Supporting Information for: "Physical mechanisms driving oxygen subduction in the global ocean"

Esther Portela[1]*, Nicolas Kolodziejczyk[1], Clément Vic[1] and Virginie Thierry[1]

[1]Univ. Brest, CNRS, IRD, Ifremer, Laboratoire d'Oceanographie Physique et Spatiale (LOPS), Plouzane, France

S1 Uncertainty Computation

In this section we describe and estimate the uncertainty associated to the computations performed in this study. The following sources of uncertainty were not considered in this analysis due to either the lack of the necessary elements to compute them or because of the little relative contribution that they have to the total error:

1. The different time lapse for each dataset. While ECCOv4 extends from 1992 to 2015, the WOA18 oxygen climatology includes data collected from 1955 to 2018 (with almost no data before 1970). This inconsistency is a source of uncertainty that we were not able to overcome. However, due to the relative little role that the distribution of $[O_2]$ plays in the total oxygen subduction (S^{ox}) we consider that this uncertainty is negligible.
2. As the WOA18 data only uses oxygen data obtained by chemical Winkler titration methods, the sampling error (critical in the CTD oxygen captors) is not considered in this study.
3. The error linked to the subduction computation will be neglected as it was performed from the ECCOv4 outputs with no associated sampling or interpolation error.

In the following, we provide estimations on the uncertainty linked to the historically sparse oxygen sampling (Figure S1a, b), and to the interannual variability of $[O_2]$ and the mass subduction flux. The interannual is the most important timescale non resolved by the monthly climatology fields used to compute S^{ox}, thus, it is thought to be the main source of uncertainty. Finally we have propagated the error associated to each variable to obtain the final uncertainty linked to the S^{ox} computation.

The oxygen data distribution in Figure S1a shows all data collected between 1955 and 2018, although sampling was very scarce before 1970. Oxygen sampling has been historically uneven; the northern hemisphere concentrate most of the data, mainly near the cost. The North Pacific and the northern North Atlantic count with approximately four times more data than the southern Hemisphere basins (Figure S1b) where observations are sparse.

The uncertainty associated to the interannual variability of oxygen data is expressed by the coefficient of variation (C.V. in %) of $[O_2]$ (Figure S1c) (C.V $= 100 \cdot \sigma/\overline{x}$, where σ is the standard deviation and $\overline{x}$ is the mean). C.V allows to have an estimate of the data variability that is unaffected by the mean (with same C.V, higher means are associated with higher standard deviations). Taking into account that the existing data cover more than 60 years the interannual variability of the $[O_2]$ is quite low. The maximum C.V reaches 40% only in very localized tropical regions but, globally, the interannual variability represents less than 15% of the mean. In addition to the tropical regions as the North Indian and the

Corresponding author: Esther Portela, `eportelanh@gmail.com`

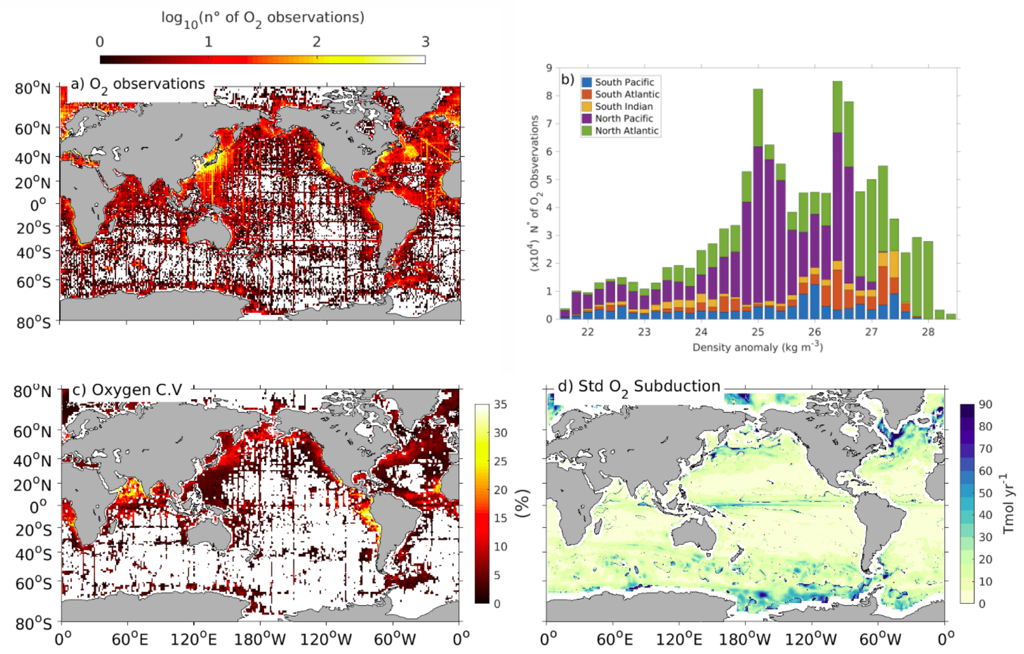

Figure S1. Uncertainty associated to the oxygen subduction components and to the oxygen sampling. a) Map of the data distribution in logarithmic scale. b) Number of historical oxygen observations contained in each density class and basin. c) Geographic distribution of the coefficient of variance (C.V, (%)) for the oxygen as obtained from WOA18. d) Geographic distribution of the interannual standard deviation of the S^{ox}

eastern tropical Pacific, relatively high interannual signal is found in the North Pacific and subtropical Atlantic. The C.V somehow reflects the data distribution (Figure S1a) but the blanked area is larger. This indicates that a big part of the ocean has been poorly sampled, making impossible the computation of the standard deviation. The Southern Ocean is one of the less sampled regions and it also shows low oxygen variability.

The uncertainties associated to the interannual standard deviation of the $[O_2]$ and the mass subduction flux were propagated to obtain the final standard deviation of the S^{ox} following the typical equation of uncertainty propagation:

$$\sigma_{(S^{ox})} = |S^{ox}| \sqrt{\left(\frac{\sigma_{(O_2)}}{[O_2]}\right)^2 + \left(\frac{\sigma_{(Sub)}}{S}\right)^2 + \frac{2 \cdot cov([O_2], S)}{[O_2] \cdot S}} \tag{1}$$

Where S is the mass subduction, σ is the standard deviation of each variable, and cov is the covariance between the $[O_2]$ and the mass subduction. Assuming that these two variables are not correlated, the covariance term can be neglected.

The distribution of the uncertainty associated to the interannual S^{ox} variability is shown in Figure S1d. The C.V was not used in this case since this metric does not work well for variables with values crossing zero as the S^{ox}. The distribution of the standard deviation of the mass subduction flux is not shown here because it approximates very much (only with different units) that of the S^{ox}. This indicates that the uncertainty of the oxygen flux across the mixed layer, as the S^{ox} itself, is driven by the physical mass flux.

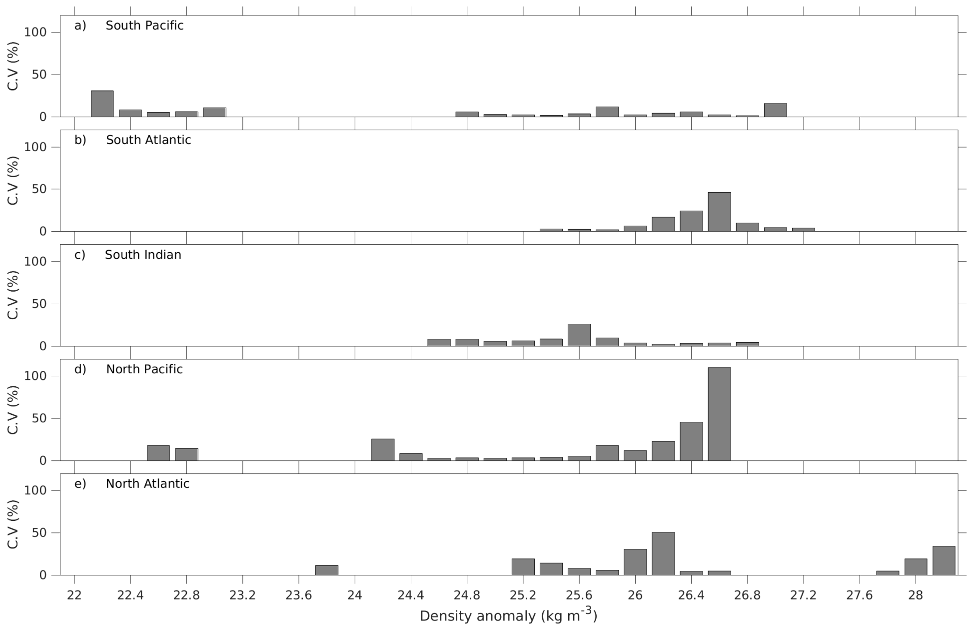

Figure S2. Coefficient of variation (C.V (%)) of the S^{ox} that illustrates the uncertainty due to the interannual variability of of the S^{ox} and the sparse oxygen sampling for every density class and basin. Density classes with S^{ox} lower than 1Tmol/Yr^{-1} where not taken into account because their C.V results in abnormally high values

In the North Atlantic, the equatorial strip and within the ACC limits, high uncertainty (*i.e.* high interannual variability) is associated to intense mean S^{ox} rates. The standard deviation represent between 30-50% of the mean value with local spots reaching the 100%. In contrast, this is not the case of the southernmost latitudes of the Southern Ocean where the high standard deviations could be explained by different factors: (i) it could represent actual interannual variability produced by the different ice coverage (ii) it might reflect the relative scarcity of data constraining the ECCOv4 reanalysis in the highest latitudes and (iii) it could be due to the fact that the net S^{ox} is nearly zero in this region. Given the impossibility of unraveling the source of uncertainty, it would be convenient to consider the S^{ox} in this region carefully. Relatively high variability is also found in the Northern North Pacific, associated with the northern edge of the subtropical gyre.

To link the reported uncertainty maps with our results, we show the C.V associated to the S^{ox} integrated in density classes (Figure S2). To obtain the C.V, the standard deviation of the S^{ox} as obtained from Equation (1) was propagated following the next equation:

$$[\sigma_m]_{(S^{ox})} = \sqrt{\sum_{i=1}^{n} (a_i^2 \cdot \sigma_i^2) + 2ab \cdot cov_{(i=1:n)}} \tag{2}$$

Where σ_m represents the standard deviation within a given density class m, a represents the area of each corresponding grid cell and σ^2 is the interannual S^{ox} variance.

Here, we assume that the errors in the $[O_2]$ and the mass subduction have no spatial correlation. Then the covariance is neglected and the error propagation associated to the integration in density classes can be expressed as the sum of the individual S^{ox} uncertainties at every given grid point. We know that this assumption is incorrect, however, we do not have a reliable way to estimate the correlation scales and such assumptions have been made in similar studies. Given that limitation, we believe that the interannual variability in density classes showed here might be underestimated.

Since the C.V expresses the percentage of variability as compared with the mean value, density classes with small mean S^{ox} will result in abnormally high C.V. To avoid this artifact, we have neglected the density classes with mean S^{ox} smaller than 1 Tmol/Yr^{-1}. We can note that in most cases, the C.V does not overpasses 50% and that the largest uncertainties are not associated with the strong S^{ox} fluxes. Instead, the maximum interannual variability is found in the northern North Pacific (>100%). High C.V. values are also associated with the STMW of the Atlantic Ocean and tropical waters in the Pacific Ocean.

The data distribution by density class and basin (Figure S2) suggests that the relative contribution of the different water masses to the global oxygen uptake is well represented. However, the uneven distribution of the oxygen sampling constitutes an important source of uncertainty, which restricts the interpretation of the results of this, and any other study on oxygen at global scale.

S2 Validation

To validate the results obtained with the reanalysis ECCOv4, we provide an alternative computation of S^{ox} (Figure S3). The lateral induction and vertical velocity terms

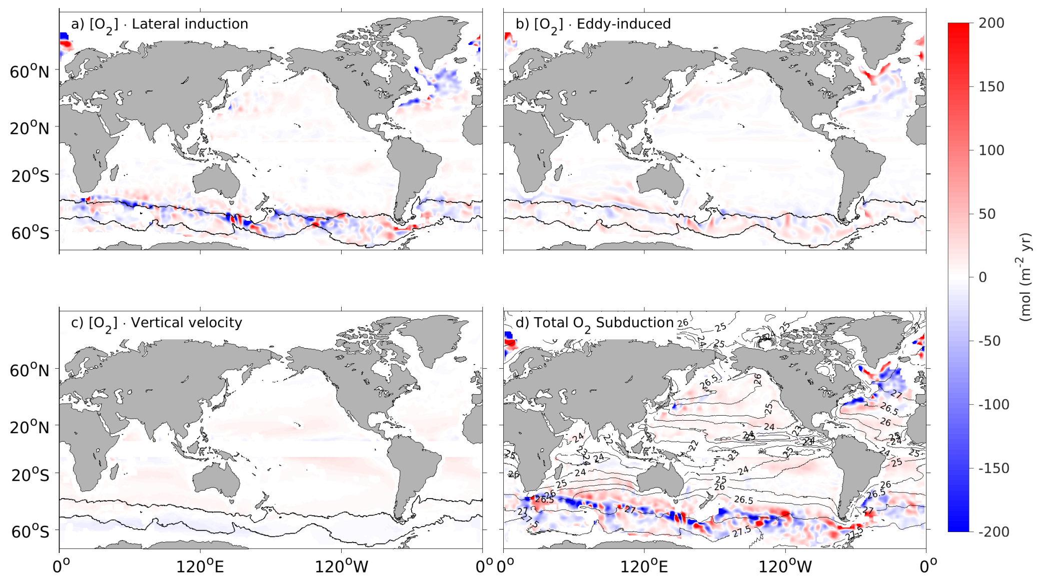

Figure S3. S^{ox} terms as computd from ISAS. a) Lateral (O_2) induction, b) (O_2) eddy-induced subduction computed from ECCOv4 bolus velocity, c) (O_2) vertical velocity and d) Total (O_2)subduction. Contours in a) indicate the mean ACC limits represented by the outermost closed streamlines through the Drake Passage. Contours in d) represent the mean position of the isopycnals on the deepest climatological MLD over the period 2006-2015.

were obtained from the Argo-gridded product "In situ analysis System" (ISAS) (Figure S3)(doi:http://doi.org/10.17882/52367). ISAS is an optimal interpolated product of the Argo global data set. The data used here comprise the period 2006 - 2015, when the data coverage is globally satisfying. All variables are reconstructed on 152 depth levels ranging from 0 to 2000 m. Due to the impossibility to obtain consistent and reliable estimates of the bolus velocity from in-situ observations, we used the bolus velocity from the ECCOv4 outputs to compute the eddy-induced term showed in Figure S3. Since this term is not critical to the total S^{ox} we consider our computation to be suitable for the validation purposes.

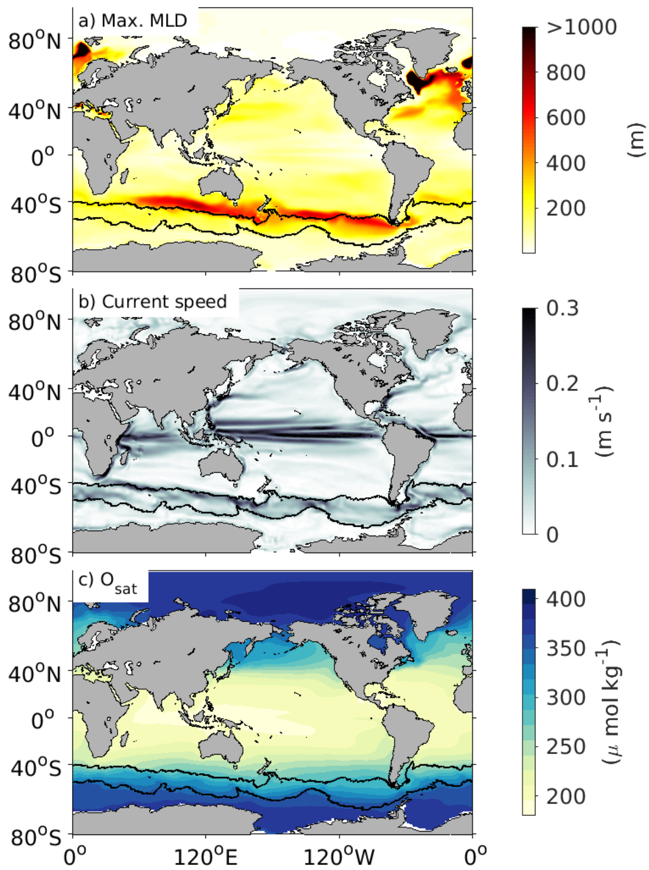

Figure S4. a) Maximum climatological MLD, b) Global geostrophic currents speed c) O_{sat}

To assess lateral induction from ISAS, we previously computed the geostrophic velocity field relative to 1000 m from hydrographic data. The mean reference velocity at that depth level, was obtained from ANDRO (doi:10.17882/47077), an Argo-based deep displacement dataset.

Following (Marshall et al., 1993) The vertical velocity was approximated by using the linear vorticity balance (Sverdrup balance) as follows:

$$w_H = w_{Ek} + \frac{\beta}{f} \int_{-H}^{0} v dz \tag{3}$$

Where w_{Ek} is the Ekman Pumping, v is the meridional component of velocity and β is the gradient of the planetary vorticity (f). Since the Ekman pumping cannot be computed within the equatorial strip, the surface between 5°S and 5°N was blanked.

Figure S3 shows a general agreement between the S^{ox} as computed with both ECCOv4 and ISAS. The main differences are due to the small-scale structures that arise, mainly in the Southern Ocean in the S^{ox} computed from ISAS. However, the main hot-spots and global features are well represented by both products with similar magnitude.

S3 Mixed layer, current speed and oxygen distribution

S^{ox} is a complex mechanism that results from the contribution of different processes (Equation 1 of the main manuscript). In this section, we provide more insight on the drivers of $S^{ox}$by showing the average distribution of (i) the late winter mixed layer depth (MLD), (ii) the global horizontal current speed at the ML base and (iii) the oxygen concentration in equilibrium with the atmosphere (O_{sat}) (Figure S4). The late winter MLD and the current

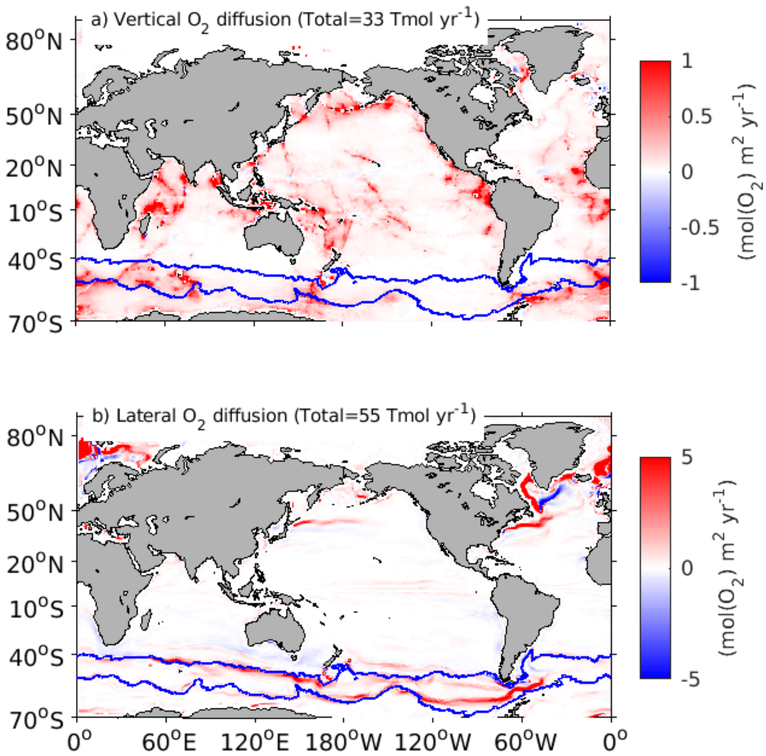

Figure S5. a) Vertical and b) lateral oxygen diffusion. The total amount of oxygen subducted by vertical and lateral diffusion are indicated in the panel's title.

speed are key components of lateral induction (Equation 1). This term shapes the total S^{ox} and is responsible of the main hot-spots of oxygen uptake and release in the global ocean. Figure S4(a, b) shows that the combination between large MLD gradients and strong currents (the ACC) explains the strong lateral induction in the Southern Ocean. In contrast, in the North Atlantic, (Labrador and Irminger seas) and in the Nordic Seas, the currents are less intense than the ACC, but the MLD gradient is the largest of the entire ocean, resulting in the highest subduction rates.

The distribution of O_{sat} (Figure S4c) is largely driven by the sea surface temperature. Consequently, the largest O_{sat} values are found at mid-high latitudes where the seawater is colder.

S4 Vertical and Lateral oxygen diffusion

Vertical oxygen diffusion was computed by using a geographically-variable vertical diffusion coefficient (k_v) based on a parametrisation of tidally-driven mixing (de Lavergne et al., 2020). k_v is determined at the base of the mixed layer.

Lateral diffusion depends on data resolution and on the effect of eddies and it shows great variability between datasets. In this study we have used a constant mean value for the lateral diffusion coefficient (k_l) to compute the lateral oxygen diffusion (Figure S5). With the coefficients utilised in this study, lateral oxygen diffusion represents almost two thirds of the total oxygen diffusion, while the vertical component provides about a third of the oxygen diffused into the ocean interior.

S5 Particular case of the Subpolar Gyre

Application of the subduction concept to a regions were the large scale flow does not support the general shallowing of mixed layers in the direction of flow (As is the case of the Subplolar Gyre in the North Atlantic) pose a high degree of complexity.

To get more insight on the S^{ox} in this complicated region, we show a section that follows a contourline along the circulation in the Subpolar Gyre (Figure S6)

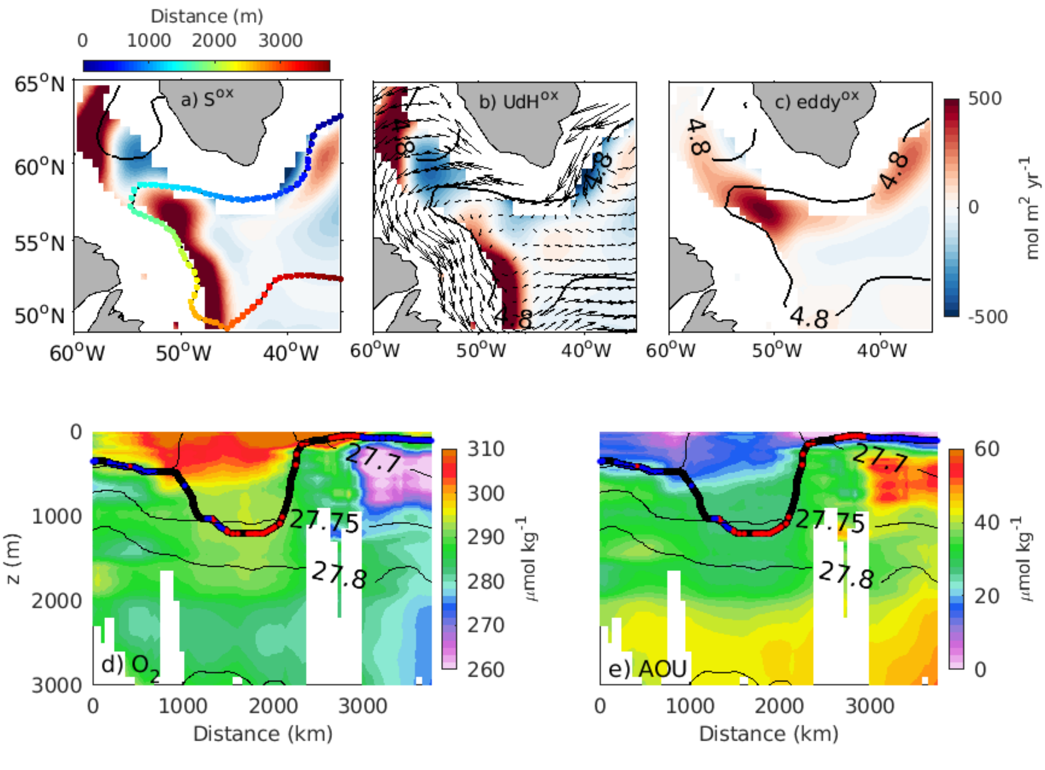

Figure S6. Sections following the circulation in the subpolar Gyre. a) S^{ox} and the contour followed where the distance is color coded. b)(O_2) lateral induction with the currents superimposed, c) (O_2) eddy-induced subduction. d) oxygen and e) AOU sections. thin contours are the isopycnals while thick contour represent the ML base. The colours (red and blue) on the ML base indicate the subductive and obductive regions respectively

References

de Lavergne, C., Vic, C., Madec, G., Roquet, F., Waterhouse, A. F., Whalen, C. B., ... Hibiya, T. (2020). A parameterization of local and remote tidal mixing. *Journal of Advances in Modeling Earth Systems*. doi: 10.1029/2020ms002065

Marshall, J., Williams, R. G., & Nurser, A. J. G. (1993). *Inferring the Subduction Rate and Period over the North Atlantic* (Vol. 23) (No. 7). doi: 10.1175/1520-0485(1993)023⟨1315:ITSRAP⟩2.0.CO;2