https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

Ocean Science
Discussions
Open Access
EGU

A clustering-based approach to ocean model-data comparison around Antarctica

Qiang Sun[1], Christopher M. Little[1], Alice M. Barthel[2] and Laurie Padman[3]

1 Atmospheric and Environmental Research, Inc., Lexington, MA 02421, USA
2 Los Alamos National Laboratory, Los Alamos, NM 87545
3 Earth and Space Research, 3350 SW Cascade Ave., Corvallis, OR 97333, USA
Correspondence to: Qiang Sun, (qsun@aer.com)

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

Ocean Science
Discussions
Open Access
EGU

Abstract

The Antarctic Continental Shelf Seas (ACSS) are a critical, rapidly-changing element of the Earth system. Analyses of global-scale general circulation model (GCM) simulations, including those available through the Coupled Model Intercomparison Project, Phase 6 (CMIP6), can help reveal the origins of observed changes and predict the future evolution of the ACSS. However, an evaluation of ACSS hydrography in GCMs is vital: previous CMIP ensembles exhibit substantial mean-state biases (reflecting, for example, misplaced water masses) with a wide inter-model spread. Here, we demonstrate the utility of clustering tools for the identification and model-data comparison of hydrographic regimes. In this proof-of-concept analysis, we apply K-means clustering to hydrographic metrics from one GCM (Community Earth System Model version 2; CESM2) and one observation-based product (World Ocean Atlas 2018; WOA), focusing on the Amundsen, Bellingshausen, and Ross Seas. When applied to WOA temperature and salinity profiles, clustering identifies "source" and "mixed" regimes that have a physically interpretable basis. For example, meltwater-freshened coastal currents in the Amundsen Sea, and high salinity shelf water formation regions in the southwestern Ross Sea, emerge naturally from the algorithm. Both regions also exhibit clearly differentiated inner- and outer-shelf regimes. The same analysis applied to CESM2 demonstrates that, although mean-state model bias can be substantial, using a clustering approach highlights that the relative differences between regimes, and the locations where each regime dominates, are well represented in the model. CESM2 is generally fresher and warmer than WOA and lacks a clearly defined fresh-water-enriched coastal current. Given the sparsity of observations on the ACSS, this technique is a promising tool for the evaluation of a larger model ensemble (e.g., CMIP6) on a circum-Antarctic basis.

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

1. Introduction

The Antarctic Continental Shelf Seas (ACSS, defined here as the ocean regions adjacent to Antarctica with water depth shallower than 2,500 m) are critical components of the climate system, playing an essential role in ice sheet mass balance, sea ice formation, and ocean circulation (Rignot et al., 2008; Hobbs et al., 2016; Bindoff, Rosenberg and Warner, 2000). ACSS ocean state, and the climate system components that are coupled to it, are changing rapidly. In the Amundsen-Bellingshausen Seas sectors, the atmosphere (Bromwich, et al., 2013) and subsurface ocean (Schmidtko et al., 2014) are warming, the sea ice-free period is rapidly increasing (Stammerjohn et al., 2012), ice shelves are thinning (Rignot et al., 2013; Paolo, Fricker and Padman, 2015), and the grounded portion of the ice sheet is losing mass at an unprecedented rate (Shepherd et al., 2018; Sutterley et al., 2014; Gardner et al., 2018). The Ross Sea has also experienced long-term changes in fresh water content (Jacobs and Giulivi, 2010; Castagno et al., 2019) and an increase in sea ice production and extent (Parkinson, 2019; M. Holland et al., 2017).

Assessing the causes of observed changes in climate and the coastal cryosphere, and their future evolution, requires coupled, global, atmosphere-ocean general circulation models (GCMs). However, recent GCMs exhibit large biases relative to modern observations, and a wide inter-model spread (Agosta, Fettweis and Datta, 2015; Sallée et al., 2013; Rickard and Behrens, 2016; Hosking et al., 2016; Little and Urban, 2016; Barthel et al., 2019). These errors may influence the future rate of regional warming, its vertical and horizontal distribution (e.g., Sallée et al., 2013; Agosta, Fettweis and Datta, 2015), and its consequences for the Antarctic Ice Sheet. For example, DeConto and Pollard (2016) projected extreme rates of 21st-century ice sheet mass loss from the Pacific sector for a high-emission scenario. However, their projections were forced using a single GCM (CCSM4) that required a +3°C correction to subsurface water temperatures in the Amundsen Sea to match observed hydrography and modern ice shelf melt rates. This significant bias correction indicates an underlying mean-state error (e.g., a misplaced water mass) that indicates substantial uncertainty in projected future ocean state.

The first step toward identifying the physical processes underlying GCM representation errors is assessing the magnitude and spatial distribution of biases. However, such a strategy must account for strong horizontal and vertical gradients in ACSS hydrographic properties, and the sparseness and variable quality of available observations. Strong gradients are evident in the Amundsen, Bellingshausen, and Ross Seas (ABRS) sector of the ACSS. There, the time-mean ocean state of the objectively analyzed temperature and salinity field, as represented in the 0.25-degree World Ocean Atlas version 2018 (WOA hereafter), suggests that the ABRS can be roughly

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

separated into two geographical regions, the Amundsen-Bellingshausen Seas and the Ross Sea (Figure 1a). In the Ross Sea, dense water formation occurs locally, through brine rejection from winter sea ice formation in coastal polynyas, resulting in regionally averaged water well below 0°C at water depths of 100m to 700m (Figure 1c). At the same depth range in the Amundsen-Bellingshausen Seas, water temperatures can reach +1.2°C due to the presence of Circumpolar Deep Water (CDW).

In addition to these stark contrasts in regional mean temperature (and salinity), there is also significant spatial variability within each region of the ABRS, and across the continental shelf break. For example, Figure 1 indicates a high standard deviation in ocean temperature on the continental shelf with water depth shallower than 700m (0.5°C in the Amundsen-Bellingshausen Seas and 1.4°C in the Ross Sea). Much of this variability is attributable to the lateral temperature gradient from the subsurface layer of CDW over the continental slope to the modified (cooled) water masses inshore. In the alongshore direction, vertical water properties in the Amundsen-Bellingshausen Seas are similar, with cold and fresh water overlying relatively warm and salty water. In the Ross Sea, water properties are different on its southwestern and eastern sides, mainly distinguished by their salinity (Figure 1d).

The sparseness of measurements on the ACSS also aggravates errors associated with gridded observational products. Coastal regions, in particular, are subject to substantial errors. Sun et al. (2019) showed that salinity biases between WOA objective analysis and the World Ocean Database increase toward coastlines. The gridded objective analysis field neglects the dynamical processes governing water mass modifications and circulations induced by complex continental shelf bathymetry (Dunn and Ridgway, 2002; Schmidtko, Johnson and Lyman, 2013).

Previous model-data comparisons on the ACSS have employed strategies such as selecting the closest grid points to the model's southern boundary (Little and Urban, 2016) or averaging over a-priori defined regions (e.g., Barthel et al., 2019). In addition to predefining the region of comparison, such methods are ill-equipped to assess model biases resulting from misplaced water masses. An alternative method is objective clustering, which can be used to identify regions of similar hydrographic properties. For example, Hjelmervik and Hjelmervik (2013) demonstrated application of a clustering-based approach using ARGO profiles to segregate the North Atlantic into groups with similar vertical *T* and *S* profiles separated by fronts.

A clustering approach requires choosing metrics to represent water mass properties. Schmidtko et al. (2014) partitioned water masses in the Southern Ocean into Winter Water (WW), CDW, and Antarctic Shelf Bottom Water

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

(ASBW) using only temperature. However, their metrics of subsurface water temperature maxima and minima are ineffective on the continental shelf, where temperature profiles are often complex and show strong lateral variability in water properties (Figure 1d). Sallée et al. (2013) proposed a method to use potential vorticity evaluated from density profiles and the local salinity minimum at 30ºS to distinguish vertical water masses in the Southern Ocean. However, their metrics are also ill-suited to identifying different hydrographic regimes on the ACSS. Here, informed by these previous studies, we develop new metrics targeted at ACSS hydrography and assess the utility of a clustering-based approach for model-data comparison.

2. Methods

In this paper, we identify hydrographic regimes and their *T*-*S* properties using metrics derived from three-dimensional grids of measured and modeled temperature and salinity (section 2.1) using a K-means clustering method (section 2.2). We then apply a clustering algorithm based on data density to exclude outliers (2.3) from the resulting "groups".

2.1 Data and processing

We use decadal-mean, objectively analyzed *T* and *S* fields from WOA for 1995-2004, with 0.25-degree resolution in both latitude and longitude. This study focuses on the domain from the west of Cape Adare (163ºE) on the western side of the Ross Sea to the southern end of Alexander Island (76ºW), at depths between 0 and 2,500 m. The landward limit of the study domain is the Antarctic coast and the ice shelf edges as identified in Figure 1a.

We compare the Community Earth System Model version 2 (CESM2; Danabasoglu et al., 2019) to WOA for the same period and domain. The time-mean model salinity and temperature fields over the 1995-2004 period are calculated from the monthly output of the Coupled Model Intercomparison Project Phase 6 (CMIP6) historical simulation (experiment tag r1i1p1f) (Eyring et al., 2016) at the native ocean model resolution (roughly 1-degree in longitude and 0.5-degree in latitude).

We used the Gibbs SeaWater (GSW) Oceanographic Toolbox of TEOS–10 (McDougall and Barker, 2011) to calculate seawater properties. The absolute salinity (S_A) has unit of g/kg, and conservative temperature (Θ) is in ºC. All seawater temperatures are referenced to the sea surface.

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

2.2 Prototype-based clustering technique (K-means)

The K-means clustering analysis used in this study is an unsupervised learning technique that classifies data into meaningful groups based on their similarity. Here, we are concerned with the hydrographic structure of a water column, and we wish to identify regions that exhibit a similar vertical structure (i.e., "hydrographic regimes"). In this study, the similarity is defined by two metrics of the water column: 1) salinity at the temperature minimum; and 2) salinity at the temperature maximum. The rationale for these choices is discussed in section 3.1.

The K-means algorithm is initialized by randomly selecting data in N dimensions (here, N=2, for the two specified metrics) for a specified number of groups (K). For each group (k_i), the Sum of Squared Distance (SSD) of each data point ($\boldsymbol{\xi}$) to the group's centroid ($\boldsymbol{c_i}$) is calculated:

$$SSD = \sum_{i=1}^{K} \sum_{\boldsymbol{\xi} \in k_i} dist(\boldsymbol{c_i}, \boldsymbol{\xi})^2 \quad with \quad \boldsymbol{c_i} = \frac{1}{m_i} \sum_{\boldsymbol{\xi} \in k_i} \boldsymbol{\xi} \qquad \text{Eqn. 1}$$

where $dist$ is the standard distance between data and centroid in N-dimensional Euclidean data space and m_i is the total number of data points in group k_i. The algorithm iterates to minimize SSD by adjusting the centroids and rearranging the data nearest to them. The K-means algorithm will have multiple solutions because it is initialized with randomly selected data. We apply the K-means 1,000 times and choose the solution with the lowest SSD for analysis.

The K-means algorithm requires specification of the number of groups (K). We use Silhouette scores $s_i(n)$ (Eqn. 2) to assess the appropriate values of K.

$$s_i(n) = \frac{b(\boldsymbol{\xi}) - a(\boldsymbol{\xi})}{max\{a(\boldsymbol{\xi}), b(\boldsymbol{\xi})\}} \qquad \text{Eqn. 2}$$

In Eqn. 2, n represents the number of data points in group k_i, $a(\boldsymbol{\xi})$ is the mean $dist$ from a data point $\boldsymbol{\xi}$ to all other data points within the group k_i, and $b(\boldsymbol{\xi})$ is mean $dist$ from $\boldsymbol{\xi}$ to all other data points outside the group k_i. Silhouette scores are evaluated for each data point $\boldsymbol{\xi}$ in the group k_i and range between -1 and 1. If $\boldsymbol{\xi}$ lays perfectly at the centroid of group k_i, then $s_i(n)$=1.

A rigid interpretation of the Silhouette algorithm would choose the value of K that corresponds to the highest mean value of $s_i(n)$. However, the optimal K value can vary with different clustering evaluation methods

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020


Ocean Science
Discussions
Open Access
EGU

(e.g., Elbow method: Thorndike, 1953) and different domains. The selection of K is thus based not only on the results of Silhouette assessment but also on the ability to interpret the groups as representative of different underlying physical processes (see section 3).

2.3 Density-based clustering technique

In subsequent sections, we use a T-S diagram to compare the properties of groups given by the K-means algorithm. We applied a data density-based clustering technique (DBSCAN) (Ester et al., 1996) to define the "core" of a group and to exclude outliers on the T-S diagram. Note that DBSCAN is only used to highlight the core of a given group and facilitate comparisons of water properties between WOA and CESM2.

The T-S core of each hydrographic regime identified by the K-means clustering is determined by the DBSCAN algorithm using two parameters: a radius (ε), and a minimum number of neighboring points (MinPts). The DBSCAN algorithm builds up pools of data by initially choosing a random data point. If the initially chosen data point has less than MinPts within ε, then it is defined as an outlier. If this data point has more than MinPts within ε, then a pool of data is initialized consisting of the initial point and the points within ε (neighbors). The pool grows by continually clustering neighboring points until these points have fewer than MinPts within ε. The algorithm continues until all data points are either clustered into pools of data or labeled as outliers. In the current study, we choose MinPts = 10 and $\varepsilon = \sqrt{S^2 + T^2}$. The value of ε is then selected (Table 1) so that the largest pool of data contains at least 97% of non-outlier points (Table 2). This pool of data constitutes the core of each group.

3 Results

3.1 Defining water column metrics

Our goal in this analysis is to utilize key features of local water columns to identify regions with similar hydrographic properties. Such metrics must be able to capture the changes in T and S in both along- and cross-shelf directions. The metrics should include salinity because it is the dominant factor influencing water column stability and reflects processes such as fresh water input and brine rejection. However, salinity poorly represents the vertical water mass changes since it increases monotonically with water depth (Figure 1); salinity alone is insufficient to identify regimes with sub-surface heat reservoirs, which are characteristic of regions with high ice shelf basal melt rates (Rignot et al., 2013; Dinniman et al., 2016; Holland et al., 2020). The metrics we use in this study – salinity at

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

the vertical temperature minimum and salinity at the vertical temperature maximum – are similar to those used by Timmermans et al. (2014) to segregate surface water from Alaska coastal water in the Central Canada Basin of the Arctic Ocean.

Along-shelf variations of water properties are evident in salinity at the vertical temperature minimum (Figure 2b). In the Amundsen-Bellingshausen Seas, the depth of minimum temperature (Figure 2c) is commonly above 200 m, where salinity is often less than 34.2 g/kg. In contrast, in the southwestern Ross Sea, the minimum temperature is usually located below 350m and coincides with much higher salinity (>34.8 g/kg). The northwestern Ross Sea contains a regime with a local temperature minimum at shallower depths approaching the shelf break, but its salinity (between 34.2 to 34.6 g/kg) is higher than near-surface water in the Amundsen-Bellingshausen Seas.

The salinity at the vertical temperature maximum shows pronounced variations in the cross-shelf direction (Figure 2d-f). The maximum water temperature (Figure 2d) is commonly found at depths above 200 m close to the coast and ice shelves (Figure 2f), and deeper toward the shelf break and over the continental slope where the water depth increases. The salinity at the vertical maximal temperature (Figure 2e) shows similar variations in the cross-shelf direction, with lower salinity (<34.7 g/kg) near the coast and ice shelves and higher salinity (>34.8 g/kg) on the continental shelves and near the shelf break.

3.2 Evaluating the optimum number of groups

We used the mean value of Silhouette score $s_i(n)$ in Eqn. 2 to evaluate an appropriate number of groups (K) for WOA and CESM2 with $2 \leq K \leq 13$ (Figure 3). For WOA, the highest value of s_i occurs when K=3; for CESM2, K=6 has the highest Silhouette score (Figure 3a-b). The spatial distribution of groups 3, 5 and 6 in the ABRS are shown in Figure 3c-h.

When WOA data are clustered into three groups (Figure 3c), the K-means algorithm segregates the water close to the Antarctic coast from the water on the shelf and continental slope. The coastal domains are further distinguished into Amundsen-Bellingshausen coastal waters and Ross coastal waters. By increasing the number of groups to five (Figure 3e), a narrow domain between coastal and shelf waters emerges. In the Ross Sea, waters on the shelf and across the shelf break are segregated into two groups. For K=6 (Figure 3g), the southeastern coastal domain of the Ross Sea (orange) is further separated from the narrow domain between coastal and shelf waters in the Amundsen-Bellingshausen Seas, while the locations of the other groups are generally unchanged.

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020


Ocean Science
Discussions
Open Access
EGU

191 Examining the groups with respect to the two metrics used in the K-means clustering (Figure 4) shows that,
192 when K=3, the groups are separated by the perpendicular lines from the incenter of the triangular T-S distributions
193 (Figure 4a). As the total number of groups increases, data points are progressively divided into smaller subsets, with
194 an asymmetry that is influenced by their original distribution in our two-metric parameter space, as well as gaps and
195 discontinuities (Figure 4c and e).
196 In CESM2, the waters in the ABRS are clustered differently. For K=3 (Figure 3d), the entire Amundsen-
197 Bellingshausen Seas region is segregated from the Ross Sea, while the southwestern Ross Sea is still recognized as
198 an independent group. With K=5 (Figure 3f), a coastal group emerges in the Amundsen-Bellingshausen Seas,
199 although its areal extent is much smaller than in WOA. In the Ross Sea, the water on the continental shelf is
200 separated from the water on the continental slope, similar to WOA. For K=6, the Amundsen Sea is segregated from
201 the Bellingshausen Sea. In Figure 4, CESM2 shows a similar range to WOA in metric space, although with much
202 larger gaps. In particular, CESM2 has substantially fewer data points with intermediate and low salinities (Figure
203 4b). Increasing K for clustering analysis of CESM2 output subdivides high salinity regimes at T_{max} based on the
204 distribution of salinity at T_{min} (Figure 4d and f).
205 For model-data comparison, the properties and general locations of the groups must be comparable.
206 Although the Silhouette scores suggest the best option is 6 groups in CESM2, the WOA data does not support 6
207 distinct groups, as the water properties below the surface layers are nearly indistinguishable between the Amundsen
208 and Bellingshausen Seas (Figure 1d). Figure 4 also indicates that the segregation of Amundsen-Bellingshausen Seas
209 regions in CESM2 is a result of discontinuities between groups 1 and 5 (Figure 4f). We thus choose to use 5 groups
210 for the rest of the study. Our findings from analyzing the temperature and salinity properties in the following
211 sections further support this decision.

212 **3.3 Physical interpretation of WOA groups**

213 Vertical profiles of temperature and salinity are shown for each WOA group in Figure 5. The mean vertical
214 structure of each group is clearly different; furthermore, the standard deviations at each depth within groups are
215 much smaller than those of regional mean profiles (Table 3). With these vertical structures as context, we examine
216 T-S properties at all depths from each WOA group in Figure 6. The DBSCAN algorithm is used to identify the "core"
217 of non-outlier data in each group, shown with dark shading in Figure 6.

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

Group 1, which occupies the inshore regions of the Amundsen-Bellingshausen Seas (Figure 3e), is characterized by weak vertical gradients in both *T* and *S* over the ~400 m water column (Figure 5a). The water in this group has relatively low salinity (33.8 to 34.5 g/kg), temperature close to the freezing point (generally lower than -1°C) and low density (26.9 and 27.5 kg/m^3) (Figure 6a), which suggests that the water in this regime is strongly influenced by coastal fresh water input (Moffat et al., 2008; Jacobs and Giulivi, 2010; Jourdain et al., 2017).

Group 2, which is spatially located between the coastal waters (groups 1 and 5) and outer continental shelf waters (groups 3 and 4), represents a narrow domain of mixing (Figure 3e). This regime is characterized by relatively high standard deviations in salinity and temperature at depths between 100 m and 700 m, indicating that the location and shape of the thermocline and halocline above the typical depth of the shelf break vary within this group (Figure 5b). Below 700 m, the range of salinity and temperature are relatively small, due to reduced water mass variability and/or the limited amount of data at these depths over the relatively narrow continental slope. In the upper ocean, group 2 has a salinity from 33.8 to 34.7 g/kg, temperature -2 to -0.5°C and density 27.1 to 27.8 kg/m^3 (Figure 5b and Figure 6b), lying between the properties of surface waters in groups 1 and 5. In the subsurface, group 2 has a temperature above -0.5°C and salinity above 34.5 g/kg, which represents modified CDW on the shelf (Carmack, 1970; Orsi and Wiederwohl, 2009; Emery, 2011).

Group 3, which is found on the outer continental shelf and the continental slope of the Ross Sea (Figure 3e), shows high standard deviations in temperature above 700 m (Figure 5c), similar to group 2. However, the water in this regime is generally denser than group 2. The surface water in group 3 is fresher than that of group 5 (Figure5c, Figure 6c and f), which may result from sea ice melt and/or lateral mixing with fresher shelf water originating in the Amundsen-Bellingshausen Seas (Assmann, Hellmer and Jacobs, 2005; Porter, et al., 2019). The subsurface water (between 100 and 600 m) of group 3 (Figure 5c and Figure 6c) does not have a clear vertical water mass transition, and denser water exhibits a wide temperature range (-1.5 to +1.5°C) with relatively high salinity (34.6 to 35 g/kg), suggestive of mixing between High Salinity Shelf Water (HSSW) and CDW.

Group 4, on the shelf of the Amundsen-Bellingshausen Seas and along most of the continental slope of the ABRS (Figure 3e), exhibits properties consistent with off-shelf Southern Ocean water as noted by Schmidtko et al. (2014). It has a well-defined vertical temperature structure with limited spatial variability (Figure 5d). In this region, Winter Water (WW) with salinity 33.8-34.5 g/kg, temperature -2 to -0.5°C and density 27 to 27.5 g/m^3, overlays

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

CDW (salinity 34.6 to 36.8 g/kg, temperature 0 to +2°C and density 27.8 to 27.9 g/m^3), with a mean profile showing a clear transition between them (Figure 6d).

Group 5, in the southwestern Ross Sea with some extensions to the southeast (Figure 3e), has higher salinity than other groups (Figure 6). The almost uniform vertical temperature profile (Figure 5e) is identified as HSSW. It is characterized by salinity 34.3 to 35.1 g/kg, temperature close to the freezing point, and density of 27.5 to 28.1 kg/m^3 (Figure 6e), resulting from brine rejection in the polynyas along the coast and Ross Ice Shelf front (Foster and Carmack, 1976). The surface portion of the waters in group 5 with salinity lower than 34.62 g/kg is often defined as Low Salinity Shelf Water (LSSW) in the Ross Sea shelf, but we generally refer to group 5 as HSSW because its volume is much higher than the LSSW (Orsi and Wiederwohl, 2009).

Overall, groups 1 and 5 (Figures 6a and 6e) show relatively homogeneous salinity and temperature, while group 4 has a pronounced thermocline and halocline at shallow depth. These three groups (1, 4 and 5) represent the three "source" ABRS hydrographic regimes. In contrast to these source regimes, groups 2 and 3 have more complex vertical structures, more spatial variability in thermocline at depths above about 600 m (roughly the shelf break) and can be considered as "mixed" regimes.

3.4 Assessing groups in CESM2

To identify hydrographic regimes in CESM2, we conduct the same analyses as described for WOA in the previous section, focusing on results for K=5 (Figure 3f). The T-S properties of each group in CESM2 are shown in Figure 7. CESM2 results are similar to WOA's in that three "source" waters are present (group 1, coastal fresh-water-enriched; group 4, off-shelf; and group 5, HSSW), but they show differences in their spatial extent (Figure 3e vs. f), volume (Table 4), and T-S properties (Figure 8).

As in WOA, HSSW (group 5) of CESM2 is localized in the southwestern Ross Sea, but its eastward extension into the southeastern Ross Sea is missing in CESM2 (Figure 3e and f), resulting in a reduced HSSW volume (Table 4). The coastal fresh water-enriched regime (group 1) is mostly absent in CESM2 and is replaced by the off-shelf regime in the Amundsen Sea.

Mismatches between CESM2 and WOA are also evident in the T-S properties of these source regimes. In general, HSSW in CESM2 has a fresh and warm bias relative to WOA (Figure 8d). Combined with its reduced volume relative to WOA, this bias in CESM2 HSSW properties suggests that weak katabatic winds in the southwestern Ross Sea may limit sea ice production and export. Group 4 (the off-shelf regime) exhibits a fresh bias

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

in WW in the upper water column, but the densest off-shelf water in group 4, i.e. CDW, is saltier and warmer (Figure 8c).

The mixed regimes shift geographic location in CESM2. The narrow mixing zone (group 2) between coastal fresh-water-enriched and off-shelf regimes in the Amundsen-Bellingshausen Seas in not evident in CESM2 (Figure 3e and f). In the Ross Sea, groups are separated into on-shelf (group 2) and off-shelf (group 3) approximately along the 1,000 m isobath (Figure 3f). CESM2 fails to show the path of export of Ross on-shelf water (group 2, Figure 3f) along the northwestern continental slope (Orsi, Johnson and Bullister, 1999), as it is seen in the WOA (group 3, Figure 3e). The core of on-shelf water (group 2) also has less overlap with HSSW (group 5) in the *T-S* diagram in CESM2 (Figure 7f) compared to WOA (Figure 6f). These results suggest that CESM2 may misrepresent the on- and off-shelf exchange of waters in the Ross Sea.

3.5 Assessing clustering over the ACSS

As the K-means algorithm is based on purely statistical criteria (centroid and minimized *SSD* in Eqn. 1) applied to specific metrics, it is valuable to assess whether clustering results are sensitive to different study domains. As a test case, we apply the same algorithm to WOA over the entire circumpolar ACSS with water depth above 2,500 m. The metrics used as input for the K-means analysis, as well as the total number of groups (K=5), are unchanged. The use of the uniformly-gridded WOA product, rather than observational data, avoids the possibility that the comparison is biased by regional variations in data density.

The location of five clustered water groups over the entire ACSS is shown in Figure 9a. Within the ABRS domain, the geographic locations of all groups are almost unchanged, indicating the clustering results in the ABRS are insensitive to substantial enlargement of the domain. The region identified as HSSW (group 5), in the southwestern Ross Sea, remains. Outside the ABRS, the clustering approach identifies water of similar properties to group 5 in the Weddell Sea near the Filchner-Ronne Ice Shelf, the George V Coast near the Mertz Glacier tongue, and Bransfield Strait and south of Trinity Peninsula (regions marked on Figure 9b). The southern Weddell Sea experiences similar conditions to the southwestern Ross Sea, with HSSW formation in winter due to brine rejection from sea ice formation enhanced by katabatic winds and tides driving a narrow but persistent along-ice-front polynya (Nicholls et al., 2009). Along the George V Coast, HSSW is also generated by similar processes acting near the Mertz Glacier ice tongue (Bindoff, Rosenberg and Warner, 2000; Post et al., 2011).

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

The waters in the subsurface of Bransfield Strait and south of the Trinity Peninsula are also grouped with the HSSW regions, although their surface water is warmer and fresher than that of other HSSW regions around Antarctica. Cook et al. (2016) showed that the regional water properties around the tip of the Antarctica Peninsula, based on the World Ocean Database, are very similar to HSSW. Gordon et al. (2000) also noted that the water properties in the center of Bransfield Strait are similar to HSSW in the Weddell Sea; they inferred that these waters are formed in western Weddell Sea coastal polynyas and flow into Bransfield Strait.

4 Discussion

We have shown that the ABRS can be clustered into different regions based on salinities at the vertical water temperature minimum and maximum. This technique can help identify regions, in model and observational datasets, in which water properties are controlled by similar physical processes. It contrasts with traditional grid point-based comparisons, which do not adequately account for misplaced water masses.

In this study, WOA has been employed to assess CESM2 results. However, the hydrographic regimes identified in WOA may be misleading if they result from interpolation/extrapolation artefacts associated with non-uniform sampling of data in time and space, or if the water column structures are not adequately represented in WOA. One source of uncertainty in WOA arises from differences between true and gridded bathymetry, complicating interpolation and extrapolation of sparely sampled data into deeper portions of the water column. In Figure 10, we compare the depths of the deepest available data in WOA and CESM2 with water depths in the International Bathymetric Chart of the Southern Ocean (IBCSO, Arndt, et al., 2013). WOA has a clear misrepresentation of the Amundsen-Bellingshausen Seas continental shelf bathymetry (Figure 10b). First, the 1,000 m isobath is shifted substantially landward in the Amundsen Sea. Second, deep across-shelf troughs (e.g., in Figure 10a) are not represented in the inner shelf of WOA, which possibly affects the value of salinity at the temperature maximum because the CDW is missing in these regions of the Amundsen-Bellingshausen Seas.

It is, therefore, unclear whether groups 1 and 2 are separated from the shelf and continental slope waters of group 4 in WOA (Figure 3e) due to their upper-ocean fresh water enrichment relative to other groups, or if the groups are influenced by under-sampling of hydrography in deep troughs of the Amundsen-Bellingshausen Seas. We note that the bathymetry of CESM2 has similar issues as WOA in the Amundsen-Bellingshausen Seas (Figure 10c). Neither WOA nor CESM2 represents the water in deep troughs below about 300 m in these regions, so the

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

differences in the groups between WOA and CESM2, i.e., the missing group 1 in the Amundsen coast and narrow group 2 in the Bellingshausen Sea, are unlikely to be due to the bathymetric misrepresentation (Figure 3e and f). We suggest, instead, that the mismatch of water properties is likely to be induced by the misrepresentation of fresh water input, or unresolved coastal currents, in CESM2 (Tseng, Bryan and Whitney, 2016; Sun et al., 2017).

We have highlighted a key advantage to assessing models with clustering-based approaches compared to traditional grid point-based methods; the ability to identify geographic displacements of hydrographic regimes and to distinguish these displacements from biases in water mass *T*-*S* properties. In addition, this approach minimizes potential biases introduced during gridding or re-gridding of data and models to a common grid for comparison studies. For example, it is possible to circumvent interpolation-related issues associated with using scattered and/or sparse data. Such datasets might include individual observations, or model output on a native grid. For example, the deepest observational temperature measurements in the World Ocean Database 2018 (WOD), even at a 1-degree resolution, show that the observations are available in coastal Amundsen-Bellingshausen Seas troughs that are not present in IBCSO (compare Figure 10d with Figure 10a); see, also, Padman, et al. (2010). More broadly, the WOD-based salinity and temperature climatology of Sun et al. (2019) reveals its use can avoid biases created by spatial interpolation of shelf water with off-shelf water.

The success of this technique at identifying locations and properties of HSSW regimes at other locations on the Antarctic continental shelf suggests that it might be used to evaluate other global and/or regional models on a circum-Antarctic basis. Other metrics might be employed depending on specific research goals. For example, the pycnocline depth, or the mean or maximum temperature below a fixed depth, may be better metrics of subsurface water masses. However, we note that comparisons of the locations of groups could become complex if the approach is applied to multiple models.

5 Conclusions

We have demonstrated the utility and sensitivity of a clustering-based approach to assessing water mass properties on the Antarctic continental shelf, using the World Ocean Atlas objective analysis product (WOA) and numerical model output from the Community Earth System Model version2 (CESM2). We segregated the waters in the ABRS into 5 physically interpretable groups using the salinity at the minimum and maximum temperature of each water column in the domain. The method identifies High Salinity Shelf Water (HSSW), coastal fresh-water-

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

354 enriched, and off-shelf hydrographic regimes in observations and the model. Water on the continental shelf and
355 upper continental slope in the ABRS generally show a warm bias in CESM2 compared to WOA. The near-surface
356 ocean in CESM2 is generally fresher than in WOA but lacks a well-defined fresh-water-enriched coastal current. In
357 the subsurface, CESM2 is saltier in regions of Circumpolar Deep Water, but fresher than WOA in HSSW formation
358 regions. Our comparison suggests that mean-state biases of CESM2 on the ACSS result from both local and remote
359 processes. A more specific investigation of the coastal processes, Southern Ocean dynamics, and atmospheric
360 forcing will help further identify the cause of these biases.

361 The clustered hydrographic regimes in the ABRS are largely unchanged when our method is applied to the
362 entire circum-Antarctic Continental Shelf Seas. HSSW-characterized regimes emerge in WOA in the southern
363 Weddell Sea, near Mertz Glacier tongue, and in Bransfield Strait. Future work will focus on applying this approach
364 to a wider range of models (e.g., CMIP6 output and circum-Antarctic simulations) and establishing techniques to
365 work with scattered observational data.

366 **Acknowledgments**

367 QS and CL were supported by NSF grant 1744789. AB was supported by the U.S. Department of Energy
368 (DOE) Office of Science Regional and Global Model Analysis (RGMA) component of the Earth and Environmental
369 System Modeling (EESM) program (HiLAT-RASM project). LP was supported by NSF grant 1744789 and NASA
370 grant NNX17AG63G. The authors thank the NCAR climate modeling groups for producing and making available
371 CESM2 output.

372

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

References

Agosta, C., Fettweis, X., & Datta, R. (2015). Evaluation of the CMIP5 models in the aim of regional modelling of the Antarctic surface mass balance. The Cryosphere, 9(6), 2311-2321.

Arndt, J. E., Schenke, H. W., Jakobsson, M., Nitsche, F. O., Buys, G., Goleby, B., . . . Wigley, R. (2013). The International Bathymetric Chart of the Southern Ocean (IBCSO) Version 1.0—A new bathymetric compilation covering circum-Antarctic waters. Geophysical Research Letters, 40(12), 3111-3117.

Arthur, D., & Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms (pp. 1027-1035). Society for Industrial and Applied Mathematics.

Assmann, K. M., Hellmer, H. H., & Jacobs, S. S. (2005). Amundsen Sea ice production and transport. Journal of Geophysical Research: Oceans, 110(C12).

Barthel, A., Agosta, C., Little, C. M., Hatterman, T., Jourdain, N. C., Goelzer, H., . . . Bracegirdle, T. J. (2019). CMIP5 model selection for ISMIP6 ice sheet model forcing: Greenland and Antarctica. The Cryosphere Discuss. Retrieved from https://doi.org/10.5194/tc-2019-191

Bindoff, N. L., Rosenberg, M. A., & Warner, M. J. (2000). On the circulation and water masses over the Antarctic continental slope and rise between 80 and 150 E. Deep Sea Research Part II: Topical Studies in Oceanography, 47(12-13), 2299-2326.

Bromwich, D. H., Nicolas, J. P., Monaghan, A. J., Lazzara, M. A., Keller, L. M., Weidner, G. A., & Wilson, A. B. (2013). Central West Antarctica among the most rapidly warming regions on Earth. Nature Geoscience, 6(2), 139.

Carmack, E. C. (1970). Water characteristics of the Southern Ocean south of the Polar Front. In M. Angel , A voyage of Discovery: George Deacon 70th Anniversary (pp. 15-42).

Cook, A. J., Holland, P. R., Meredith, M. P., Murray, T., Luckman, A., & Vaughan, D. G. (2016). Ocean forcing of glacier retreat in the western Antarctic Peninsula. Science, 353(6296), 283-286.

Danabasoglu, G., Lamarque, J. F., Bachmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., . . . Strand, W. G. (2019). The Community Earth System Model version 2 (CESM2). Journal of Advances in Modeling Earth Systems, submitted.

DeConto, R. M., & Pollard, D. (2016). Contribution of Antarctica to past and future sea-level rise. Nature, 531(7596), 591.

Dunn, J. R., & Ridgway, K. R. (2002). Mapping ocean properties in regions of complex topography. Deep Sea Research Part I: Oceanographic Research Papers, 49(3), 591-604.

Emery, W. J. (2011). Water types and water masses. Encyclopedia of ocean sciences, 6, 3179-3187.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd, 96(34), 226-231.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9, 1937-1958. doi:10.5194/gmd-9-1937-2016

Foster, T. D., & Carmack, E. C. (1976). Frontal zone mixing and Antarctic Bottom Water formation in the southern Weddell Sea. Deep Sea Research and Oceanographic Abstracts, 23(4), 301-317.

Gardner, A. S., Moholdt, G., Scambos, T., Fahnstock, M., Ligtenberg, S., van den Broeke, M., & Nilsson, J. (2018). Increased West Antarctic and unchanged East Antarctic ice discharge over the last 7 years. Cryosphere, 12(2), 521-547.

Gordon, A. L., Mensch, M., Zhaoqian, D., Smethie Jr., W. M., & De Bettencourt, J. (2000). Deep and bottom water of the Bransfield Strait eastern and central basins. Journal of Geophysical Research: Oceans, 105(C5), 11337-11346.

Hjelmervik, K. T., & Hjelmervik, K. (2013, June). Improved estimation of oceanographic climatology using empirical orthogonal functions and clustering. In 2013 MTS/IEEE OCEANS-Bergen, 1-5.

Hobbs, W. R., Massom, R., Stammerjohn, S., Reid, P., Williams, G., & Meier, W. (2016). A review of recent changes in Southern Ocean sea ice, their drivers and forcings. Global and Planetary Change, 143, 228-250.

Holland, M. M., Landrum, L., Raphael, M., & Stammerjohn, S. (2017). Springtime winds drive Ross Sea ice variability and change in the following autumn. Nature communications, 8(1), 1-8.

Hosking, J. S., Orr, A., Bracegirdle, T. J., & Turner, J. (2016). Future circulation changes off West Antarctica: Sensitivity of the Amundsen Sea Low to projected anthropogenic forcing. Geophysical Research Letters, 43(1), 367-376.

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

Jacobs, S. S., & Giulivi, C. F. (2010). Large multidecadal salinity trends near the Pacific–Antarctic continental margin. Journal of Climate, 23(17), 4508-4524.

Jourdain, N. C., Mathiot, P., Merino, N., Durand, G., Le Sommer, J., Spence, P., . . . Madec, G. (2017). Ocean circulation and sea-ice thinning induced by melting ice shelves in the Amundsen Sea. Journal of Geophysical Research: Oceans, 122(3), 2550-2573.

Little, C. M., & Urban, N. M. (2016). CMIP5 temperature biases and 21st century warming around the Antarctic coast. Annals of Glaciology, 57(73), 69-78.

McDougall, T. J., & Barker, P. M. (2011). Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox. SCOR/IAPSO WG, 127, 1-28.

Moffat, C., Beardsley, R. C., Owens, B., & Van Lipzig, N. (2008). A first description of the Antarctic Peninsula Coastal Current. Deep Sea Research Part II: Topical Studies in Oceanography, 55(3-4), 277-293.

Nicholls, K. W., Østerhus, S., Makinson, K., Gammelsrød, T., & Fahrbach, E. (2009). Ice-ocean processes over the continental shelf of the southern Weddell Sea Antarctica: A review. Reviews of Geophysics, 47(3).

Orsi, A. H., & Wiederwohl, C. L. (2009). A recount of Ross Sea waters. Deep Sea Research Part II, 56(13-14), 778-795.

Orsi, A. H., Johnson, G. C., & Bullister, J. L. (1999). Circulation, mixing, and production of Antarctic Bottom Water. Progress in Oceanography, 43(1), 55-109.

Padman, L., Costa, D. P., Bolmer, S. T., Goebel, M. E., Huckstadt, L. A., Jenkins, A., . . . Shoosmith, D. R. (2010). Seals map bathymetry of the Antarctic continental shelf. Geophysical Research Letters, 37(21).

Paolo, F. S., Fricker, H. A., & Padman, L. (2015). Volume loss from Antarctic ice shelves is accelerating. Science, 348(6232), 327-331.

Parkinsona, C. L. (2019). A 40-y record reveals gradual Antarctic sea ice increases followed by decreases at rates far exceeding the rates seen in the Arctic. Proceedings of the National Academy of Sciences, 116(29), 14414-14423.

Porter, F. D., Springer, S. R., Padman, L., Fricker, H. A., Tinto, K. J., Riser, S. C., . . . the ROSETTA-Ice Team. (2019). volution of the Seasonal Surface Mixed Layerof the Ross Sea, Antarctica, ObservedWith Autonomous Profiling Floats. Journal of GeophysicalResearch: Oceans, 124(7), 4934–4953.

Porter, S. E., Parkinson, C. L., & Mosley-Thompson, E. (2016). Bellingshausen Sea ice extent recorded in an Antarctic Peninsula ice core. Journal of Geophysical Research: Atmospheres, 121(23), 13-886.

Post, A. L., Beaman, R. J., O'Brien, P. E., Eléaume, M., & Riddle, M. J. (2011). Community structure and benthic habitats across the George V Shelf, East Antarctica: trends through space and time. Deep Sea Research Part II: Topical Studies in Oceanography, 58(1-2), 105-118.

Rickard, G., & Behrens, E. (2016). CMIP5 Earth system models with biogeochemistry: A Ross Sea assessment. Antarctic Science, 28(5), 327-346.

Rignot, E., Bamber, J. L., Van Den Broeke, M. R., Davis, C., Li, Y., Van De Berg, W. J., & Van Meijgaard, E. (2008). Recent Antarctic ice mass loss from radar interferometry and regional climate modelling. Nature geoscience, 1(2), 106-110.

Rignot, E., Jacobs, S., Mouginot, E., & Scheuchl, B. (2013). Ice-shelf melting around Antarctica. Science, 341(6143), 266-270.

Sallée, J. B., Shuckburgh, E., Bruneau, N., Meijers, A. J., Bracegirdle, T. J., Wang, Z., & Roy, T. (2013). Assessment of Southern Ocean water mass circulation and characteristics in CMIP5 models: Historical bias and forcing response. Journal of Geophysical Research: Oceans, 118(4), 1830-1844.

Schmidtko, S., Heywood, K. J., Thompson, A. F., & Aoki, S. (2014). Multidecadal warming of Antarctic waters. Science, 346(6214), 1227-1231.

Schmidtko, S., Johnson, G. C., & Lyman, J. M. (2013). MIMOC: A global monthly isopycnal upper-ocean climatology. Journal of Geophysical Research: Oceans, 118, 1658–1672.

Shepherd, A., Ivins, E. R., Geruo, A., Barletta, V. R., Bentley, M. J., Bettadpur, S., . . . Horwath, M. (2012). A reconciled estimate of ice-sheet mass balance. Science, 338(6111), 1183-1189.

Shepherd, A., Ivins, E., Rignot, E., Smith, B., van den Broeke, M., Velicogna, I., . . . Wouters, B. (2018). Mass balance of the Antarctic Ice Sheet from 1992 to 2017. Nature(558), 219-222.

Stammerjohn, S., Massom, R., Rind, D., & Martinson, D. (2012). Regions of rapid sea ice change: An inter-hemispheric seasonal comparison. Geophysical Research Letters, 39(6).

Sun, Q., Whitney, M. M., Bryan, F. O., & Tseng, Y.-h. (2017). A box model for representing estuarine physical processes in Earth system models. Ocean Modelling, 112, 139-153.

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

481 Sun, Q., Whitney, M. M., Bryan, F. O., & Tseng, Y.-h. (2019). Assessing the skill of the improved treatment of
482 riverine freshwater in the Community Earth System Model relative to a new salinity climatology. Journal
483 of Advances in Modeling Earth Systems. doi:10.1029/2018MS001349
484 Sutterley, T. C., Velicogna, I., Rignot, E., Mouginot, J., Flament, T., Van Den Broeke, J. M., . . . Reijmer, C. H.
485 (2014). Mass loss of the Amundsen Sea Embayment of West Antarctica from four independent techniques.
486 Geophysical Research Letters, 41(23), 8421-8428.
487 Thorndike, R. L. (1953). Who Belongs in the Family? Psychometrika, 18(4), 267–276.
488 Timmermans, M. L., Proshutinsky, A., Golubeva, E., Jackson, J. M., Krishfield, R., McCall, M., . . . Nishino, S.
489 (2014). Mechanisms of Pacific summer water variability in the Arctic's Central Canada Basin. Journal of
490 Geophysical Research: Oceans, 119(11), 7523-7548.
491 Tseng, Y.-H., Bryan, F. O., & Whitney, M. M. (2016). Impacts of the representation of riverine freshwater input in
492 the community earth system model. Ocean Modelling, 105, 71-86.
493
494

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

Ocean Science
Discussions
Open Access
EGU

Tables and Figures

Table 1 The radius ε used in the DBSCAN for WOA and CESM2

	Group 1	Group 2	Group 3	Group 4	Group 5
WOA	0.05	0.05	0.04	0.03	0.03
CESM2	0.045	0.04	0.06	0.035	0.04

Table 2 The coverage (%) of the majority group of DBSCAN in the total non-outlier data

	Group 1	Group 2	Group 3	Group 4	Group 5
WOA	99.6	97.9	99.9	100	100
CESM2	100	97.3	99.5	99.9	99.7

Table 3 The salinity and temperature standard deviation of WOA (at depth of 500 m if not specified)

		Salinity (g/kg)		Temperature (°C)	
		Amundsen & Bellingshausen	Ross	Amundsen & Bellingshausen	Ross
Geography		0.16 (200 m)	0.11	0.84 (200 m)	1.37
		0.10		1.42	
K-means groups	1	0.10 (200 m)	N/A	0.22 (200 m)	N/A
	2	0.07		1.34	
	3	N/A	0.08	N/A	0.97
	4	0.08		0.44	
	5	N/A	0.10	N/A	0.17

Table 4 The percentage of clustered water in the total ocean volume in the ABRS.

	Group 1	Group 2	Group 3	Group 4	Group 5
WOA	1.0	3.6	21.0	62.1	12.3
CESM2	0.3	7.2	33.2	50.4	8.9

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

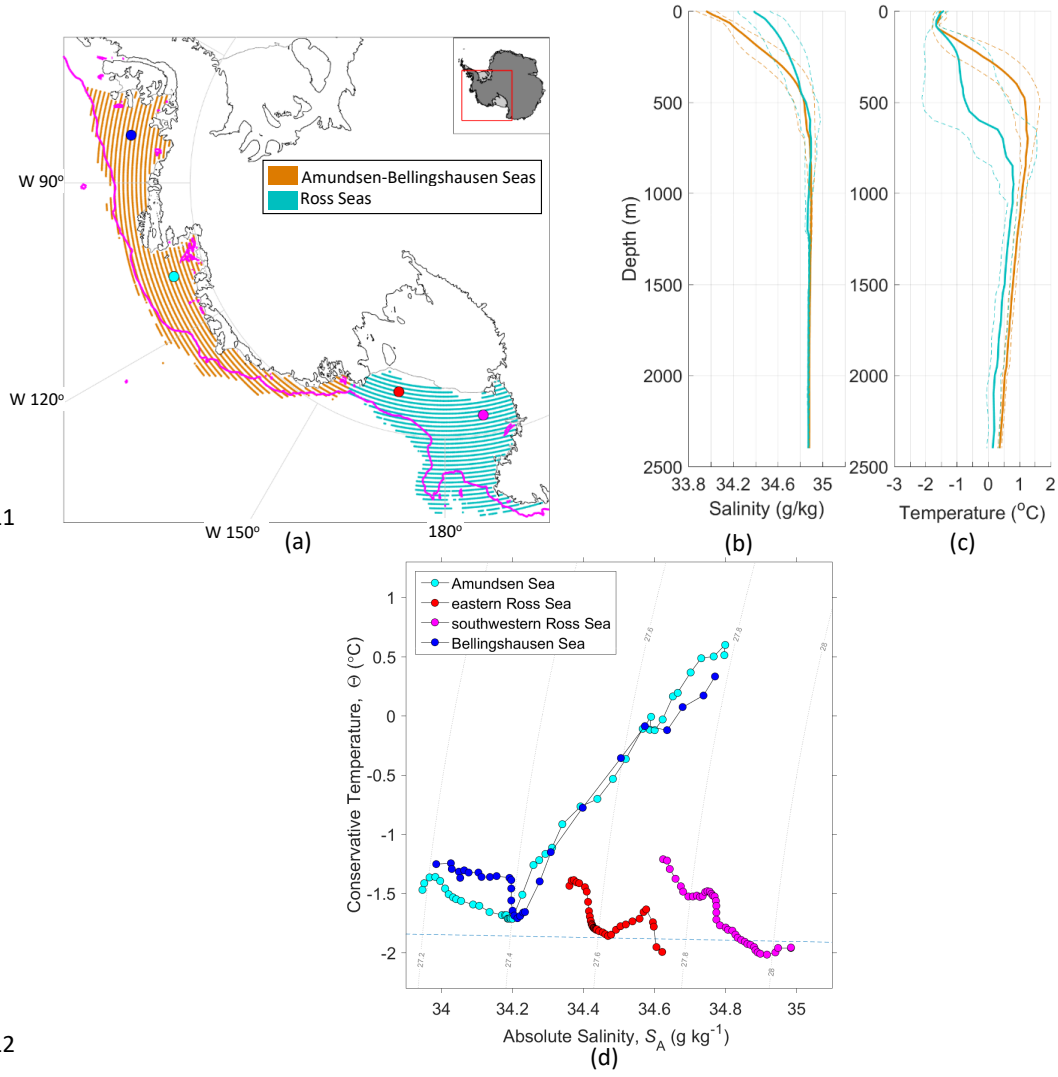


Figure 1 (a) the study domain of Amundsen-Bellingshausen Seas and Ross Sea with bathymetry above 2,500 m. The magenta line indicates the 1,000 m IBCSO depth contour. (b) and (c) show geographically averaged decadal (1995-2004) WOA salinity and temperature profiles in the Amundsen-Bellingshausen Seas (orange; corresponding to the orange stippled region in (a)) and the Ross Sea (cyan; corresponding to the cyan stippled region in (a)). Dashed lines indicate ±1 standard deviation of values at each depth in each region. d) T-S properties of selected water columns (corresponding to colored circles in panel (a)).

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

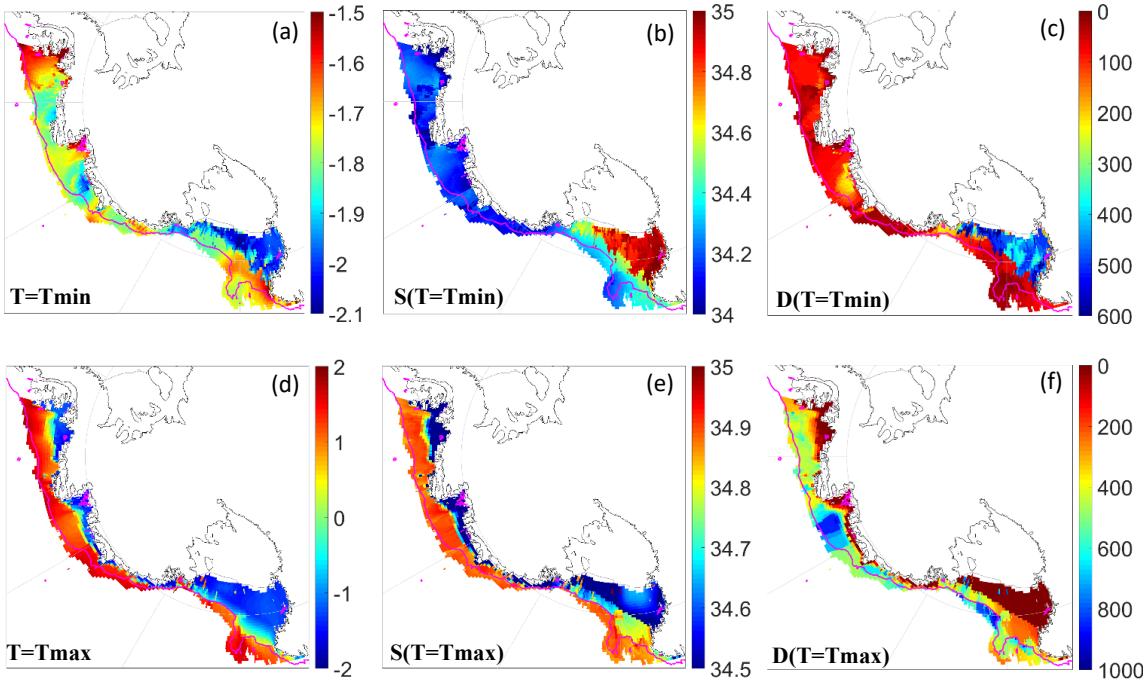


Figure 2 Clustering metrics in WOA. Minimum temperature at each grid point (a), and the salinity (b) and water depth at the minimum temperature. d-f) as a-c, but for quantities at the temperature maximum.

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

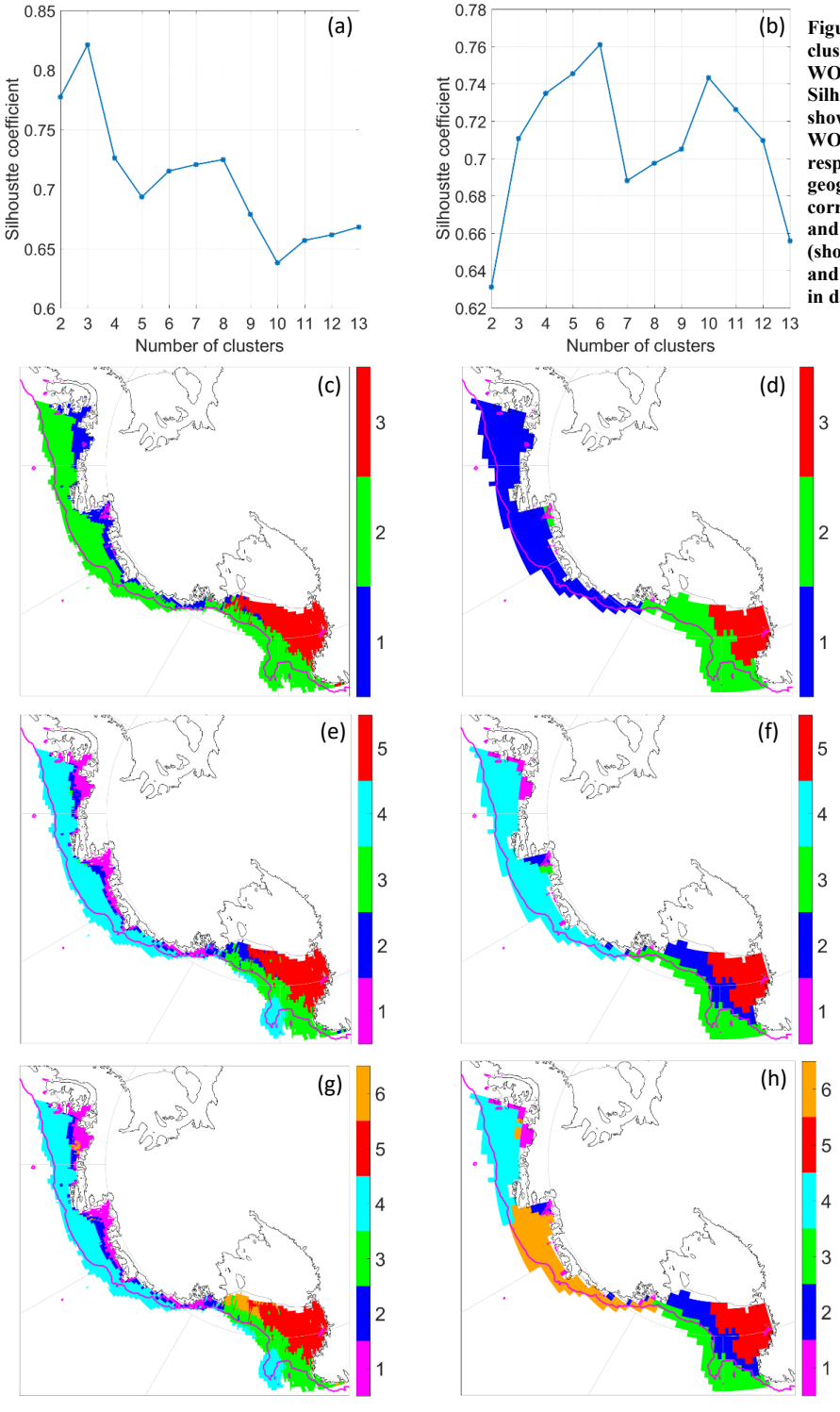


Figure 3 K-means clustering evaluation for WOA and CESM. Silhouette analysis is shown in (a) and (b) for WOA and CESM, respectively. The geographic regions corresponding to 3, 5 and 6 groups for WOA, (shown in c, e and g). and for CESM (shown in d, f and h).

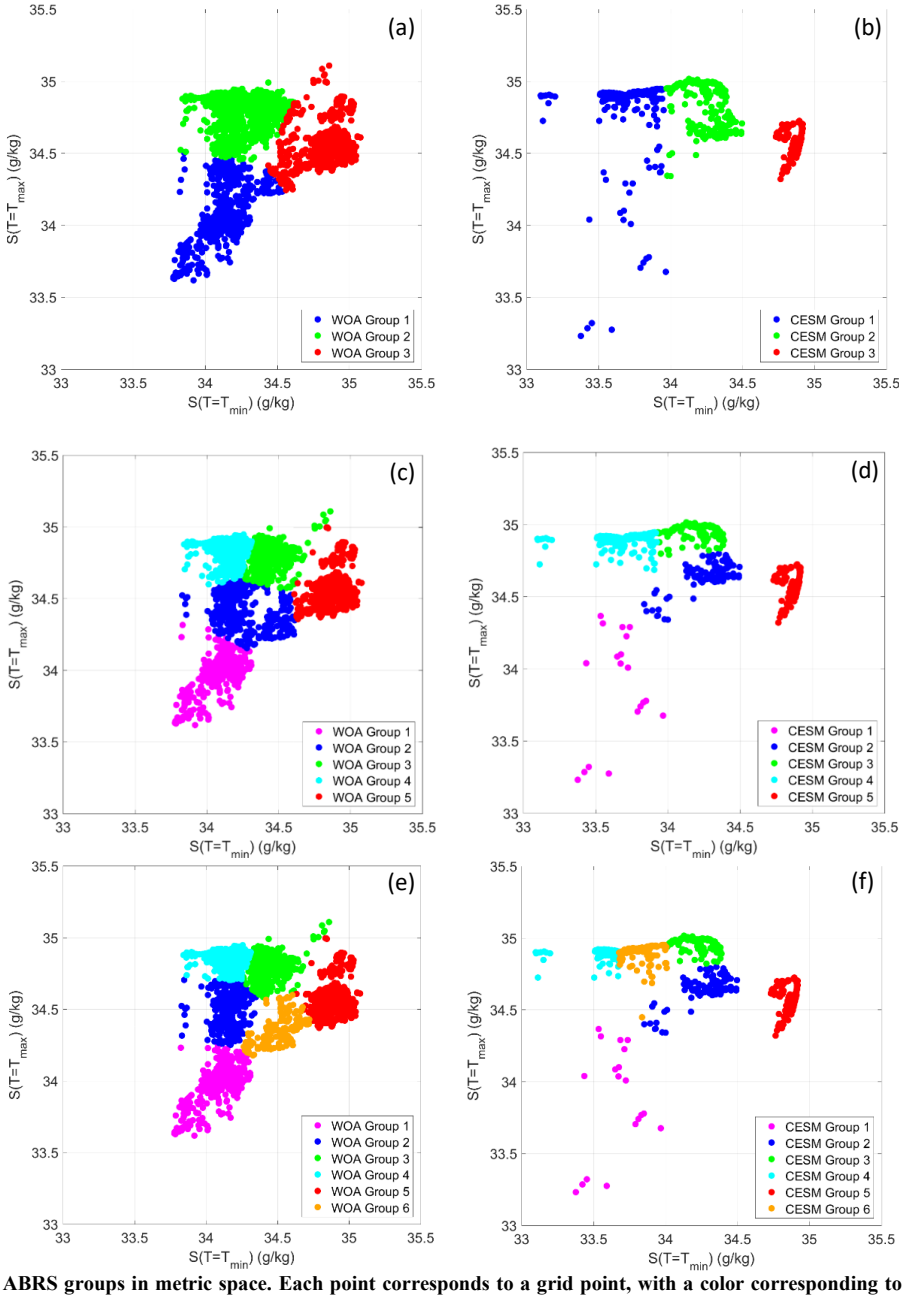
https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

Ocean Science
Discussions
Open Access
EGU

560

561

562

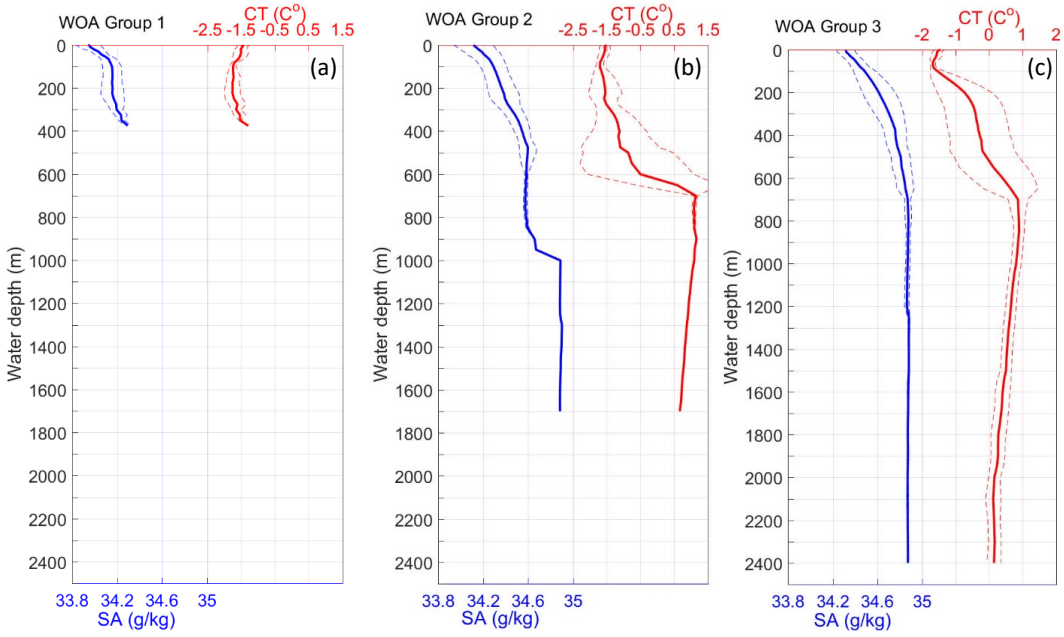


563
564 **Figure 4. ABRS groups in metric space. Each point corresponds to a grid point, with a color corresponding to its group**
565 **number, for K=3, 5 and 6, for WOA, (shown in a, c and e). and for CESM2 (shown in b, d and f).**

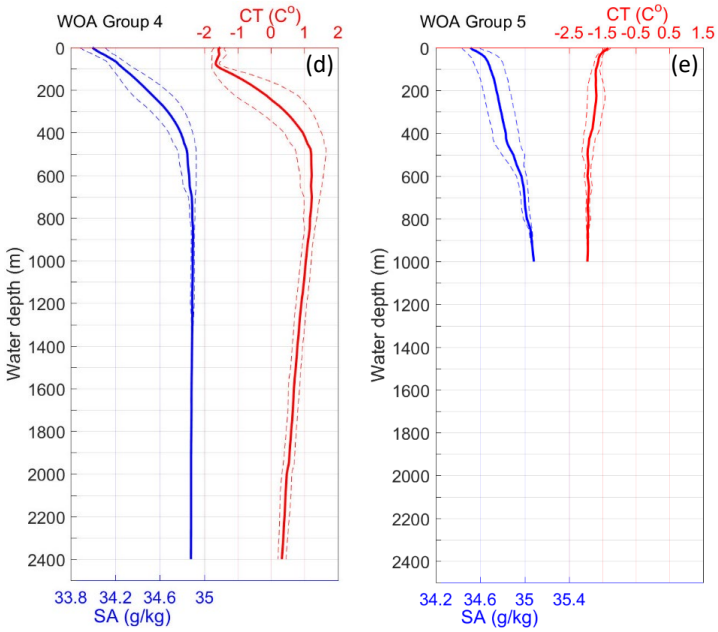
566
567

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

568



569



570

571 **Figure 5 Mean (solid lines) WOA salinity (in blue) and temperature (in red) profiles for five groups (from a to e) shown in**
572 **Figure 3 (e). ±1 standard deviation at each depth is shown with dashed lines.**

573

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

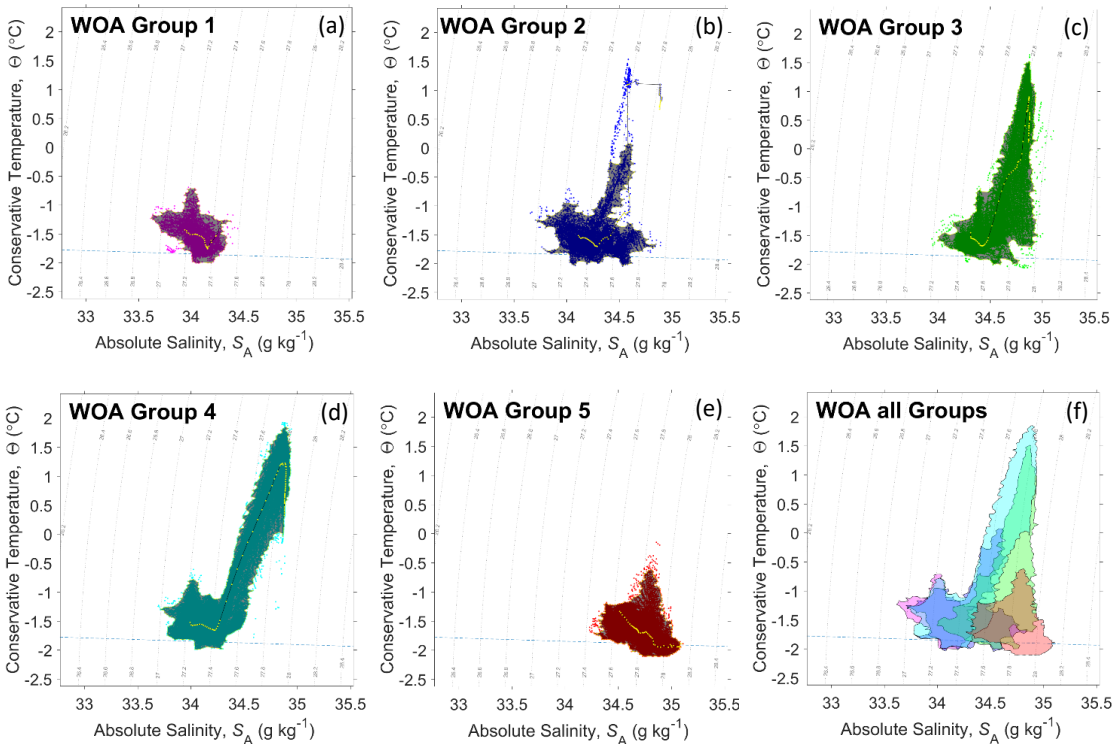


Figure 6 *T-S* properties for the five WOA groups (from a to e) shown in Figure 3 (e). The yellow dotted lines show the profile of mean temperature and salinity in each group, and the dark shaded areas are the cores of water property from the density-based clustering results. The cores of all five groups are overlaid on the same plot in (f).

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

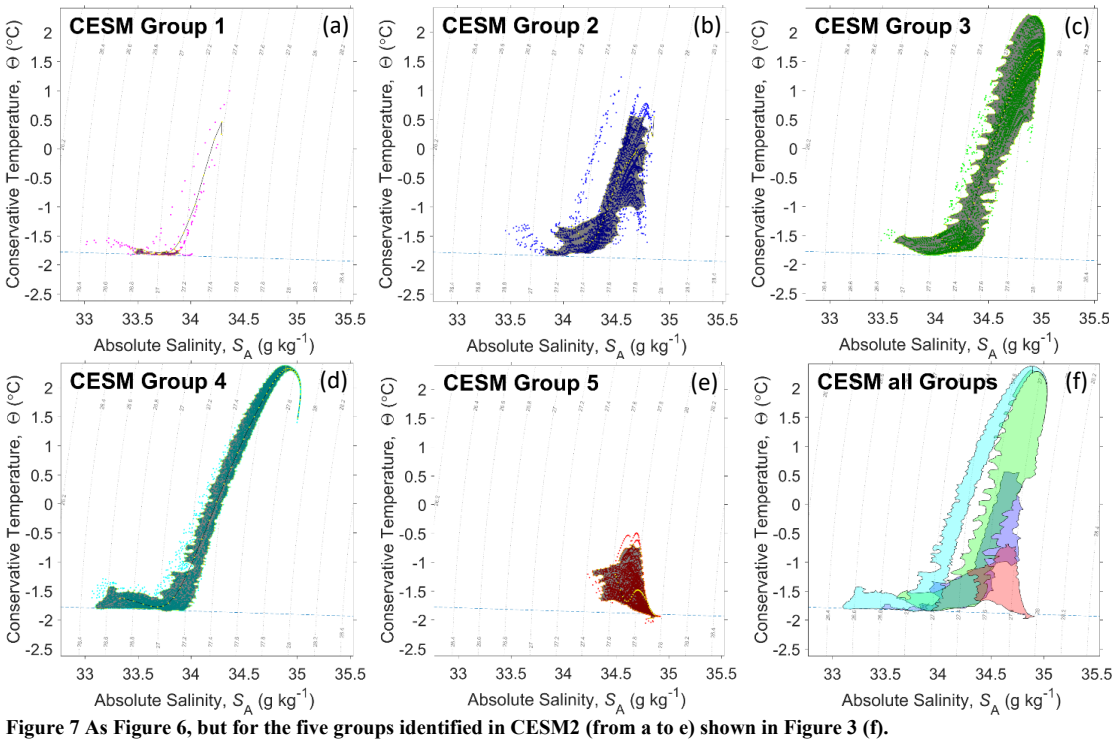
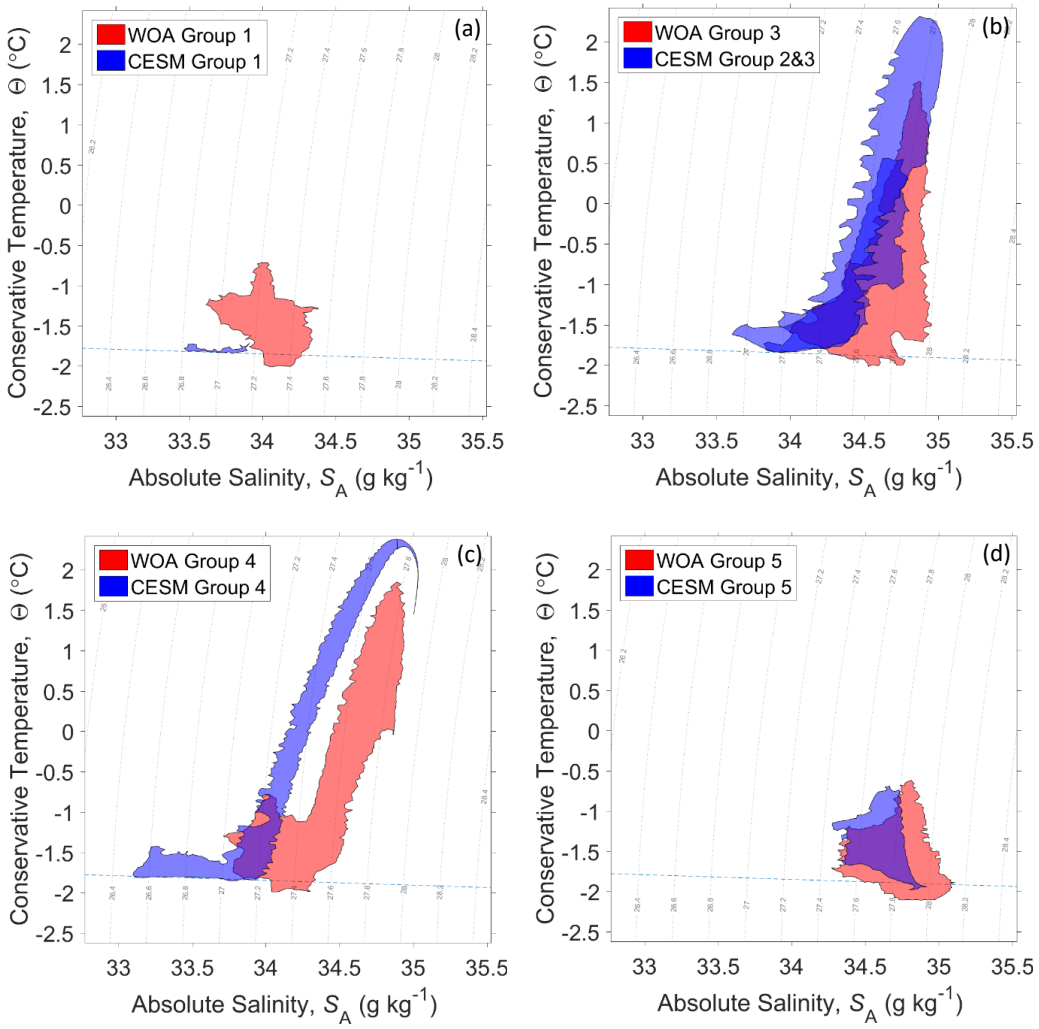


Figure 7 As Figure 6, but for the five groups identified in CESM2 (from a to e) shown in Figure 3 (f).

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

Ocean Science
Discussions
Open Access
EGU

632



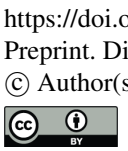
633

634

635 **Figure 8 The properties of core water masses in WOA (red) and CESM2 (blue). Note that groups 2 and 3 have been**
636 **combined for CESM2.**

637
638
639
640
641
642
643
644
645
646
647

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020



Ocean Science
Discussions
Open Access
EGU

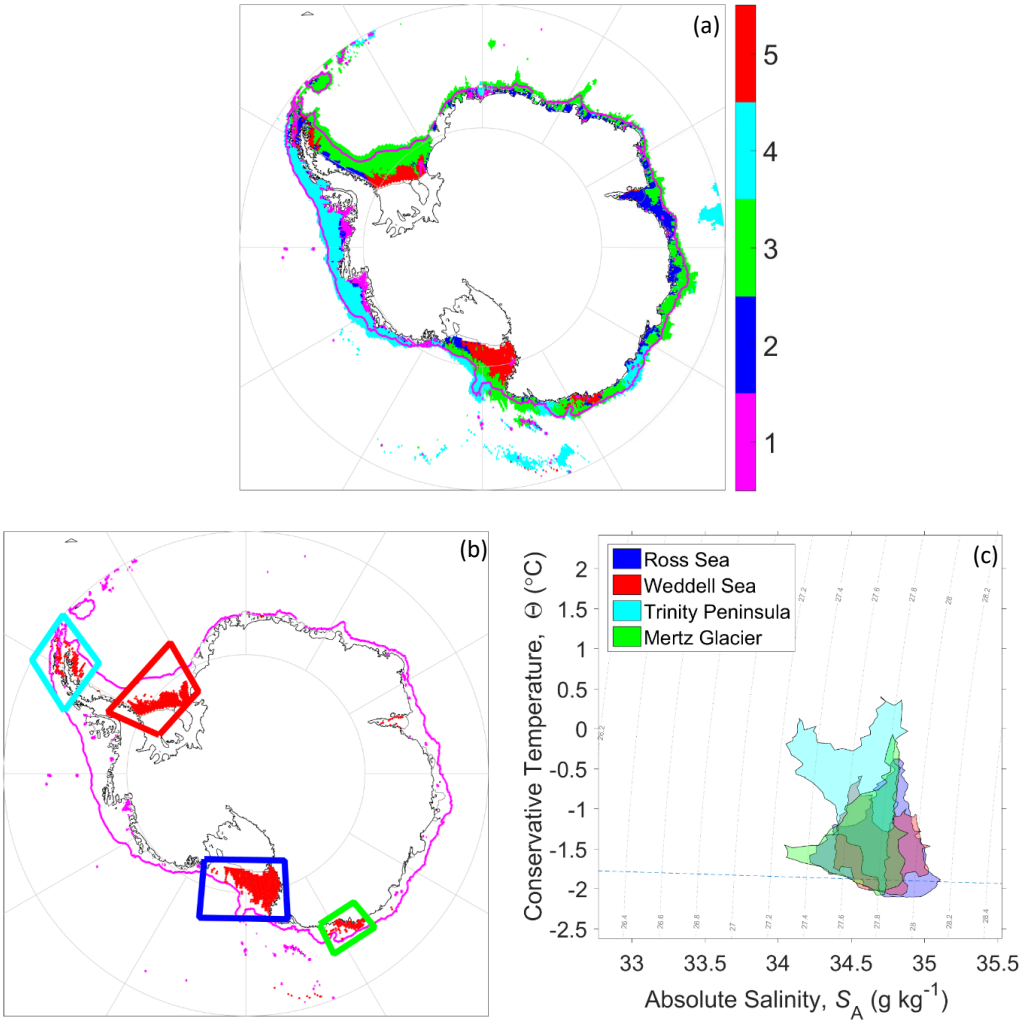


Figure 9 (a) WOA-based groups on the entire ACSS (same color code as Figure 3e). (b) Four places are identified as HSSW regime with color codes blue: southwestern Ross Sea; red: Weddell Sea near the Filchner-Ronne Ice Shelf, the George V Coast; cyan: Bransfield Strait and south of Trinity Peninsula; and green: Mertz Glacier tongue. (c)*T-S* properties of group 5 (HSSW) regions, with their geographic location and color code matched in (b).

https://doi.org/10.5194/os-2020-51
Preprint. Discussion started: 4 June 2020

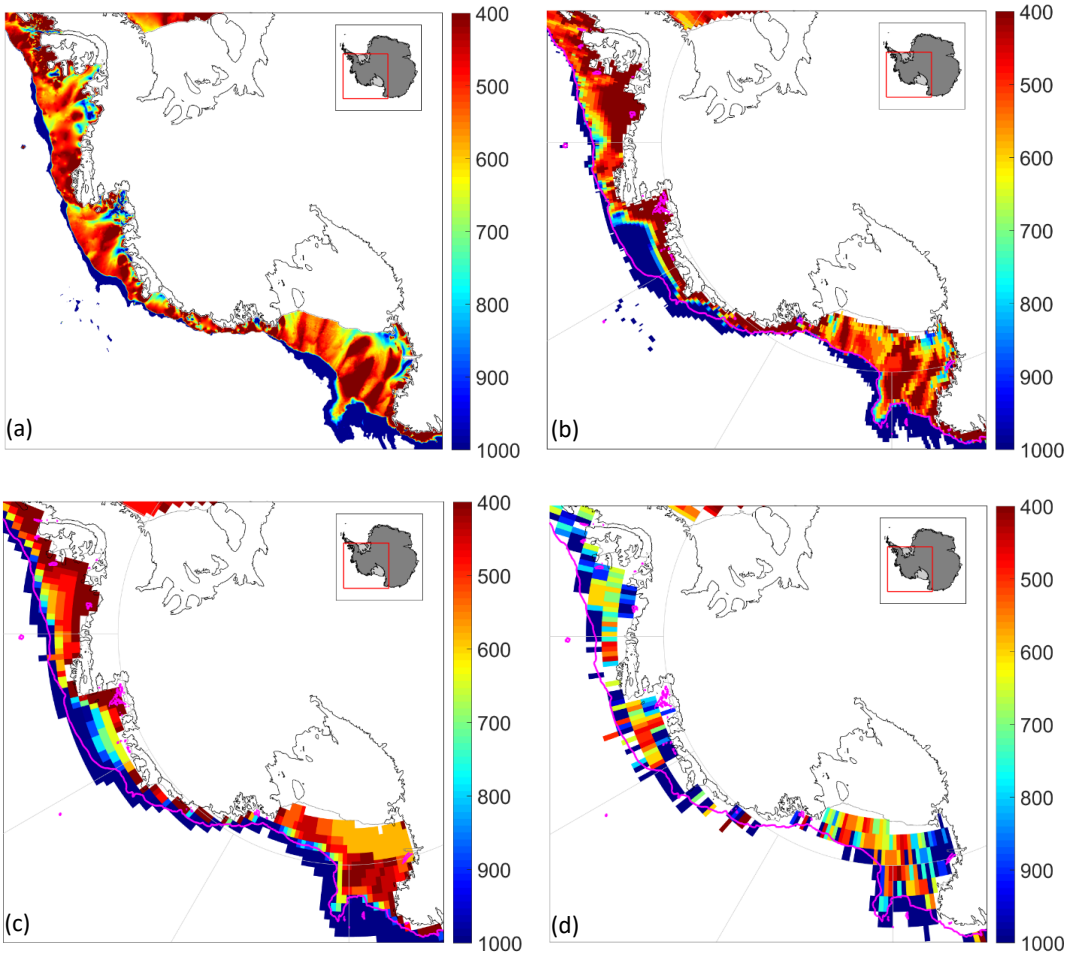


Figure 10 Bathymetry between 400 m and 1,000 m of: (a) IBCSO (500 m horizontal resolution), (b) WOA (0.25-degree horizontal resolution), (c) CESM2 (1 x 0.5-degree lon/lat resolution), and (d) the WOD bin-averaged into 1-degree horizontal resolution with all types of instrument with temperature measurements). The magenta line indicates the 1,000m IBCSO depth contour.