

1 Objectively classified dynamics: A global assessment of barotropic vorticity

Maike Sonnewald*, Carl Wunsch

³ Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA and
⁴ Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA.

Patrick Heimbach

*The University of Texas at Austin, 201 East 24th Street, Austin, TX 78712, USA and
Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA*

⁸ *Corresponding author address: Maike Sonnewald, Massachusetts Institute of Technology, 77
⁹ Massachusetts Ave, Cambridge, MA 02139, USA.

¹⁰ E-mail: maike.s@mit.edu

ABSTRACT

11 The closure of the barotropic vorticity budget is demonstrated in the
12 ECCOv4r2 State Estimate for the 1992-2013 average, with a residual of the
13 order 10^{-12} , three orders of magnitude smaller than the stress (surface and
14 bottom), Coriolis, bottom pressure torque, non-linear and viscous terms. The
15 bottom pressure torque term is largely seen as the remainder of the Coriolis
16 and stress terms, with notable exceptions where the non-linear terms are
17 significant. The objective K-Means clustering algorithm is used to search for
18 dynamical patterns in the global ocean based on the terms in the barotropic
19 vorticity budget. We find large coherent regions, confirming the classification
20 technique is a robust identification method of dynamical regimes. In 56%
21 of the world ocean, the 20 year average is quite stable. The surface and
22 bottom stress terms are balanced by the bottom pressure torque and the
23 non-linear terms. The subtropical gyres cover 11% of the global ocean,
24 and are characterized by a “quasi-Sverdrupian” regime where the barotropic
25 vorticity input by the Coriolis term is balanced by the wind and bottom stress
26 term. The subpolar gyre is split into two, with the northern extension being
27 an amplified version of the southern. An almost identical regime is seen in
28 large areas of the Southern Ocean, but opposite in sign. The shift from the
29 subpolar to subtropical gyres is explained in terms of a shift from strong
30 interior flow to one dominated by the Western Boundary Currents. The $\approx 6\%$
31 of the ocean in the remaining area is strongly influenced by the non-linear
32 contribution to the barotropic vorticity, and found in key areas in the Southern
33 ocean and along the Western Boundaries. These areas are not well captured
34 and characterized by the barotropic vorticity and merit further study.

³⁶ **1. Motivation**

³⁷ The global ocean is characterized in terms of dynamical regions using the depth-integrated,
³⁸ barotropic vorticity (BV) equation, and the adequacy of this simplified view of the ocean is
³⁹ assessed for decadal timescales. Characterizing the global ocean and reducing the overwhelming
⁴⁰ dimensionality of observed dynamics has a long history in oceanography. Stommel (1948) and
⁴¹ Munk (1950) assume the topography is not key in the subtropics, while topography has been
⁴² key to Southern Ocean dynamics Munk and Palmén (1951). General Circulation Models (GCM)
⁴³ solve the primitive equations in full detail, and using this the success of simple characterizations
⁴⁴ of ocean features is assessed. The existence of dynamical regimes would indicate what circulation
⁴⁵ features are key, and what forces are active to allow their emergence. With this, the need for the
⁴⁶ complicated GCM approach can be assessed in terms of the success of a reduced approach.

⁴⁷

⁴⁸ The success of work such as Munk (1950) relies on a depth-integrated view, where lateral
⁴⁹ friction provides closure in what is often referred to as Munk boundary layer structure for Western
⁵⁰ Boundary Currents. A quasi-Sverdrupian framework is invoked, where the depth-integrated
⁵¹ meridional flow is given by the local wind stress curl. The potential success of such frame-
⁵² works suggests predictive skill based purely on the wind field, and suggests that a barotropic,
⁵³ depth-integrated view, is appropriate to characterize the subtropical ocean. GCM studies and
⁵⁴ observational work present a more intricate ocean, where baroclinic structure and subsequent
⁵⁵ non-linearities are key to ocean dynamics. If this is the case for large regions of the ocean, or
⁵⁶ if they govern key dynamics on relevant timescales, the simplified theoretical view would be
⁵⁷ inadequate.

⁵⁸

59 The presence of dynamical regimes is suggested by the zonal nature of the wind stress forcing
60 and the geometry of the ocean basins with underlying topography. Classifying and identifying
61 regions in the world ocean is done using the BV equation, closing the global budget. Yeager
62 (2015); Schoonover *et al.* (2016); Le Bras *et al.* (in review) have assessed the dynamics of the
63 BV budget focusing on the North Atlantic. The link to the overturning and gyre circulation, as
64 well as to the importance of mechanisms such as buoyancy forcing have been assessed. A global
65 view allows the interpretation dynamics in a global sense, which we see illuminates similarities in
66 governing dynamics.

67

68 Objective classification from machine learning has the potential to powerfully simplify the inter-
69 pretation of ocean dynamics. The objective aspect allows an unbiased assessment of the success
70 of the theoretical framework used, and the dimensionality reduction offered by using a classifier is
71 an exciting and novel approach in physical oceanography. Ardyna *et al.* (2017) applied a similar
72 method to identify regions with distinct biological activity, and Liang *et al.* (in review) used K-
73 Means clustering to identify key regions for data collection to build maps of Nitrate in the Southern
74 Ocean.

75 **2. Barotropic Vorticity in the ECCOv4 State Estimate and K-Means clustering**

76 The BV equation is used to assess the extent to which we can close this budget in the EC-
77 COv4 global bi-decadal state estimate, as described by Wunsch and Heimbach (2013); Forget *et*
78 *al.* (2015) and others (see also ECCO Consortium (2017a,b)). The state estimate is global 1° with
79 tropical mesh refinement. A least-squares with Lagrange multipliers approach is used to obtain
80 the state estimate. The result is an adjusted, yet *free-running* version of the MIT General Circula-
81 tion Model (MITgcm, Adcroft *et al.* (2004)) solving the primitive equations. In contrast to most

82 “reanalysis” products, the ECCO oceanic state satisfies basic conservation laws for enthalpy, salt,
 83 volume, and momentum remaining largely within error estimates of a diverse set of global data
 84 (Wunsch and Heimbach 2007, 2013; Stammer *et al.* 2016). Regions without data are filled in a
 85 dynamically consistent way using the dynamics, avoiding the use of untested statistical hypotheses
 86 e.g., Reynolds *et al.* (2013).

87 To classify the ocean’s dynamical regions the BV equation is used, based on the solutions to
 88 the primitive equations from the ECCOv4 State Estimate. Illustrating the simplification, take the
 89 momentum equations of the ocean as a thin fluid on a rotating sphere:

$$\partial_t \mathbf{u} + f \mathbf{k} \times \mathbf{u} = -\frac{1}{\rho_0} \nabla p + \frac{1}{\rho_0} \partial_z \tau + \mathbf{a} + \mathbf{b}, \partial_z p = -g\rho, \nabla \cdot \mathbf{v} = 0. \quad (1)$$

90 The pressure, acceleration due to gravity, density and vertical shear stress are denoted p , g , ρ
 91 and τ respectively, with ρ_0 the reference density; the three dimensional velocity field $\mathbf{v} = (\mathbf{u},$
 92 $\mathbf{v}, \mathbf{w}) = (\mathbf{u}, \mathbf{w})$; the unit vector is denoted \mathbf{k} ; planetary vorticity as a function of latitude ϕ in
 93 $f\mathbf{k} = (0, 0, 2\Omega \sin \phi)$; the viscous forcing by vertical shear is denoted $\partial_z \tau$; the non-linear terms
 94 are \mathbf{a} and the horizontal viscous forcing \mathbf{b} . Assuming steady state, the vertical integral from
 95 $z = \eta(x, y, t)$ to $z = H(x, y)$ gives

$$\beta V = \frac{1}{\rho_0} \nabla p_b \times \nabla H + \frac{1}{\rho_0} \nabla \times \tau + \nabla \times \mathbf{A} + \nabla \times \mathbf{B}, \quad (2)$$

96 where $\nabla \cdot \mathbf{U} = 0, \mathbf{U} \cdot \nabla f = \beta V$, the bottom pressure is denoted p_b , $\mathbf{A} = \int_H^\eta \mathbf{a} dz$ and $\mathbf{B} = \int_H^\eta \mathbf{b} dz$.
 97 The LHS of equation 2 is the planetary vorticity advective term, while the RHS of equation 2 is
 98 the bottom pressure torque, the wind and bottom stress curl, the non-linear torque and the viscous
 99 torque, respectively. The non-linear torque is composed of three terms:

$$\nabla \times \mathbf{A} = \nabla \times \left[\int_{-H}^0 \nabla \cdot (\mathbf{u}\mathbf{u}) dz \right] + [w\zeta]_{z=H}^{z=0} + [\nabla w \times \mathbf{u}]_{z=H}^{z=0}. \quad (3)$$

100 The RHS on equation 3 represents the curl of the vertically integrated momentum flux divergence,
 101 the non-linear contribution to vortex tube stretching and the conversion of vertical shear to
 102 barotropic vorticity. The horizontal viscous forcing includes that induced by subgrid-scale
 103 parameterizations. We work on the twenty year averaged fields of the BV equation after a
 104 laplacian smoother is applied.

105

106 Assessing the presence of dominant global patterns in the BV equation, the K-Means clustering
 107 method determines if there are patterns in the variability. Patterns of variability constitute
 108 areas in the hyperspace given by the terms of the BV equation that are more densely populated
 109 than others. The clustering of configurations of terms and associated variability is identified by
 110 the K-Means method. Referred to as an unsupervised learning algorithm (MacQueen 1967), a
 111 minimization of the sum of squares (i.e variance) given by the partitioning of the hyperspace is
 112 iteratively approached:

$$J = \sum_{j=1}^K \sum_{i=1}^n \|x_i^j - c_j\|^2 \quad (4)$$

113 where the number of K clusters (c_j) is fixed a priori, randomly scattered among the n data
 114 points (x). The distance between a data point x_i^j and c_j is given by $\|x_i^j - c_j\|^2$. Initially, each
 115 data point is associated with the closest K cluster. The position of c_j is recalculated, and the
 116 association reassessed. This is iteratively repeated. The problem is computationally NP-hard, and
 117 sensitive to the initialization and choice of K (see appendix A1). From visual inspection, and
 118 using the Akaike and Bayesian Information Criteria (AIC and BIC), clusters have robust stable

¹¹⁹ geographical extents with $K > 35$. The AIC and BIC indicate robust asymptotic regimes, and we
¹²⁰ use $K=50$ for the remaining analysis. The data was scaled to have zero mean and unit variance.

¹²¹

¹²² 3. Results

¹²³ Figure 1 illustrates the closure in ECCOv4r2 for the 20 year average of the BV budget
¹²⁴ described in equation 2. The residual is smaller than $\pm 10^{-12}$, well within an acceptable range
¹²⁵ of closure compared to a magnitude $\pm 10^{-9}$ of terms in the BV equation. On this basis, we
¹²⁶ proceed with further analysis with confidence. Some numerical issues on the shelf/shallow water
¹²⁷ have larger residuals, but these only amount to 2.9% of the global ocean and are attributed to noise.

¹²⁸

¹²⁹ Figure 1b illustrates the Coriolis term in ECCOv4r2 for the 20 year average from equation 2.
¹³⁰ This is balanced well by the wind and bottom stress BV terms shown in Figure 1d (the bottom
¹³¹ stress term is small) and bottom pressure torque shown in Figure 1c. The remainder is largely
¹³² found in the non-linear BV contributions seen in Figure 1e, with the lateral viscous dissipation
¹³³ largely being an order of magnitude smaller, apart from in localized regions in the Southern Ocean.

¹³⁴

¹³⁵ The wind and bottom stress BV term in Figure 1d are largely zonally symmetric in the 20
¹³⁶ year average, with large patterns of negative BV to the south in the Southern Ocean, large gyre
¹³⁷ patterns visible in the Pacific and Atlantic basins. The bottom pressure torque term in Figure
¹³⁸ 1c is associated with interactions with steep bathymetry, e.g. in the Southern Ocean we see a
¹³⁹ large positive patch leading up to the Antarctic-Pacific ridge, and a negative patch following.
¹⁴⁰ This is consistent with vortex stretching as the ridge is crossed, and we see a similar feature
¹⁴¹ associated with the Mid-Atlantic ridge in the Atlantic sector of the Southern Ocean. Along

142 Western Boundaries, the BPT is positive to the west and negative just adjacent to the east,
143 consistent with studies such as Myers *et al.* (1996). The BV of the non-linear terms is similarly
144 concentrated along the western edge of basins where western boundary currents are found, but
145 is less spatially coherent than the BPT term. The Southern Ocean stands out as a region of high
146 activity, particularly the Atlantic sector. The contribution of the lateral viscous dissipation to the
147 BV is comparatively small, but also concentrated in the Southern Ocean.

148

149 Figure 4 illustrates the spatial extent of the dynamical regimes picked out by the K-Means
150 algorithm. The numbering is arbitrary, and the variability is mainly found in a handful of
151 geographically coherent patches. There is a zonal pattern suggesting the wind stress contribution
152 to the BV is key, but also much spatial structure in the Southern Ocean and along the western
153 boundaries. Figures 6 and 7 isolate the geographical area and the associated area averaged
154 magnitudes of the terms in the BV equation, ordered in decreasing percentage coverage of the
155 global ocean. Table 1 summarizes the area covered, corresponding figure and dominant terms in
156 the BV equation found in the dominant clusters.

157

158 Figure 6a illustrates the reasonable quiet ocean covering 56% of the global ocean. This is found
159 in zonal streaks in the Tropics, a thin ribbon in the Southern Ocean notably in the Pacific sector
160 where the BPT and non-linear terms are small. In the Northern Hemisphere, the areas outline the
161 edges of the subtropical and subpolar gyres. Large areas of the Arctic are also covered. Figure
162 ?? demonstrates the balance of terms that made the K-Means algorithm pick out these regions.
163 The wind and bottom stress term is the major source of negative depth integrated vorticity, while
164 the BPT adds positive vorticity. The non-linear terms add negative vorticity. Figure 6c illustrates
165 the spatial region covering the next dynamical region that covers 18% of the ocean area. In

the Northern Hemisphere, these cover the southern region of the subpolar gyres. There is a zonal streak across the equator in both the Atlantic and Pacific, absent in the Indian Ocean. The Southern Hemisphere show large expanses in both the Pacific and Atlantic, but the are covered in the Indian ocean is limited. The bar chart in Figure 6d highlight the wind as the major source of barotropic vorticity, with sinks in the Coriolis term and BPT. There is a small sink in the non-linear terms. Figure 6e illustrates the 11% of the ocean area selected by the next cluster. The subtropical gyres in the Northern Hemisphere Atlantic and Pacific stand out, together with thin streaks on the Equator. Isolated streaks are seen in the Southern Ocean, and a large area of the Southern Hemisphere Indian Ocean in the tropics. Figure 6f illustrates that these areas have a clear balance between the input of barotropic vorticity from the Coriolis term and a sink in the wind and bottom stress. This corresponds to a quasi-Sverdrupian regime in that we look at the depth integrated ocean. Figure 7a shows the area covered by the cluster covering 7% of the ocean. This is largely a complimentary poleward extension to the cluster covering 18% of the ocean seen in Figure 6c. In the Northern Hemisphere, the cluster largely represents the Northern edge of the subpolar gyre. In the Southern Hemisphere, the cluster is found on the Eastern edge of the Pacific and Atlantic basins, just to the south and flaring out westwards of the continental barrier. In the Indian Ocean, this barrier could be seen to be New Zealand or Australia, and the area filled by this dynamical regime fills the subtropical Indian Ocean down to the border with the Southern Ocean. This regime is absent from the Southern Ocean. The bar chart in Figure 7b illustrates that it is an amplified version of the barcharts seen in figure 6d, being an order of magnitude larger, but still having the wind as the major source of barotropic vorticity, with sinks in the Coriolis term and BPT. There is a small source in the non-linear terms. The Southern Ocean is better represented in the area covering 2% of the global world ocean seen in Figure 15a. This is seen mainly in a series of streaks in the Southern Ocean, mainly staying around 60°S where there is

190 no continental block. Isolated areas are also seen in the Northern Hemisphere. The bar charts in
191 Figure 15b illustrate that these are practically an inverse of the barcharts for the area representing
192 the Northern extension of the subtropical gyres covering 7% of the world ocean. The wind as the
193 major sink of barotropic vorticity, with sources in the Coriolis term and BPT. There is a small sink
194 in the non-linear terms.

195

196 The final map in Figure 7e is a summary of the areas terms that account for > 0.2%, these sum
197 to 6.3% of the world ocean, and individual plots can be seen in appendix A1. These are mainly
198 found in the Southern Ocean and along the Western Boundaries. Areas of rough bathymetry
199 stand out, such as the Pacific-Antarctic Ridge and the Drake Passage area. The bar chart shown in
200 Figure 7f illustrates that the non-linear contribution to the barotropic vorticity stands out, together
201 with the Coriolis term. The different constituents are quite varied, but strong contributions from
202 the non-linear terms are consistent. Unpicking these specifically is outside the current scope, but
203 will be discussed in a following paper.

204

205 **4. Discussion and Conclusions**

206 The present t BV equation is found to be a reasonable approximation for large areas of the
207 world ocean. Figure 1 demonstrated that the global ocean has large coherent regions, but that the
208 dynamical balances implicit in the 20 year average are more nuanced than a balance between the
209 wind stress, Coriolis and BPT terms. Areas where the non-linear terms are small suggest that
210 the BV is a good approximation to what the ECCOv4r2 State Estimate sees with the primitive
211 equations, but some areas are influenced by the non-linear terms, suggesting that the depth
212 integrated view is not adequate. Areas where the non-linear terms are important dominate in the

213 western boundaries of continents, as well as the Southern Ocean where the ACC interacts with
214 bathymetric obstacles.

215

216 Yeager (2015) describes the North Atlantic circulation in terms of its BV equation breakdown.
217 Following this, we interpret the regions in the North Atlantic as a shift where strong interior flow
218 (baroclinic meridional, North Atlantic Current, NAC, and North Atlantic deep water, NADW,
219 flow over the Mid Atlantic Ridge) transitions to more barotropic flow (deep western boundary
220 current, DWBC). The geographical extent and transition between the 18% (subpolar gyre I) and
221 11% (subtropical) areas make sense in light of this, particularly as they are mirrors of each other,
222 but with the difference in wind sign and BPT interactions. The 7% (subpolar gyre II) region
223 complements the subtropical region, being an intensified extension and both with the addition
224 of small contributions of non-linear terms. Extrapolating this to the global picture, this suggests
225 that strong interior flow is present in vast expanses of the Southern Hemisphere, as well as
226 the North Pacific. The quasi-Sverdrupian regime we see in the Northern Hemisphere subtrop-
227 ics is virtually absent in the South Pacific, suggesting the absence of a strictly barotropic flow
228 structure. This is found in the Southern and Indian Ocean, indicating that the wind field dominates.

229

230 The sign and spatial distribution of the wind stress term suggests the importance of Ekman
231 pumping(negative)/suction(positive). The Equatorial and Southern Ocean regions have Ekman
232 pumping, while the subpolar gyre areas have Ekman suction where modewaters are created. The
233 BPT term mirrors the wind stress term, suggesting it acts as either a source or a sink in opposite
234 compliment to the wind stress. A full description of the dynamics is outside the scope of the
235 current study.

236

237 The area that is dominantly non-linear illustrates that key ocean regions may not conform to
238 the simplification presented by the twenty year averaged BV equation. This region is a collection
239 of smaller terms. Interactions with topographic barriers, as well as WBC regions stand out,
240 suggesting that the vertically integrated momentum flux divergence, the non-linear contribution to
241 vortex tube stretching and the conversion of vertical shear to barotropic vorticity are active and
242 key to the dominant dynamics. A detailed description of balances present here will be the subject
243 of a future study.

244

245 The assumed variance structure by the K-Means clustering could mean a more computationally
246 efficient solution is available by exploring other methods such as the more generalised gaussian
247 mixture models. Doing this other covariance shapes can be explored, and we present initial results
248 in figure 2. This was not found to impact results, but will likely be key for a robust scaling of the
249 method to higher resolution data.

250

251 The present study found that the BV budget closes in the ECCOv4r2 State Estimate over
252 twenty years. The balance observed deviated from a straightforward interpretation of the three
253 dominant terms as the wind, Coriolis and BPT terms, with significant covariance and input from
254 the non-linear terms in dynamically key regions. These regions were identified using the K-Means
255 clustering method, which was a good method for characterization. Large areas of the ocean are
256 well described by the BV equations, and five regions were identified covering 93% of the world
257 ocean: A stable ocean (56%), subpolar gyre I (18%), subtropical gyre (11%), subpolar gyre II
258 (7%) and the Southern Ocean gyre (2%). The final area is dominated by the non-linear terms and
259 will be the subject of a future study.

260

261 **5. Acknowledgments**

262 This work was funded by the US National Aeronautics and Space Administration Sea Level
263 Change Team (contract NNX14AJ51G) and through the ECCO Consortium funding via the Jet
264 Propulsion Laboratory. This work would not have been possible without the support of Anne
265 Reinarz, Edward Doddridge, Roosa Tikkainen and Katherine Rosenfeld.

266 APPENDIX

267 **K-Means and influence of Information Criteria**

268 **A1. K-Means and influence of Information Criteria and covariance assumptions**

269 The Akaike and Bayesian Information Criteria (AIC and BIC) are used; Minimizing the expec-
270 tation of the prediction error where k is the number of parameters/clusters:

$$\text{AIC} = 2k - 2\ln(\mathcal{L}),$$

$$\text{BIC} = k\ln(n) - 2\ln(\mathcal{L}),$$

271 where n is the number of datapoints, k the number of parameters and \mathcal{L} is the likelihood:

$$\mathcal{L} = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\sum_{i=1}^N \frac{(\zeta_i - \hat{\zeta}_i)^2}{2\sigma^2}\right).$$

272 ζ_i is the observed, and $\hat{\zeta}_i$ is the prediction, so $(\zeta_i - \hat{\zeta}_i)^2$ are the prediction residuals. In the estimate,
273 the AIC value is minimized, which determines the smallest appropriate order to represent the
274 time-series. As discussed by Priestley (1981) and Yang (2005), the AIC can overestimate the
275 orde.

277 We demonstrate that we are in an asymptotic regime in Figure 12. The Gaussian Mixture
278 Models are a more generalised form of the K-Means clustering, this is used to assess the impact
279 of assumptions relating to the covariance structure; spherical, diagonal, tied or full covariance as
280 seen in Figure 2. Using the BIC, this is found to likely not impact results, but could be key for
281 scaling up the method for high resolution applications where computational efficiency becomes
282 more important.

283

284 References

- 285 Adcroft, A., Hill, C., Campin, J. M., Marshall, J., & Heimbach, P. (2004). Overview of the formulation and numerics of the MIT GCM (pp.
286 139-150). Presented at the ECMWF Conference Proceedings, Shinfield Park, Reading, UK.
- 287 Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B.N. ; Csiki, F., 2nd International
288 Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akademiai Kiad, p. 267-281.
- 289 Ardyna, M., H. Claustre, J. Sallee, F. D'ovidio, B. Gentili, G. van Dijken, F. D'Ortenzio, and K. R. Arrigo, 2017: Delineating environmental control
290 of phytoplankton biomass and phenology in the southern ocean. *Geophys. Res. Lett.*, 44, 50165024, doi:10.1002/2016GL072428.
- 291 Church, J.A., et al., 2013: Sea Level Change. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the
292 Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J.
293 Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Chapter 13, pp. 1137-1216, Cambridge University Press.
- 294 ECCO Consortium, 2017a, A Twenty-Year Dynamical Oceanic Climatology: 1994-2013. Part 1: Active Scalar Fields: Temperature, Salinity,
295 Dynamic Topography, Mixed-Layer Depth, Bottom Pressure.
- 296 ECCO Consortium, 2017b, A Twenty-Year Dynamical Oceanic Climatology: 1994-2013. Part 2: Velocities, Property Transports, Meteorological
297 Variables, Mixing Coefficients.
- 298 Forget, G., J.-M. Campin, P. Heimbach, C. N. Hill, R. M Ponte, and C. Wunsch, ECCO version 4: an integrated framework for non-linear inverse
299 modeling and global ocean state estimation, *Geo. Sci. Model Dev.*, 8, 2015
- 300 Hirschi, J. J.-M., Blaker, A. T., Sinha, B., Coward, A., de Cuevas, B., Alderson, S., and Madec, G.: Chaotic variability of the meridional overturning
301 circulation on subannual to interannual timescales, *Ocean Sci.*, 9, 805-823, doi:10.5194/os-9-805-2013, 2013.

- 302 Le Bras, I., Toole, J.M., Sonnewald, M.: A bulk Potential Vorticity budget for the western North Atlantic based on observations. In review.
- 303 Liang, Y.-C. *et al.*, in review.
- 304 J. B. MacQueen (1967): "Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium
305 on Mathematical Statistics and Probability", Berkeley, University of California Press, 1:281-297.
- 306 Munk, W., 1950: On the wind-driven ocean circulation. *J. Meteor.*, 7, 7993, doi:10.1175/1520-0469(1950)007;0080:OTWDOC_2.0.CO;2.
- 307 Munk, W.H., and Palmén, E.: Note on the dynamics of the Antarctic Circumpolar Current, *Tellus*, 3, 53-55, 1940.
- 308 Myers, Paul G., Augustus F. Fanning, Andrew J. Weaver, 1996: JEBAR, Bottom Pressure Torque, and Gulf Stream Separation. *J. Phys. Oceanogr.*,
309 26, 671683. doi: [http://dx.doi.org/10.1175/1520-0485\(1996\)026<0671:JBPTAG>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1996)026<0671:JBPTAG>2.0.CO;2)
- 310 Priestley, M. B. (1981) Spectral Analysis and Time Series, London: Academic Press.
- 311 Reynolds, R.W., D.B. Chelton, J. Roberts-Jones, M.J. Martin, D. Menemenlis, and C.J. Merchant, 2013: Objective Determination of Feature
312 Resolution in Two Sea Surface Temperature Analyses. *J. Climate*, 26, 2514-2533, <https://doi.org/10.1175/JCLI-D-12-00787.1>
- 313 Schoonover, J., Dewar, W., Wienders, N., Gula, J., McWilliams, J.C., Molemaker, M.J., Bates, S.C., Danabasoglu, G. and Yeager, S.: North
314 Atlantic Barotropic Vorticity Balances in Numerical Models, *Journal of Physical Oceanography* 2016 46:1, 289-303
- 315 Stainforth, D., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D., Kettleborough, J., Knight, S., Martin, A., Murphy, J., Piani, C., Sexton,
316 D., Smith, L., Spicer, R., Thorpe, A. and Allen, M.: 2005, Uncertainty in predictions of the climate response to rising levels of greenhouse
317 gases, *Nature* 433(7024), 403-406. url:<http://oro.open.ac.uk/5025/>
- 318 D. Stammer, M. Balmaseda, P. Heimbach, A. Koehl, and A. Weaver, 2016: Ocean Data Assimilation in Support of Climate Applications: Status
319 and Perspectives. *Ann. Rev. Mar. Sci.*, 8, 491-518.
- 320 Stommel, H. (1948), The westward intensification of wind-driven ocean currents, *Eos Trans. AGU*, 29(2), 202206, doi:10.1029/TR029i002p00202.
- 321 Yang, Y. (2005), "Can the strengths of AIC and BIC be shared?", *Biometrika*, 92: 937-950, doi:10.1093/biomet/92.4.937.
- 322 Yeager, S.: Topographic coupling of the Atlantic overturning and gyre circulations, *J. Phys. Oceanogr.*, 45, 1258-1284. doi: <http://dx.doi.org/10.1175/JPO-D-14-0100.1>, 2015.
- 324 Wunsch, C., and P. Heimbach, 2007: Practical global oceanic state estimation. *Physica D*, 230, 197-208.
- 325 Wunsch, C. and P. Heimbach, 2013, Dynamically and kinematically consistent global ocean circulation and ice state estimates. In *Ocean Circulation
326 and Climate* 2nd Edition, Siedler et al., Eds.

327 Wunsch, C. The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. Bull. Am. Met. Soc. 80,
328 245-255 (1999).

329 Wunsch, C. (2013), Covariances and linear predictability of the Atlantic Ocean, Deep Sea Res., Part II, 85, 228-243.

330 Wunsch, C. (2015), Modern Observational Physical Oceanography: Understanding the Global Ocean. Princeton University Press.

331 **LIST OF TABLES**

332 Table 1.	Percentage of area covered by specific balance and corresponding figure. Lead-	19
333	ing order terms are sorted by magnitude, colors indicating if barotropic vorticity	
334	is added (red) or removed (blue) by the leading order term.	

Area (Figure nr)	Leading terms
55.6%, Stable ocean (Fig. 6a)	$0 \approx \nabla \times \tau_{sb} + \nabla \times \mathbf{A} + \nabla p_b \times \nabla H$
17.6%, Subpolar gyre I (Fig. 6c)	$0 \approx \nabla \times \tau_{sb} + \nabla p_b \times \nabla H + \nabla \cdot (\mathbf{fU})$
11.2%, Subtropical gyre (Fig. 6e)	$0 \approx \nabla \times \tau_{sb} + \nabla \cdot (\mathbf{fU})$
6.6%, Subpolar gyre II (Fig. 7a)	$0 \approx \nabla \times \tau_{sb} + \nabla \cdot (\mathbf{fU}) + \nabla p_b \times \nabla H$
2.0%, Southern Ocean gyre (Fig. 15a)	$0 \approx \nabla \times \tau_{sb} + \nabla \cdot (\mathbf{fU}) + \nabla p_b \times \nabla H$
$\approx 6.3\%$, Dominantly non-linear (Fig. 7e)	$0 \approx \nabla \cdot (\mathbf{fU}) + \nabla \times \mathbf{A} + \nabla \times \tau_{sb}$

TABLE 1: Percentage of area covered by specific balance and corresponding figure. Leading order terms are sorted by magnitude, colors indicating if barotropic vorticity is added (**red**) or removed (**blue**) by the leading order term.

335 LIST OF FIGURES

336	Fig. 1. The breakdown of the barotropic vorticity budget over 1992-2013 in the EC-	
337	COv4r2 State Estimate.	21
338	Fig. 2. Trying different covariance models to check the convergence.	22
339	Fig. 3. The AIC and BIC asymptote and we choose a K of 50 for our analysis. Error bars represent	
340	2σ , capturing the stochastic start seen of the algorithm.	23
341	Fig. 4. The area selected by the clusters. The colors represent the clusters, and are in arbitrary order.	24
342	Fig. 5. The percentage of global ocean area covered by the clusters.	25
343	Fig. 6. Maps of the selected locations (left) and corresponding area averaged his-	
344	togram (right) of the terms in the BV equation. The colorbar is kept, but the	
345	color/ordering of the map are arbitrary. Colors in the barchart indicate if BV is	
346	added (red) or removed (blue).	26
347	Fig. 7. Maps of the selected locations (left) and corresponding area averaged his-	
348	togram (right) of the terms in the BV equation. The colorbar is kept, but the	
349	color/ordering of the map are arbitrary. Colors in the barchart indicate if BV is	
350	added (red) or removed (blue).	27
351	Fig. 8. The Little Ones: Maps of the selected locations (left) and corresponding area	
352	averaged histogram (right) of the terms in the BV equation. The colorbar is kept,	
353	but the color/ordering of the map are arbitrary. Colors in the barchart indicate if	
354	BV is added (red) or removed (blue).	28
355	Fig. 9. The Little Ones: Maps of the selected locations (left) and corresponding area	
356	averaged histogram (right) of the terms in the BV equation. The colorbar is kept,	
357	but the color/ordering of the map are arbitrary. Colors in the barchart indicate if	
358	BV is added (red) or removed (blue).	29
359	Fig. 10. The Little Ones: Maps of the selected locations (left) and corresponding area	
360	averaged histogram (right) of the terms in the BV equation. The colorbar is kept,	
361	but the color/ordering of the map are arbitrary. Colors in the barchart indicate if	
362	BV is added (red) or removed (blue).	30
363	Fig. 11. The area selected by the clusters. The colors represent the clusters, and are in arbitrary order.	31
364	Fig. 12. The percentage of global ocean area covered by the clusters.	32
365	Fig. 13. Maps of the selected locations (left) and corresponding area averaged histogram	
366	(right).	33
367	Fig. 14. Maps of the selected locations (left) and corresponding area averaged histogram	
368	(right).	34
369	Fig. 15. Maps of the selected locations (left) and corresponding area averaged histogram	
370	(right).	35

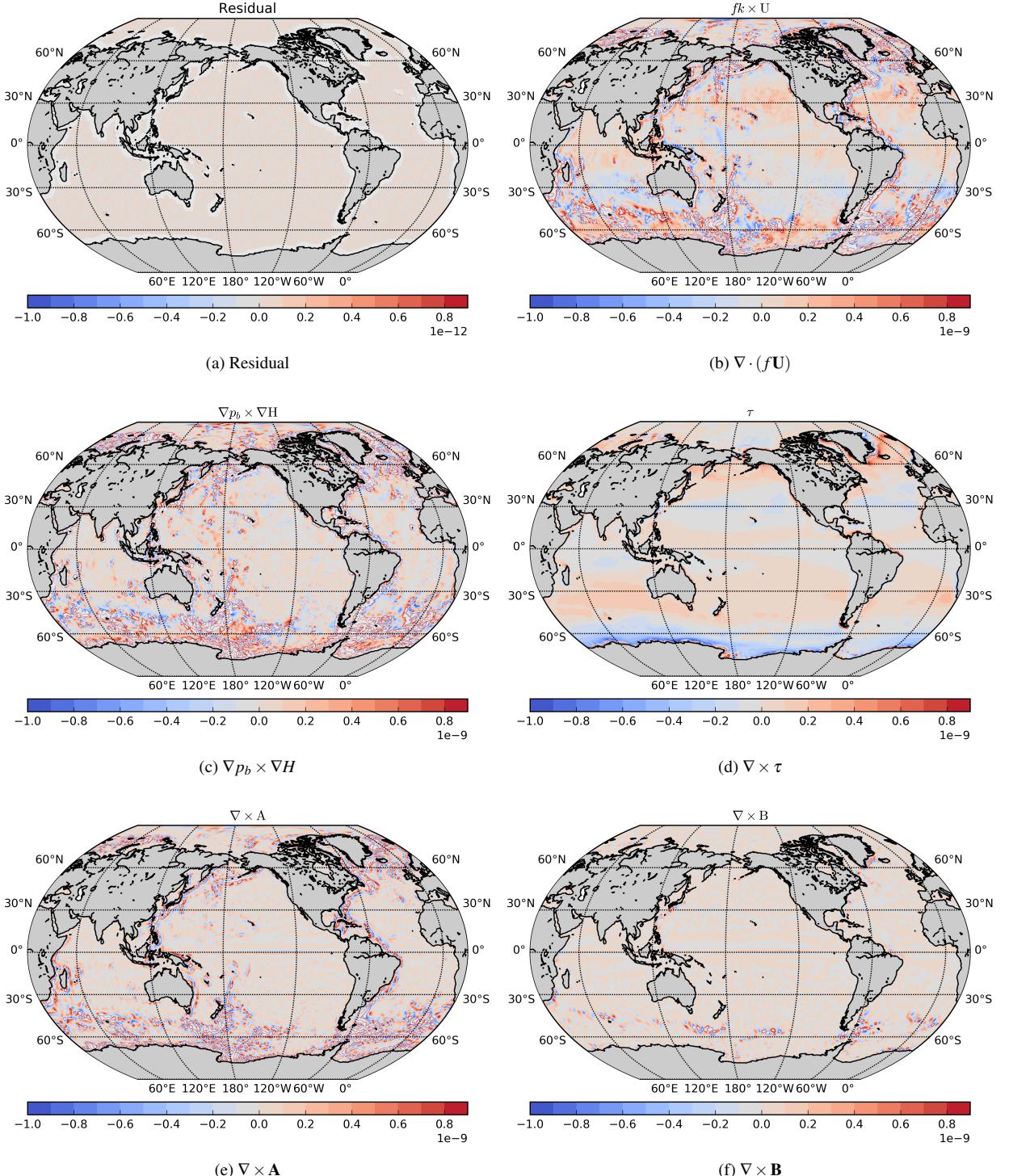


FIG. 1: The breakdown of the barotropic vorticity budget over 1992-2013 in the ECCOv4r2 State Estimate.

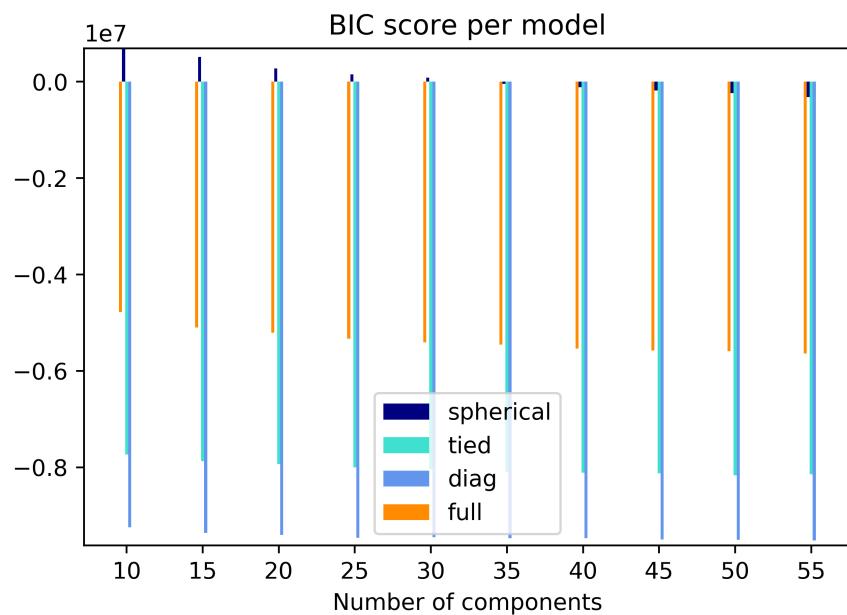


FIG. 2: Trying different covariance models to check the convergence.

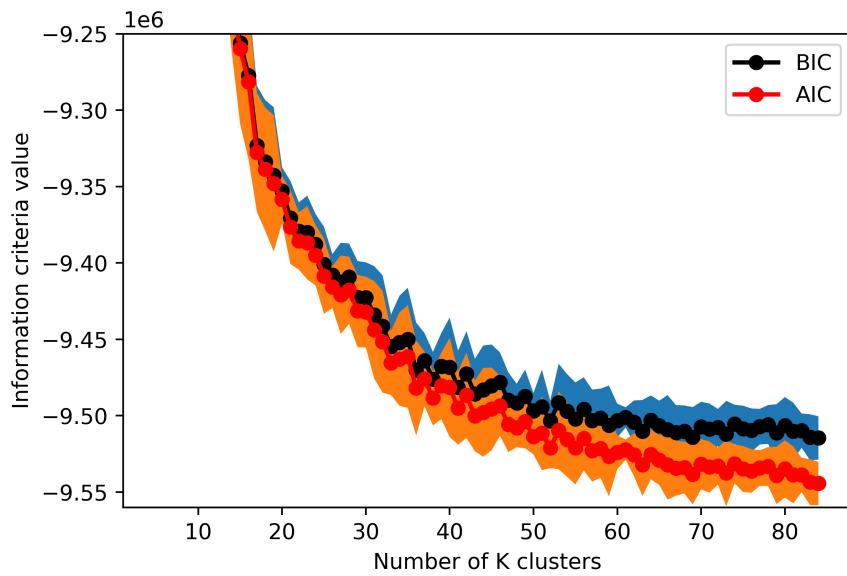


FIG. 3: The AIC and BIC asymptote and we choose a K of 50 for our analysis. Error bars represent 2σ , capturing the stochastic start seen of the algorithm.

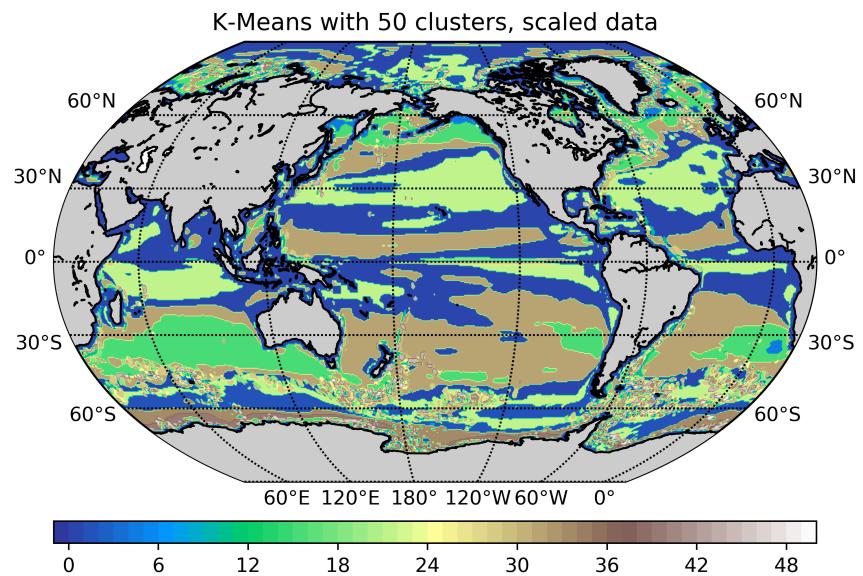


FIG. 4: The area selected by the clusters. The colors represent the clusters, and are in arbitrary order.

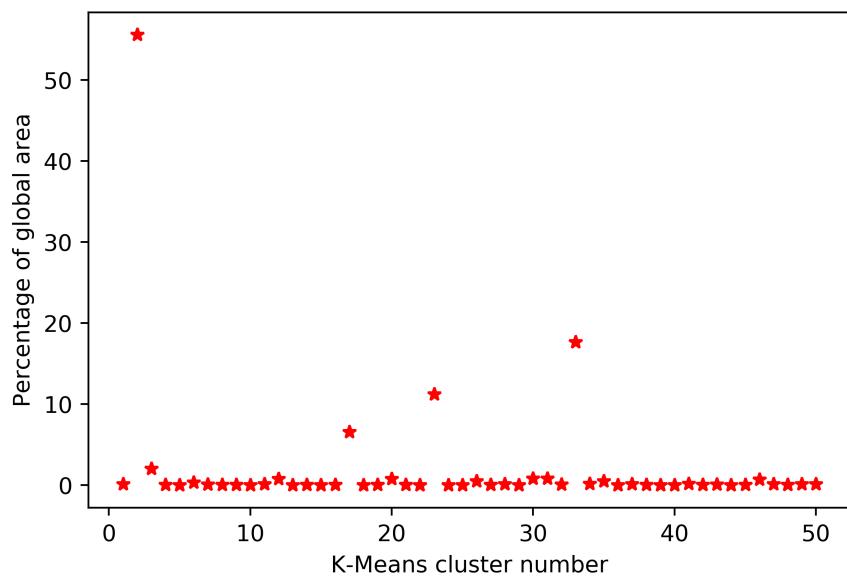


FIG. 5: The percentage of global ocean area covered by the clusters.

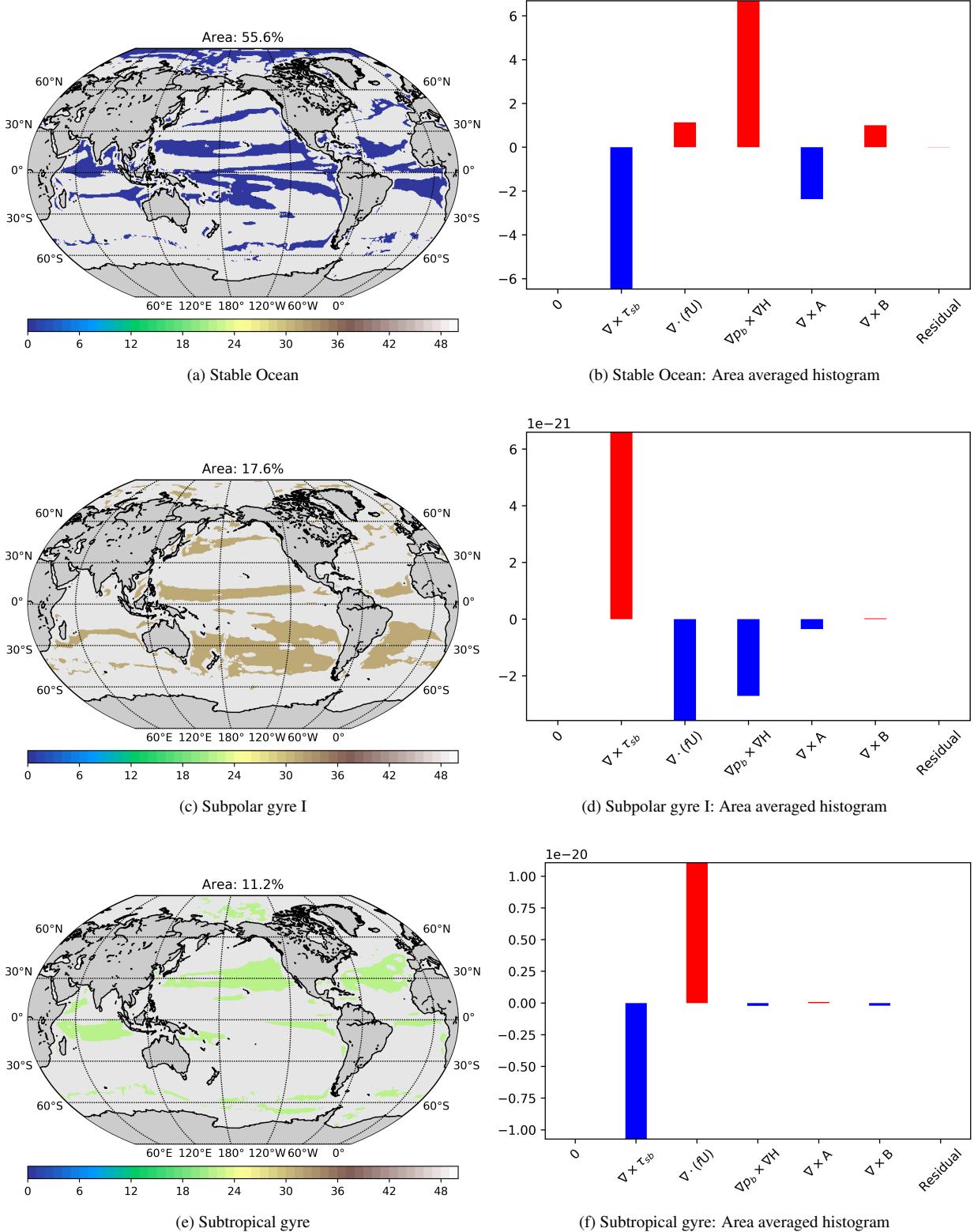
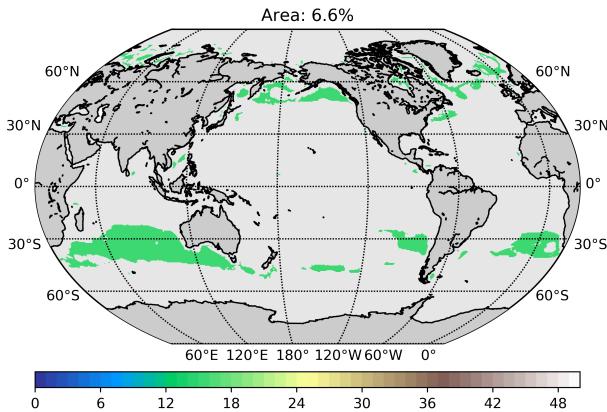
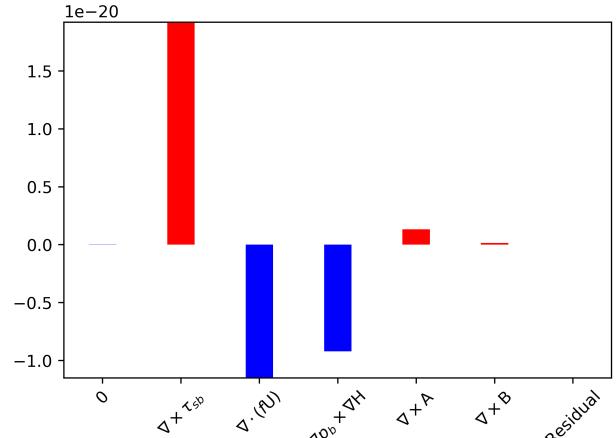


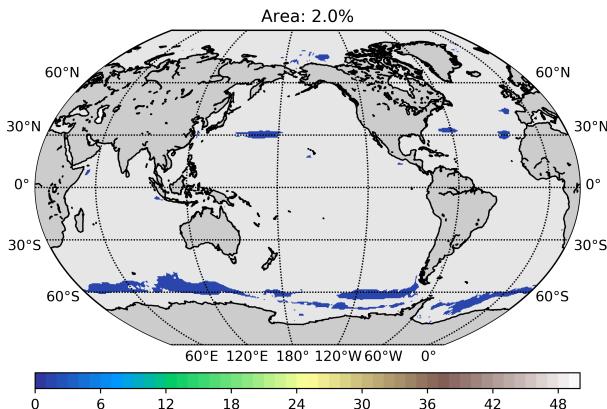
FIG. 6: Maps of the selected locations (left) and corresponding area averaged histogram (right) of the terms in the BV equation. The colorbar is kept, but the color/ordering of the map are arbitrary. Colors in the barchart indicate if BV is added (red) or removed (blue).



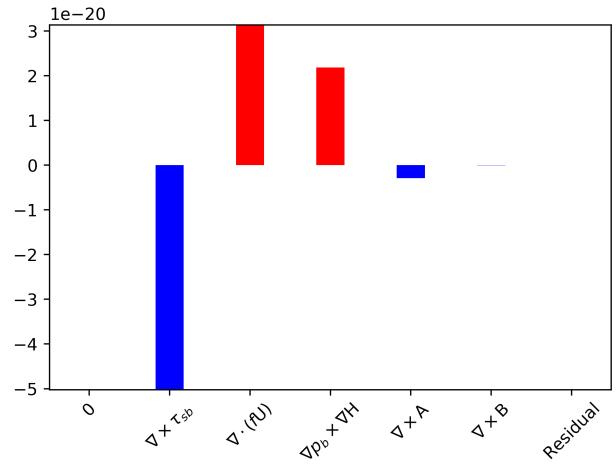
(a) Subpolar gyre II



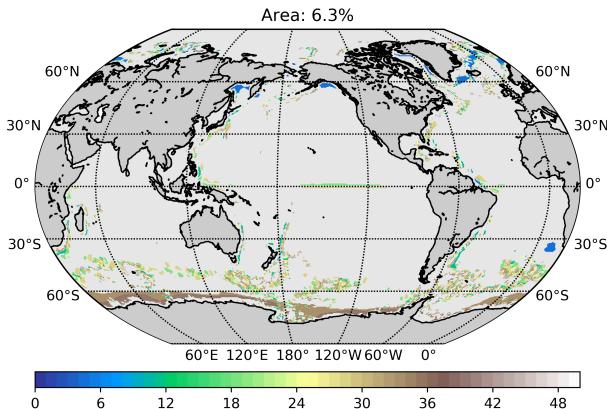
(b) Subpolar gyre II: Area averaged histogram



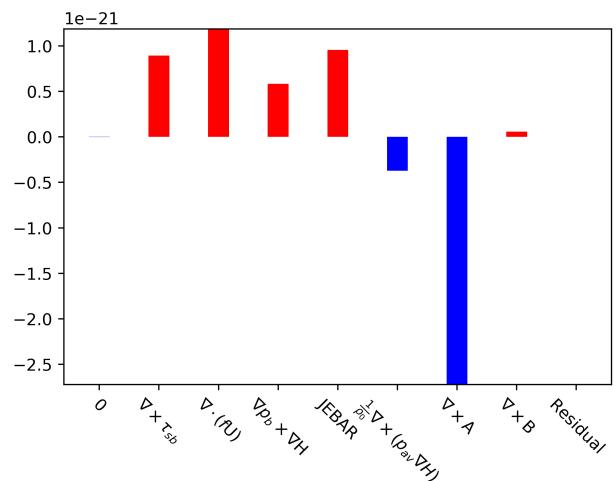
(c) Southern Ocean gyre



(d) Southern Ocean gyre: Area averaged histogram

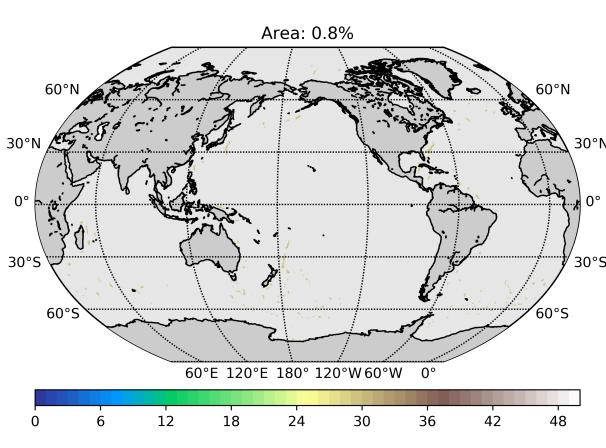


(e) Dominantly non-linear

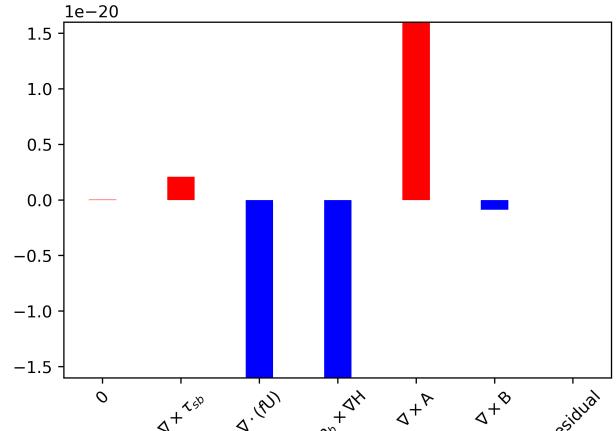


(f) Dominantly non-linear: Area averaged histogram

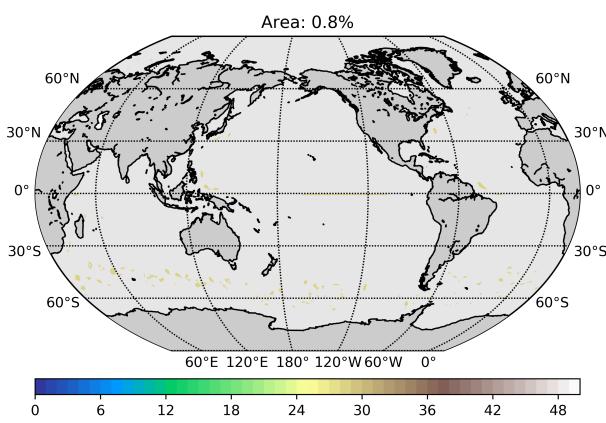
FIG. 7: Maps of the selected locations (left) and corresponding area averaged histogram (right) of the terms in the BV equation. The colorbar is kept, but the color/ordering of the map are arbitrary. Colors in the barchart indicate if BV is added (red) or removed (blue).



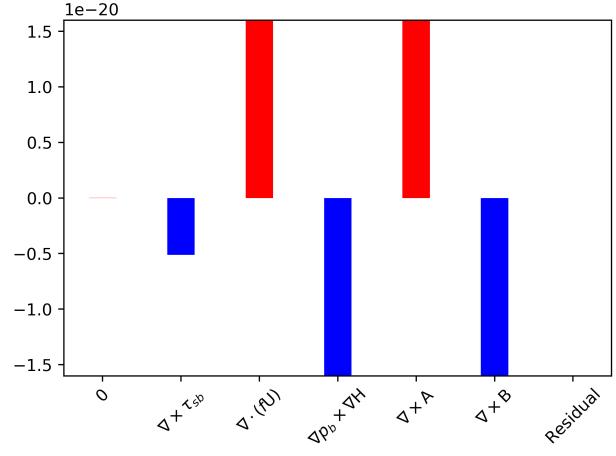
(a)



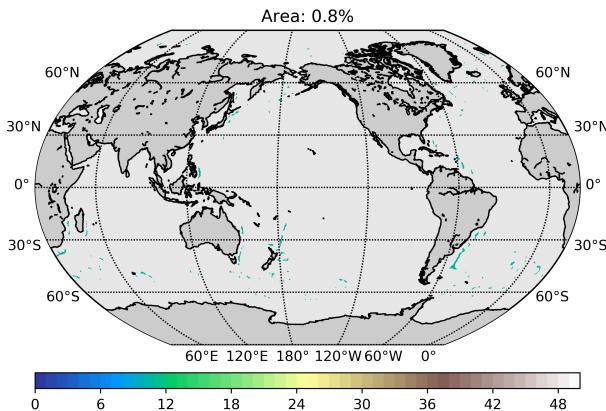
(b) Area averaged histogram



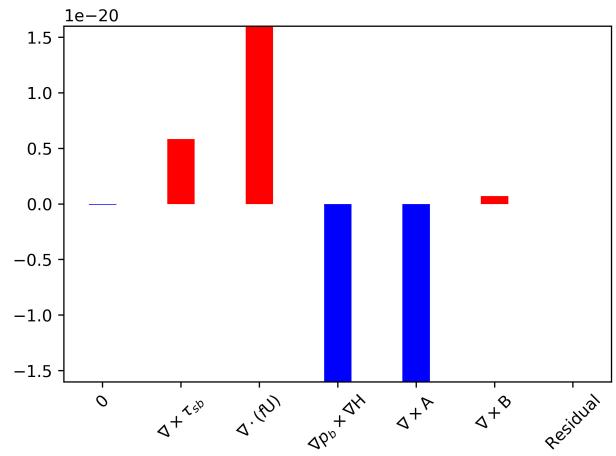
(c)



(d) Area averaged histogram

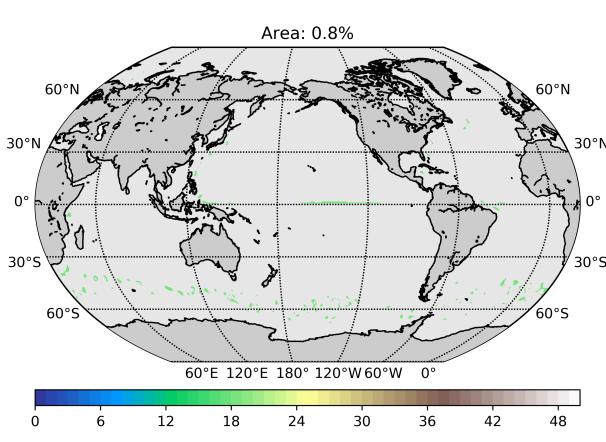


(e)

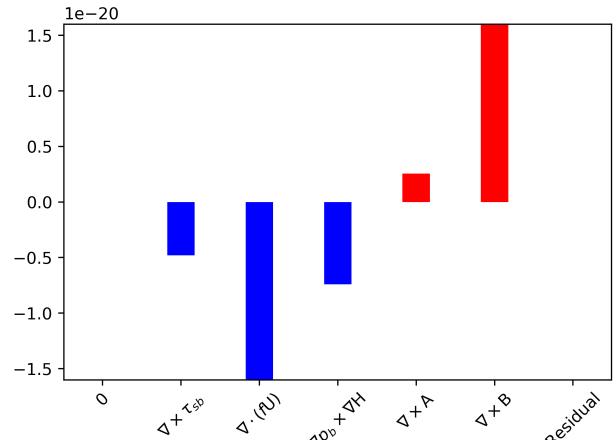


(f) Area averaged histogram

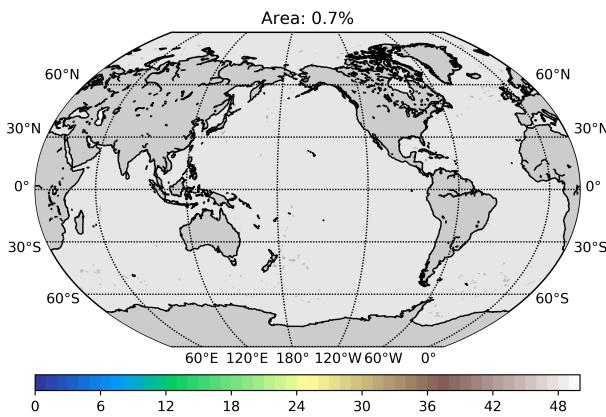
FIG. 8: The Little Ones: Maps of the selected locations (left) and corresponding area averaged histogram (right) of the terms in the BV equation. The colorbar is kept, but the color/ordering of the map are arbitrary. Colors in the barchart indicate if BV is added (red) or removed (blue).



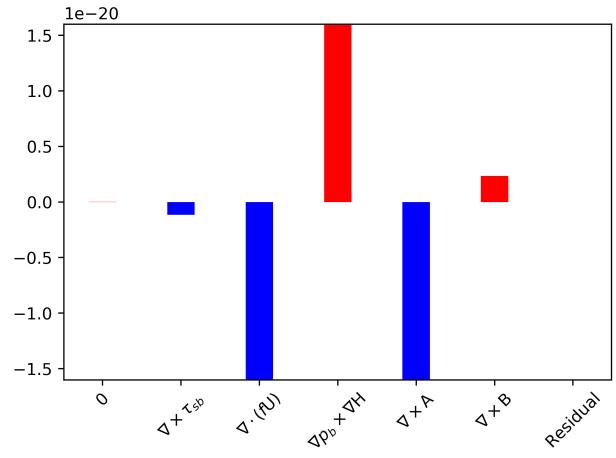
(a)



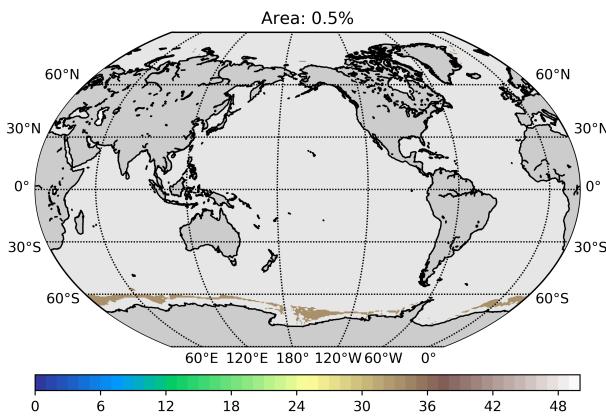
(b) Area averaged histogram



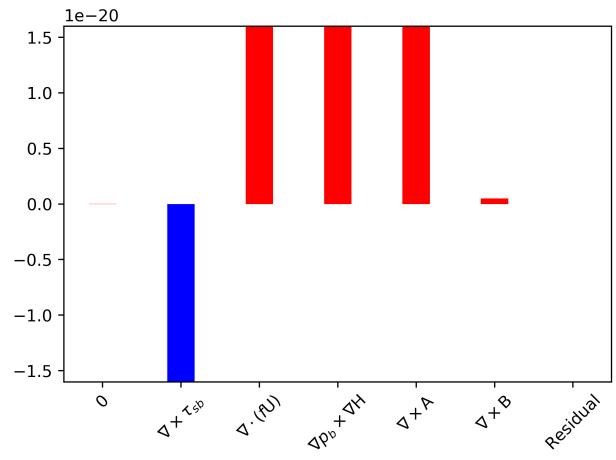
(c)



(d) Area averaged histogram



(e)



(f) Area averaged histogram

FIG. 9: The Little Ones: Maps of the selected locations (left) and corresponding area averaged histogram (right) of the terms in the BV equation. The colorbar is kept, but the color/ordering of the map are arbitrary. Colors in the barchart indicate if BV is added (red) or removed (blue).

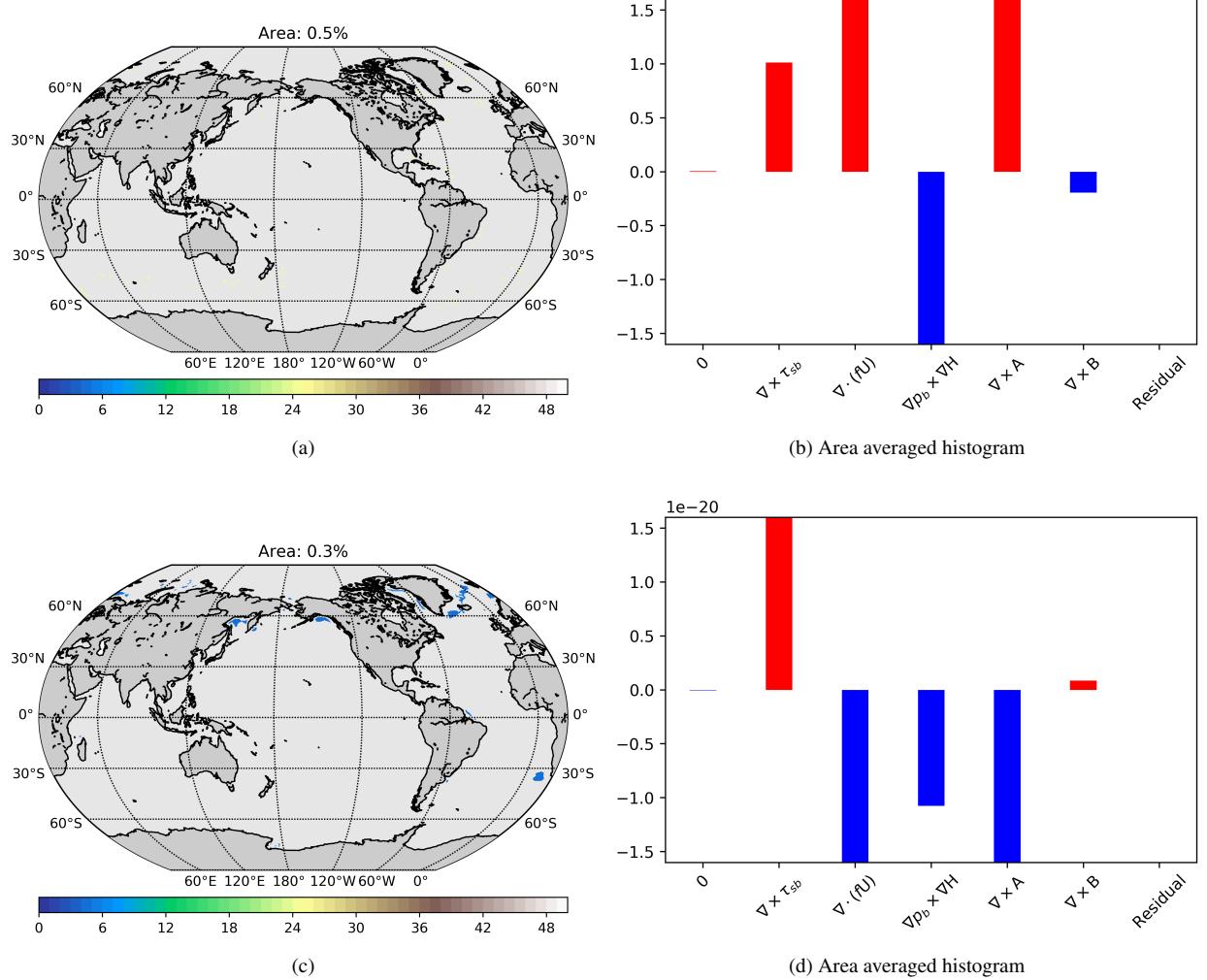


FIG. 10: The Little Ones: Maps of the selected locations (left) and corresponding area averaged histogram (right) of the terms in the BV equation. The colorbar is kept, but the color/ordering of the map are arbitrary. Colors in the barchart indicate if BV is added (red) or removed (blue).

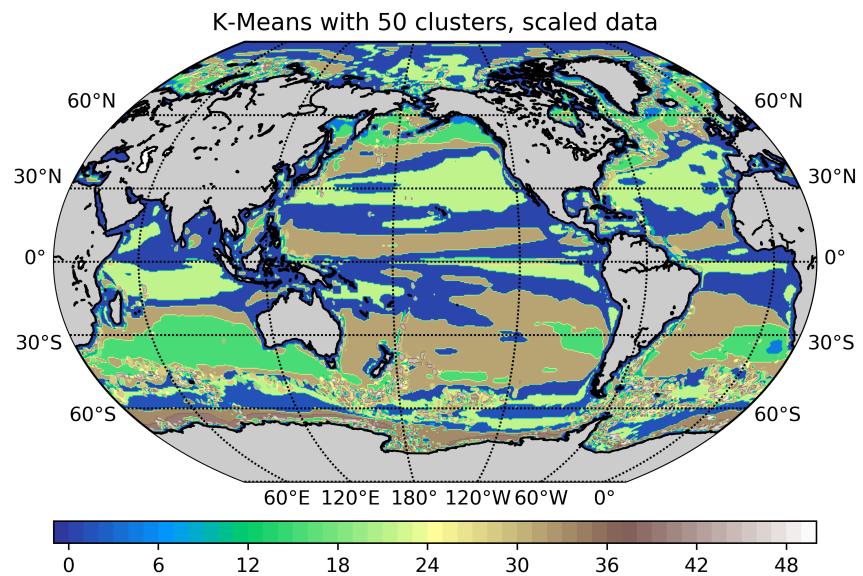


FIG. 11: The area selected by the clusters. The colors represent the clusters, and are in arbitrary order.

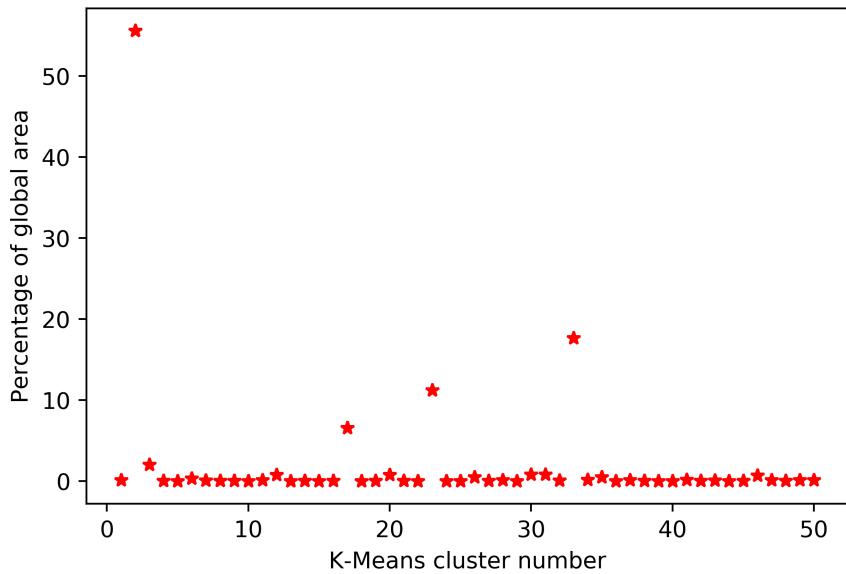


FIG. 12: The percentage of global ocean area covered by the clusters.

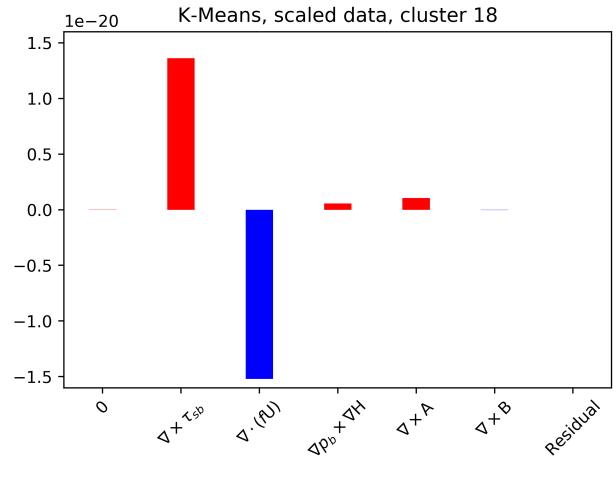
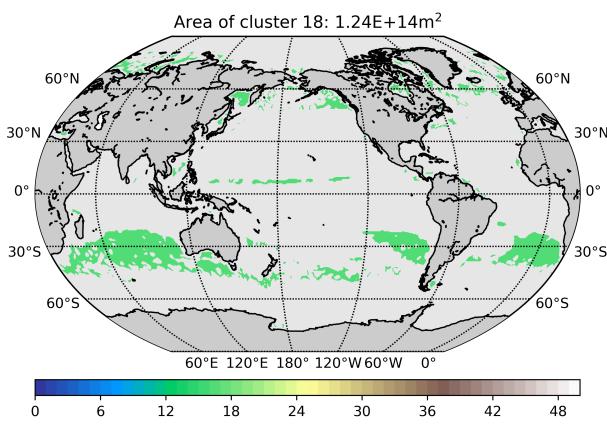
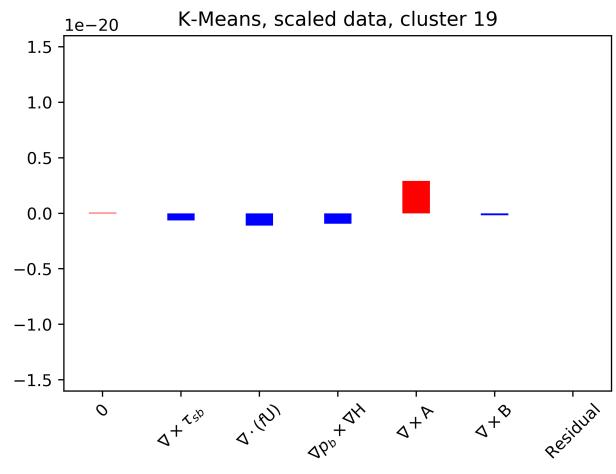
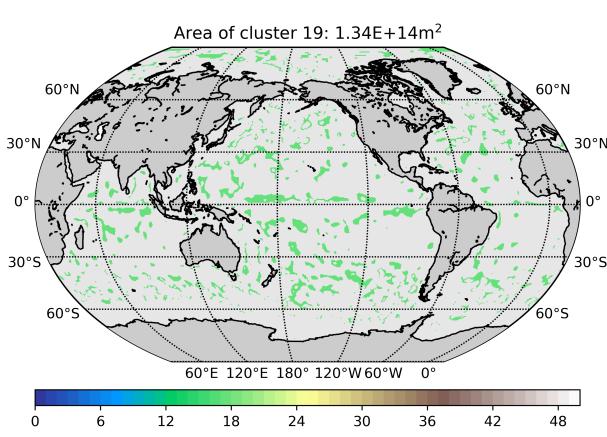
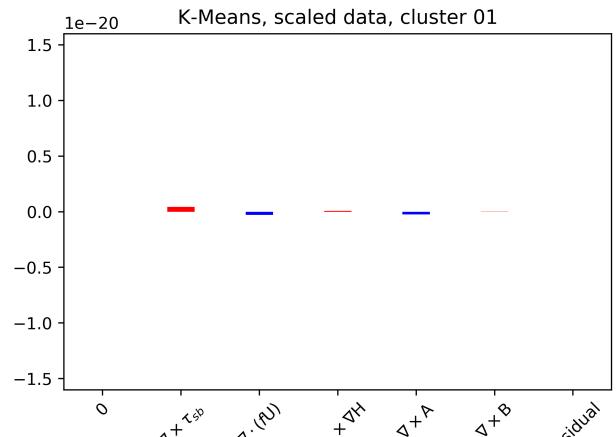
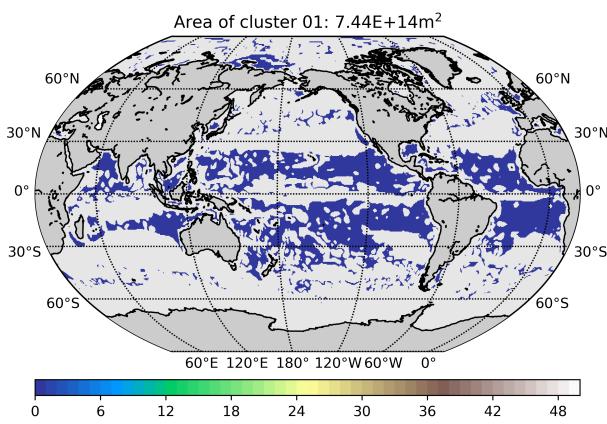


FIG. 13: Maps of the selected locations (left) and corresponding area averaged histogram (right).

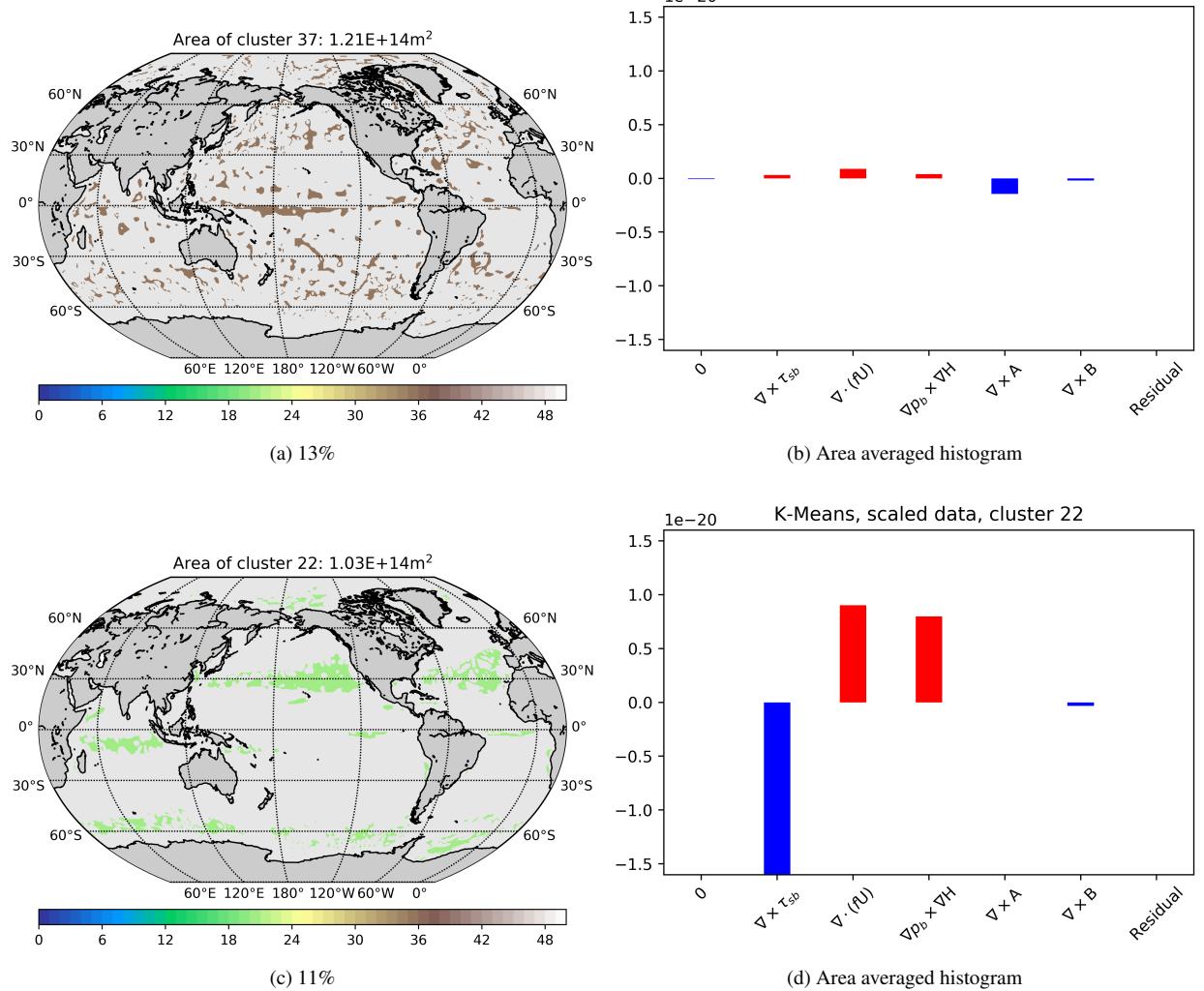


FIG. 14: Maps of the selected locations (left) and corresponding area averaged histogram (right).

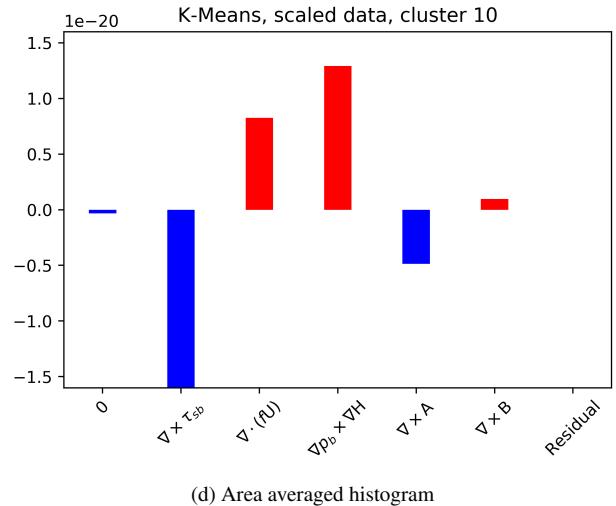
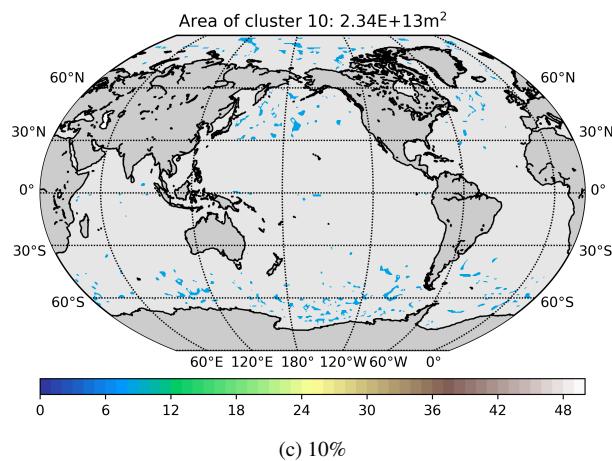
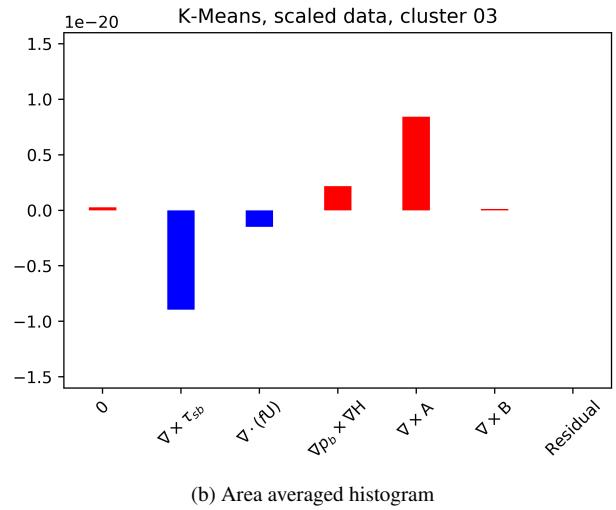
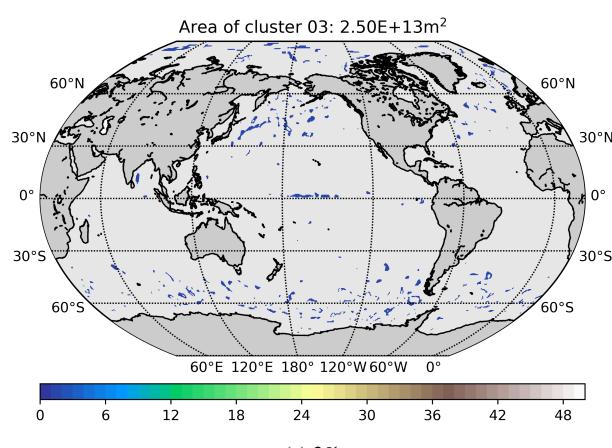


FIG. 15: Maps of the selected locations (left) and corresponding area averaged histogram (right).