# UNSUPERVISED CLASSIFICATION OF THE ANTARCTIC MARGINAL ICE ZONE

A PREPRINT

**Noah S. Day**
University of Adelaide, Australia
noah.day@adelaide.edu.au

**Luke G. Bennetts**
University of Adelaide, Australia
luke.bennetts@adelaide.edu.au

**Siobhan P. O'Farrell**
CSIRO Environment, Australia
siobhan.o'farrell@csiro.au

**Alberto Alberello**
University of East Anglia, UK
a.alberello@uea.ac.uk

**Fabien Montiel**
University of Otago, New Zealand
fabien.montiel@otago.ac.nz

**Key Points:**

- Statistical clustering applied to sea ice model outputs including floe sizes to define the Antarctic marginal ice zone
- Predicted marginal ice zone widths are similar to satellite observations of wave penetration distances
- Sea ice melt rates and drift speeds shown to be greater in the marginal ice zone than the interior ice pack

## ABSTRACT

The Antarctic marginal ice zone is the regularly wave-affected outer band of the sea ice covered Southern Ocean. The ice cover in the marginal ice zone is typically unconsolidated and contains smaller, thinner ice floes than the inner ice pack, which makes it a highly dynamic region and susceptible to rapid expansion or contraction. Here, an unsupervised statistical method is used to cluster sea ice data from 2010–2019 simulated by a global sea ice model (CICE6 combined with a waves propagation module), such that it defines a sea ice region with marginal ice zone characteristics. Floe size is shown to be the key variable in classifying the marginal ice zone in the statistical method. The method is shown to give marginal ice zone widths similar to those derived from satellite observations of wave penetration distances, but contrasts with those using the 15–80% areal sea ice concentration definition, particularly during austral winter. Using the proposed definition, the marginal ice zone is found to undergo a seasonal transition due to new ice formation in winter, increased drift in spring, and increased rates of wave-induced breakup and melting in summer. The study motivates incorporation of wave and floe-scale processes in sea ice models, and the methods are available for application to outputs from high-resolution and coupled sea ice–ocean–wave models for more detailed studies of the marginal ice zone (in both hemispheres).



## Plain language summary

The record setting responses of Antarctic sea ice extent to climate change over recent decades has generated intense research interest in the marginal ice zone at the outskirts of the ice cover. The presence of ocean waves in this region creates a granular ice cover composed of small floes that affects heat fluxes between the atmosphere and ocean, whilst shielding the inner pack ice, fast ice, and ice shelves from waves. Marginal ice zone studies are restricted by the absence of a pragmatic definition that is compatible with its description as the wave-affected sea ice region. Most studies use concentration-based definitions, despite these overlooking large wave-affected areas. We propose a statistical definition informed by sea ice properties, including floe sizes, and apply the definition to model outputs, leveraging on advances in global sea ice models. The resulting marginal ice zone is composed of young-small floes during winter and older-fractured floes during summer. The predicted marginal ice zone widths are consistent with observations of wave penetration distances and contrast with the concentration-based definition during winter. The proposed approach to defining the marginal ice zone may provide the basis for future studies of its contribution to Antarctic sea ice variability.


## 1 Introduction

The Antarctic marginal ice zone (MIZ) is the region at the outer margins of the sea ice covered Southern Ocean, where the ice cover is regularly affected by ocean-surface waves (as well as other open-ocean processes) (Wadhams, 1986; Bennetts et al., 2022b). Intense atmospheric systems and circumpolar-scale fetches of the Southern Ocean generate large swell waves (Bennetts et al., 2023), which create a considerably larger MIZ than in the Arctic Ocean (Weeks, 2010) (although MIZ conditions are becoming increasingly common in the Arctic, see Stopa et al., 2016; Thomson, 2022). Wave energy attenuates with distance travelled through the MIZ (Meylan et al., 2014; Montiel et al., 2016), so that the MIZ protects the inner ice pack (Squire, 2007), fast ice (Fraser et al., 2023), and ice shelves (Massom et al., 2018; Teder et al., 2022) from waves. Waves regulate the properties of the ice cover in the MIZ, such that the ice floes are relatively small (in size) and are typically unconsolidated (Toyota et al., 2011, 2016; Alberello et al., 2019). The nature of the MIZ ice cover promotes heat exchanges and drag between both the atmosphere and ocean (McPhee et al., 1987; Vihma et al., 2014), enhancing sea ice drift (Alberello et al., 2020) and melting (Horvat et al., 2016). Indeed, trends in wave heights and sea ice extents are linked in the Southern Ocean (Kohout et al., 2014), suggesting the rapid expansion and contraction of the MIZ in response to dynamic and thermodynamic forcing (Massom et al., 2008; Vichi et al., 2019) plays a role in the long-term variability of Antarctic sea ice (J. Turner & Comiso, 2017). Thus, the properties of the MIZ have potentially important implications for the changing state of Antarctic sea ice (Eayrs et al., 2021; Bennetts et al., 2022a), and it has experienced over a decade of increasing research activity motivated by the response of both Arctic and Antarctic sea ice to climate change (Bennetts et al., 2022b).

Studies of the MIZ (in both hemispheres) are prohibited by the practical challenge of defining the MIZ as the (regularly) wave-affected region (Squire, 2022). An areal sea ice concentration (SIC) threshold of 15–80% has been the standard MIZ definition (10–80% has also been used by NSIDC, 2023), where the lower and upper limit correspond with definitions of sea ice extent (SIE) and “close ice" respectively (Comiso, 2006; WMO, 2014), although these thresholds have no apparent relation to MIZ dynamics (such as those from waves Dumont, 2022). In particular, during austral winter when the Antarctic sea ice extent is greatest, the outer margins of the ice cover contains large areas near 100%

SIC, in the form of pancake ice and interstitial frazil (Wadhams et al., 1987; Alberello et al., 2019), whilst hosting large waves (Alberello et al., 2022). Creating a pragmatic definition of the Antarctic MIZ, in keeping with its underpinning characteristic as the wave-affected region, is the first priority for the research field (Squire, 2022).

The SIC-threshold definition has been applied to satellite data in both hemispheres (Strong, 2012; Strong & Rigor, 2013; Stroeve et al., 2016; Strong et al., 2017) and compared to model data in the Arctic (Rolph et al., 2020). The SIC-defined Antarctic MIZ increases linearly in area from austral summer to spring, before peaking in December and returning to a minimum in February (Stroeve et al., 2016; Vichi, 2022). In response to observations of Antarctic sea ice drift in regions where the SIC surpasses 80% (Alberello et al., 2020; Womack et al., 2022), Vichi (2022) proposed a method based on the standard deviation of daily SIC anomalies from a monthly-climatological mean (using satellite data) to capture these areas and identify the likelihood of encountering MIZ conditions. Vichi (2022) showed that the SIC-anomaly definition produced a greater rate of growth in Antarctic MIZ area between summer–early-winter then remained relatively constant until November where it peaked to reach a maximum in December. Soleymani and Scott (2023) applied the method to Arctic sea ice.

Techniques have recently been developed to capture snapshots of wave heights within the sea ice cover from ICESat-2 altimetry data in the Antarctic (Horvat et al., 2020; Brouwer et al., 2022), leading to wave-based MIZ definitions (Horvat et al., 2020, also produced results for the Arctic). ICESat-2 orbits the Earth 15 times per day and provides an along-track altimeter dataset (ATL07 product) with global coverage from late-2018 to present. Brouwer et al. (2022) applied additional selection criteria to negate uncertainties in wave heights arising from cloud cover, which enabled wave amplitudes smaller than ice freeboard variability to be detected, and resulted in a wider Antarctic MIZ than was found by Horvat et al. (2020). However, both methods produced a maximum MIZ area (and width) during austral winter and a minimum in summer. The seasonality of the wave-defined MIZ area (or width) and significant wave heights in the Southern Ocean (Young et al., 2020) broadly agree (Brouwer et al., 2022), contrasting with results using SIC-based definitions.

A floe size criterion has been proposed to define the MIZ (Squire, 2022; Dumont, 2022), since the sizes of floes in MIZ are typically orders of magnitude smaller than floes in the interior pack. Application of such a definition to observations is restricted by the scarcity of floe size data, e.g., Dumont (2022) used floe size to determine the MIZ width over the southern Fram Strait (Arctic). Floe size distributions (FSD) have recently been implemented into global- and hemispheric-scale sea ice models (e.g., CICE, LIM, neXtSIM). The strong relationship between waves and the FSD (Dumont et al., 2011; Williams et al., 2013a, 2013b; Montiel & Squire, 2017; Mokus & Montiel, 2022) motivated the inclusion of wave impacts within sea ice models (Horvat & Tziperman, 2015; Zhang et al., 2015), and the integration into global sea ice models (Bennetts et al., 2017; Roach, Horvat, et al., 2018; Roach et al., 2019; Bateson et al., 2020; Boutin et al., 2020). Standalone global sea ice models were used to limit computational cost and more easily identify impacts of the FSD on sea ice area and volume. These studies either prescribed the shape of the FSD or used a prognostic FSD (i.e., the shape arises from model physics). Bennetts et al. (2017) and Bateson et al. (2020) both used the waves-in-ice module (Williams et al., 2013a, 2013b) to model floe breakup, and prescribed power laws to the FSD, with simple schemes to represent the welding between floes. Roach, Horvat, et al. (2018) implemented the prognostic joint floe size and ice thickness distribution (FSTD) from Horvat and Tziperman (2015) with the rate of welding determined by the geometric probability that two floes overlap. Roach et al. (2019) furthered this model by letting waves determine the size of new floes (Shen et al., 2001) and coupled CICE to a wave model (WAVEWATCH III). Bateson et al. (2022) tested the effects of these two FSD methods in the Arctic, and found the power law produced better agreement with FSD observations. However, the model with a power-law FSD could not replicate the heterogeneity of floe sizes within the ice cover or rate at which small floes would melt as seen when using the prognostic FSD. Boutin et al. (2018) implemented a power law FSD within WAVEWATCH III, before coupling the wave model to LIM3 and including additional physics that increased ice drift in the Arctic MIZ (Boutin et al., 2020, 2021). Boutin, Williams, Horvat, and Brodeau (2022) found agreement between MIZ areas when using a floe-size threshold and a wave-height threshold on Arctic model output, and were comparable to the observations of Horvat et al. (2020). The FSTD has been included within CICE6 (the Community Ice CodE version 6 Hunke et al., 2017), i.e., the sea ice component of CESM2 (Community Earth System Model version 2). CICE6 allows the simulation of a comprehensive sea ice dataset to investigate which sea ice variables are most indicative of MIZ dynamics.

In this article, we propose a statistical definition of the Antarctic MIZ based on a set of (static) sea ice properties, including floe size. The definition is applied to a decade of daily outputs from the CICE6 model combined with a wave-propagation module, leveraging on the FSTD's capability (which was introduced into the latest version of the CICE model). The definition utilises an unsupervised machine learning algorithm ($k$-means clustering) to classify the data into distinct sea ice regions. Unsupervised methods have been used to separate different sea ice types (e.g., Massom et al., 1999, used satellite data), and the $k$-means algorithm has been shown to be appropriate for climate science applications (Wilks, 2011) including sea ice data retrieved from satellites (Farooq et al., 2023) and models (LIM3 Moreno-Chamarro et al., 2020). We specify the number of clusters to $k$-means based on a wave heuristic, such that

the outer cluster is the regularly wave-affected sea ice region, i.e., the description of the MIZ, without explicitly including waves properties in the clustered dataset. MIZ widths produced by the $k$-means MIZ definition are compared with those from the standard SIC-threshold definition, and assessed against satellite observations of wave penetration distances (Brouwer et al., 2022). Further, the seasonal physical processes that drive the summer and winter MIZ are investigated.

sectionMethods

### 1.1 Numerical Model

#### 1.1.1 CICE6 Sea Ice Model

We use sea ice data outputs from the CICE6 sea ice model (v6.2.0) on a tripolar grid with a horizontal resolution of $1^\circ$ (latitude–longitude). CICE6 is a global continuum sea ice model that simulates the evolution of sea ice thickness, $h$, and floe size (radius), $r$, as a joint floe size and thickness distribution (FSTD), $f(\mathbf{x}, t; r, h)$, for spatial locations on the ocean surface, $\mathbf{x}$, at times, $t$. The FSTD is defined such that the areal sea ice concentration within a predefined floe size and thickness category is given by $f(r, h)\,\mathrm{d}r\,\mathrm{d}h$ (Roach, Horvat, et al., 2018), where $\mathrm{d}r$ and $\mathrm{d}h$ are the widths of the respective categories. The floe size distribution (FSD) is returned by integrating over thickness, the ice thickness distribution (ITD) is returned by integrating over floe size, and the areal sea ice concentration, $a_i$, is achieved by integrating over both thickness and floe size.

The equation for evolution of the FSTD is (Horvat & Tziperman, 2015)

$$\frac{\partial f(r,h)}{\partial t} = -\nabla \cdot (f(r,h)\mathbf{u}) + \mathcal{L}_{\mathcal{T}} + \mathcal{L}_{\mathcal{M}} + \mathcal{L}_{\mathcal{W}}. \tag{1}$$

The terms on the right-hand side of Eq. (1) denote advection, thermodynamics, mechanical deformation processes and wave-induced breakup of sea ice, respectively. Horizontal ice velocity, $\mathbf{u}$, is determined by the output of an elastic-viscous-plastic (EVP) rheological model (Hunke & Dukowicz, 1997). The representative floe size, $r_a$, is defined by (Roach, Horvat, et al., 2018)

$$r_a = \frac{1}{a_i}\int_{\mathcal{R}}\int_{\mathcal{H}} r f(r,h)\,\mathrm{d}h\,\mathrm{d}r, \tag{2}$$

and, similarly, the mean ice thickness, $h_i$, is defined by (Thorndike et al., 1975)

$$h_i = \frac{1}{a_i}\int_{\mathcal{R}}\int_{\mathcal{H}} h f(r,h)\,\mathrm{d}h\,\mathrm{d}r, \tag{3}$$

Thermodynamics that affect the FSD include lateral growth and melt, as well as new floe formation and welding of existing floes. Ocean surface waves directly change the floe size dimension of the FSTD by (i) limiting the size of newly formed floes (promoting the formation of pancake ice), or (ii) fracturing existing floes via a flexural strain breakup criterion (Horvat & Tziperman, 2015). Given a regular wave (single direction, period and amplitude), the maximum floe size of newly formed floes, $r_{\max}$, is determined by a tensile failure limit (Shen et al., 2001)

$$r_{\max} \approx \sqrt{2\frac{C_2\lambda^2}{\pi^3 A g \rho_i}}, \tag{4}$$

where $\lambda$ is wavelength, $A$ is wave amplitude, $g$ is gravitational acceleration, $\rho_i$ is ice density, and $C_2 = 0.167\,\mathrm{kg\,m^{-1}\,s^{-2}}$ (Roach, Smith, & Dean, 2018). In irregular sea states, the wavelength is approximated by the representative (peak) wavelength, $\lambda_p$, which is calculated using the deep-water surface gravity wave dispersion relation (Williams et al., 2013a),

$$\lambda_p = \frac{2\pi g}{\omega_p^2}, \tag{5}$$

where $\omega_p$ is peak-angular frequency (as is used by Roach et al., 2019). By assumption, the amplitude $A = H_s/2$ (Roach, Horvat, et al., 2018), where $H_s$ is significant wave height and is defined by

$$H_s = 4\sqrt{\int S(\omega)\,\mathrm{d}\omega} \tag{6}$$

for a wave spectral density function (SDF), $S(\omega)$, in terms of angular frequency. $H_s$ and $\omega_p$ are input into CICE6 from the waves-in-ice module.

The joint distribution is discretised into a specified number of bins, with default settings allocating five bins for the ITD (Bitz et al., 2001) and twelve bins for the FSD. Consistent with previous studies using the FSTD in CICE, the minimum and maximum floe sizes are set to 0.067 m and 850 m, respectively (Roach, Horvat, et al., 2018; Roach et al., 2019; Bateson et al., 2020). An exponential bin spacing scheme is applied to provide finer resolutions for smaller floe sizes. Increasing the number of FSD bins to sixteen, while offering minimal improvement in agreement with FSD observations, incurs a computational time cost of approximately 60% (Bateson et al., 2020).

### 1.1.2 Waves-in-Ice Module (CICE6-WIM)

We incorporate a waves-in-ice module (WIM) into CICE6 (to create CICE6-WIM), which is a modified version of the WIM developed by Bennetts et al. (2017) for CICE4. The breakup model included in CICE6 is used (Horvat & Tziperman, 2015), rather than directly coupling wave-induced ice breakup and wave attenuation (Bennetts et al., 2017), meaning that no wave energy is lost when wave-induced breakup occurs. The WIM prescribes a frequency–direction SDF one cell north of the 1% SIC edge, where the influence of sea ice on waves is assumed to be negligible. This is achieved by applying an idealised wave-energy spectrum and prescribing $H_s$, $\omega_p$, and mean wave direction ($\bar{\theta}$ in radians). The mean wave parameters are taken from a global WAVEWATCH III (WW3) hindcast (Smith et al., 2021).

Assuming a single-peaked directional spectrum, the frequency-directional spectrum, $E(\omega, \theta)$, is given by,

$$E(\omega, \theta) = S(\omega) D(\theta - \bar{\theta}), \tag{7}$$

where $E(\omega, \theta)$ is the directional SDF, $S(\omega)$ is the frequency SDF, and $D(\theta - \bar{\theta})$ is a frequency-independent angular spreading function centred around the mean wave direction. $E(\omega, \theta)$ is integrated over the directional spectrum, such that only southward-waves enter the sea ice (due south $\pm\pi/2$). We define due south at 0, hence, the incident frequency SDF, $S_{\text{in}}(\omega)$, is initialised by

$$S_{\text{in}}(\omega) = \int_{-\pi/2}^{\pi/2} E(\omega, \theta)\, \mathrm{d}\theta. \tag{8}$$

$S(\omega)$ is initialised using a Bretschneider spectrum (Bretschneider, 1959),

$$S(\omega) = \frac{5}{16} H_s^2 \frac{\omega_p^4}{\omega^5} e^{-\frac{5}{4}\left(\frac{\omega_p}{\omega}\right)^4}, \tag{9}$$

and directional spreading is controlled by a cosine-$2s$ model (Longuet-Higgins, 1963),

$$D(\theta - \bar{\theta}) = \begin{cases} C \cos^{2s}\left(\frac{\theta - \bar{\theta}}{2}\right) & \text{for } \theta - \bar{\theta} \in [-\frac{\pi}{2}\, \frac{\pi}{2}], \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where $C$ is a normalising constant. The degree of directional spreading by the cosine-$2s$ function is determined by $s$. Since the Southern Ocean is dominated by swell, the circular standard deviation of the directional wave energy spectrum, $\sigma_\theta$, is narrower than wind-driven sea states (Derkani et al., 2021). In accordance with Derkani et al. (2021) who reported that $\sigma_\theta \leq 30^\circ$ when the sea state is dominated by swell, we set $s = 2.5$.

The incident frequency-SDFs, $S_{\text{in}}(\omega)$, are propagated directly southward along meridional lines (i.e., 1D) with frequency-dependent wave attenuation. Wave attenuation is applied cell-wise to the incident SDF by,

$$S_i(\omega) = S_{\text{in}}(\omega) e^{-a_i(x)\alpha(\omega)x} \tag{11}$$

where $S_i$ is the SDF within the ice cover and $x$ is the distance from the ice edge. The attenuation coefficient is determined from an empirical model for Antarctic sea ice, which is scaled by SIC ($a_i$) (Meylan et al., 2014; Bennetts et al., 2017), such that

$$\alpha(\omega) = \frac{1}{a_{i,\text{MBK}}} (\beta_2 \omega^2 + \beta_4 \omega^4), \tag{12}$$

where $\beta_2 = 5.38 \times 10^{-5}$, $\beta_4 = 2.95 \times 10^{-5}$ (units of $\text{s}^2\text{m}^{-1}$ and $\text{s}^4\text{m}^{-1}$, respectively), and $a_{i,\text{MBK}} = 0.7$ is the SIC observed in Meylan et al. (2014). CICE6-WIM is forced by the WW3 hindcast and waves are propagated on an hourly basis (matching the model's thermodynamic timestep) and steady-state conditions are assumed for propagation.

### 1.1.3 Model Configuration

We conduct a CICE6-WIM simulation spanning from 2005 to 2019, with the first five years of simulation treated as spin-up, and, thus, discarded from analysis. The simulation was forced inter-annually by atmospheric (Tsujino et al., 2018), oceanic (Kiss et al., 2020), and ocean-surface-wave data (Smith et al., 2021). CICE6 includes a slab ocean to

model the mixed-layer ocean, which was forced by sea surface temperature, and salinity (surface current velocities were also input into CICE6-WIM). We use the CICE6 default restoring time of 90 days and fixed mixed-layer depth of 20 m. Previous, Arctic-focused studies have used a variety of constant restoring times (e.g., 5 or 20 days; Schröder et al., 2019; Bateson et al., 2020) and a restoring time dependent on the mixed-layer temperature (Tsamados et al., 2015). Petty, Holland, and Feltham (2014) incorporated a mixed-layer ocean model in CICE, but this was shown to be most influential in Antarctic ice shelf seas (rather than in the MIZ). A simulation of ACCESS-OM2 (a global coupled ocean–sea ice model using MOM5-CICE5) which was submitted to CMIP6 was used as oceanic forcing (Kiss et al., 2020). For consistency with ACCESS-OM2, CICE6-WIM runs on the same model grid, and uses four ice layers with one snow layer to resolve the thermodynamics. Initial conditions were adopted from ACCESS-OM2, however as ACCESS-OM2 used CICE5 (with a constant floe size of 300 m), the FSD was initialised globally using a fixed power law distribution (Perovich & Jones, 2014).

### 1.2 Data Analysis

#### 1.2.1 Data Overview and Preparation

We perform a $k$-means cluster analysis on daily CICE6-WIM model outputs of Antarctic sea ice data spanning from 2010–2019. Given a specified number of clusters, $k$, $k$-means classifies data into $k$ distinct clusters by minimising the Euclidean distance between each data vector and the cluster centroids (MacQueen, 1967). We select five variables to describe different sea ice types: areal sea ice concentration ($a_i$), mean ice thickness ($h_i$), mean snow thickness, representative floe size ($r_a$), and ice age. Ice age is the average age of the ice within each grid cell and is computed as an ice volume-weighted tracer. These variables have been used to describe sea ice types in previous studies, e.g., unconsolidated pancake or brash ice, young to first-year ice with little snow cover, thicker first-year ice with a thicker snow cover, etc. (Massom et al., 1999). We apply a 15% SIC mask (as is used by Comiso & Zwally, 1984; NSIDC, 2023) to remove open-ocean cells, so that each cluster can be interpreted as a different sea ice type. To complete $k$-means on data vectors with elements of varying units, the variables must be non-dimensionalised to limit the chance of a single variable dominating the clustering (Wilks, 2011). For example, floe size (which we now use interchangeably with the representative radius, $r_a$) has a standard deviation of 300 m, whereas all other variables have standard deviations less than 1 unit. Therefore, we complete min-max scaling on all variables, so that their range is limited between 0 and 1. Since this is a linear transformation, the shapes of the distributions are not affected.

#### 1.2.2 Cluster Number Selection

Figures 1(a–c) show snapshots from 2019 of the $k$-means clustered sea ice regions during the mean sea ice extent maximum over 2010–2019 (26th of August) for $k = 2$–4, respectively. Regardless of the number of clusters specified to $k$-means, the clusters produce a pattern of zonal sea ice regions with the southermost cluster being confined to the Ross and Weddell seas. The outer band of sea ice resembles a MIZ (Cluster 1; shown in red) as it separates the open ocean from the inner pack, however the number of clusters specified control its width. When $k = 2$, the sea ice is divided into two zonal bands of similar areas. The outer band tends to be the widest east of the Weddell Sea and along the Antarctic Peninsula ($\sim 800$ km), but narrows in the Ross and Weddell seas ($\sim 400$ km). When $k$ is increased from 2 to 3, the exterior cluster for $k = 2$ is split into two regions, with the exterior cluster for $k = 3$ (i.e., Cluster 1 in Fig. 1) narrower than for $k = 2$ (mostly between 100–250 km). The inner part of Cluster 1 for $k = 2$ (i.e., Cluster 2 for $k = 3$) is an intermediary region between the exterior cluster and the interior ice pack. The addition of a fourth cluster ($k = 4$) predominantly separates the innermost cluster (Cluster 3 for $k = 3$) into two regions, and pushes out the outer two regions (Cluster 1 and Cluster 2 for $k = 3$). Consequently, the MIZ for $k = 4$ is typically restricted to only the outermost cell of the ice cover, and no cells for some longitudes.

We determine an appropriate number of clusters for the present study, such that the MIZ (Cluster 1) agrees (indirectly) with its description as the “actively wave-affected" area of sea ice (Wadhams, 1986). We interpret this description of the MIZ to be the region where incoming waves from the open ocean remain energetic ($H_s$ order tens of centimetres to metres), with only occasional high-energy waves ($H_s > 1$ m) able to reach (and potentially breakup) the interior sea ice pack (Kohout et al., 2014). We apply a wave-statistic heuristic based on the significant wave height ($H_s$) at the interior (southern) boundary of the MIZ (Cluster 1) to quantify the distribution of wave energy transmitted into the interior sea ice pack (i.e., typically the northern boundary of Cluster 2). Histograms of $H_s$ at the MIZ's interior boundary for $k = 2$, 3, and 4 (over 2010–2019) are displayed in Fig. 1(d–f), respectively. The mean $H_s$ along this boundary for $k = 2$–4 are 0.08 m, 0.46 m, and 1.01 m, respectively. Therefore, we discount $k = 2$, as the waves have become vanishingly small at the southern boundary of its MIZ. The exceedance probability of large waves entering the interior ice pack (i.e., $H_s > 1$ m) is 0.16 for $k = 3$ and 0.35 for $k = 4$. These large waves have a corresponding mean representative wavelength of 250 m for $k = 3$, and 150 m for $k = 4$. Therefore, $k = 3$ is most closely aligned with the requirement for

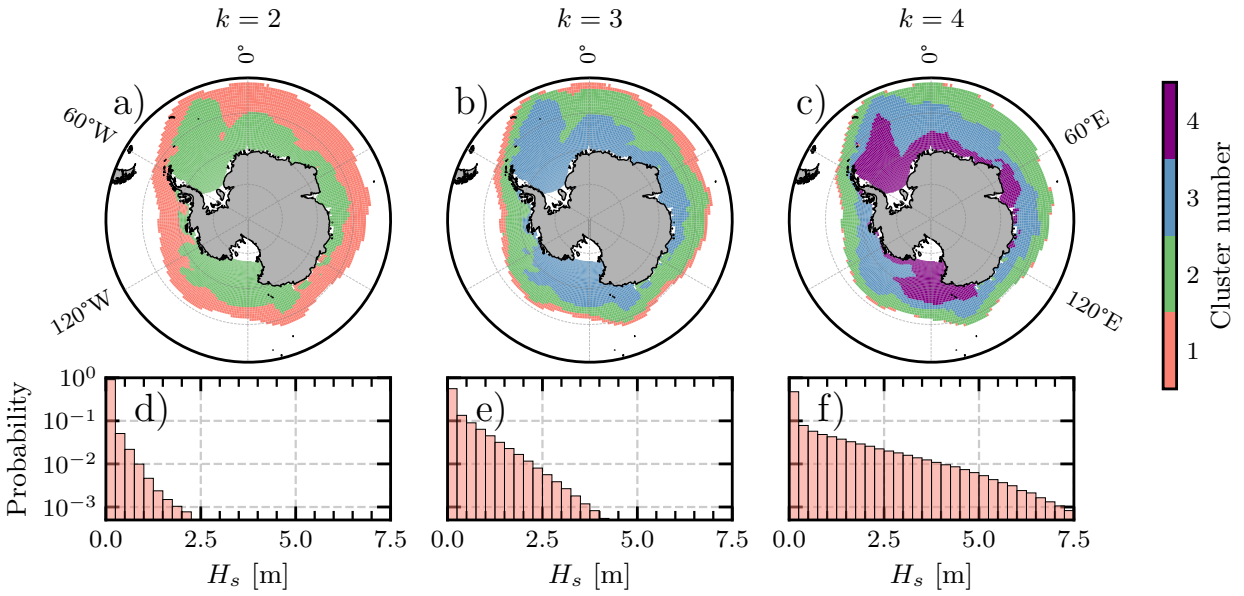


Figure 1: Changes in Antarctic sea ice classification by the number of clusters ($k$) specified to $k$-means. a–c) Snapshot maps of the sea ice clusters from 2019 at the average sea ice extent maximum over 2010–2019 (26th August), where a) $k = 2$, b) $k = 3$, and c) $k = 4$. A longitudinal-latitudinal grid (grey-dashed lines; in 60° and 10° increments, respectively) is overlaid onto the maps. d–f) Histograms of significant wave heights ($H_s$) at the circumpolar southern boundary of Cluster 1 (red) over 2010–2019 with a semi-log scale, for d) $k = 2$, e) $k = 3$, and f) $k = 4$. The numbers labelling the sea ice clusters are ordered from low–high latitudes.

the MIZ to only occasionally allow energetic, long-period waves into the interior pack, hence, three clusters will be used for the rest of the study.

We also applied statistical tests to determine the optimal clusters (standard silhouette and Calinski-Harabasz test Rousseeuw, 1987; Calinski, 1968) for $k = 2$–5, but there was no clear outcome.

## 2 Results

Results are produced from $k$-means clustering on 23 million data vectors of scaled CICE6-WIM output variables, with $k = 3$ specified clusters across the ice-covered Southern Ocean domain (south of 50°S with $a_i > 15\%$) from 1st January 2010 to 31st December 2019. The labels of the sea ice regions are kept consistent throughout the section and ordered by low–high latitudes, with Cluster 1 denoting the MIZ.

### 2.1 Cluster Analysis

Figure 2a shows the unscaled means of each of the sea ice clusters (i.e., back transforming the $k$-means centroids to original units). Floe size, ice and snow thickness are the most clearly separated between the clusters and tend to increase with cluster number. Ice age in Cluster 1 and Cluster 2 are similar, indicating that these floes are also mostly seasonal (Wadhams, 1986). Sea ice concentrations within Cluster 2 and Cluster 3 are both very high ($> 95\%$), with Cluster 1 covering relatively lower concentrations ($80\%$), although $63\%$ of the datapoints in Cluster 1 exceed $80\%$. The average ice floe within Cluster 1 has a radius of 27 m, is 0.37 m thick, and 3.6 months old, with little snow cover ($< 0.04$ m). Cluster 3 generally consists of the largest ($> 600$ m), thickest ($> 1$ m), and oldest (5 months) floes with the most snow cover ($> 0.2$ m). Cluster 2 is an intermediary ice type between the types found in Cluster 1 and Cluster 3, and includes floes of medium size and thickness that found in the highest concentrations (on average) and have a similar age to those in Cluster 1.

Figures 2(b–d) show the distributions (histograms with kernel density estimates overlaid) of floe size (semi-log scale), SIC (semi-log scale), and ice thickness (linear scale), respectively. Cluster 1 is characterised by floes smaller than 50 m, with only of $5.6\%$ of floes exceeding 100 m. Cluster 2 is comprised of mostly medium to large floes (100–500 m), and Cluster 3 is dominated by large floes ($> 550$ m). The SIC distributions of all three clusters have a mode greater than

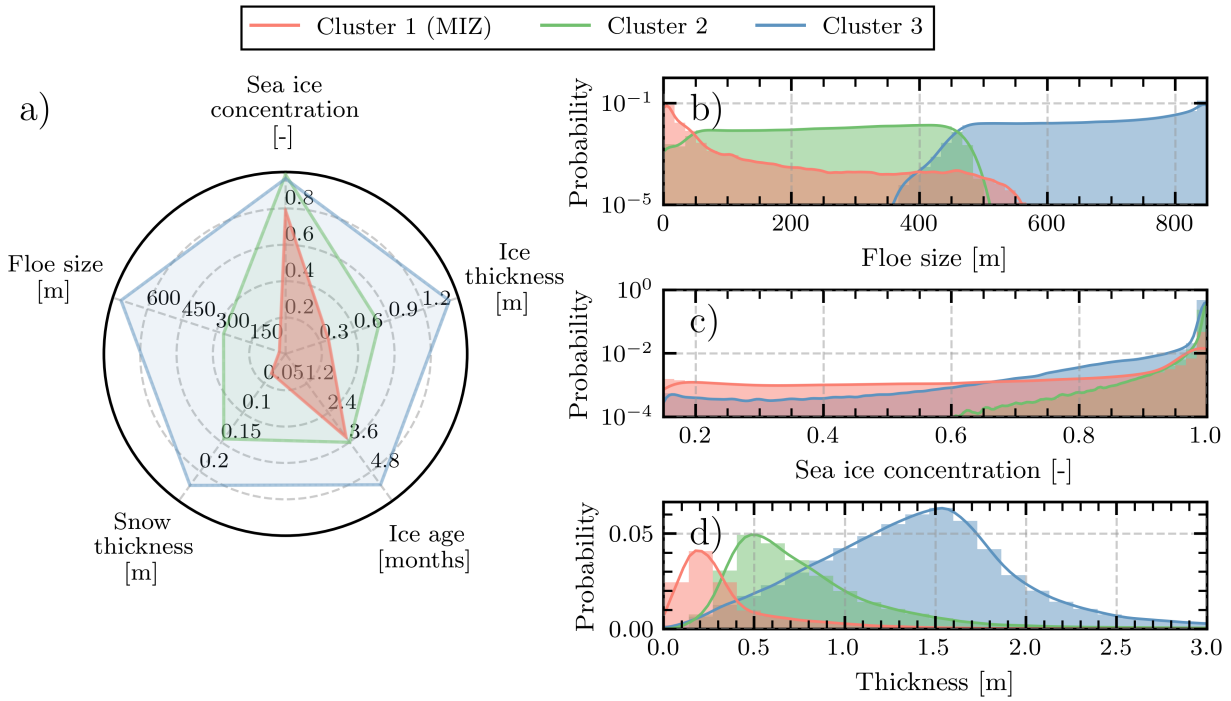


Figure 2: Antarctic sea ice (unscaled) cluster centroid locations and distributions of three key sea ice variables. a) Radar plot of the centroid locations (i.e., the mean sea ice properties) of each cluster. b–d) Probability density functions grouped by cluster number, of b) floe size (semi-log scale), c) sea ice concentration (semi-log scale), and d) ice thickness (linear scale). The distributions of each variable are normalised, such that the the histograms across the three clusters sum to unity.

95%, resulting in large amounts of overlap, although the areas of lower concentrations tend to be classified as Cluster 1. Cluster 1 (MIZ) contains the thinnest ice, of which 78.5% is less than 0.5 m thick. The variance of ice thickness in Clusters 2 and 3 are greater than in Cluster 1; both are skewed towards thicker ice, with Cluster 3 containing the majority of ice thicker than 1 m.

### 2.2 Cluster Identification Dependence on Floe Size

Following the results of Figure 2, a test is completed to determine the influence of floe size on $k$-means clustering. The test includes a comparison of (i) the original dataset (as a reference), (ii) a dataset of only floe size, and (iii) a dataset of the variables excluding floe size. Fig. 3a shows a snapshot (in 2019) of the clusters fit to the original dataset, and is shown at the average SIE maximum (equivalent to Fig. 1b). The equivalent maps of the new clusters generated from only floe size, and the original dataset excluding floe size, are shown in Fig. 3(b–c), respectively (see supporting information Fig. S1 for maps at the SIE minimum).

All three maps display an outer cluster (Cluster 1), an innermost cluster (Cluster 3), and an intermediary cluster (Cluster 2). The maximum width of Cluster 1 when all variables are used for clustering is 253 km and occurs in June. When only floe size data is used (Fig. 3b), Cluster 1 widens to a maximum of 520 km (in June) and restricts the two remaining clusters towards the coastline (see supporting information Fig. S2 for the separation of floe size amongst the clusters). Finally, when floe size is excluded from the clustering (Fig. 3c) the maximum width of Cluster 1 during winter is reduced to 17 km (in September) and is restricted to cells along the ice edge. Cluster 2 then absorbs large amounts of area from both the previous Cluster 1 and Cluster 3. Thus, floe size is the key variable for $k$-means to determine a Cluster 1 which resembles a MIZ (i.e., the wave affected region).

### 2.3 Evolution of the Antarctic Sea Ice Clusters

Figure 4(a–d) show snapshot maps of the clusters in four months 2019 (March, June, September, and December, respectively). Cluster 1 is located along the ocean-sea ice boundary and consistently surrounds the other two clusters, typically with Cluster 2 along its interior border. In March (Fig. 4a), the majority of sea ice resides in the Ross and

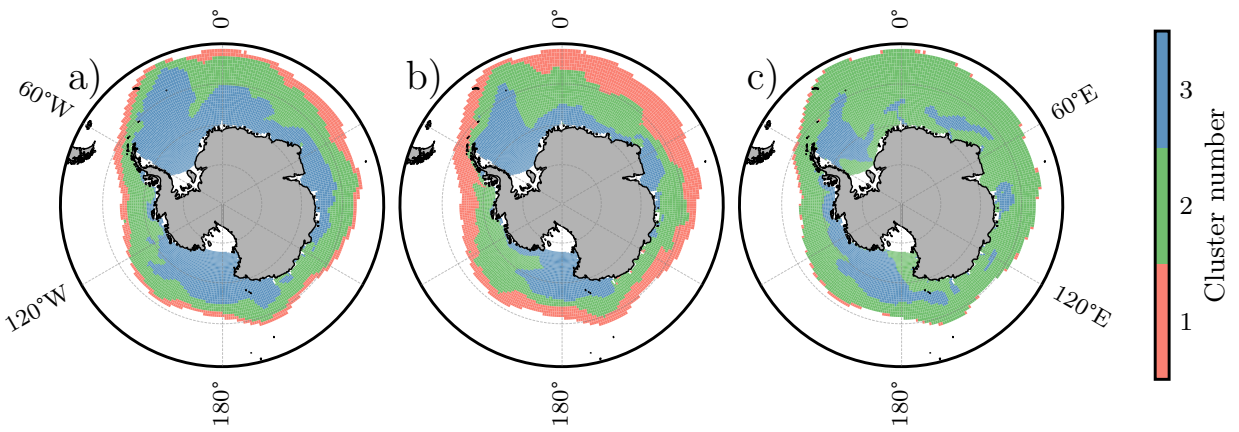


Figure 3: The influence of floe size data on the $k$-means clustering of sea ice. Snapshot maps taken on the 26th August 2019 of the sea ice cluster from $k$-means using a) the original dataset, b) only floe size data, c) data from all variables excluding floe size.

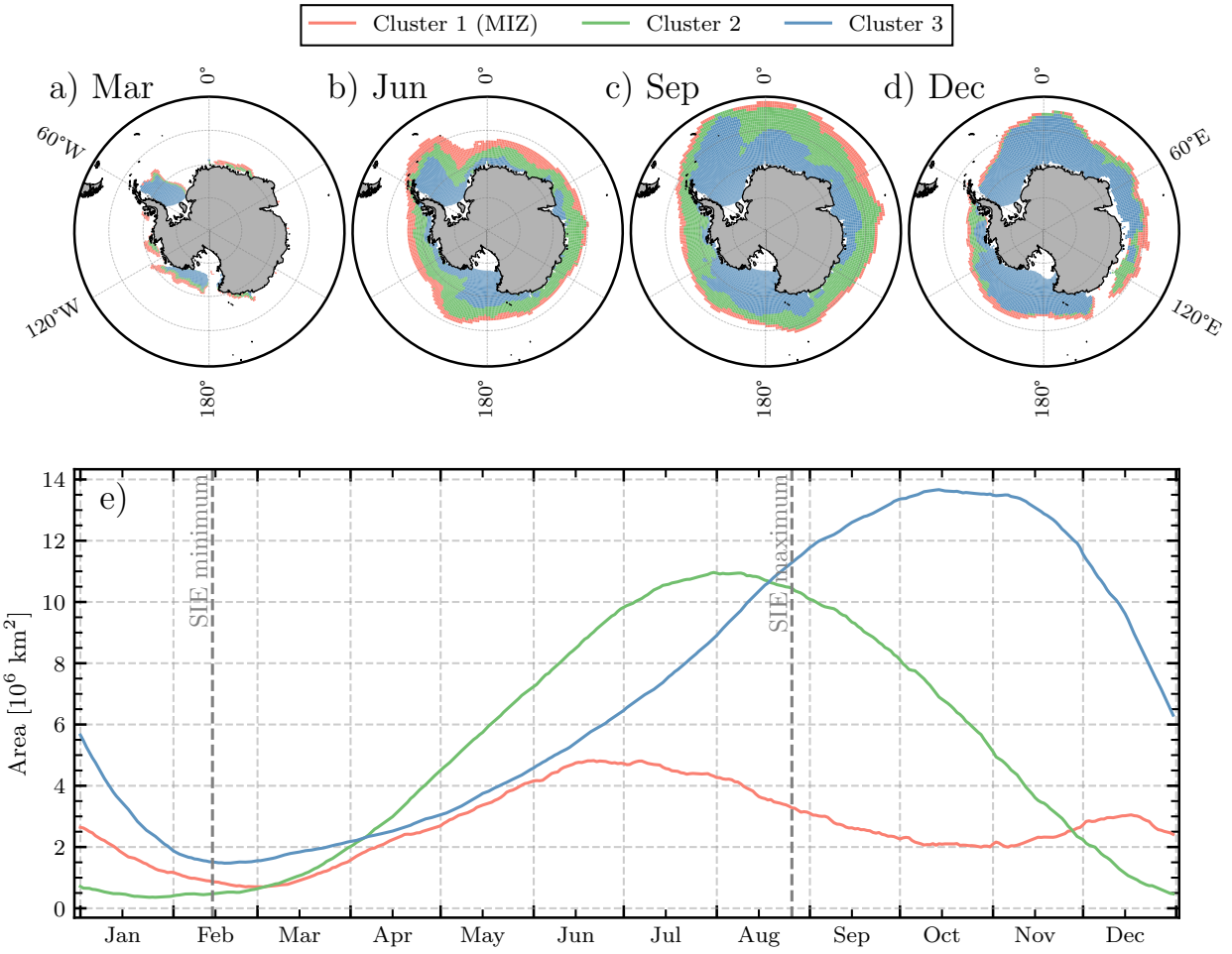


Figure 4: Seasonal spatial and areal evolution of the sea ice clusters. a–d) Daily snapshots of the sea ice clusters taken on the 1st of a) March, b) June, c) September, and d) December in 2019. e) Daily time series of the area within each sea ice cluster (average over 2010–2019) with the average date of minimum (14th of February) and maximum (26th of August) sea ice extent (SIE) shown in dashed-grey lines.

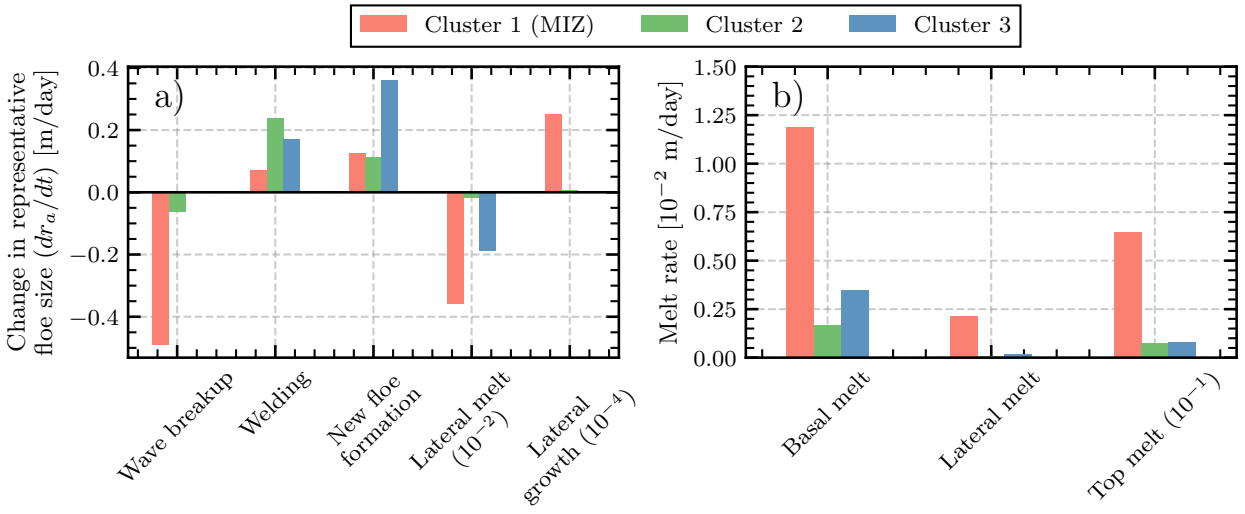


Figure 5: Mean floe size and melt rate tendencies over 2010–2019, grouped by sea ice cluster. a) Change in representative floe size by each individual processes. b) Basal, lateral, and top melt rate of sea ice.

Weddell seas, consisting mostly of large, thick floes (Cluster 3) with thinner, smaller floes (Cluster 1) skirting the ice edge. During the peak of Antarctic sea ice expansion (April–June; Fig. 4b), pockets of Cluster 2 appear along the Antarctic coastline, signifying the formation of new ice (as large, thin floes) in coastal polynyas, where waves from CICE6-WIM do not reach. The new ice increases the area of Cluster 2, resulting in it covering the largest fraction of sea ice from May–July. Cluster 3 expands in mid-winter (September; Fig. 4c), driven by its increased coverage of the Antarctic coastline, due to the new ice from early winter in Cluster 2 transitioning to Cluster 3 as it consolidates and thickens. This culminates in Cluster 3 dominating the ice cover over October–December (Fig. 4d).

Figure 4e shows the seasonal areal evolution of each of the clusters (averaged daily over 2010–2019). The area of Cluster 1 reaches two peaks annually, with a maximum in June–July of $\sim 5$ million km$^2$ and the second peak of $\sim 3$ million km$^2$ in December. Cluster 2 has an annual maximum area of $\sim 11$ million km$^2$ (in July) and Cluster 3 has a maximum area of $\sim 14$ million km$^2$ (in October). As the growth season commences, all three clusters increase in area, with Cluster 2 (i.e., younger, medium–large floes) increasing the most rapidly and becoming dominant in area over winter. The area of Cluster 1 and Cluster 3 grow at similar rates during early winter, although Cluster 1 reaches a maximum area in June–July before decreasing over August–September (coinciding with the decrease of Cluster 2). Cluster 3 continues to grow until October (past the date of maximum SIE), before sharply declining in late spring. The rapid loss of area in Cluster 3 occurs during the second growth phase of Cluster 1 in November. Since new ice formation is limited over summer, this suggests that interior sea ice becomes unconsolidated and re-classified as Cluster 1, leading to it displaying the least intraseasonal variability. During the summer a sharper transition of floe sizes and thickness is formed between the exterior to the interior pack, as Cluster 1 now interfaces with Cluster 3 (since the area of Cluster 2 is diminished).

### 2.4 Floe Size Processes and Melt Rates Across the Antarctic Sea Ice Clusters

Figure 5a shows the average tendencies (over 2010–2019) of the physical processes that affect floe size across the sea ice clusters. The impact on floe size is measured as the change in representative floe size ($r_a$) per day, where bar heights indicate the mean tendency and are coloured corresponding to cluster number. Wave-induced breakup of floes mainly occurs within Cluster 1 (the MIZ), although Cluster 2 experiences some breakup. It is the dominant process for floe size evolution in Cluster 1, with its mean tendency reducing floe size by $0.49\,\mathrm{m\,day^{-1}}$. Welding and new floe formation increase floe size by $0.07$ and $0.12\,\mathrm{m\,day^{-1}}$, respectively. This means that new floes within the MIZ are of comparable sizes to representative floe size, it does not indicate the volume of new ice produced (see supporting information Fig. 4a for the new floe formation seasonal time series). The tendencies of lateral melt and growth change the floe size by $-3.59 \times 10^{-3}$ and $2.5 \times 10^{-5}$ ($\mathrm{m\,day^{-1}}$). Welding is the dominant process in Cluster 2, increasing the floe size by $0.24\,\mathrm{m\,day^{-1}}$. Cluster 3 experiences a similar rate of lateral melt to Cluster 1, as it is the dominant ice cluster over the summer (Fig. 4), although, overall new floe formation is the driver of floe size change within Cluster 3.

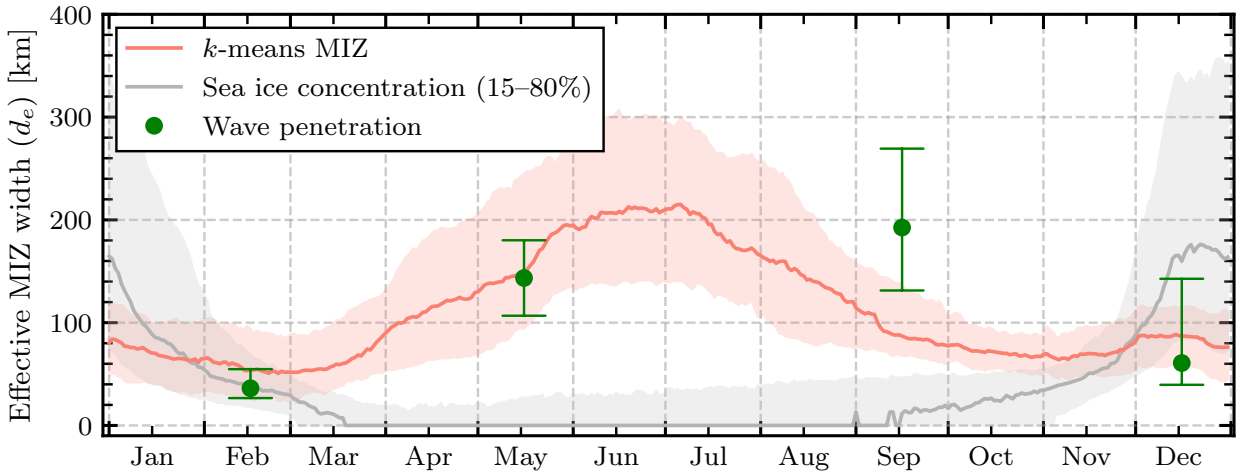


Figure 6: Assessment of MIZ widths from CICE6-WIM outputs against satellite observations. Seasonal time series (averaged over 2010–2019) of effective MIZ widths of Cluster 1 from $k$-means (red) and from the 15–80% sea ice concentration range definition (grey), shown as daily median values (curves) and interquartile ranges (corresponding shaded regions). Monthly medians of satellite-derived wave penetration distances from February, May, September, and December 2019 (Brouwer et al., 2022) are overlaid (green bullets) with their respective interquartile ranges (error bars).

Figure 5b shows the mean rate of the three modes of sea ice melt (basal, lateral and top) across the sea ice clusters. All three clusters are driven by basal melt, with it being approximately a factor of 10 greater than top melt. Cluster 1 exhibits the greatest melting in all three types, since it is the northernmost cluster and borders the open ocean. Cluster 3 displays the second most melting with Cluster 2 exhibiting the least. As the area of Cluster 2 reduces to less than $\sim 1$ million km$^2$ over the summer months (Fig. 4e), there is only a short period of time for it to experience the increased temperatures, thus limiting the opportunity for melting.

### 2.5 Assessment of Antarctic MIZ Widths Against Satellite Observations

To compare with the results of Brouwer et al. (2022), in which effective MIZ widths, $d_e$, are derived from satellite observations of wave penetration distances into the sea ice cover, we define the equivalent for our definition by (Wadhams, 1975),

$$d_e = \int_0^d a_i(x)dx, \tag{13}$$

where $d$ is the distance from the ice edge to the interior MIZ boundary (i.e., the southern boundary of Cluster 1). The distance $d_e$ is the the total length of ice floes in the MIZ along a meridional transect. We also compare with the effective MIZ width given by the traditional 15–80% SIC definition, i.e., the total length of floes in between the 15%- and 80%-SIC contours.

Figure 6 shows the seasonal daily-median effective MIZ widths from $k$-means (red curve) and the SIC-threshold definition (grey curve) over 2010–2019, plus their respective interquartile ranges (corresponding shaded areas). The effective MIZ width from wave observations by Brouwer et al. (2022) over 2019 are presented as monthly medians and interquartile ranges, for February, May, September and December. The $k$-means derived effective MIZ width has a similar seasonal shape to the area of Cluster 1 (Fig. 4e), with a maximum in July ($d_e = 215$ km) and a secondary peak in December ($d_e = 89$ km), although it is less pronounced as the effective MIZ width incorporates an increasingly reduced sea ice concentration from spring to summer. The minimum effective MIZ width occurs during the SIE minimum (February; $d_e = 51$ km). A secondary minimum occurs in spring (November 6th) where the MIZ width reduces to $d_e = 64$ km. In contrast, the SIC-derived effective MIZ width shrinks to a minimum $d_e \approx 0$ km during April–September, when the $k$-means effective MIZ width is greatest, as the SICs at the outer margins of the ice cover often exceeds 80% during winter. The maximum occurs in December ($d_e = 176$ km), which is twice as wide than that of the $k$-means defined MIZ in summer. When the clustering is completed on the dataset excluding floe size, we see strong agreement with the SIC definition (see supporting information Fig. S3). The effective MIZ width seasonality implied by Brouwer et al. (2022)'s observations is similar to that of the $k$-means MIZ, with a minimum in February (36 km), and a maximum in September (192 km) . The $k$-means effective MIZ width strongly agrees with (Brouwer et

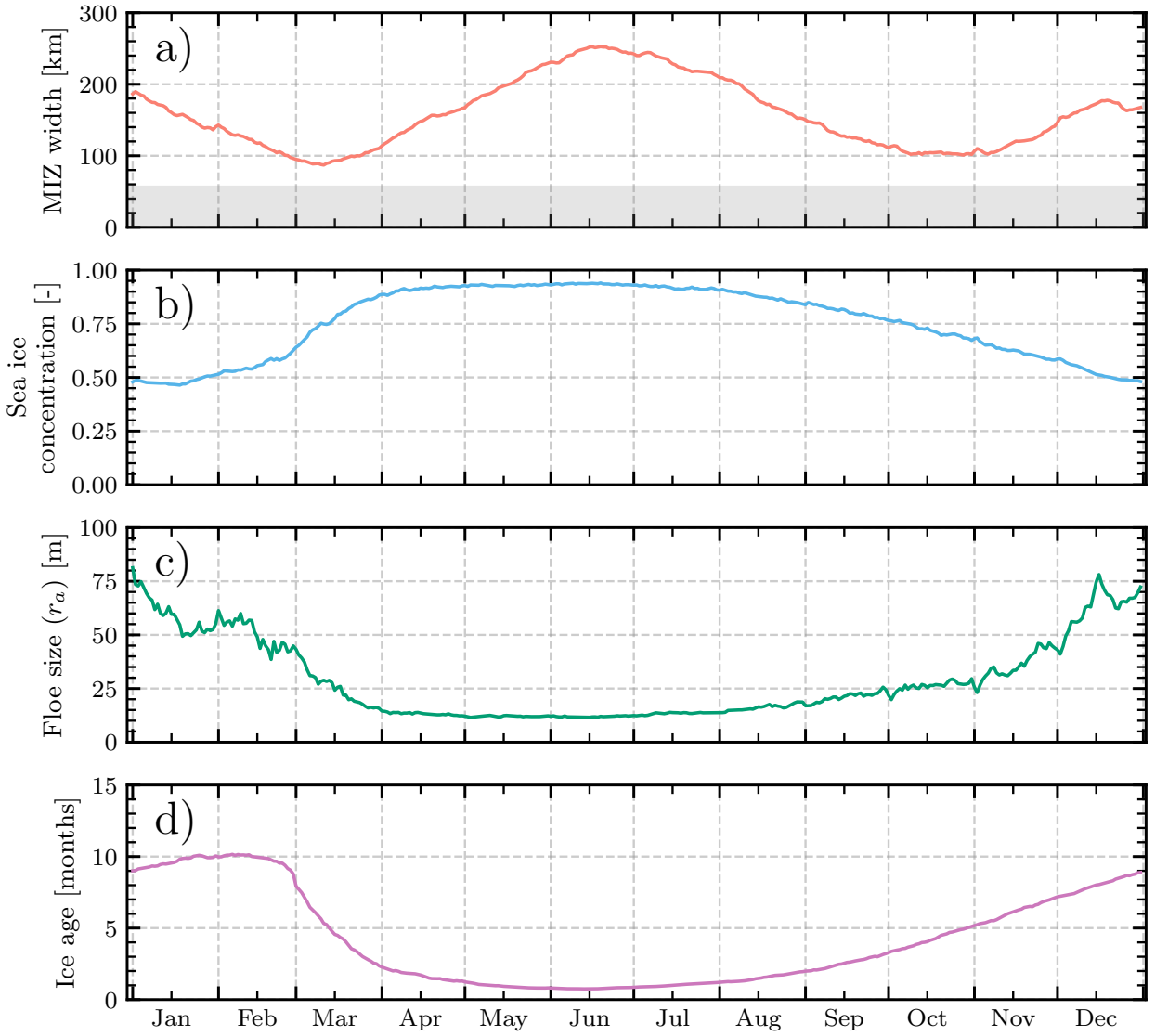


Figure 7: Seasonal time series of Antarctic MIZ (Cluster 1) properties averaged over 2010–2019. Seasonal a) MIZ width (width of Cluster 1 regardless of ice concentration), b) sea ice concentration, c) representative floe size, and d) ice age. The average (north–south) width of CICE6-WIM grid cell ($\sim$ 56 km) is indicated in panel a) with light grey shading.

al., 2022)'s observations in May, and shows good agreement in February and December. Differences are apparent in September, but both have a seasonal maximum of $\sim$ 200 km.

### 2.6 Antarctic MIZ Seasonality

Figure 7a shows the seasonal time series (averaged over 2010–2019) of the absolute MIZ width, $d$ (not to be confused with the effective distance, $d_e$). For reference, the grey shaded area represents the mean north–south width of a CICE6–WIM grid-cell in the Antarctic domain. The panels below show seasonal time series over the same period for the b) average SIC, c) representative floe size, and d) ice age in the MIZ. As expected, the seasonal shape of the mean MIZ width is similar to its mean area (Fig. 4e) and median effective MIZ width (Fig. 6). It has a maximum of $d = 253$ km in June, and a second peak of $d = 178$ km in December. The summer and spring minima are $d = 87$ km in March and $d = 101$ km in late-October, respectively.

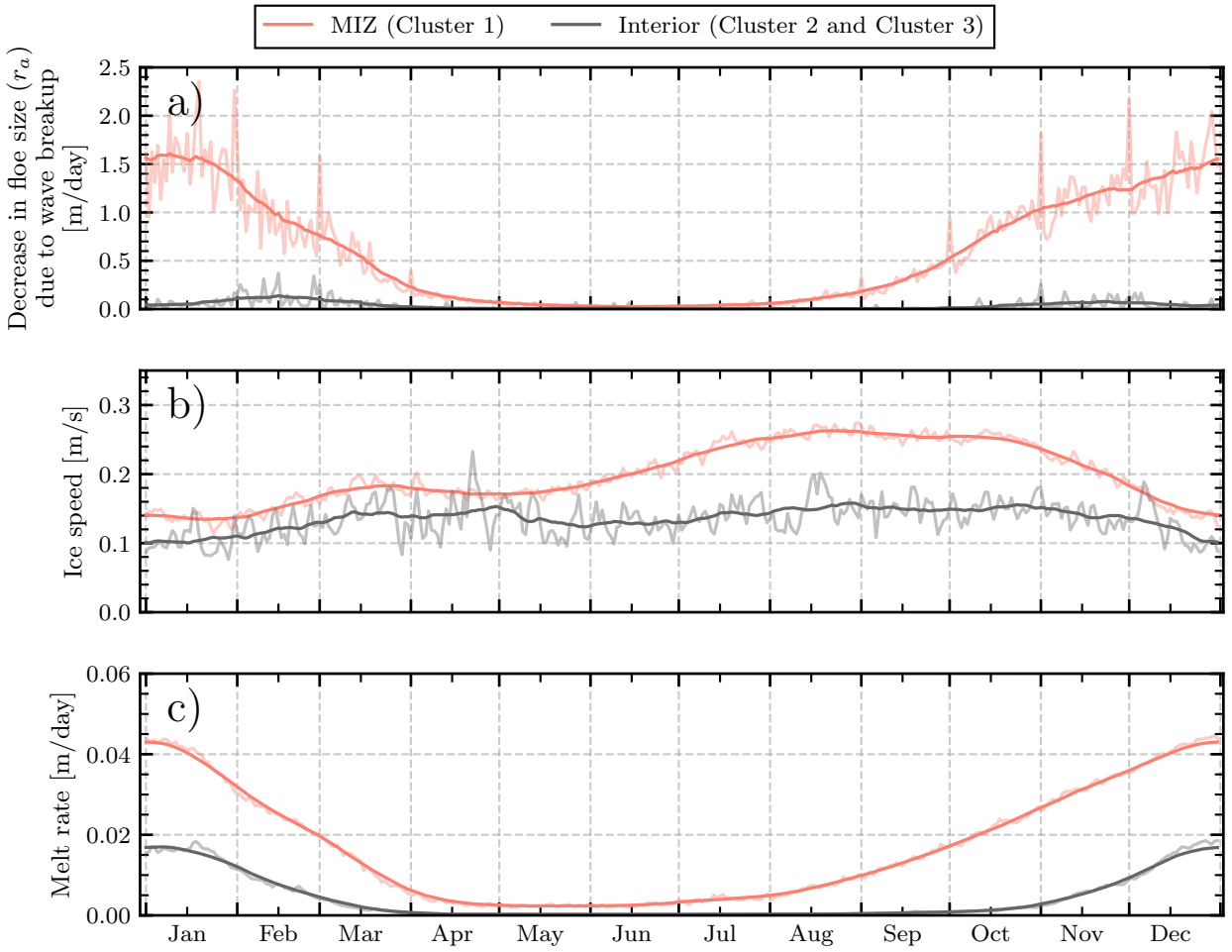


Figure 8: Seasonality of Antarctic sea ice dynamics averaged over 2010–2019. Daily averaged seasonal time series are shown for the MIZ (light red) and interior sea ice (i.e., Clusters 2 and 3 combined in light grey). A 30-day moving average is overlaid for both the MIZ (red) and the interior pack (dark grey). a) Decrease in floe size from wave-induced breakup, b) ice speed, and c) total melt rate.

The mean ice age follows a similar seasonal cycle to floe size in the MIZ, with a peak in summer and a low in winter (Fig. 7c–d). In early winter, the decrease in ice age and floe size simultaneously with the rapid increase in SIC suggests the formation of high-concentration pancake ice fields. Floe sizes and ice age reduce to $\sim 20$ m and less than a month old which coincides with the annual MIZ width maximum. The spring contraction of the MIZ coincides with a decrease in SIC, and an increase in both floe size and ice age. This transition shifts the floe composition in the MIZ from newly-formed pancakes in high concentrations to lower concentrations of older-larger floes.

Figure 8 shows seasonal (daily-mean) time series of a) the decrease in representative floe size by wave-induced breakup, b) sea ice drift speed, and c) melt rate. The average values within the MIZ (light red) are shown alongside those of the interior sea ice (i.e., Clusters 2 and 3 combined; light grey). A 30-day moving average for both clusters is overlaid (red and dark grey, respectively). The MIZ experiences wave-induced breakup primarily over the summer, and consistently at greater rates than the interior clusters (as expected from Fig. 5). Floes in the winter MIZ experience less breakup as their mean representative floe size is limited to $\sim 20$ m (i.e., within the first few categories of the FSD). The moving average of drift speed in the interior remains mostly between 0.1–0.15 $\mathrm{ms}^{-1}$. The MIZ has comparable speeds to the interior during summer, however it becomes more mobile over July–November with drift speeds are $\sim 70\%$ greater than that of the interior ice, as the MIZ width is contracting (Fig. 7a). The rate of melting in the MIZ is greater than the interior pack over the whole annual cycle. As expected, the melt cycle in both the MIZ and interior display strong seasonality with maxima during summer and minima during winter. The MIZ maximum melt rate of $4.35 \times 10^{-2}$ m day$^{-1}$ is almost triple that of the interior pack ($1.49 \times 10^{-2}$ m day$^{-1}$). The second MIZ width peak in

December coincides with increased rates of breakup and melting, resulting in the MIZ containing larger broken floes in lower concentrations than in winter.

## 3 Conclusions and Discussion

Unsupervised classification was used to reduce multivariate snapshots of the Antarctic sea ice cover properties into statistically defined clusters. The sea ice dataset comprised of SIC, ice thickness, ice age, representative floe size and snow thickness from standalone CICE6-WIM model outputs over 2010–2019. A wave heuristic was used to show that $k = 3$ clusters produced a MIZ-like outer cluster, as the wave-affected region that typically shields the inner ice pack (the two interior clusters) but allows occasional waves to reach the outskirts of the inner pack, e.g., large-amplitude waves created by polar storms and cyclones (Kohout et al., 2014; Vichi et al., 2019). Floe sizes were found to have the greatest influence on the classification method, but did not completely determine the clustering results, as shown by the differences in Fig. 3(a–b), and in Fig. S1(a–b). The cluster corresponding with the MIZ was shown to be typically composed of small, thin floes in relatively low concentrations with little snow cover (Fig. 2).

Daily outputs of the $k$-means clustering were used to determine the seasonal behaviour of Antarctic MIZ area (Fig. 4). The area of the MIZ in the standalone model reaches has an annual minimum in mid-summer (1 million $\mathrm{km}^2$), a maximum in winter (5 million $\mathrm{km}^2$), before it contracts in spring (2 million $\mathrm{km}^2$), and finally reaches a second peak in early-summer (3 million $\mathrm{km}^2$). The winter peak coincides with the expansion of both interior clusters, whereas the summer maximum occurs whilst the area of the interior clusters are decreasing. The average effective MIZ widths displayed similar behaviour (although with a reduced summer peak), and were shown to agree well with Brouwer et al. (2022)'s observations based on wave propagation distances, particularly in February, May, and December (Fig. 6). Both methods produced a minimum width in summer (February), and a maximum width of approximately 200 km in winter. The largest discrepancy occurred in September, which may be a result of the standalone model not capturing the timing of the sea ice maximum extent (e.g., the average sea ice extent maximum in CICE6-WIM occurred on the 26th of August, which is over a month prior to the satellite observed maximum on 30th of September in 2019 NSIDC, 2023). As sea ice thermodynamics are sensitive to mixed-layer depth (Tsamados et al., 2015), the areal sea ice evolution would likely be improved with a variable depth mixed-layer model, whether coupled to a full-ocean model or a prognostic mixed layer model (Petty et al., 2014).

For the CICE6-WIM model outputs, $k$-means clustering was shown to produce markedly different MIZ widths to those obtained from the classic 15–80% SIC definition (Fig. 6). In contrast to $k$-means, the SIC-defined MIZ widths are at their minimum over the winter months ($\sim 0$ km). The SIC definition does not identify the MIZ-type driven by the pancake cycle (Lange et al., 1989), which contains large waves generated by winter cyclones deep into concentrations close to 100% (Alberello et al., 2022), albeit the coarse spatial resolution of the model grid (1° latitude-longitude) limits the representation of the steep gradient in SIC between the open ocean and the ice pack (e.g., as shown in ASMR2 satellite images Alberello et al., 2022; Pitt et al., 2022). Thus, the CICE6-WIM outputs underestimate the SIC-defined MIZ (especially when it is less than the average grid-cell length of 56 km). The SIC-defined MIZ sharply increases in width and spread (interquartile range) during summer, which coincides with the $k$-means MIZ increasing in floe size (in mean; Fig. 7c) and becoming disperse (mean SIC $\sim 50\%$; Fig. 7b).

Vichi (2022) found a seasonal MIZ area using the SIC definition applied to satellite observations that contrasts with our SIC-defined MIZ widths. In particular, he found that the MIZ area grows at a linear rate from 2 million $\mathrm{km}^2$ in February to October, before peaking to 8 million $\mathrm{km}^2$ in December, meaning that it is non-zero over winter (broadly agreeing with the results of Stroeve et al., 2016). There is significant variability when comparing MIZ areas between satellite products or against model data (Stroeve et al., 2016; Rolph et al., 2020). Although, the lack of sea ice–ocean feedbacks in our model potentially exacerbate the amount of frazil ice produced from the thermodynamic scheme (mushy-layer, see A. K. Turner et al., 2013) within CICE (Bailey et al., 2020), resulting in increased SICs across the cover, decreasing the SIC-threshold MIZ widths during winter. Further, Vichi (2022) showed that the MIZ area defined by a temporal standard deviation of SIC initially grew at a greater rate than the SIC threshold area (over February–May), and then remained between 5–6 million $\mathrm{km}^2$ over winter (similar to the $k$-means winter area; Fig. 4). Both of the SIC definitions Vichi (2022) applied to satellite data identified areas deeper into the pack, such as coastal and open-ocean polynyas. These features are not identified as MIZ by the $k$-means definition (Fig. 3 and Fig. S1a), which instead returned the outer band of ice cover affected by open-ocean waves (Fig. 1 and Fig. 8), since no wave generation occurs within the sea ice cover or polynyas in CICE6-WIM.

The bimodal seasonality of the MIZ area (maxima during the middle of the austral winter and summer; e.g., Fig. 4e) indicates a seasonal transition in the floe composition and regime shift in the dominant physical processes. During the Antarctic sea ice expansion, new floes within the MIZ are limited in size due to incoming waves from the open ocean, resulting in the proliferation of floes in the smallest floe size and ice thickness categories, which are interpreted as

pancake ice formation with interstitial frazil ice (Wadhams et al., 2018; Alberello et al., 2019). This results in a winter MIZ in which the average ice floe is less than 1 month old and 20 m in length, as well as the SIC being close to 100% (Fig. 7b–d). The winter MIZ retreats in the mid–late-winter to early–mid-spring as the ice edge stops advancing before the Antarctic ice cover rapidly retreated. However, the MIZ had a second growth period over late-spring–early-summer when large floes in the inner ice pack are fractured by waves and are re-classified into the MIZ. Thus, the summer MIZ was characterised as previously consolidated ice, rather than new floes (as was the case in winter). Sea ice concentrations within the MIZ continued to reduce into January, with the age of the floes reaching a maximum in February. The melt season decreased the summer MIZ width from mid-December–mid-March (occurring near the SIE minimum) where it reduced to 100 km wide. After the SIE minimum, new ice formation began, reducing the floe size and ice age in the MIZ, while increasing SIC, completing the seasonal cycle.

The Antarctic MIZ has been observed to contain unconsolidated sea ice cover consisting of small floes (Toyota et al., 2011, 2016; Alberello et al., 2019), due to frazil and pancake ice formation (Wadhams et al., 2018; Skatulla et al., 2022), and wave breakup of larger floes (Kohout et al., 2016). These small floes are more mobile, e.g., faster drift speeds (Alberello et al., 2020; Womack et al., 2022), and more susceptible to melting (especially lateral melt Steele, 1992; Horvat et al., 2016; Bateson et al., 2020). The $k$-means cluster analysis identified a MIZ with these properties in a parsimonious manner, i.e., the clustering used no variables on sea ice dynamics or thermodynamics. Wave-induced breakup of floes was shown to predominantly occur within the MIZ during summer (including re-classifying ice which previously resided in an interior cluster to the MIZ). Wave-induced breakup in the summer MIZ was approximately a factor of eight greater than that of the interior pack ($\sim 1.6\,\mathrm{m\,day^{-1}}$ and $\sim 0.2\,\mathrm{m\,day^{-1}}$ in the MIZ and interior, respectively; Fig. 8a). The late-winter to spring months coincided with the average drift speed within the MIZ to be around $0.1\,\mathrm{ms^{-1}}$ greater than that of the interior, whilst floes were in near free-drift conditions (Fig. 8b and Fig. S4b). Similar to the wave-breakup of floes, melt rates were found to be dominant over the summer season, with the total melt rate having (on average) twice the impact on MIZ ice than interior ice (annual maxima of $4.5 \times 10^{-2}\,\mathrm{m\,day^{-1}}$ in the MIZ and $2.1 \times 10^{-2}\,\mathrm{m\,day^{-1}}$ in the interior; Fig. 8c).

In conclusion, $k$-means classification with three clusters has been shown as a suitable method to identify the Antarctic MIZ, in a manner consistent with its description as the regularly wave-affected region of the ice cover (Wadhams, 1986), using sea ice model outputs, with the requirement that the outputs include floe sizes. Future studies should test the approach on model outputs for Arctic sea ice, in which pancake ice is less prevalent (although suggested to increase with a changing wave climate Thomson et al., 2017) and the MIZ is generally smaller (Weeks, 2010) as waves rarely penetrate more than 100 km into the MIZ (Cooper et al., 2022). Moreover, it should be tested in a coupled model setting as standalone sea ice models struggle to resolve feedbacks between sea ice and the ocean (including surface waves), and with a more sophisticated wave propagation model, e.g., where the attenuation rate changes in response to floe size (Meylan et al., 2021). It will then be available for detailed studies on the unique properties of the MIZ and its influence on the responses of Arctic and Antarctic sea ice covers to the changing climate.

## Open Research Section

Model output data from CICE6-WIM used to produce this manuscript are available at `https://zenodo.org/records/10162321` (Day et al., 2023). Statistical analysis was completed using Scikit-learn (version 1.1.3, available under the BSD-3-Clause license Pedregosa et al., 2011). The data wrangling and cleaning performed used Xarray (version 2022.11.0, available under the Apache-2.0 license Hoyer & Hamman, 2017) and Pandas (version 1.5.0, available under the BSD-3-Clause license Pandas development team, 2020). Figures were made using Matplotlib (version 3.4.3, license available at `https://matplotlib.org/stable/users/project/license.html` Hunter, 2007) and maps were created using Cartopy (version 0.20.2, available under the GNU GPLv3 license Met Office, 2010–2015).

## Acknowledgments

NSD and LGB are supported by Australian Research Council grants FT190100404, DP200102828 and LP200100406. NSD is supported by the Consortium for Ocean-Sea Ice Modelling in Australia. NSD was supported by the Quarterly Journal of Mechanics and Applied Mathematics, DR Stranks Travelling Fellowship, and the Adelaide University Graduate Centre to complete an extended visit with AA at the University of East Anglia. The authors thank the National Computational Infrastructure (NCI) for use of their high performance computer (Gadi) to conduct model simulations. The authors also thank the Isaac Newton Institute for Mathematical Sciences for hosting meetings which conceptualised this study.

## References


Alberello, A., Bennetts, L. G., Heil, P., Eayrs, C., Vichi, M., MacHutchon, K., … Toffoli, A. (2020). Drift of pancake ice floes in the winter Antarctic marginal ice zone during polar cyclones. *Journal of Geophysical Research: Oceans*, *125*(3), e2019JC015418.

Alberello, A., Bennetts, L. G., Onorato, M., Vichi, M., MacHutchon, K., Eayrs, C., … Toffoli, A. (2022). Three-dimensional imaging of waves and floes in the marginal ice zone during a cyclone. *Nature Communications*, *13*(1), 4590.

Alberello, A., Onorato, M., Bennetts, L. G., Vichi, M., Eayrs, C., MacHutchon, K., & Toffoli, A. (2019). Pancake ice floe size distribution during the winter expansion of the Antarctic marginal ice zone. *The Cryosphere*, *13*(1), 41–48.

Bailey, D. A., Holland, M. M., DuVivier, A. K., Hunke, E. C., & Turner, A. K. (2020). Impact of a new sea ice thermodynamic formulation in the CESM2 sea ice component. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002154.

Bateson, A. W., Feltham, D. L., Schröder, D., Hosekova, L., Ridley, J. K., & Aksenov, Y. (2020). Impact of sea ice floe size distribution on seasonal fragmentation and melt of Arctic sea ice. *The Cryosphere*, *14*(2), 403–428.

Bateson, A. W., Feltham, D. L., Schröder, D., Wang, Y., Hwang, B., Ridley, J. K., & Aksenov, Y. (2022). Sea ice floe size: its impact on pan-Arctic and local ice mass and required model complexity. *The Cryosphere*, *16*(6), 2565–2593.

Bennetts, L. G., Bitz, C. M., Feltham, D. L., Kohout, A. L., & Meylan, M. H. (2022a). Marginal ice zone dynamics: future research perspectives and pathways. *Philosophical Transactions of the Royal Society A*, *380*(2235), 20210267.

Bennetts, L. G., Bitz, C. M., Feltham, D. L., Kohout, A. L., & Meylan, M. H. (2022b). Theory, modelling and observations of marginal ice zone dynamics: multidisciplinary perspectives and outlooks. *Philosophical Transactions of the Royal Society A*, *380*(2235), 20210265.

Bennetts, L. G., O'Farrell, S., & Uotila, P. (2017). Impacts of ocean-wave-induced breakup of Antarctic sea ice via thermodynamics in a stand-alone version of the CICE sea-ice model. *The Cryosphere*, *11*(3), 1035–1040.

Bennetts, L. G., Shakespeare, C. J., Vreugdenhil, C. A., Foppert, A., Gayen, B., Meyer, A., … Stanley, G. J. (2023). Closing the loops on Southern Ocean dynamics: From the circumpolar current to ice shelves and from bottom mixing to surface waves. *Authorea Preprints*.

Bitz, C. M., Holland, M. M., Weaver, A. J., & Eby, M. (2001). Simulating the ice-thickness distribution in a coupled climate model. *Journal of Geophysical Research: Oceans*, *106*(C2), 2441–2463.

Boutin, G., Ardhuin, F., Dumont, D., Sévigny, C., Girard-Ardhuin, F., & Accensi, M. (2018). Floe size effect on wave-ice interactions: Possible effects, implementation in wave model, and evaluation. *Journal of Geophysical Research: Oceans*, *123*(7), 4779–4805.

Boutin, G., Lique, C., Ardhuin, F., Rousset, C., Talandier, C., Accensi, M., & Girard-Ardhuin, F. (2020). Towards a coupled model to investigate wave–sea ice interactions in the Arctic marginal ice zone. *The Cryosphere*, *14*(2), 709–735.

Boutin, G., Williams, T., Horvat, C., & Brodeau, L. (2022). Modelling the Arctic wave-affected marginal ice zone: a comparison with ICESat-2 observations. *Philosophical Transactions of the Royal Society A*, *380*(2235), 20210262.

Boutin, G., Williams, T., Rampal, P., Olason, E., & Lique, C. (2021). Wave–sea-ice interactions in a brittle rheological framework. *The Cryosphere*, *15*(1), 431–457.

Bretschneider, C. L. (1959). *Wave variability and wave spectra for wind-generated gravity waves* (No. 118). The Board.

Brouwer, J., Fraser, A. D., Murphy, D. J., Wongpan, P., Alberello, A., Kohout, A., … Williams, G. D. (2022). Altimetric observation of wave attenuation through the Antarctic marginal ice zone using ICESat-2. *The Cryosphere*, *16*(6), 2325–2353.

Calinski, T. (1968). A dendrite method for cluster analysis. In *Biometrics* (Vol. 24, p. 207).

Comiso, J. C. (2006). Abrupt decline in the Arctic winter sea ice cover. *Geophysical Research Letters*, *33*(18).

Comiso, J. C., & Zwally, H. J. (1984). Concentration gradients and growth/decay characteristics of the seasonal sea ice cover. *Journal of Geophysical Research: Oceans*, *89*(C5), 8081–8103.

Cooper, V. T., Roach, L. A., Thomson, J., Brenner, S. D., Smith, M. M., Meylan, M. H., & Bitz, C. M. (2022). Wind waves in sea ice of the western Arctic and a global coupled wave-ice model. *Philosophical Transactions of the Royal Society A*, *380*(2235), 20210258.

Day, N. S., Bennetts, L. G., O'Farrell, S. P., Alberello, A., & Montiel, F. (2023). *Model output: CICE6-WIM Antarctic sea ice data (2010–2019) [Dataset].* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.10162320` doi: doi:10.5281/zenodo.10162320

Derkani, M. H., Alberello, A., Nelli, F., Bennetts, L. G., Hessner, K. G., MacHutchon, K., … Toffoli, A. (2021). Wind, waves, and surface currents in the Southern Ocean: observations from the Antarctic circumnavigation expedition. *Earth System Science Data*, *13*(3), 1189–1209.

Dumont, D. (2022). Marginal ice zone dynamics: history, definitions and research perspectives. *Philosophical Transactions of the Royal Society A*, *380*(2235), 20210253.

Dumont, D., Kohout, A. L., & Bertino, L. (2011). A wave-based model for the marginal ice zone including a floe breaking parameterization. *Journal of Geophysical Research: Oceans*, *116*(C4).

Eayrs, C., Li, X., Raphael, M. N., & Holland, D. M. (2021). Rapid decline in Antarctic sea ice in recent years hints at future change. *Nature Geoscience*, *14*(7), 460–464.

Farooq, U., Rack, W., McDonald, A., & Howell, S. (2023). Representation of sea ice regimes in the western Ross Sea, Antarctica, based on satellite imagery and AMPS wind data. *Climate Dynamics*, *60*(1-2), 227–238.

Fraser, A. D., Wongpan, P., Langhorne, P. J., Klekociuk, A. R., Kusahara, K., Lannuzel, D., … Wienecke, B. (2023). Antarctic landfast sea ice: A review of its physics, biogeochemistry and ecology. *Reviews of Geophysics*, *61*(2), e2022RG000770.

Horvat, C., Blanchard-Wrigglesworth, E., & Petty, A. A. (2020). Observing waves in sea ice with ICESat-2. *Geophysical Research Letters*, *47*(10), e2020GL087629.

Horvat, C., & Tziperman, E. (2015). A prognostic model of the sea-ice floe size and thickness distribution. *The Cryosphere*, *9*(6), 2119–2134.

Horvat, C., Tziperman, E., & Campin, J.-M. (2016). Interaction of sea ice floe size, ocean eddies, and sea ice melting. *Geophysical Research Letters*, *43*(15), 8083–8090.

Hoyer, S., & Hamman, J. (2017). xarray: N-D labeled arrays and datasets in Python [Software]. *In revision, J. Open Res. Software*.

Hunke, E. C., & Dukowicz, J. K. (1997). An elastic–viscous–plastic model for sea ice dynamics. *Journal of Physical Oceanography*, *27*(9), 1849–1867.

Hunke, E. C., Lipscomb, W. H., Jones, P., Turner, A. K., Jeffery, N., & Elliott, S. (2017). *CICE, the Los Alamos sea ice model* (Tech. Rep.). Los Alamos, NM (United States): Los Alamos National Lab.(LANL).

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment [Software]. *Computing in Science & Engineering*, *9*(3), 90–95. doi: doi:10.1109/MCSE.2007.55

Kiss, A. E., Hogg, A. M., Hannah, N., Boeira Dias, F., Brassington, G. B., Chamberlain, M. A., … Zhang, X. (2020). ACCESS-OM2 v1.0: a global ocean–sea ice model at three resolutions. *Geoscientific Model Development*, *13*(2), 401–442.

Kohout, A. L., Williams, M. J., Dean, S. M., & Meylan, M. H. (2014). Storm-induced sea-ice breakup and the implications for ice extent. *Nature*, *509*(7502), 604–607.

Kohout, A. L., Williams, M. J., Toyota, T., Lieser, J., & Hutchings, J. K. (2016). In situ observations of wave-induced sea ice breakup. *Deep Sea Research Part II: Topical Studies in Oceanography*, *131*, 22–27.

Lange, M. A., Ackley, S. F., Wadhams, P., Dieckmann, G. S., & Eicken, H. (1989). Development of sea ice in the Weddell Sea. *Annals of Glaciology*, *12*, 92–96.

Longuet-Higgins, M. S. (1963). The effect of non-linearities on statistical distributions in the theory of sea waves. *Journal of fluid mechanics*, *17*(3), 459–480.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 281–297).

Massom, R. A., Comiso, J. C., Worby, A. P., Lytle, V. I., & Stock, L. V. (1999). Regional classes of sea ice cover in the East Antarctic pack observed from satellite and in situ data during a winter time period. *Remote Sensing of Environment*, *68*(1), 61–76.

Massom, R. A., Scambos, T. A., Bennetts, L. G., Reid, P., Squire, V. A., & Stammerjohn, S. E. (2018). Antarctic ice shelf disintegration triggered by sea ice loss and ocean swell. *Nature*, *558*(7710), 383–389.

Massom, R. A., Stammerjohn, S. E., Lefebvre, W., Harangozo, S. A., Adams, N., Scambos, T. A., … Fowler, C. (2008). West Antarctic Peninsula sea ice in 2005: Extreme ice compaction and ice edge retreat due to strong anomaly with respect to climate. *Journal of Geophysical Research: Oceans*, *113*(C2).

McPhee, M. G., Maykut, G. A., & Morison, J. H. (1987). Dynamics and thermodynamics of the ice/upper ocean system in the marginal ice zone of the Greenland Sea. *Journal of Geophysical Research: Oceans*, *92*(C7), 7017–7031.

Met Office. (2010–2015). *Cartopy: a cartographic python library with a Matplotlib interface [Software]*. Exeter, Devon. Retrieved from `https://scitools.org.uk/cartopy`

Meylan, M. H., Bennetts, L. G., & Kohout, A. L. (2014). In situ measurements and analysis of ocean waves in the Antarctic marginal ice zone. *Geophysical Research Letters*, *41*(14), 5046–5051.

Meylan, M. H., Horvat, C., Bitz, C. M., & Bennetts, L. G. (2021). A floe size dependent scattering model in two-and three-dimensions for wave attenuation by ice floes. *Ocean Modelling*, *161*, 101779.

Mokus, N. G., & Montiel, F. (2022). Wave-triggered breakup in the marginal ice zone generates lognormal floe size distributions: a simulation study. *The Cryosphere*, *16*(10), 4447–4472.

Montiel, F., & Squire, V. A. (2017). Modelling wave-induced sea ice break-up in the marginal ice zone. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *473*(2206), 20170258.

Montiel, F., Squire, V. A., & Bennetts, L. G. (2016). Attenuation and directional spreading of ocean wave spectra in the marginal ice zone. *Journal of Fluid Mechanics*, *790*, 492–522.

Moreno-Chamarro, E., Ortega, P., & Massonnet, F. (2020). Impact of the ice thickness distribution discretization on the sea ice concentration variability in the NEMO3.6–LIM3 global ocean–sea ice model. *Geoscientific Model Development*, *13*(10), 4773–4787.

NSIDC. (2023). *Sea ice index daily and monthly image viewer.* Retrieved 2023-11-01, from `https://nsidc.org/data/seaice_index`

Pandas development team. (2020). *pandas-dev/pandas: Pandas [Software].* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.3509134` doi: doi:10.5281/zenodo.3509134

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, É. (2011). Scikit-learn: Machine learning in Python [Software]. *Journal of Machine Learning Research*, *12*, 2825–2830.

Perovich, D. K., & Jones, K. F. (2014). The seasonal evolution of sea ice floe size distribution. *Journal of Geophysical Research: Oceans*, *119*(12), 8767–8777.

Petty, A. A., Holland, P. R., & Feltham, D. L. (2014). Sea ice and the ocean mixed layer over the Antarctic shelf seas. *The Cryosphere*, *8*(2), 761–783.

Pitt, J. P., Bennetts, L. G., Meylan, M. H., Massom, R. A., & Toffoli, A. (2022). Model predictions of wave overwash extent into the marginal ice zone. *Journal of Geophysical Research: Oceans*, *127*(10), e2022JC018707.

Roach, L. A., Bitz, C. M., Horvat, C., & Dean, S. M. (2019). Advances in modeling interactions between sea ice and ocean surface waves. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4167–4181.

Roach, L. A., Horvat, C., Dean, S. M., & Bitz, C. M. (2018). An emergent sea ice floe size distribution in a global coupled ocean-sea ice model. *Journal of Geophysical Research: Oceans*, *123*(6), 4322–4337.

Roach, L. A., Smith, M. M., & Dean, S. M. (2018). Quantifying growth of pancake sea ice floes using images from drifting buoys. *Journal of Geophysical Research: Oceans*, *123*(4), 2851–2866.

Rolph, R. J., Feltham, D. L., & Schröder, D. (2020). Changes of the Arctic marginal ice zone during the satellite era. *The Cryosphere*, *14*(6), 1971–1984.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, *20*, 53–65.

Schröder, D., Feltham, D. L., Tsamados, M., Ridout, A., & Tilling, R. (2019). New insight from CryoSat-2 sea ice thickness for sea ice modelling. *The Cryosphere*, *13*(1), 125–139.

Shen, H. H., Ackley, S. F., & Hopkins, M. A. (2001). A conceptual model for pancake-ice formation in a wave field. *Annals of Glaciology*, *33*, 361–367.

Skatulla, S., Audh, R. R., Cook, A., Hepworth, E., Johnson, S., Lupascu, D. C., ... Vichi, M. (2022). Physical and mechanical properties of winter first-year ice in the Antarctic marginal ice zone along the Good Hope Line. *The Cryosphere*, *16*(7), 2899–2925.

Smith, G. A., Hemer, M., Greenslade, D., Trenham, C., Zieger, S., & Durrant, T. (2021). Global wave hindcast with Australian and Pacific island focus: From past to present. *Geoscience Data Journal*, *8*(1), 24–33.

Soleymani, A., & Scott, A. (2023). Arctic marginal ice zone interannual variability and change point detection using two definitions (1983-2022). *Environmental Research Letters*.

Squire, V. A. (2007). Of ocean waves and sea-ice revisited. *Cold Regions Science and Technology*, *49*(2), 110–133.

Squire, V. A. (2022). A prognosticative synopsis of contemporary marginal ice zone research. *Philosophical Transactions of the Royal Society A*, *380*(2235), 20220094.

Steele, M. (1992). Sea ice melting and floe geometry in a simple ice-ocean model. *Journal of Geophysical Research: Oceans*, *97*(C11), 17729–17738.

Stopa, J. E., Ardhuin, F., & Girard-Ardhuin, F. (2016). Wave climate in the Arctic 1992–2014: Seasonality and trends. *The Cryosphere*, *10*(4), 1605–1629.

Stroeve, J. C., Jenouvrier, S., Campbell, G. G., Barbraud, C., & Delord, K. (2016). Mapping and assessing variability in the Antarctic marginal ice zone, pack ice and coastal polynyas in two sea ice algorithms with implications on breeding success of snow petrels. *The Cryosphere*, *10*(4), 1823–1843.

Strong, C. (2012). Atmospheric influence on Arctic marginal ice zone position and width in the Atlantic sector, February–April 1979–2010. *Climate dynamics*, *39*, 3091–3102.

Strong, C., Foster, D., Cherkaev, E., Eisenman, I., & Golden, K. M. (2017). On the definition of marginal ice zone width. *Journal of Atmospheric and Oceanic Technology*, *34*(7), 1565–1584.

Strong, C., & Rigor, I. G. (2013). Arctic marginal ice zone trending wider in summer and narrower in winter. *Geophysical Research Letters*, *40*(18), 4864–4868.

Teder, N. J., Bennetts, L. G., Reid, P. A., & Massom, R. A. (2022). Sea ice-free corridors for large swell to reach Antarctic ice shelves. *Environmental Research Letters*, *17*(4), 045026.

Thomson, J. (2022). Wave propagation in the marginal ice zone: connections and feedback mechanisms within the air–ice–ocean system. *Philosophical Transactions of the Royal Society A*, *380*(2235), 20210251.

Thomson, J., Ackley, S., Shen, H. H., & Rogers, W. E. (2017). *The balance of ice, waves, and winds in the Arctic autumn, Eos, 98.*

Thorndike, A. S., Rothrock, D. A., Maykut, G. A., & Colony, R. (1975). The thickness distribution of sea ice. *Journal of Geophysical Research*, *80*(33), 4501–4513.

Toyota, T., Haas, C., & Tamura, T. (2011). Size distribution and shape properties of relatively small sea-ice floes in the Antarctic marginal ice zone in late winter. *Deep Sea Research Part II: Topical Studies in Oceanography*, *58*(9-10), 1182–1193.

Toyota, T., Kohout, A., & Fraser, A. D. (2016). Formation processes of sea ice floe size distribution in the interior pack and its relationship to the marginal ice zone off East Antarctica. *Deep Sea Research Part II: Topical Studies in Oceanography*, *131*, 28–40.

Tsamados, M., Feltham, D. L., Petty, A., Schröder, D., & Flocco, D. (2015). Processes controlling surface, bottom and lateral melt of arctic sea ice in a state of the art sea ice model. *Philosophical Transactions of the Royal Society A*, *373*(2052), 20140167.

Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., ... Yamazaki, D. (2018). JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do). *Ocean Modelling*, *130*, 79–139.

Turner, A. K., Hunke, E. C., & Bitz, C. M. (2013). Two modes of sea-ice gravity drainage: A parameterization for large-scale modeling. *Journal of Geophysical Research: Oceans*, *118*(5), 2279–2294.

Turner, J., & Comiso, J. C. (2017). Solve Antarctica's sea-ice puzzle. *Nature*, *547*(7663), 275–277.

Vichi, M. (2022). An indicator of sea ice variability for the Antarctic marginal ice zone. *The Cryosphere*, *16*(10), 4087–4106.

Vichi, M., Eayrs, C., Alberello, A., Bekker, A., Bennetts, L. G., Holland, D. M., ... Toffoli, A. (2019). Effects of an explosive polar cyclone crossing the Antarctic marginal ice zone. *Geophysical Research Letters*, *46*(11), 5948–5958.

Vihma, T., Pirazzini, R., Fer, I., Renfrew, I. A., Sedlar, J., Tjernström, M., ... Gascard, J. C. (2014). Advances in understanding and parameterization of small-scale physical processes in the marine Arctic climate system: a review. *Atmospheric Chemistry and Physics*, *14*(17), 9403–9450.

Wadhams, P. (1975). Airborne laser profiling of swell in an open ice field. *Journal of Geophysical Research*, *80*(33), 4520–4528.

Wadhams, P. (1986). The seasonal ice zone. In *The geophysics of sea ice* (pp. 825–991). Springer.

Wadhams, P., Aulicino, G., Parmiggiani, F. F., Persson, P. P. O. G., & Holt, B. (2018). Pancake ice thickness mapping in the Beaufort Sea from wave dispersion observed in SAR imagery. *Journal of Geophysical Research: Oceans*, *123*(3), 2213–2237.

Wadhams, P., Lange, M. A., & Ackley, S. F. (1987). The ice thickness distribution across the Atlantic sector of the Antarctic Ocean in midwinter. *Journal of Geophysical Research: Oceans*, *92*(C13), 14535–14552.

Weeks, W. (2010). *On sea ice*. University of Alaska Press.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.

Williams, T. D., Bennetts, L. G., Squire, V. A., Dumont, D., & Bertino, L. (2013a). Wave–ice interactions in the marginal ice zone. Part 1: Theoretical foundations. *Ocean Modelling*, *71*, 81–91.

Williams, T. D., Bennetts, L. G., Squire, V. A., Dumont, D., & Bertino, L. (2013b). Wave–ice interactions in the marginal ice zone. Part 2: Numerical implementation and sensitivity studies along 1D transects of the ocean surface. *Ocean Modelling*, *71*, 92–101.

WMO. (2014). *Sea-ice nomenclature.* Retrieved 2023-11-06, from `https://library.wmo.int/idurl/4/41953`

Womack, A., Vichi, M., Alberello, A., & Toffoli, A. (2022). Atmospheric drivers of a winter-to-spring Lagrangian sea-ice drift in the Eastern Antarctic marginal ice zone. *Journal of Glaciology*, *68*(271), 999–1013.

Young, I. R., Fontaine, E., Liu, Q., & Babanin, A. V. (2020). The wave climate of the Southern Ocean. *Journal of Physical Oceanography*, *50*(5), 1417–1433.

Zhang, J., Schweiger, A., Steele, M., & Stern, H. (2015). Sea ice floe size distribution in the marginal ice zone: Theory and numerical experiments. *Journal of Geophysical Research: Oceans*, *120*(5), 3484–3498.

# Supplemental Materials: Unsupervised classification of the Antarctic marginal ice zone

## Introduction

This Supporting Information provides additional context for the results of the accompanying manuscript. Figures of the $k$-means-derived Antarctic marginal ice zone when floe size was removed from the analysis dataset (Fig. S1–S2), and additional time series of MIZ dynamics (rate of new ice formation and internal stress) are included (Fig. S4). When $k$-means clustering is only completed on floe size data, it returns a set of thresholds to differentiate clusters (Fig. S2). The summer MIZ is more similar to when $k$-means clustering was completed without floe size data (Fig. S1), however, floe size data eliminates the algorithm identifying cells in coastal polynyas. The MIZ width in 2019 is comparable to the 2010–2019 average (Fig. S3). When floe size is excluded from the clustering the resulting MIZ is similar to the area with sea ice concentrations between 15–80%. Internal stresses and the change in floe size from new ice formation are consistently lower in the MIZ than in the interior clusters (Fig. S4).

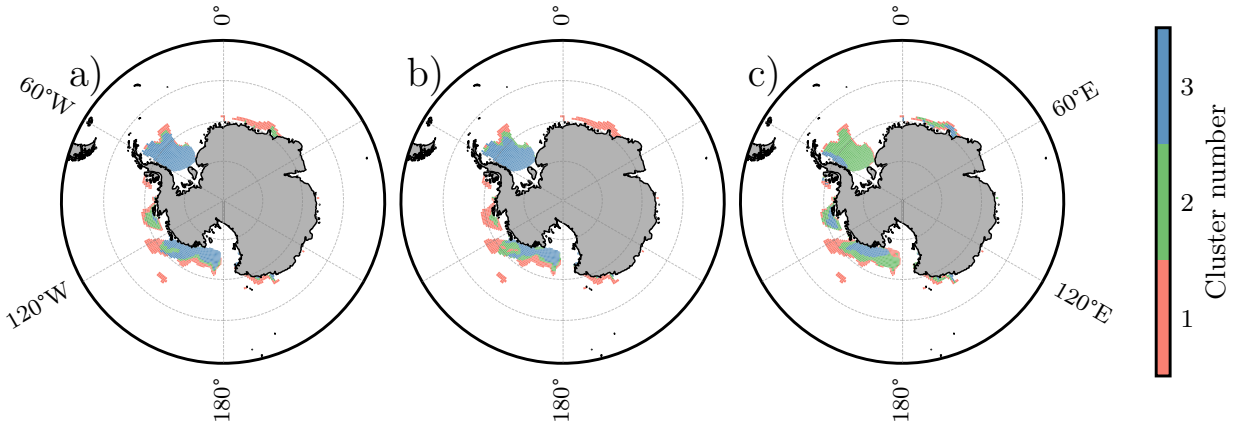


Figure S1: Same as Figure 3 but at the average summer sea ice extent minimum (14th of February 2019).

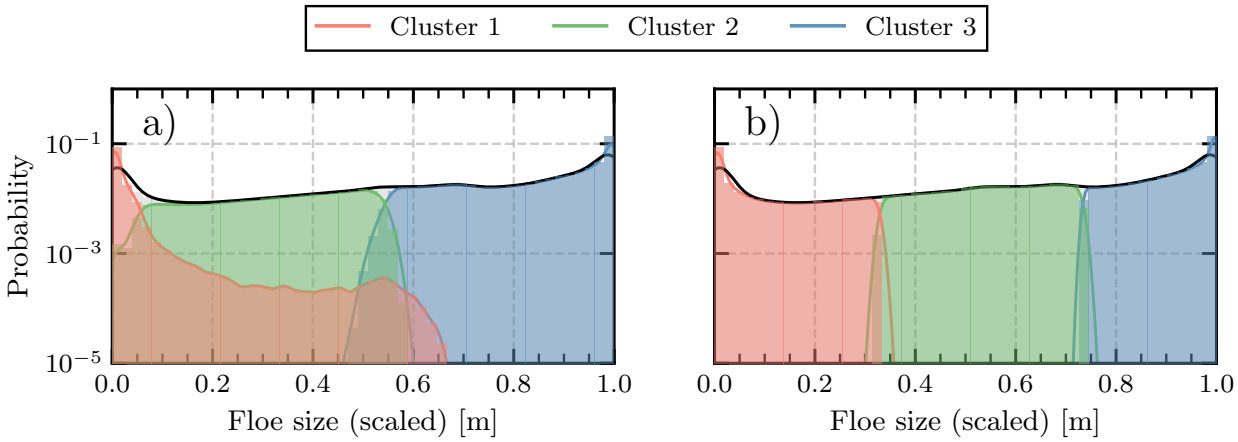


Figure S2: Floe size distribution of each of the sea ice clusters from $k$-means using a) the full dataset and b) only floe size data. The combined FSD is overlaid in the thick black line. a) is the same as Fig. 2b.

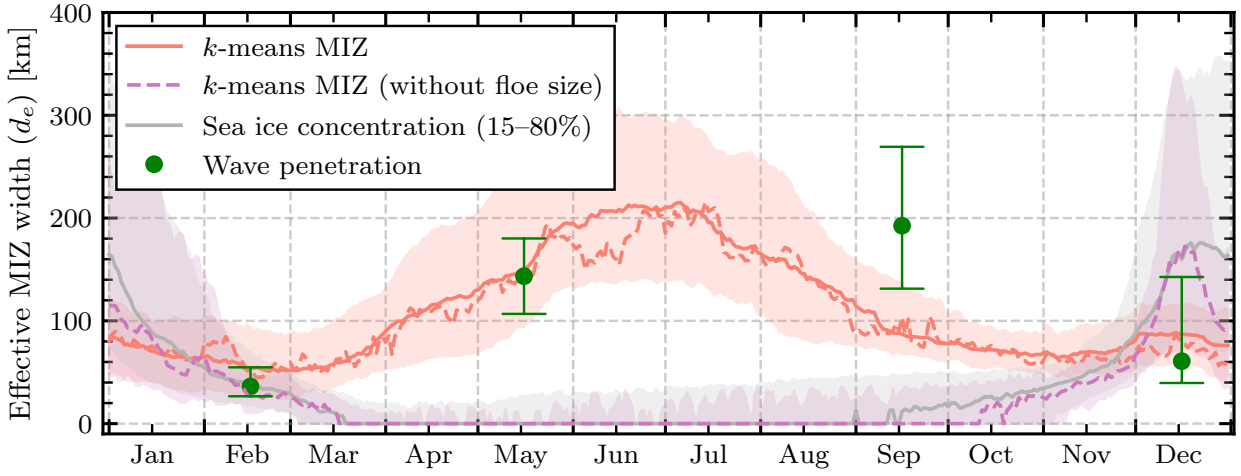


Figure S3: Same as Figure 6, but with the MIZ width result (dashed-red line) and MIZ width from when $k$-means was completed without floe size data (dashed-pink line) for 2019 overlaid.

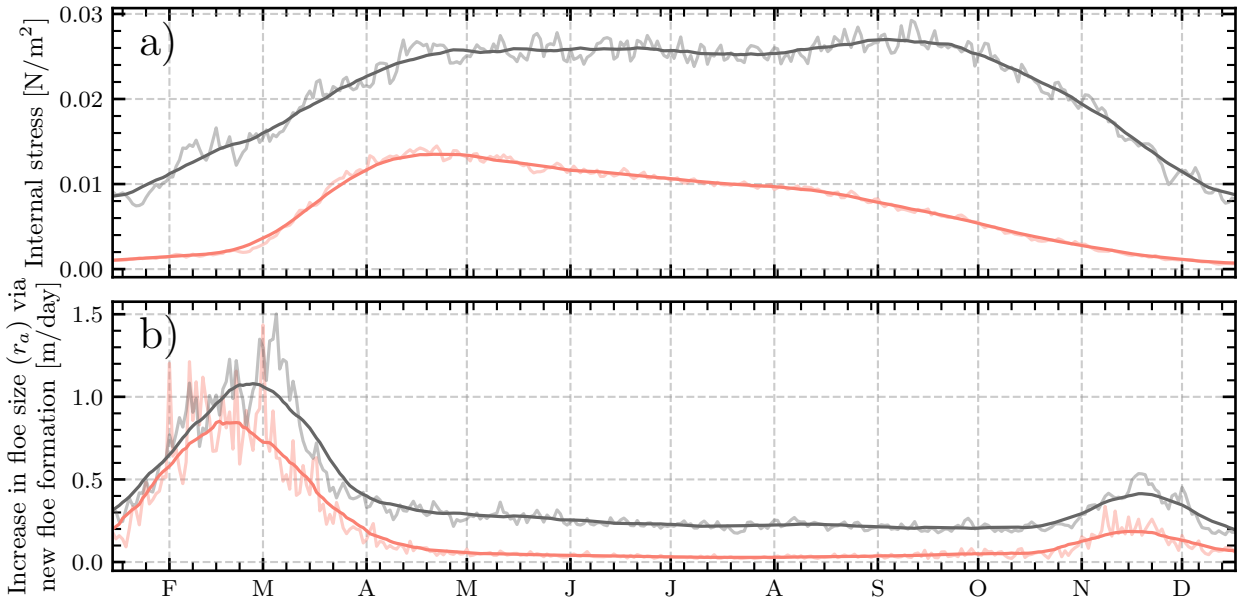


Figure S4: Same as Figure 8 (MIZ in red, interior pack ice in grey) but for a) internal stress and b) the increase in representative floe size from new ice floe formation.