computed and those less than 62 pixels were discarded. The skewness of the distribution of solidity for the remaining CCs was calculated. The optimal threshold for the scene was determined as the point at which the global minimum of skewness was achieved across all thresholds (Inset C, Fig. 2). Within the broad region of low skewness in inset C the classification results change very little (hence skewness is stable), and the final performance is minimally sensitive to threshold variations within this region. This broad minimum also implies that it may be possible to achieve the same results with fewer runs in the ensemble, although optimising the number of runs to minimise computational expense remains for future work. The full 0–1 range of thresholds was not used because very occasionally the clustering algorithm, being stochastic, produces outlier predictions of either near-complete iceberg coverage or very sparse iceberg coverage. These result in small numbers of either very large or very small objects with high solidity being identified leading to unrepresentative global skewness minima at the extremes of the frequency range. The selected 0.16–0.84 frequency interval effectively mitigates for the effects of these edge-cases and reduces the sensitivity of results to the stochasticity of individual runs by requiring agreement between at least eight ensemble members to determine a pixel as being either iceberg or not iceberg.

2.3. Validation

An independent validation was conducted against a subset representing the central portion of eight scenes of the Thwaites glacier area (outlined in blue on Fig. 1). Scenes were spaced throughout one calendar year from October 2019 to September 2020 in order to represent seasonal variability and to allow for assessment of performance across contrasting surface temperature and melt state conditions. All icebergs above 0.1 km^2 were manually vectorised from the SAR scenes while they were displayed at a consistent scale of 1:25000. A combined total of 4106 icebergs were identified across the eight scene subsets. Manually vectorised maps were rasterised and compared to predictions.

Pixel-wise validation statistics were computed as producer's (recall) and user's (precision) accuracies alongside the F1 score (the harmonic mean of precision and recall) and overall accuracy. These statistics are valid within the SAR domain since complementary data sources (e.g., optical) were not used. Connected components were again computed for manually-derived rasters and predictions compared to provide precision, recall and F1 score at an intersection over union (IOU, based on bounding boxes) of 0.5. IOU is the ratio of the overlap between bounding boxes (intersection) and their combined extent (union), with a value of 1 implying perfect alignment of size and position. IOU = 0.5 is a widely used threshold for evaluating object detection methods. Object-based validation metrics were calculated for all objects and within each size class A1, A2 and A3 (Wesche and Dierking, 2015). Manually-derived and prediction objects were also assessed for one-to-many mappings with their counterparts.

A post-hoc validation procedure was also used for four of the scene subsets. This was somewhat analogous to the Barbat et al. (2019) assessment for false positives. Outlines of each predicted iceberg object were overlaid on the original cross-polarised SAR scene and manually allocated to one of five categories, as follows: Category 1 - correctly identified iceberg; Category 2 - Not iceberg or part of iceberg; Category 3 - Iceberg containing holes; Category 4 - Icebergs that need splitting into multiple objects; Category 5 - part of an iceberg.

3. Results

3.1. Classifier outputs

The DP clustering algorithm, run 50 times on each of eight scenes, produced maps of frequency with which pixels were identified as iceberg. The distribution of values within the frequency rasters varies between scenes on a seasonal basis (Fig. 4).

Dynamic thresholding of these rasters based on the skewness of the solidity distribution for connected components produced binary rasters of icebergs. The resulting iceberg classification for a Spring scene is illustrated in Fig. 5 (other scenes in supplementary material S2)

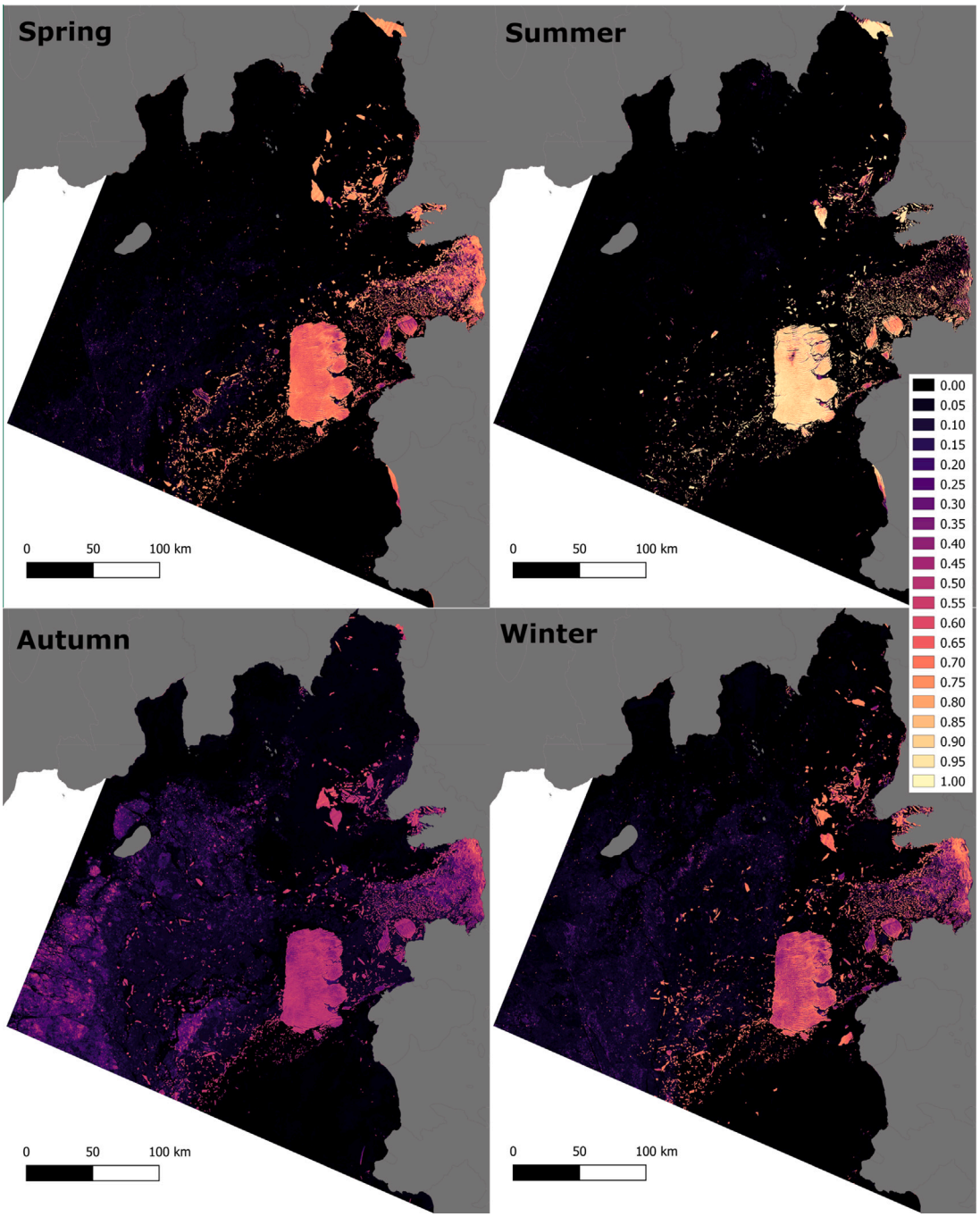

Fig. 4. Frequencies of iceberg prediction within 50 run ensembles for four austral seasons. Note contrasting ranges of values. Contains modified Copernicus Sentinel data 2019–2020.

3.2. Classifier performance

Classifier performance, assessed against independently manually vectorised iceberg outlines, is detailed for both pixel-level and object-level in Table 1. The post-hoc validation method, when applied to the same subsets, furnished precision values. We report a 'strict' precision which counts only category 1 allocations as true positives, and a 'lenient' precision, which counts all but category 2 as true positives since these are all valid detections of current or former land ice even if the object representation is imperfect. Performance varies between austral seasons, as does the nature of the errors. Pixel-level user accuracy across all seasons is 0.981 ± 0.015, producer accuracy is 0.939 ± 0.032, F1 score is 0.960 ± 0.018. Object-level performance is somewhat lower and more variable with average precision of 0.694 ± 0.146, recall of 0.783 ± 0.040 and F1 of 0.729 ± 0.086. This is largely driven by errors of commission for small objects in Autumn and Winter and errors of omission in Summer, particularly for intermediate-sized objects (A2). In some cases, the three large A3 class objects were predicted as being a number of smaller objects (see Table 2). Post-hoc precision, particularly if considering the 'lenient' interpretation, tends to be somewhat higher than derived from the independent validation (0.81 ± 0.13).

3.3. Nature of errors

The counts of intersecting objects for both manually-derived (MD) and predicted (Pred.) maps illustrate the occurrence of false negatives (misses, 0), but also several one-to-many object mappings where one MD object is represented by multiple predicted objects or *vice versa*.

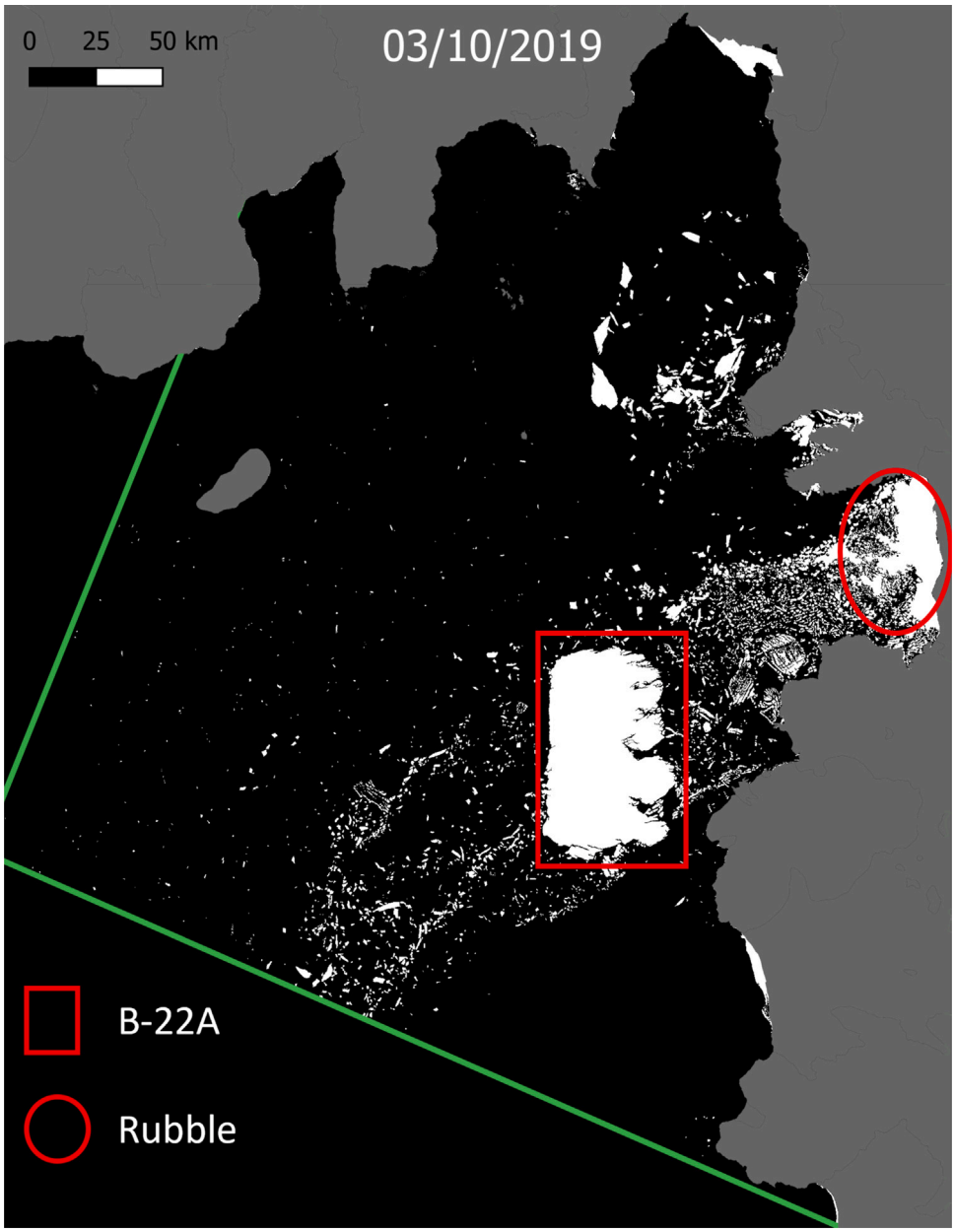

Fig. 5. Thresholded, binary map of iceberg detections for austral Spring scene (3rd Oct 2019). Red outlines indicate location of tabular iceberg B-22A and the rubble of small icebergs adjacent to the Thwaites calving tongue referred to in the text. Green line denotes extent common to all scenes. Contains modified Copernicus Sentinel data 2019. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

One-to-one mappings, however, dominate. While predicted objects tend to conflate relatively small numbers of manually-derived objects (maximum of 4), manually-derived objects may contain large numbers of predicted objects (up to 55 in March 2020) (Table 2). Within the independent validation metrics (Table 1) these one-to-many mappings will be represented as false positives in the case of manually-derived objects and false negatives in the case of predicted objects and consequently account, in large part, for the relatively low precision values in March and April 2020. A total of 4773 icebergs were detected by the classifier in the eight subset scenes compared to 4106 validation objects, indicating a tendency to overestimate the number of icebergs.

For four example scenes spaced throughout the year, the intensity values associated with True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN) are illustrated in Fig. 6. Most striking is the substantial overlap in both HH and HV channels between the TN and FN pixels in summer, and the greater separability of FN and FP distributions in HV than HH polarisation.

3.4. Post-hoc validation

Mean and standard deviation of size (pixels) for objects within each post-hoc category illustrate that false positives (category 2) across all scenes are overwhelmingly dominated by small objects (Table 3). The lower size limit for objects assessed was 63 pixels (approximately 0.1 km^2)

3.5. Seasonality of performance

The performance of our classifier can be seen to vary seasonally (Fig. 7). The higher variability seen in the object-level classifications

Table 1
Summary of validation performance of classifier across eight dates. Top: Object level performance resolved by object size classes A1–A3 following Wesche and Dierking (2015) and aggregated across all objects. Middle: Equivalent overall object-level precision from 'strict' (S) and 'lenient' (L) post-hoc validation. Bottom: Independent pixel-level performance. User accuracy equates to precision while producer accuracy equates to recall.

Object-level										
Date (dd/mm/yyyy)	03/10/2019	27/11/2019	19/01/2020	14/03/2020	19/04/2020	11/06/2020	17/08/2020	15/09/2020	Mean	SD
A1 precision	0.788	0.795	0.728	0.479	0.379	0.509	0.543	0.730	0.619	0.160
A1 recall	0.748	0.795	0.745	0.780	0.760	0.784	0.851	0.762	0.778	0.034
A1 F1 score	0.768	0.795	0.736	0.594	0.505	0.617	0.663	0.746	0.678	0.101
A2 Precision	0.836	0.908	0.933	0.798	0.744	0.832	0.878	0.922	0.856	0.066
A2 Recall	0.764	0.857	0.554	0.632	0.702	0.662	0.783	0.787	0.718	0.099
A2 F1 Score	0.798	0.882	0.695	0.705	0.722	0.737	0.828	0.849	0.777	0.071
A3 Precision	0.000	1.000	1.000	1.000	1.000	0.667	1.000	1.000	0.833	0.356
A3 Recall	0.000	0.667	0.333	0.500	1.000	1.000	1.000	1.000	0.688	0.383
A3 F1 Score	0.000	0.800	0.500	0.667	1.000	0.800	1.000	1.000	0.721	0.342
Overall Precision	0.831	0.840	0.825	0.589	0.459	0.601	0.630	0.810	0.698	0.146
Overall Recall	0.779	0.829	0.692	0.790	0.788	0.800	0.777	0.806	0.783	0.040
Overall F1 Score	0.804	0.834	0.753	0.675	0.580	0.686	0.695	0.808	0.729	0.086
Post-hoc										
Precision (S)	0.767		0.796		0.508		0.673		0.686	0.130
Precision (L)	0.918		0.931		0.650		0.748		0.812	0.136
Pixel-level										
User Accuracy	0.995	0.988	0.996	0.986	0.955	0.972	0.968	0.991	0.981	0.015
Producer Accuracy	0.935	0.972	0.918	0.884	0.912	0.956	0.971	0.964	0.939	0.032
Overall Accuracy	0.992	0.996	0.991	0.985	0.985	0.991	0.993	0.994	0.991	0.004
F1 Score	0.964	0.980	0.956	0.933	0.933	0.964	0.969	0.977	0.960	0.018

Table 2
Numbers of objects intersecting each manually-derived (MD) and predicted (Pred.) object for eight validation dates.

Date (dd/mm/yyyy)	03/10/2019	27/11/2019	19/01/2020	14/03/2020	19/04/2020	11/06/2020	17/08/2020	15/09/2020
MD objects	664	432	436	448	443	521	622	540
Intersecting objects								
0	108	31	87	41	32	40	88	54
1	545	396	340	384	402	467	528	477
2	9	4	5	14	6	12	5	8
3		1	1	4	1	2	1	1
4			2	3				
5			1	1				
6					1			
11	1							
37	1							
42					1			
55				1				
Pred. objects	622	426	366	601	760	694	767	537
Intersecting objects								
0	22	33	5	119	309	226	249	60
1	589	381	356	469	441	449	498	461
2	11	10	5	12	7	12	17	13
3		2		1	2	5	3	3
4					1	1		
5						1		

combined with the validation set that spans only a single year make assessing this seasonality difficult. At a pixel level, however, classifier performance is consistent enough to provide an initial demonstration of how performance varies on a seasonal basis. A sine function fitted to the pixel-level F1 scores estimated the period to be 1.41 years, with an RMSE of 0.005. This predicts the best classifier performance during the austral winter, and worst classifier performance around February. We consider the seasonality in more detail in the discussion.

3.6. Iceberg statistics

For the intersection of all eight scenes covering the Amundsen Sea embayment (see Fig. 4), the total annual mean area of icebergs was estimated as 7263 ± 861 km^2, with an average of 3722 ± 872 individual icebergs detected. Iceberg area is dominated by large icebergs (A5), of which one is identified in the majority of classifications (B-22A, see Fig. 5). In October and November 2019, and June 2020 the rubble of icebergs adjacent to the Thwaites glacier calving front are not separable into individual objects and become classified as a second class A5 iceberg (see Fig. 5). Small icebergs in classes A1 and A2 account for approximately 32% of the iceberg surface area and 99% of the number of detections. Intermediate icebergs in classes A3 and A4 account for only about 19% of the total area. A small number and area of A0 class (<0.1 km^2) were identified, resulting from bisection when scenes were clipped to the intersection area (Table 4). No clear secular trend or seasonal pattern has been observed in the time-series of iceberg areas or numbers (Fig. 8)

4. Discussion

Iceberg detection from remotely sensed imagery is a long-standing problem that has attracted considerable attention over several decades, particularly in an Arctic context (Rawson et al., 1979; Wadhams, 1988; Partington et al., 1994; Sephton and Partington, 1998; Heiselberg et al., 2022; Xiao et al., 2020) with applications to Antarctica and the Southern Ocean proliferating relatively recently (Wesche and Dierking, 2015; Mazur et al., 2017; Barbat et al., 2019; Koo et al., 2021, 2023). If the provenance of icebergs is to be properly constrained, then it is

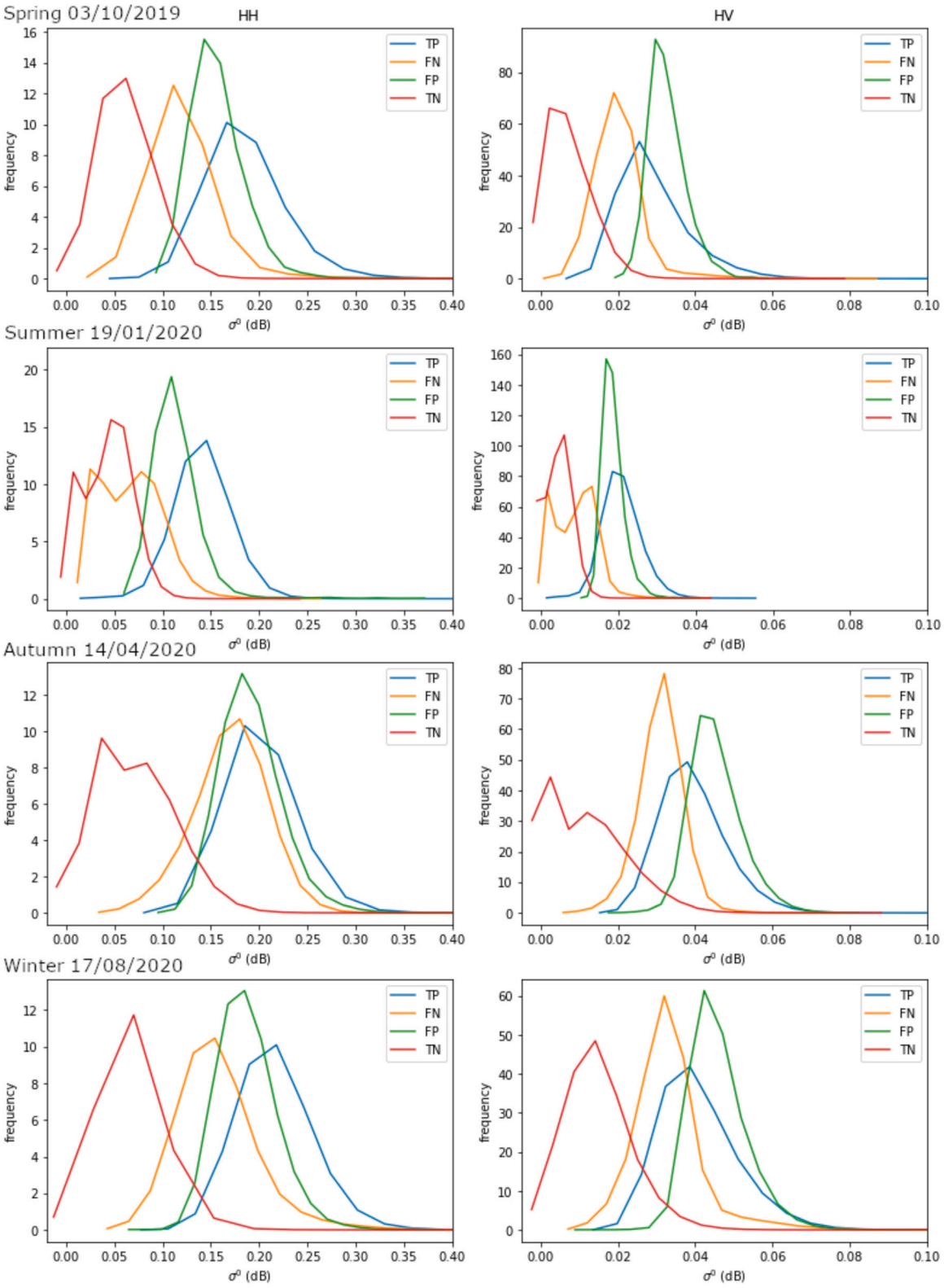

Fig. 6. Histograms of intensity values in HH (left) and HV (right) channels for True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN) observations from four validation subsets.

necessary to identify them as close to their source as possible, however we are not aware of any previous studies seeking to detect and segment icebergs within a mixture of sea ice and open water and in locations that include major glacier calving fronts. Furthermore, all previous attempts to develop algorithms for iceberg detection of which we are aware require a greater or lesser degree of operator involvement either in threshold determination or provision of training data. In contrast, we have presented a novel approach based on non-parametric Dirichlet Process clustering that is conceptually distinct from previous methods in that it is inherently adaptive to scene composition and structure and

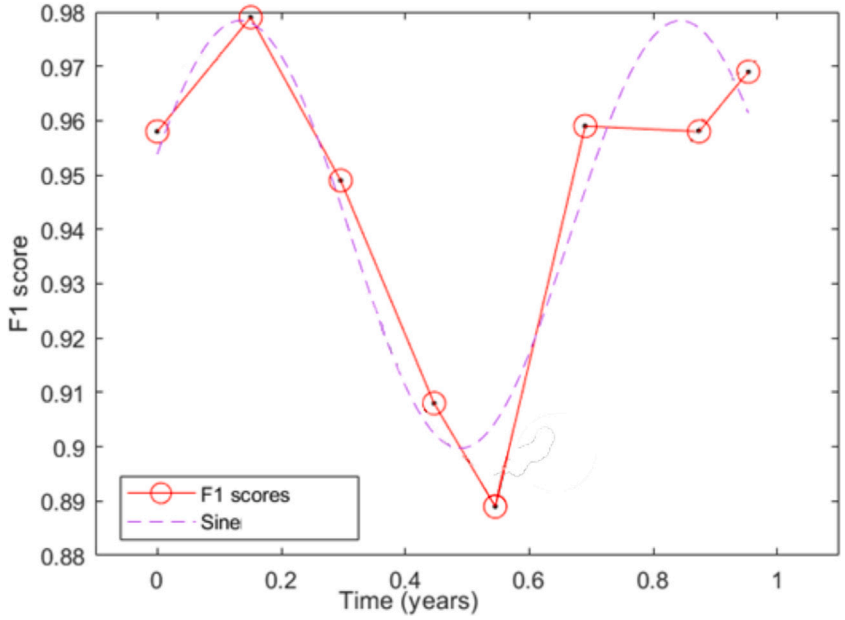

Fig. 7. Seasonality of pixel-level F1 scores derived from validation (red), sine function (dashed). Zero point for time axis is 3rd Oct 2019.

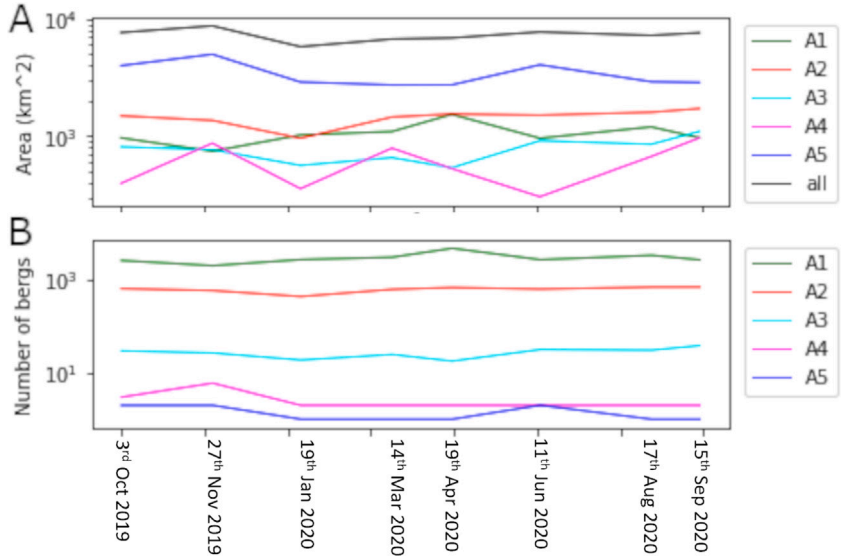

Fig. 8. Time-series of iceberg area (A) and numbers of icebergs (B) from equal-area scenes for five size classes (after Wesche and Dierking, 2015) commencing on 3rd Oct 2019.

requires no *a priori* training data or threshold determination. As such, our approach represents the basis of a fully automatable, and therefore scalable, method for iceberg detection and monitoring in an operational context.

4.1. Classifier performance

The method presented here demonstrates excellent performance when validated for eight scenes of the Amundsen Sea embayment, with mean F1 score of 0.960 ± 0.018 when considering pixels and therefore the total area of icebergs. User accuracy (equivalent to precision) tends to exceed producer accuracy (recall), suggesting that false positive detections are few, while false negatives are somewhat more prevalent, although still typically rare (Table 1). Our mean producer accuracy of 93.9% is slightly higher than that reported by Mazur et al. (2017) of 93.2 and exceeds the proportion of area correctly identified by Karvonen et al. (2021) (92%), Williams et al. (1999) (84%), Silva and Bigg (2005) (70%) and Wesche and Dierking (2015) (86.7% for the 46% of icebergs < 5 km^2 that were detected), all of whom employ methods requiring manual intervention. Our validation approach of comparison against icebergs manually delineated from the underlying SAR data is common to these previous studies. Our method does exhibit sensitivity to seasonal effects, with a minimum producer accuracy of 84.4% during the austral Summer, which is comparable to existing approaches, but with a maximum of 97.2% in late Spring. Our method is comparable with the best-performing algorithm of which we are aware, that of Barbat et al. (2019), although we argue that our validation approach is more robust and less likely to overestimate performance. Barbat et al. (2019) report a false positive rate of 2.3 +−0.4% compared to our mean pixel-wise equivalent of 1.9 +−0.15%. Their miss rate of 3.3 +−0.4% however compares favourably to the 6.1 +−0.32% reported here. The Barbat et al. (2019) approach also represents the most fully automated alternative method available although it does still require some initial training data, although much of the information gain is achieved through their incremental learning approach.

Table 3
Proportions of objects allocated to the five post-hoc categories. Category 1 - correctly identified iceberg; Category 2 - Not iceberg or part of iceberg; Category 3 - Iceberg containing holes; Category 4 - Icebergs that need splitting into multiple objects; Category 5 - part of an iceberg. Mean and SD values are provided to illustrate the distribution of object sizes that are correctly or incorrectly classified in different ways. Units are in pixels with 625 pixels per km^2. Where fewer than three objects contribute to the calculation SD is reported as N/A.

	Category	1	2	3	4	5
3rd Oct 2019	Proportion	0.77	0.08	0.00	0.06	0.08
	Mean	431	112	986910	1026	358
	SD	545	68	N/A	971	404
19th Jan 2020	Proportion	0.80	0.07	0.02	0.03	0.09
	Mean	4074	158	483	769	326
	SD	60800	104	383	792	427
19th Apr 2020	Proportion	0.51	0.35	0.00	0.02	0.12
	Mean	529	175	964804	1193	340
	SD	828	247	N/A	722	507
17th Aug 2020	Proportion	0.67	0.25	0.01	0.04	0.03
	Mean	468	139	257297	2010	358
	SD	758	109	442821	1932	636

Table 4
Annual Iceberg area and count across eight equal-area scenes for five size classes (after Wesche and Dierking, 2015). SD = standard deviation.

	Mean area (km^2)	SD area (km^2)	Mean count	SD count
Total	7263.41	860.89	3721.9	872.4
Class A0	0.14	0.15	4.5	5.3
Class A1	1049.51	229.92	3047.5	834.8
Class A2	1446.51	224.75	638.3	87.9
Class A3	771.20	187.36	27.6	7.0
Class A4	607.64	250.46	2.6	1.4
Class A5	3388.42	840.99	1.4	0.5

When considering objects detected we applied an intersection over union of bounding boxes of 0.5 to define a correct detection. Object-level performance is worse than when considered at pixel level, with average precision of 0.698 ± 0.146. As with the pixel-level performance a seasonal effect is observable. Precision tends to be low in late Summer and Autumn, reflecting a proliferation of typically very small (<0.2 km^2) iceberg objects (Table 1). Many of these, however, fall within areas of the large, tabular B22A that are not identified as a single contiguous iceberg object as a result of the considerable surface texture. The fact that most false-positive objects represent portions of larger objects is reflected in the excellent pixel-level performance. Peak precision occurs in November, at 0.840. Our object-level precision, while temporally variable, is comparable to the 0.72 reported by Karvonen et al. (2021).

Object-level recall is typically higher than precision and much less seasonally variable at 0.783 ± 0.040. Our results are comparable to those of Williams et al. (1999) (0.84), while comparing favourably to Silva and Bigg (2005) (0.70) and Wesche and Dierking (2015) (0.46). Mazur et al. (2017), however report considerably higher object-level recall of 0.96, although their method included a manual step for the detection of missed icebergs and false alarm reduction. Karvonen et al. (2021), working in the Northern hemisphere, report recall of 0.92.

Our validation shows that precision is lowest for the smallest objects, reflecting false positives primarily in the A1 class, while recall tends to diminish with larger objects, since these often have considerable surface texture that results in them not being detected as contiguous regions of pixels. This aligns to the behaviour of the Mazur et al. (2017) method in the Amundsen Sea, which they note contrasts with the other studies (Silva and Bigg, 2005; Wesche and Dierking, 2015; Williams et al., 1999) where recall tends to improve with increasing iceberg size. Barbat et al. (2019) report exceptional precision and recall (0.977 and 0.967), as the complements of their false alarm and miss parameters respectively. As previously noted, however, they employ a post hoc validation approach of visually assessing results for false alarms and obvious misses. As we demonstrate by applying an analogous procedure to our data (Table 3), this tends to result in overestimates of performance relative to the method we employ. Nevertheless, their approach appears to outperform ours in terms of representation of individual objects. It is important to note, however, that previous methods have been developed and validated on scenes and regions with less challenging conditions for object delineation. The ambition of this study is to provide a method that works in congested near-calving-front areas where icebergs may be touching or embedded in fast ice. Both Barbat et al. (2019) and Mazur et al. (2017, 2019) recognised the challenges associated with such regions.

Fig. 6 illustrates the intensity values associated with true positive, false positive, true negative and false negative pixels in the original HH and HV bands of the SAR image. The overlap between these distributions highlights the value that the recursive DPMM implementation adds to the procedure since a single-stage classifier, once membership probability thresholds had been applied, would result in abrupt transitions in the intensity domain. These histograms also provide insight to the performance of the classifier and potential avenues for improvement. Notably, in Summer, there is considerable overlap for both HH and HV channels between the TN and FN distributions, suggesting that the FN pixels will be very challenging to separate from TN pixels using intensity information alone. Any attempt to do so may result in improved recall but probably at substantial expense in terms of precision (false positives will proliferate). For the other seasons there is substantial overlap in the distributions of FN and FP pixel values in the HH channel, suggesting that there is little information remaining there for discrimination. In contrast, however, FN and FP distributions tend to be highly distinct in the HV channel, with FP pixel values being considerably higher than FN ones and often higher even than TP values (Spring, Autumn, Winter). This could arise from high-backscatter from snow cover on areas of sea ice outside the summer season and suggests that it may be challenging to reduce the FP rate further based on the HV channel but that some information remains there that may reduce FNs and thereby improve recall/producer accuracy. This is, however, beyond the scope of this study and Fig. 6 illustrates that the ensembled, recursive DPMM classification that we present here efficiently exploits the attributes of dual-polarisation SAR imagery.

Classifier performance varies on a seasonal basis being best in Winter and Spring and worst in Summer and Autumn. Producer and user accuracy do not, however, vary in phase with each other. The seasonal variability is systematic. The sine function (Fig. 7) shows a promising goodness of fit for a single year and relatively small number of data points. The systematically seasonal nature of classifier performance implies robustness of the approach. Furthermore, in future work applying our method to a longer, denser time series of images, more robust performance functions will be developed to inform the estimation and communication of temporally varying uncertainties.

4.2. Iceberg populations

Our classifier identified a total of 29739 icebergs between 0.1 km^2 and approximately 2800 km^2 across an annual time-series of eight scenes covering the Amundsen Sea Embayment between 2019 and 2020. The iceberg population is dominated by smaller icebergs, with class A1 representing 81% of all observations. Iceberg area, however, is dominated by the large tabular iceberg B22A, which persists in almost the same location throughout the observation period. This iceberg accounts for approximately 30% of the total iceberg area of the embayment. There is no obvious seasonal pattern, nor secular trend within the time-series observations overall, or for any of the individual size classes, either in terms of iceberg count or area (Fig. 8). This remains the same if B22A (size class A5) is excluded from consideration. It should be noted, however, that indicators of a 'state change' in calving regime would probably only be detectable over a longer (decadal) period.

This will be the focus of future work. We have not attempted any statistical analysis of the dataset as it is limited in scope and intended for demonstration purposes only. Nevertheless, the data illustrate the potential utility of the method and may yield interesting observations when applied more extensively.

4.3. Future work

While the object-level performance of our classifier is good and exceeds the capabilities of some of the existing studies, it is not as good as might be expected from the pixel-level performance which improves upon that of other methods. This points to a direction for improvement in further work. Markov Random Fields are often used in conjunction with DPMMs to improve prediction smoothness in image segmentation tasks in other domains such as medical imaging (e.g. Albughdadi et al., 2017; Balafar, 2012; Forbes et al., 2018; Song et al., 2017). Such an approach, given the high pixel-level accuracy of our method, could considerably improve object-level representations but is beyond the scope of the current study. Similarly, it may be possible to use more sophisticated pixel-to-object aggregation methods than connected components labelling, which has been shown here to have limitations. It should also be noted that previous methods for iceberg detection, with the exception of Koo et al. (2023), have relied upon coarser 75 m Envisat Advanced Synthetic Aperture Radar or 100 m RAMP AMM-1 data. The 40 m Sentinel 1 data we use here provides opportunities for detection of smaller objects (in our case down to 0.1 km^2), but the increased resolution implies a reduced smoothness, which may exacerbate the challenge of object delineation somewhat. Notwithstanding the influences of spatial resolution, the clustering method that we employ, because of its adaptive nature, should be agnostic to the SAR sensor producing the images. Future work will therefore apply this approach to existing (e.g. Envisat ASAR, AMM-1, SAOCOM, NiSAR, TerraSAR-X, NovaSAR, RadarSat etc.) and forthcoming (e.g. Rose-L, Harmony) radar missions with different wavelengths, revisit schedules and spatial resolutions to develop near-real-time monitoring across a multi-sensor array of data streams using a consistent methodology. The DPMM that we present is univariate, since we use PCA to condense the information from the HH and HV channels to a single dimension prior to clustering. We do this for computational reasons but, in theory, the DPMM clustering is extensible to an arbitrary number of dimensions simultaneously. Improvements to the classifier may therefore be possible by incorporating additional information, such as the Gray-Level Co-occurrence Matrix (Haralick et al., 1973) textures that Barbat et al. (2019) found improved their segmentation. Further improvements to object-level representations may be achieved through morphological operations to, for example, separate touching icebergs that are currently treated as a single object, and connect multiple objects that actually represent a single, larger iceberg. There is also substantial, unexploited information within the frequency rasters developed at the ensembling stage that may improve object discrimination. Such post-processing would add particular value in the most congested areas closest to calving fronts.

5. Conclusions

We have presented a novel, fully unsupervised approach to iceberg detection that outperforms most existing methods despite being applied to a spatial context in which iceberg discrimination is likely to be more challenging than those hitherto addressed. Average classifier performance is high, with pixel-level false alarms at 1.9% and misses at 6.1%. The method overcomes the problem of high backscatter variability between SAR scenes by its use of a non-parametric Dirichlet Process clustering algorithm. We will upscale the application in future work. Classifier training is conducted automatically within the workflow on a per-scene basis. The unsupervised nature of the algorithm, which does not require any operator input except for validation purposes, theoretically makes it ideally suited to deployment at scale. The method is also applicable close to calving fronts, which enables determination of iceberg provenance, and thereby glacier contributions to iceberg and associated meltwater dynamics. Future work will develop automated tracking algorithms exploiting the ability to track individual icebergs and their lineages from calving, through fragmentation events, to their disappearance below classifier resolution, thereby informing multiple scientific priorities.

CRediT authorship contribution statement

Ben Evans: Conceptualization, Methodology, Formal analysis, Investigation, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Anita Faul:** Conceptualization, Methodology, Software, Writing – review & editing. **Andrew Fleming:** Conceptualization, Methodology, Supervision, Writing – review & editing. **David G. Vaughan:** Conceptualization, Writing – review & editing. **J. Scott Hosking:** Supervision, Writing – review & editing, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Acknowledgements

We dedicate this paper to David Vaughan, a visionary and much missed colleague whose role in this work was pivotal. This work is supported by Wave 1 of The UKRI Strategic Priorities Fund under the EPSRC Grant EP/T001569/1, particularly the ‘AI for Science’ theme within that grant and The Alan Turing Institute.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.rse.2023.113780.

References

Albughdadi, M., Chaari, L., Tourneret, J.Y., Forbes, F., Ciuciu, P., 2017. A Bayesian non-parametric hidden Markov random model for hemodynamic brain parcellation. Signal Process. 135, 132–146. http://dx.doi.org/10.1016/J.SIGPRO.2017.01.005.

Arrigo, K.R., Van Dijken, G.L., Ainley, D.G., Fahnestock, M.A., Markus, T., 2002. Ecological impact of a large antarctic iceberg. Geophys. Res. Lett. 29 (7), 8–1–8–4. http://dx.doi.org/10.1029/2001GL014160.

Balafar, M.A., 2012. Gaussian mixture model based segmentation methods for brain MRI images. Artif. Intell. Rev. 41 (3), 429–439. http://dx.doi.org/10.1007/S10462-012-9317-3.

Bamber, J.L., Riva, R.E., Vermeersen, B.L., Lebrocq, A.M., 2009. Reassessment of the potential sea-level rise from a collapse of the west antarctic ice sheet. Science 324 (5929), 901–903. http://dx.doi.org/10.1126/SCIENCE.1169335.

Barbat, M.M., Rackow, T., Wesche, C., Hellmer, H.H., Mata, M.M., 2021. Automated iceberg tracking with a machine learning approach applied to SAR imagery: A weddell sea case study. ISPRS J. Photogramm. Remote Sens. 172, 189–206. http://dx.doi.org/10.1016/J.ISPRSJPRS.2020.12.006.

Barbat, M.M., Wesche, C., Werhli, A.V., Mata, M.M., 2019. An adaptive machine learning approach to improve automatic iceberg detection from SAR images. ISPRS J. Photogramm. Remote Sens. 156, 247–259. http://dx.doi.org/10.1016/J.ISPRSJPRS.2019.08.015.

Barnes, D.K., 2017. Iceberg killing fields limit huge potential for benthic blue carbon in antarctic shallows. Global Change Biol. 23 (7), 2649–2659. http://dx.doi.org/10.1111/GCB.13523.

Bauer, P., Stevens, B., Hazeleger, W., 2021. A digital twin of earth for the green transition. Nature Clim. Change 11 (2), 80–83. http://dx.doi.org/10.1038/S41558-021-00986-Y.

Biddle, L.C., Kaiser, J., Heywood, K.J., Thompson, A.F., Jenkins, A., 2015. Ocean glider observations of iceberg-enhanced biological production in the northwestern weddell sea. Geophys. Res. Lett. 42 (2), 459–465. http://dx.doi.org/10.1002/2014GL062850.

Bigg, G.R., Cropper, T.E., O'Neill, C.K., Arnold, A.K., Fleming, A.H., Marsh, R., Ivchenko, V., Fournier, N., Osborne, M., Stephens, R., 2018. A model for assessing iceberg hazard. Nat. Hazards 92 (2), 1113–1136. http://dx.doi.org/10.1007/S11069-018-3243-X/FIGURES/15.

Blair, G.S., 2021. Digital twins of the natural environment. Patterns 2 (10), 100359. http://dx.doi.org/10.1016/J.PATTER.2021.100359.

Budge, J.S., Long, D.G., 2018. A comprehensive database for antarctic iceberg tracking using scatterometer data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11, 434–442. http://dx.doi.org/10.1109/JSTARS.2017.2784186.

Castro, D.C., Glocker, B., 2016. Dirichlet process mixture models : Application to brain image segmentation. In: NIPS 2016 Workshop on Practical Bayesian Nonparametrics.

Cheng, W., Ma, L., Yang, T., Liang, J., Zhang, Y., 2016. Joint lung CT image segmentation: A hierarchical Bayesian approach. PLoS One 11 (9), e0162211. http://dx.doi.org/10.1371/JOURNAL.PONE.0162211.

Depoorter, M.A., Bamber, J.L., Griggs, J.A., Lenaerts, J.T., Ligtenberg, S.R., Van Den Broeke, M.R., Moholdt, G., 2013. Calving fluxes and basal melt rates of antarctic ice shelves. Nature 502 (7469), 89–92. http://dx.doi.org/10.1038/nature12567.

Derraz, F., Peyrodie, L., Pinti, A., Taleb-Ahmed, A., Chikh, A., Hautecoeur, P., 2010. Semi-automatic segmentation of multiple sclerosis lesion based active contours model and variational Dirichlet process'. Comput. Model. Eng. Sci. 67 (2), 95–117. http://dx.doi.org/10.3970/CMES.2010.067.095.

Duprat, L.P., Bigg, G.R., Wilton, D.J., 2016. Enhanced southern ocean marine productivity due to fertilization by giant icebergs. Nat. Geosci. 9 (3), 219–221. http://dx.doi.org/10.1038/ngeo2633, URL https://www.nature.com/articles/ngeo2633.

Faul, A.C., 2019. A Concise Introduction To Machine Learning, first ed. Chapman and Hall/CRC, http://dx.doi.org/10.1201/9781351204750.

Ferreira da Silva, A.R., 2007. A Dirichlet process mixture model for brain MRI tissue classification. Med. Image Anal. 11 (2), 169–182. http://dx.doi.org/10.1016/J.MEDIA.2006.12.002.

Ferreira da Silva, A.R., 2009. Bayesian mixture models of variable dimension for image segmentation. Comput. Methods Programs Biomed. 94 (1), 1–14. http://dx.doi.org/10.1016/J.CMPB.2008.05.010.

Forbes, F., Lu, H., Arbel, J., 2018. Non parametric Bayesian priors for hidden Markov random fields: application to image segmentation. In: BNPSI 2018 : Workshop on Bayesian Non Parametrics for Signal and Image Processing. Bordeaux, France, URL https://hal.archives-ouvertes.fr/hal-01941687/.

Gerrish, L., Fretwell, P., Cooper, P., 2021. High Resolution Vector Polygons of the Antarctic Coastline (7.4). UK Polar Data Centre, Natural Environment Research Council, UK Research & Innovation, http://dx.doi.org/10.5285/cdeb448d-10de-4e6e-b56b-6a16f7c59095.

Gill, R.S., 2001. Operational detection of sea ice edges and icebergs using SAR. Can. J. Remote Sens. 27 (5), 411–432. http://dx.doi.org/10.1080/07038992.2001.10854884.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern. 3, 610–621. http://dx.doi.org/10.1109/TSMC.1973.4309314, URL http://ieeexplore.ieee.org/document/4309314/.

Heiselberg, P., Sørensen, K.A., Heiselberg, H., Andersen, O.B., Maimaitiyiming, M., Alom, Z., Maimaitijiang, M., Heiselberg, P., Sørensen, K.A., Heiselberg, H., Andersen, O.B., 2022. SAR ship and iceberg discrimination in arctic conditions using deep learning. Remote Sens. 14 (9), 2236. http://dx.doi.org/10.3390/RS14092236.

Impoco, G., Tuminello, L., 2015. Incremental learning to segment micrographs. Comput. Vis. Image Underst. 140, 144–152. http://dx.doi.org/10.1016/J.CVIU.2015.03.007.

Karvonen, J., Gegiuc, A., Niskanen, T., Montonen, A., Buus-Hinkler, J., Rinne, E., 2021. Iceberg detection in dual-polarized C-band SAR imagery by segmentation and nonparametric CFAR (SnP-CFAR). IEEE Trans. Geosci. Remote Sens. 1–12. http://dx.doi.org/10.1109/TGRS.2021.3070312.

Koo, Y.H., Xie, H., Ackley, S.F., Mestas-Nuñez, A.M., Macdonald, G.J., Hyun, C.U., 2021. Semi-automated tracking of iceberg B43 using sentinel-1 SAR images via Google Earth Engine. Cryosphere 15 (10), 4727–4744. http://dx.doi.org/10.5194/TC-15-4727-2021.

Koo, Y., Xie, H., Mahmoud, H., Iqrah, J.M., Ackley, S.F., 2023. Automated detection and tracking of medium-large icebergs from sentinel-1 imagery using google earth engine. Remote Sens. Environ. 296, 113731. http://dx.doi.org/10.1016/J.RSE.2023.113731, URL https://linkinghub.elsevier.com/retrieve/pii/S0034425723002821.

Lancelot, C., De Montety, A., Goosse, H., Becquevort, S., Schoemann, V., Pasquer, B., Vancoppenolle, M., 2009. Spatial distribution of the iron supply to phytoplankton in the southern ocean: A model study. Biogeosciences 6 (12), 2861–2878. http://dx.doi.org/10.5194/BG-6-2861-2009.

Liang, D., Guo, H., Zhang, L., Li, H., Wang, X., 2021. Sentinel-1 EW mode dataset for antarctica from 2014–2020 produced by the casearth cloud service platform. Big Earth Data http://dx.doi.org/10.1080/20964471.2021.1976706.

Liu, Y., Moore, J.C., Cheng, X., Gladstone, R.M., Bassis, J.N., Liu, H., Wen, J., Hui, F., 2015. Ocean-driven thinning enhances iceberg calving and retreat of antarctic ice shelves. Proc. Natl. Acad. Sci. USA 112 (11), 3263–3268. http://dx.doi.org/10.1073/PNAS.1415137112, URL www.pnas.org/cgi/doi/10.1073/pnas.1415137112.

Macdonald, G.J., Ackley, S.F., Mestas-Nuñez, A.M., Blanco-Cabanillas, A., 2023. Evolution of the dynamics, area, and ice production of the amundsen sea polynya, antarctica, 2016–2021. Cryosphere 17, 457–476. http://dx.doi.org/10.5194/TC-17-457-2023.

Mantripragada, K., Qureshi, F.Z., 2022. Evaluation of Dirichlet process Gaussian mixtures for segmentation on noisy hyperspectral images. http://dx.doi.org/10.48550/arxiv.2203.02820, arXiv preprint arXiv:2203.02820.

Marino, A., Dierking, W., Wesche, C., 2016. A depolarization ratio anomaly detector to identify icebergs in sea ice using dual-polarization SAR images. IEEE Trans. Geosci. Remote Sens. 54 (9), 5602–5615. http://dx.doi.org/10.1109/TGRS.2016.2569450.

Matsuoka, K., Skoglund, A., Roth, G., de Pomereu, J., Griffiths, H., Headland, R., Herried, B., Katsumata, K., Brocq, A.L., Licht, K., Morgan, F., Neff, P.D., Ritz, C., Scheinert, M., Tamura, T., de Putte, A.V., van den Broeke, M., von Deschwanden, A., Deschamps-Berger, C., Liefferinge, B.V., Tronstad, S., Melvær, Y., 2021. Quantarctica, an integrated mapping environment for antarctica, the southern ocean, and sub-antarctic islands. Environ. Model. Softw. 140, 105015. http://dx.doi.org/10.1016/J.ENVSOFT.2021.105015.

Mazur, A.K., Wåhlin, A.K., Kalén, O., 2019. The life cycle of small-to medium-sized icebergs in the amundsen sea embayment. Polar Res. 38, 1–17. http://dx.doi.org/10.33265/polar.v38.3313.

Mazur, A.K., Wåhlin, A.K., Krężel, A., 2017. An object-based SAR image iceberg detection algorithm applied to the amundsen sea. Remote Sens. Environ. 189, 67–83. http://dx.doi.org/10.1016/J.RSE.2016.11.013.

Orbanz, P., Buhmann, J.M., 2006. Smooth Image Segmentation by Nonparametric Bayesian Inference. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 3951 LNCS, Springer, Berlin, Heidelberg, pp. 444–457. http://dx.doi.org/10.1007/11744023-35.

Parashar, S., Stapleton, G., Worsfold, R., O'Neil, R., 1980. Potential of SAR in detecting and monitoring icebergs. Cold Reg. Sci. & Technol. 1 (3–4), 195–210. http://dx.doi.org/10.1016/0165-232X(80)90048-8.

Partington, K.C., Oddy, C.J., Sephton, A.J., Walker, N.P., Willis, C.J., Fearon, M., Hodson, W., Walker, C., 1994. A demonstrator system for monitoring sea ice from space. WIT Trans. Built Environ. 5, 410. http://dx.doi.org/10.2495/CMO940371.

Pedregosa, F., Michel, V., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Pedregosa, F., Varoquaux, G., Gramfort, A., Thirion, B., Grisel, O., Dubourg, V., Passos, A., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine learning in Python. J. Machine Learn. Res. 12, 2825–2830, URL https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf.

Power, D., Youden, J., Lane, K., Randell, C., Flett, D., 2001. Iceberg detection capabilities of radarsat synthetic aperture radar. Can. J. Remote Sens. 27 (5), 476–486. http://dx.doi.org/10.1080/07038992.2001.10854888.

Raiswell, R., Benning, L.G., Tranter, M., Tulaczyk, S., 2008. Bioavailable iron in the southern ocean: The significance of the iceberg conveyor belt. Geochem. Trans. 9 (1), 1–9. http://dx.doi.org/10.1186/1467-4866-9-7/TABLES/2.

Rawson, R., Larson, R., Shuchman, R., Worsfold, R., 1979. The capability of SAR systems for iceberg detection and characterization. Desalination 29 (1–2), 109–133. http://dx.doi.org/10.1016/S0011-9164(00)88419-2.

Robinson, N.J., Williams, M.J., 2012. Iceberg-induced changes to polynya operation and regional oceanography in the southern ross sea, antarctica, from in situ observations. Antarctic Sci. 24 (5), 514–526. http://dx.doi.org/10.1017/S0954102012000296.

Scambos, T.A., Bell, R.E., Alley, R.B., Anandakrishnan, S., Bromwich, D.H., Brunt, K., Christianson, K., Creyts, T., Das, S.B., DeConto, R., Dutrieux, P., Fricker, H.A., Holland, D., MacGregor, J., Medley, B., Nicolas, J.P., Pollard, D., Siegfried, M.R., Smith, A.M., Steig, E.J., Trusel, L.D., Vaughan, D.G., Yager, P.L., 2017. How much, how fast?: A science review and outlook for research on the instability of antarctica's thwaites glacier in the 21st century. Glob. Planet. Change 153, 16–34. http://dx.doi.org/10.1016/J.GLOPLACHA.2017.04.008.

Schloesser, F., Friedrich, T., Timmermann, A., DeConto, R.M., Pollard, D., 2019. Antarctic iceberg impacts on future southern hemisphere climate. Nat. Clim. Chang. 9 (9), 672–677. http://dx.doi.org/10.1038/s41558-019-0546-1, URL https://www.nature.com/articles/s41558-019-0546-1.

Sephton, A.J., Partington, K.C., 1998. Towards operational monitoring of arctic sea ice by SAR. In: Tsatoulis, K., Kwok, R. (Eds.), Analysis of SAR Data of the Polar Oceans. Springer, Berlin, Heidelberg, pp. 259–279. http://dx.doi.org/10.1007/978-3-642-60282-5_12.

Shapiro, L.G., 1996. Connected component labeling and adjacency graph construction. Mach. Intell. Pattern Recogn. 19, 1–30. http://dx.doi.org/10.1016/S0923-0459(96)80011-5.

Silva, T.A., Bigg, G.R., 2005. Computer-based identification and tracking of antarctic icebergs in SAR images. Remote Sens. Environ. 94 (3), 287–297. http://dx.doi.org/10.1016/J.RSE.2004.10.002.

Silva, T.A., Bigg, G.R., Nicholls, K.W., 2006. Contribution of giant icebergs to the southern ocean freshwater flux. J. Geophys. Res.: Oceans 111 (C3), http://dx.doi.org/10.1029/2004JC002843.

Slater, T., Hogg, A.E., Mottram, R., 2020. Ice-sheet losses track high-end sea-level rise projections. Nat. Clim. Chang. 10 (10), 879–881. http://dx.doi.org/10.1038/s41558-020-0893-y.

Smith, K.L., Robison, B.H., Helly, J.J., Kaufmann, R.S., Ruhl, H.A., Shaw, T.J., Twining, B.S., Vernet, M., 2007. Free-drifting icebergs: hot spots of chemical and biological enrichment in the weddell sea. Science 317 (5837), 478–482. http://dx.doi.org/10.1126/science.1142834.

Soldal, I.H., Dierking, W., Korosov, A., Marino, A., 2019. Automatic detection of small icebergs in fast ice using satellite wide-swath SAR images. Remote Sens. 11 (7), 806. http://dx.doi.org/10.3390/RS11070806.

Song, W., Li, M., Zhang, P., Wu, Y., Jia, L., An, L., 2017. Unsupervised PolSAR image classification and segmentation using Dirichlet process mixture model and Markov random fields with similarity measure. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (8), 3556–3568. http://dx.doi.org/10.1109/JSTARS.2017.2684301.

Song, L., Peters, D.K., Huang, W., Power, D., 2021. Ship-iceberg discrimination from sentinel-1 synthetic aperture radar data using parallel convolutional neural network. Concurr. Comput.: Pract. Exper. 33 (17), e6297. http://dx.doi.org/10.1002/CPE.6297.

Stammerjohn, S.E., Maksym, T., Massom, R.A., Lowry, K., Arrigo, K.R., Yuan, X., Raphael, M., Randall-Goodwin, E., Sherrell, R.M., Yager, P.L., 2015. Seasonal sea ice changes in the amundsen sea, antarctica, over the period of 1979–2014. Elementa 3, 55. http://dx.doi.org/10.12952/JOURNAL.ELEMENTA.000055/112737.

Stern, A.A., Adcroft, A., Sergienko, O., 2016. The effects of antarctic iceberg calving-size distribution in a global climate model. J. Geophys. Res.: Oceans 121 (8), 5773–5788. http://dx.doi.org/10.1002/2016JC011835.

Sun, L., Zhang, Y., Tian, G., Ma, M., 2009. SAR image segmentation using GHM-based Dirichlet process mixture models. In: Proceedings of the 2009 International Joint Conference on Computational Sciences and Optimization, Vol. 1. CSO 2009, pp. 886–888. http://dx.doi.org/10.1109/CSO.2009.371.

Topouzelis, K., Singha, S., 2016. Incidence angle normalization of wide Swath SAR data for oceanographic applications. Open Geosci. 8 (1), 450–464. http://dx.doi.org/10.1515/GEO-2016-0029/MACHINEREADABLECITATION/RIS.

Vernet, M., Smith, K.L., Cefarelli, A.O., Helly, J.J., Kaufmann, R.S., Lin, H., Long, D.G., Murray, A.E., Robison, B.H., Ruhl, H.A., Shaw, T.J., Sherman, A.D., Sprintall, J., Stephenson, G.R., Stuart, K.M., Twining, B.S., 2012. Islands of ice: Influence of free-drifting antarctic icebergs on pelagic marine ecosystems. Oceanography 25, 38–39. http://dx.doi.org/10.5670/OCEANOG.2012.72.

Wadhams, P., 1988. Winter observations of iceberg frequencies and sizes in the south atlantic ocean. J. Geophys. Res.: Oceans 93 (C4), 3583–3590. http://dx.doi.org/10.1029/JC093IC04P03583.

Wesche, C., Dierking, W., 2012. Iceberg signatures and detection in SAR images in two test regions of the Weddell Sea, Antarctica. J. Glaciol. 58, 325–339. http://dx.doi.org/10.3189/2012J0G11J020.

Wesche, C., Dierking, W., 2015. Near-coastal circum-antarctic iceberg size distributions determined from synthetic aperture radar images. Remote Sens. Environ. 156, 561–569. http://dx.doi.org/10.1016/J.RSE.2014.10.025.

Williams, R.N., Rees, W.G., Young, N.W., 1999. A technique for the identification and analysis of icebergs in synthetic aperture radar images of antarctica. 20 (15–16), 3183–3199. http://dx.doi.org/10.1080/014311699211697.

Xiao, X., Shen, X., Ke, C., Zhou, X., 2020. Comparison of machine learning algorithms based on sentinel-1A data to detect icebergs. Acta Geod. Cartogr. Sinica 49 (4), 509. http://dx.doi.org/10.11947/J.AGCS.2020.20190174.

Young, N.W., Turner, D., Hyland, G., Williams, R.N., 1998. Near-coastal iceberg distributions in east antarctica, 50-145°E. Ann. Glaciol. 27, 68–74. http://dx.doi.org/10.3189/1998AOG27-1-68-74.