Identifying Southern Ocean fronts using unsupervised classification and edge detection

Simon Thomas\textsuperscript{1}, Dan Jones\textsuperscript{2}, Anita Faul\textsuperscript{2}, Erik Mackie\textsuperscript{1}, and Etienne Pauthenet\textsuperscript{3}

\textsuperscript{1}University of Cambridge
\textsuperscript{2}NERC British Antarctic Survey
\textsuperscript{3}Sorbonne Université, LOCEAN-IPSL, CNRS/IRD/MNHN

January 20, 2023

Abstract

Fronts are ubiquitous in the climate system. In the Southern Ocean, fronts delineate water masses, which correspond to upwelling and downwelling branches of the overturning circulation. A robust understanding of Southern Ocean fronts is key to projecting future changes in overturning and the associated air-sea partitioning of heat and carbon. Classically, oceanographers define Southern Ocean fronts as a small number of continuous linear features that encircle Antarctica. However, modern observational and theoretical developments are challenging this traditional framework to accommodate more localized views of fronts [Chapman et al. 2020]. In this work, we present two related methods for calculating fronts from oceanographic data. The first method uses unsupervised classification (specifically, Gaussian Mixture Modeling or GMM) and an interclass metric to define fronts. This approach produces a discontinuous, probabilistic view of front location, emphasizing the fact that the boundaries between water masses are not uniformly sharp across the entire Southern Ocean. The second method uses Sobel edge detection to highlight rapid changes [Hjelmervik & Hjelmervik, 2019]. This approach produces a more local view of fronts, with the advantage that it can highlight the movement of individual eddy-like features (such as the Agulhas rings). The fronts detected using the Sobel method are moderately correlated with the magnitude of the velocity field, which is consistent with the theoretically expected spatial coincidence of fronts and jets. We will present our python GitHub repository, which will allow researchers to easily apply these methods to their own datasets. Figure caption Two methods for interpretable front detection. Solid lines represent classical fronts. (a) The “inter-class” metric, which indicates the probability that a grid cell is a boundary between two classes. The classes are defined by GMM of principal component values (PCs) derived from both temperature and salinity. The different colors indicate different class boundaries. (b) Sobel edge detection: approximately the magnitude of the spatial gradient of the PCs divided by each field’s standard deviation, which highlights locations of rapid change.
Identifying Southern Ocean fronts using unsupervised classification and edge detection

AGU FALL MEETING
Online Everywhere | 1–17 December 2020

Simon Thomas¹, ²; †, Daniel Jones¹, Anita Faul¹,
Erik Mackie¹, ³, Etienne Pauthenet⁴

¹British Antarctic Survey, Cambridge; ²Department of Physics, University of Cambridge; ³Cambridge Zero, University of Cambridge;
⁴Sorbonne Universités, UPMC Université, L’OCEAN-IPSL, Paris.

†sithom@bas.ac.uk

British Antarctic Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

UNIVERSITY OF CAMBRIDGE

LOCEAN
Laboratoire d’Océanographie et du Climat
Expérimentations et Approches Numériques
Global sea surface temperatures and currents

Credit: NASA
Chapman et al. 2020 challenged traditional views of fronts.

The deep reaching jets of the Antarctic Circumpolar Current are expected to correspond to sharp gradients of temperature and salinity.
<table>
<thead>
<tr>
<th></th>
<th>Global</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Methods</strong></td>
<td>Contours, Water mass criteria, GMM</td>
<td>Gradient thresholding, Sobel Edge, Skewness</td>
</tr>
<tr>
<td><strong>Pro</strong></td>
<td>Interpretable</td>
<td>Easy to define</td>
</tr>
<tr>
<td><strong>Con</strong></td>
<td>Hard to define</td>
<td>Hard to interpret</td>
</tr>
</tbody>
</table>

Table: A summary of Table 1 in Chapman et al. 2020.
Dataset & Features: BSOSE-i106

- BSOSE-i106 state estimate (Verdy & Mazloff 2017).
- Monthly mean $T$ and $S$ profiles between 300m–2000m (Rosso et al. 2020).
Preprocessing: Principal Component Analysis

- Take coefficients from 3 principal components (North et al. 1982; Pauthenet et al. 2017).

Explained Variance:

- PC1: 75%
- PC2: 16%
- PC3: 7%
Global Model: Gaussian Mixture Modelling

- Using GMM (Maze et al. 2017; Jones et al. 2019).

- $K$ clusters.

- Posterior probabilities:

$$
P (c_n = c_k) = \frac{\lambda_k \mathcal{N} (\vec{x}_n ; \vec{\mu}_k, \Sigma_k)}{\sum_{k=1}^{K} \lambda_k \mathcal{N} (\vec{x}_n ; \vec{\mu}_k, \Sigma_k)}$$  \hspace{1cm} (1)

- We define (thanks to A.F.):

$$\mathcal{I} (\vec{x}_n) = 1 - \left( P (c = c_k)_{\text{max}} - P (c = c_l)_{\text{runner-up}} \right),$$  \hspace{1cm} (2)

where $\vec{x}_n$ is the $n^{th}$ profile's principal component values.
Figure: GMM and $\mathcal{I}$-metric with $K = 5$ in 2PC space.
Figure: $\mathcal{I}$ of GMM with $K = 5$. 
Global: \( I \)-metric

Figure: \( I \) of GMM with \( K = 5 \).
Local: Sobel Edge Detection

- We build on a Sobel edge detection method (Hjelmervik & Hjelmervik 2019) on the 3 PC-fields.
- Approximately the smoothed PC 2D gradient.

Figure: Sobel edge fronts in latitude direction.
Conclusion

- We propose a probabilistic metric for defining water mass boundaries. Use a depth range, not surface information.

- This expands on the classical, time-averaged view of fronts and water mass boundaries, following Chapman *et al.* 2020.

- Paper and a GitHub repository in prep.


