Supporting Information for:
"Physical mechanisms driving the global ocean breathing"

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

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

S1 Mixed layer, current speed and oxygen distribution

The oxygen subduction (S^{ox}), is a complex mechanism resulting from the sum of different individual processes (Equation 1). In this section, to provide more insight on the drivers of S^{ox}, we show the average distribution of (i) the maximum mixed layer depth (MLD) at each grid point, (ii) the global horizontal current speed at the MLD and (iii) the oxygen concentration in equilibrium with the atmosphere (O_{sat}) (Figure S1).

The maximum MLD and the current speed are key factors for the lateral induction (Equation 1), which overall is the main driver of the total oxygen flux across the mixed layer base. Figure S1(a, b) shows how in the Southern Ocean there is a combination between MLD gradients and strong currents (the ACC) that explains the strong lateral induction. In contrast, in the North Atlantic, (in the 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 largest subduction rates.

The O_{sat} distribution (Figure S1c) is largely driven by the temperature at the sea surface. Consequently, the largest O_{sat} values are found at mid-high latitudes, together with the maximum S^{ox}. Nevertheless, the $[O_2]$ has shown to have a minor role in the total oxygen uptake which is dominated by the mass subduction flux across the mixed layer base.

S2 Validation

To validate the results obtained with the reanalysis ECCOv4, we provide a plot of the total subduction and its three main components. The lateral induction and vertical velocity terms were obtained from the Argo-gridded product "In situ analysis System" (ISAS) (Figure S2)(doi:http://doi.org/10.17882/52367). ISAS is an optimal interpolated product of the Argo global data set that we used between 2006 and 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 S2. Since this term is not critical to the total S^{ox} we consider our computation to be suitable for the validation purposes.

To compute the 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.

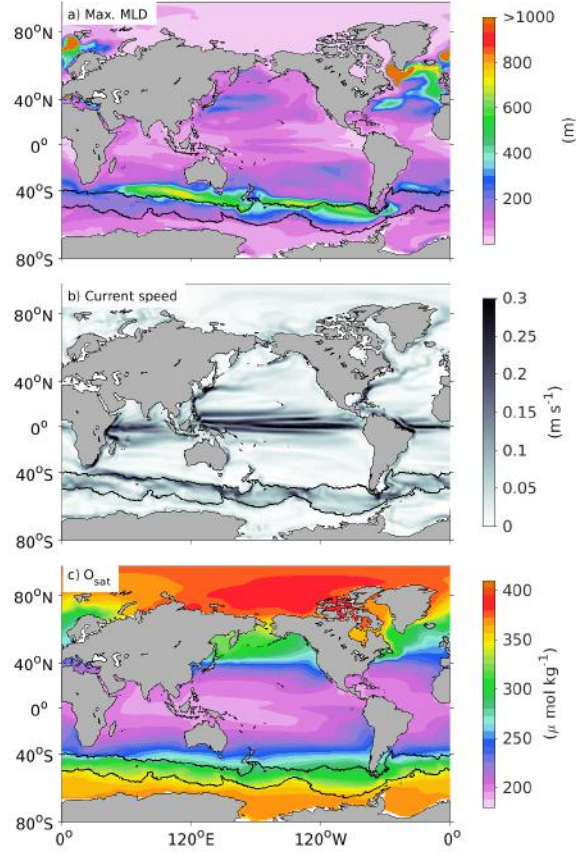Corresponding author: Esther Portela, `eportelanh@gmail.com`

Figure S1. Fields related to the oxygen subduction. a) Maximum climatological MLD, b) Global geostrophic currents speed c) O_{sat}

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

$$w_H = w_{Ek} + \frac{\beta}{f} \int_{-H}^{0} v dz \qquad (1)$$

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 S2 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 also similar magnitude.

S3 Uncertainty Computation

Some 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. The omitted error sources are the following:
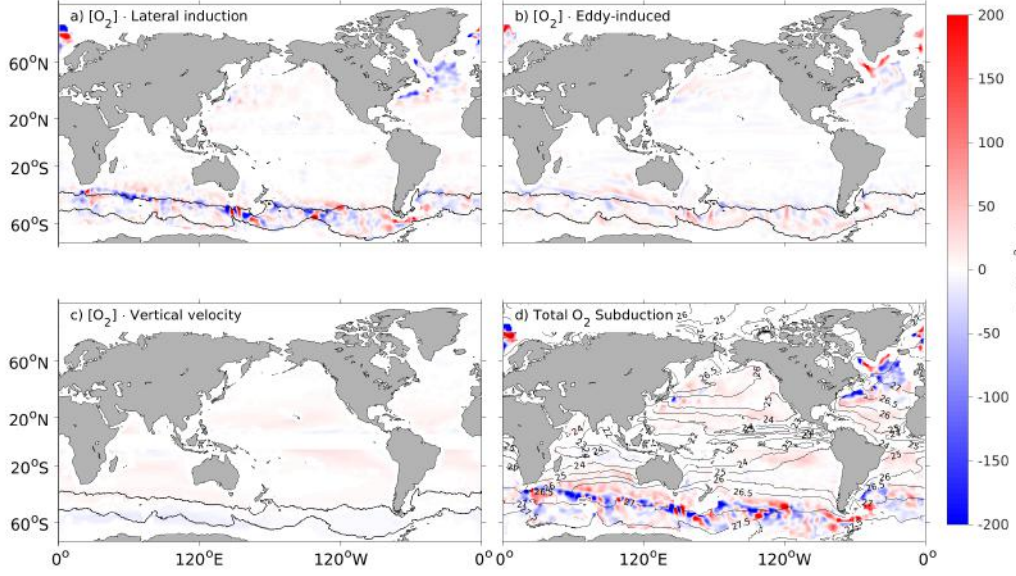
Figure S2. Oxygen subduction terms as computd from ISAS. a) Lateral (O_2) induction, b) (O_2) eddy-induced subduction computed from ECCOv4 bolus velocity, c) (O_2) vertical velocity and d) Total (O_2)subduction. Contours in a) indicate the mean ACC limits represented by the outermost closed streamlines through the Drake Passage. Contours in d) represent the mean position of the isopycnals on the deepest climatological MLD over the period 2006-2015.

1. In this study we have used two datasets which monthly climatology have different temporal coverage. While ECCOv4 extends from 1992 to 2015, the WOA18 oxygen climatology includes data collected from 1955 to 2018. This inconsistency is a source of uncertainty that we were not able to overcome and that needs to be taken in account for the interpretation of the results. However, due to the relative little role that the distribution of $[O_2]$ plays in the total 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 error associated to the O_2-sensor calibration (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 reanalysis 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 S3a, b), and to the interannual variability of the $[O_2]$ 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 a primary 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 shown in Figure S3a represents every single data collected between 1955 and 2018. We can see that the oxygen sampling has been historically uneven, typically, northern latitudes concentrate most of the data, mainly near the cost. The North Pacific and the northern North Atlantic count with approximately with four times more

data than the southern Hemisphere basins (Figure S3b) where observations are sparse. On the other hand, the distribution among density classes at each basin is acceptable suggesting that the relative contribution of the different water masses to the global oxygen uptake is well represented. Nevertheless, even if the whole ocean has been covered, the oxygen data are badly distributed which restricts the interpretation of the results of this, and any other study on oxygen at global scale.
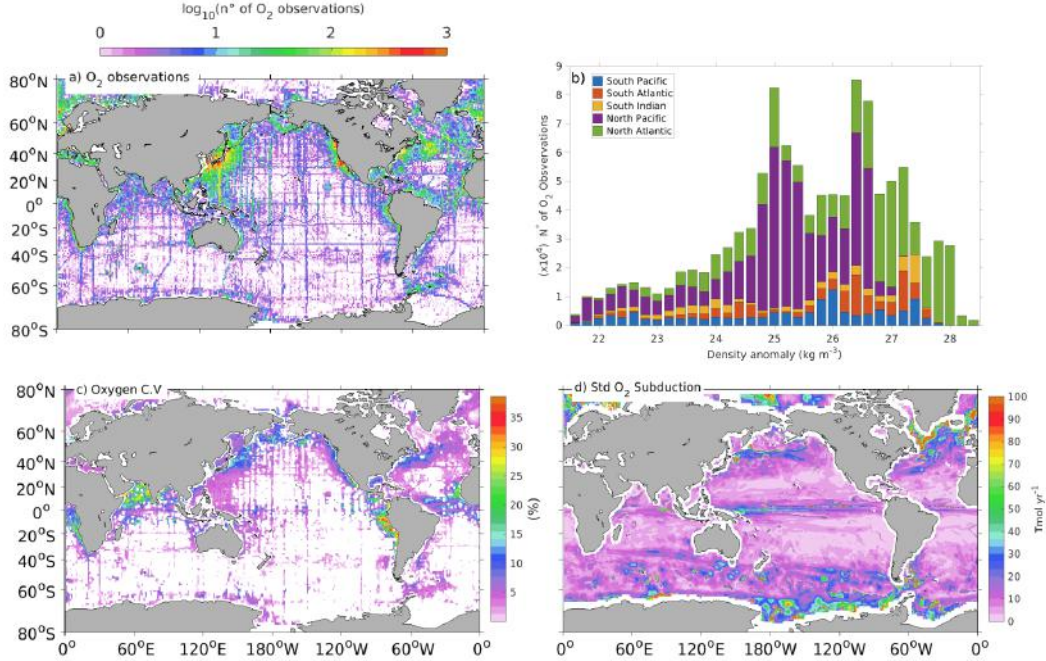

Figure S3. Uncertainty associated to the oxygen subduction components and to the oxygen sampling. a). b) Number of historical oxygen observations contained in each density class and basin. c) Maps of the coefficient of variance (C.V, (%)) for the oxygen as obtained from WOA18. d) Map of the interannual standard deviation of the S^{ox}

The uncertainty associated to the interannual variability of the sparse oxygen data is expressed by the coefficient of variance (C.V. in %) of the $[O_2]$ (Figure S3c) (C.V $= 100 \cdot \sigma/\overline{x}$, where σ is the standard deviation and $\overline{x}$ is the mean of all the available data within each grid cell between 1955 and 2018). The C.V allows to have an estimate of the data variability that is not affected by the mean value (with an equal C.V, locations with greater means will probably have larger 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 value. In addition to the tropical regions as the North Indian and the eastern tropical Pacific, relatively high interannual oxygen variability is found in the North Pacific and subtropical North and South Atlantic. The C.V somehow reflects the data distribution (Figure S3a) 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. Despite this results, due to the dominant role that the mass subduction flux plays on the total S^{ox} we believe that the sparse oxygen sampling is a minor contributor to the total S^{ox} uncertainty. 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:

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

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 S3d. 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.

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 (ii) 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.
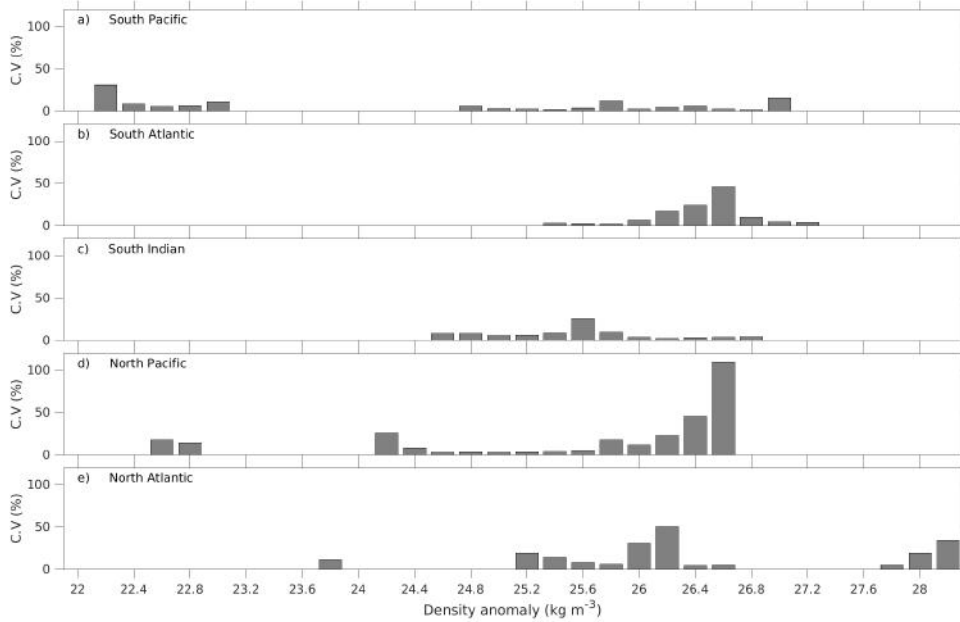

Figure S4. Coefficient of variation (C.V (%)) that illustrates the uncertainty due to the interannual variability of of the S^{ox} and the sparse oxygen sampling for every density class and basin. Density classes with S^{ox} lower than 1Tmol/Yr^{-1} where not taken into account because their C.V results in abnormally high values

To link the reported uncertainty maps with our results, we show the C.V associated to the S^{ox} integrated in density classes (Figure S4). To obtain the C.V, the standard deviation of the S^{ox} as obtained from Equation (2) was propagated following the next equation:

$$[\sigma_m]_{(S^{ox})} = \sqrt{\sum_{i=1}^{n}(a_i^2 \cdot \sigma_i^2) + 2ab \cdot cov_{(i=1:n)}} \tag{3}$$

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 term 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.
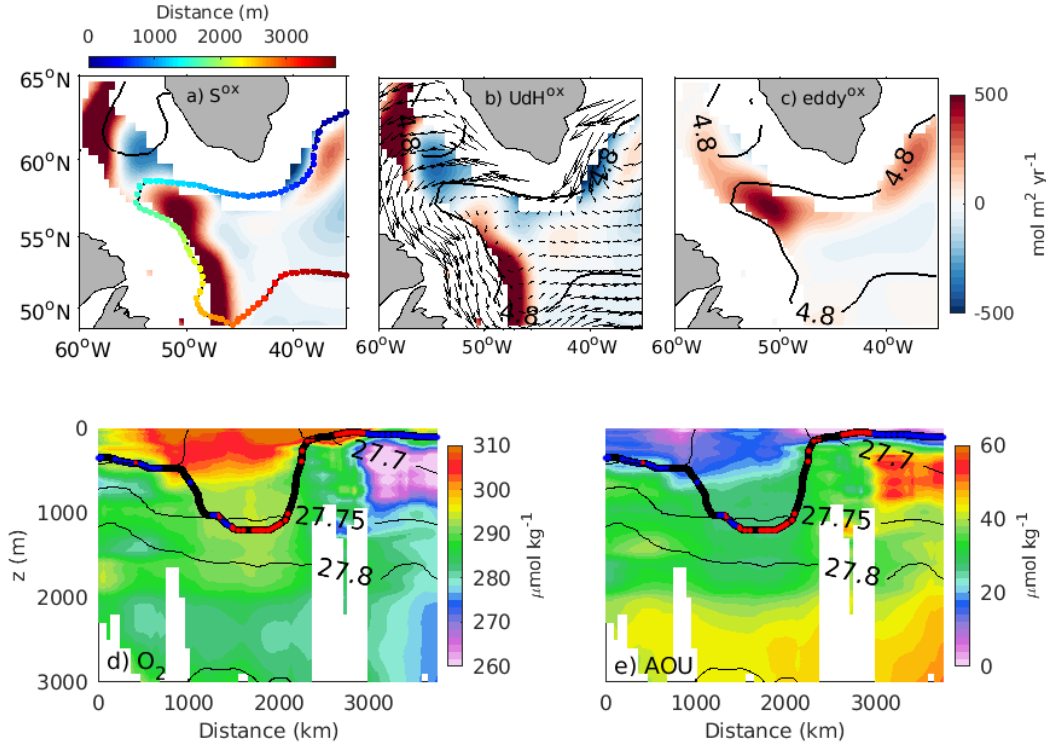

Figure S5. Sections following the circulation in the subpolar Gyre. a) S^{ox} and the contour followed with distance color coded. b)(O_2) lateral induction with the currents superimposed, c) (O_2) eddy-induced subduction. d) oxygen and e) AOU sections. thin contours are the isopycnals while thick contour is the ML base. The colours (red and blue) on the ML base indicate the subductive and obductive regions respectively

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 do 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.

S4 Particular case of the Subpolar Gyre

Application of 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 subduction estimates. To shed more light on the S^{ox} in this complicated region, we show a section that follows a contourline along the circulation in the Subpolar Gyre (Figure S5)

S5 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.
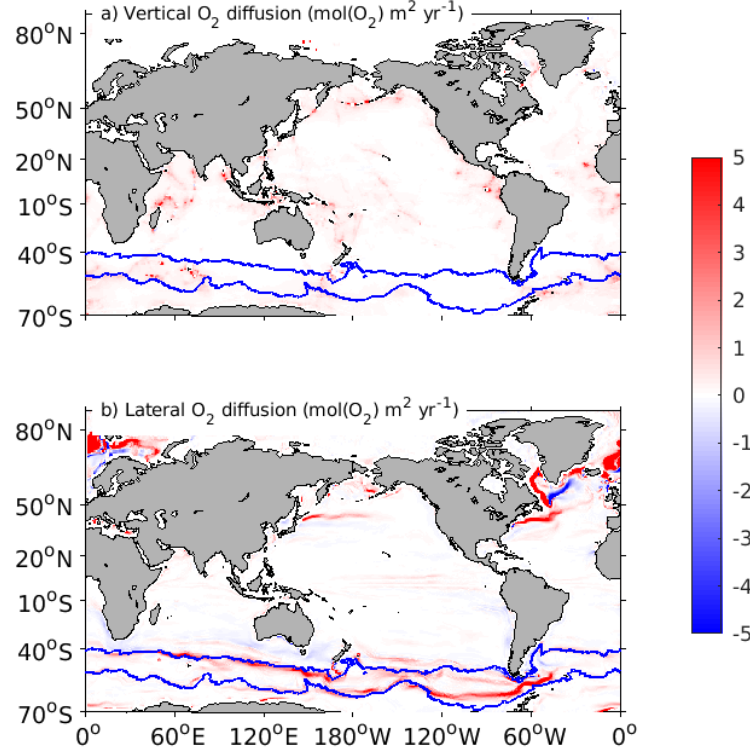

Figure S6. a) Vertical and b) lateral oxygen diffusion.

Lateral diffusion depends on data resolution and on the effect of eddies, it shows great variability between datasets and its magnitude is also related with that of the bolus velocity. In this study we used a constant mean value of the k_l coefficient to compute the lateral oxygen diffusion showed in Figure S6

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