AGU PUBLICATIONS

Global Biogeochemical Cycles

RESEARCH ARTICLE

10.1002/2015GB005354

Key Points:
- OUR, a standard method in ocean biogeochemistry, underestimates ocean respiration severely
- The underestimate is due to the uneven distribution of respiration in the ocean and its patterns
- The degree of underestimate depends on the relative importance of advection versus diffusive mixing

Supporting Information:
- Supporting Information S1

Correspondence to:
W. Koeve,
wkoeve@geomar.de

Citation:
Koeve, W., and P. Kähler (2016), Oxygen utilization rate (OUR) underestimates ocean respiration: A model study, *Global Biogeochem. Cycles*, *30*, doi:10.1002/2015GB005354.

Received 11 DEC 2015
Accepted 7 JUL 2016
Accepted article online 17 JUL 2016

Oxygen utilization rate (OUR) underestimates ocean respiration: A model study

W. Koeve[1] and P. Kähler[2]

[1]Biogeochemical Modelling, GEOMAR Helmholtz-Zentrum für Ozeanforschung, Kiel, Germany, [2]Chemical Oceanography, GEOMAR Helmholtz-Zentrum für Ozeanforschung, Kiel, Germany

Abstract We use a simple 1-D model representing an isolated density surface in the ocean and 3-D global ocean biogeochemical models to evaluate the concept of computing the subsurface oceanic oxygen utilization rate (OUR) from the changes of apparent oxygen utilization (AOU) and water age. The distribution of AOU in the ocean is not only the imprint of respiration in the ocean's interior but is strongly influenced by transport processes and eventually loss at the ocean surface. Since AOU and water age are subject to advection and diffusive mixing, it is only when they are affected both in the same way that OUR represents the correct rate of oxygen consumption. This is the case only when advection prevails or with uniform respiration rates, when the proportions of AOU and age are not changed by transport. In experiments with the 1-D tube model, OUR underestimates respiration when maximum respiration rates occur near the outcrops of isopycnals and overestimates when maxima occur far from the outcrops. Given the distribution of respiration in the ocean, i.e., elevated rates near high-latitude outcrops of isopycnals and low rates below the oligotrophic gyres, underestimates are the rule. Integrating these effects globally in three coupled ocean biogeochemical and circulation models, we find that AOU-over-age based calculations underestimate true model respiration by a factor of 3. Most of this difference is observed in the upper 1000 m of the ocean with the discrepancies increasing toward the surface where OUR underestimates respiration by as much as factor of 4.

1. Introduction

The steady state oxygen concentration at any location in the ocean is maintained by the balance of oxygen supply and consumption. Supply is by the advective and diffusive transport of surface water having picked up oxygen by exchange with the atmosphere; consumption is by the respiration of organic matter produced in surface water and reaching the interior of the ocean as sinking particulate organic matter (POM) and by convective, advective, and diffusive transport of dissolved organic matter (DOM).

In the interior ocean respiration proceeds at very low rates making its direct measurement a challenge. Current estimates rely on indirect approaches that determine the rate of oxygen consumption in subsurface water by dividing the change in apparent oxygen utilization (AOU) along a given distance on an isopycnal surface by the change in age along the same distance, see equation (1). AOU is the oxygen debt having accumulated since the last contact of the respective water with the atmosphere, i.e., it is the difference between oxygen concentration at equilibrium with air (a function of temperature and salinity) and the actual oxygen concentration: $AOU = [O_2{}^{sat}\ (\theta,\ S)] - [O_2]$, with S = salinity and θ = potential temperature [*Redfield*, 1942; *Redfield et al*., 1963; *Pytkowicz*, 1971]. Water age (t in equation (1)) is the time elapsed since this equilibration when age is set to zero. In practice, the slope of a linear regression of AOU against water age is taken to represent the (mean) oxygen utilization rate (OUR) over the distance concerned. OUR calculated in this way is considered to represent subsurface oxygen consumption. This has currently the status of textbook knowledge [*Sarmiento and Gruber*, 2006; *Emerson and Hedges*, 2008].

$$OUR = \mathrm{d}AOU/\mathrm{d}t = \mathrm{d}\left(\left[O_2{}^{sat}\right] - \left[O_2{}^{obs}\right]\right)/\mathrm{d}t \quad (1)$$

[4] OUR (and derived and related rates) are prominent supports of today's view of ocean biogeochemistry. Organic matter transfer from surface to deep water has been estimated by equating it with the vertical OUR integral, i.e., total respiration equal to export production [*Jenkins*, 1982] which itself is, on appropriate scales of space and time, equal to new production [*Eppley and Peterson*, 1979; *Oschlies and Kähler*, 2004]. The vertical attenuation of this transfer by organic matter breakdown within the ocean, i.e., the depth function of respiration, is read from the depth function of OURs on stacked isopycnals [*Jenkins*, 1982; *Sarmiento and Gruber*, 2006]. By using OUR, extremely low respiration rates are accessible.

OUR-based estimates of subsurface respiration have contributed significantly to the (ongoing) debate about the productivity and the overall autotrophy or heterotrophy of the oceans [*Jenkins*, 1982, 1987; *del Giorgio and Duarte*, 2002; *Williams et al*., 2004; *Burd et al*., 2010]. In these and many other studies, discrepancies have been pointed out between various measures of organic matter production and consumption. Since OUR is determined on large scales of space and time, many of the uncertainties associated with surface layer rate-based estimates of export production, such as the seasonality or patchiness of production, are avoided. For this reason an OUR approach is often preferred to other methods [*Kähler et al*., 2010]. Also, OUR does not exclude the contribution of dissolved organic matter to the total as it is the case with organic matter flux determinations using particle interceptor traps.

Despite these advantages, the necessity in the OUR approach of a suitable method of age determination has restricted its more widespread use. A wide range of respiration rates, covering nearly an order magnitude, can result from different age determination methods [*Sarmiento et al*., 1990]. Common age tracers are properties from which age can be calculated given certain assumptions [*Sonnerup*, 2001; *Matsumoto*, 2007; *Waugh et al*., 2003; *Koeve et al*., 2015]. In the real ocean there is no tracer, which provides an ideal clock running from the moment when water has lost contact with the atmosphere, i.e., the time from which onward AOU accumulates. However, two age tracers with independent time history may be combined to fully constrain the transit time distribution (TTD) [*Hall et al*., 2002; *Waugh et al*., 2003], which in turn allows to compute a mean age which is closer than other age tracer estimates to the steady state ideal age of ocean models. It is this mean age which has recently been applied in OUR studies from the Pacific Ocean [*Sonnerup et al*., 2015].

Another issue is that AOU has been shown to overestimate oxygen utilization because the assumption of perfect equilibration ("100% saturation") of oxygen between air and surface water, an assumption made when computing AOU, does not universally hold. AOU has therefore sometimes been substituted by TOU (true oxygen utilization) [*Ito et al*., 2004], a model tracer which accounts for oxygen saturation states other than 100% (mostly lower) in surface water at the time of water mass formation. An overestimate of AOU as compared with TOU was particularly observed in the deep ocean, at least in several global ocean models [*Duteil et al*., 2013]. Globally, this overestimate can be equivalent to up to 25% of the AOU inventory. When computing OUR from changes of AOU and age (i.e., equation (1)), however, uncertainties arising from the preformed-oxygen assumption and also some issues related to age (e.g., preformed age [*Koeve et al*., 2015]) cancel out and become irrelevant.

There is, however, a more fundamental issue regarding the calculation of OUR, which is usually overlooked. AOU, from which OUR is calculated, is not representative of total respiration, but only the fraction of oxygen consumption which has been preserved in the ocean; the other fraction having been neutralized by O_2 equilibration at the surface. An important fraction not being preserved is indicated by the well known global mean vertical distributions of AOU (decreasing toward the surface) and respiration (increasing toward the surface) [e.g., *Martin et al*., 1987, their Figure 6]. While most of the respiration takes place in the upper ocean, the larger fraction of AOU is found deeper in the ocean interior. Similarly with water age: water ages everywhere with the same rate, but the vertical distributions of water age show an increase with depth. The partial removal of the effects of respiration on AOU and of aging on water age are brought about by transport of surface water (AOU = 0, age = 0) into the ocean's interior. The gradients of AOU and age from which OUR is calculated are therefore influenced by both respiration and equilibration on the AOU side, and aging and the zeroing of age as regards water age, and the dispersion of these effects. OUR, meant to be a gross rate, is calculated from a net effect. Currently, the use of the OUR concept is based on the implicit assumption that advective and diffusive transports affect AOU and age in the same way, so as not to affect OUR estimates. But is this assumption valid?

In this study, we use a simplified representation of an isolated isopycnal in the ocean (1-D tube model, section 3) and a suite of 3-D coupled ocean circulation and biogeochemical models (section 4) to resolve this issue conceptually and quantitatively.

2. Approach

We use different ocean circulation and biogeochemical models in which oxygen is affected by air-sea exchange with the atmosphere, implicit or explicit photosynthetic production in surface water, and respiration in the interior of the ocean. The models were augmented with additional tracers which allow to follow

the accumulation and transport of the oxygen debt in the interior ocean. The tracer of oxygen consumption accumulates any O_2 deficit produced by organic matter breakdown in the interior of the ocean, is transported with the model's circulation, and is destroyed (set to zero) when the respective water reaches the surface of the model ocean. This tracer thus provides an estimate of the model's true oxygen utilization (TOU). By using a TOU tracer instead of diagnosed AOU in this study, we avoid any ambiguities potentially arising from O_2 not reaching equilibrium with the atmosphere in some ocean regions [*Duteil et al*., 2013]. The TOU tracer has the additional advantages of (a) behaving conservatively with respect to diffusive mixing and (b) that there is no artificial modification of the TOU tracer from, e.g., subsurface warming as it has been observed for AOU [*Dietze and Oschlies*, 2005].

For the determination of the time elapsed since the last contact with the atmosphere we use a tracer of water age behaving ideally (ideal age tracer [*Thiele and Sarmiento*, 1990; *England*, 1995; *Koeve et al*., 2015]). This tracer works like a clock adding up time after having been restored to zero when the respective water parcel resided at the surface. Everywhere else it ages with a rate of 1d/d and is subject to mixing and advection in the interior of the ocean. In this way uncertainties arising from the use of imperfect age tracers are avoided. The ideal age tracer behaves ideally since the age distribution derived from it is not distorted by mixing.

We use the output (TOU tracer, ideal age tracer) of the model runs after steady state is attained after 6000 years. OUR is diagnosed from regressions of TOU against age (equation (1)).

In any of the models the local oxygen consumption is known at any time and place. In our experiments with a simple 1-D model (see next section) a distribution of respiration is imposed and therefore also known. In the global biogeochemical models (section 4) local oxygen consumption is the biogeochemical oxygen sink computed at model runtime from the concentrations of organic matter (variable in time and space) and the models' rules (rate constant for organic P decay and a O_2:P ratio) as long as oxygen is available. The oxygen sink terms are added up over 1 year and stored in the three-dimensional field of what we call the "true" model-O_2-consumption rate in the following. Comparing "diagnosed" OUR (from equation (1)) with the true model-O_2-consumption rate (or the imposed respiration in the 1-D case) allows us to examine whether the OUR analysis reproduces respiration.

3. Diagnosing Respiration in a 1-D Tube Model Representing a Vertically Isolated Isopycnal

OUR is usually determined on isopycnal surfaces [*Jenkins*, 1982; *Sonnerup et al*., 1999]. To investigate the effects of isopycnal transport processes, we construct a simple 1-D tube model (Figure S1 in the supporting information) with unidirectional flow (similar to that in *Sarmiento et al*. [1990]) to represent a vertically isolated isopycnal. The tube model consists of 100 aligned compartments. Air-sea exchange of oxygen and resetting of the TOU and age tracers to zero occurs in the two outermost (outcrop) compartments. Respiration occurs in all boxes at imposed (and therefore known) rates. We test how well these imposed rates can be reproduced by diagnosing OUR from the regression of TOU and the age tracer over a number of compartments. There is no diapycnal exchange (water is enclosed like inside a tube). Recorded tracers are O_2, TOU, and ideal age. Tracer concentrations (C) in the tube are affected by advection and/or diffusion. At steady state ($dC/dt = 0$) the sink-minus-source (sms) term is balanced by advective and diffusive transports according to equation (2):

$$\mathrm{sms} = -v\, dC/dx + K\, d^2C/dx^2 \tag{2}$$

with v and K being the along-isopycnal velocity and diffusivity, respectively. For oxygen and age the sms terms are respiration and aging, respectively. In the outcrop compartments the initial oxygen concentration is set to a fixed value mimicking air-water equilibrium (at a given surface temperature and salinity) and is restored to this value at each time step (i.e., we assume instant and complete gas exchange at constant temperature and salinity).

The imposed respiration in the other boxes successively consumes oxygen and increases the value of the TOU tracer. Water ages uniformly (1d/d) in all compartments. We study four different distributions of the respiration rate which are designed in a way that total respiration (between compartments 2 and 99) is the same in all cases; it is only its spatial distribution which varies. In a first case (EVEN) the imposed respiration rate is constant; i.e., it has the same value in every compartment. For the second case (DISTANT), we prescribe a distribution of respiration, which is closer to the reality in the ocean, i.e., with high rates near

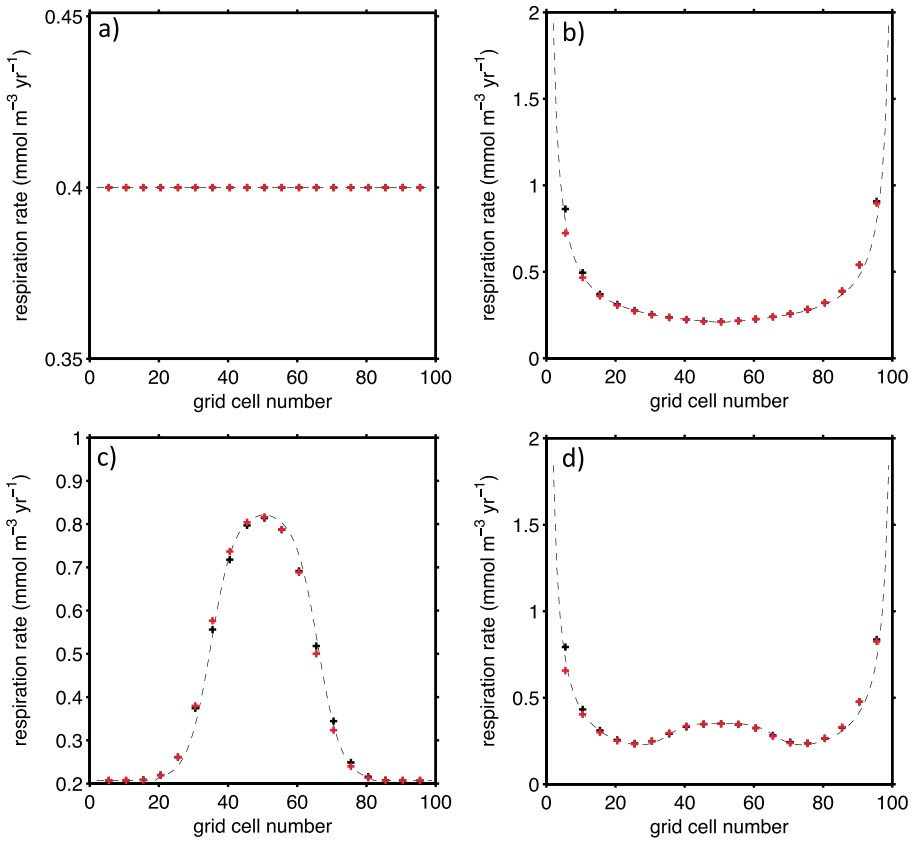


Figure 1. One-dimensional tube experiments: advection only. In this set of experiments with the tube model unidirectional advection is the only transport process. The flow is from left to right. Water subducts in grid cell 1 and upwells in cell 100 (see Figure S1 for a schematic of the tube model). Imposed (dashed line) and diagnosed (red cross) respiration rates for $K_H = 0\,m^2\,s^{-1}$, $A_H = 0.3\,cm\,s^{-1}$. Integrated respiration for cells 2 to 99 is the same in all experiments. Black cross denotes imposed respiration averaged over 10 neighboring grid cells. (a) EVEN, constant imposed respiration rate, (b) DISTANT, imposed respiration high near outcrops, (c) CENTRAL, imposed respiration high at the center of the tube, and (d) COMBINATION, imposed respiration high near the outcrops and slightly elevated near the center.

the "northern" and "southern" ends and a minimum in the center. A reverse distribution with maximum respiration in the middle and minima toward the distant boxes is imposed for the third case (CENTRAL). Finally, we combine the latter two cases in a distribution of respiration mimicking conditions on an isopycnal in the ocean yet more closely (COMBINATION), i.e., with high rates toward the outcrops (where a real isopycnal is at shallow depth and below a highly productive surface system), slightly elevated respiration in the center (where a real isopycnal is under the influence of a productive equatorial upwelling ecosystem and also shoals relative to its depth in the subtropics), and low respiration rates in between (real isopycnal deep below the oligotrophic gyres). In all experiments model output is analyzed after the model has reached steady state, which is checked by the time trajectory of the tube's O_2 minimum. We diagnose (regional) OUR for 10 (9 for outermost parts of the tube because the outcrop compartments are excluded) neighboring compartments each (e.g., for compartments 2 to 10, 6 to 15, 11 to 20, …, 86 to 95, 90 to 99) and compare these estimates with the mean of the imposed respiration rates of the respective compartments. We apply the four different cases in a series of experiments, with either advection only, diffusion only, and combinations of both.

In a first set of experiments, we consider advection to be the only transport process. This version of the tube model, with advection only and no diffusion, is in fact a representation of the textbook OUR concept [*Sarmiento and Gruber*, 2006; *Emerson and Hedges* 2008], the "box car" type [*Sarmiento et al.*, 1990]. Here OUR yields an acceptable estimate of the imposed respiration rate (Figure 1). Advection does not affect the proportion of the gradients of TOU and age since it transports TOU and age in the same way. However, it shifts the tracers in the direction of the advective transport, causing a slight misplacement of the estimate relative to the distribution of imposed respiration. This misplacement is larger at higher velocities (not shown).

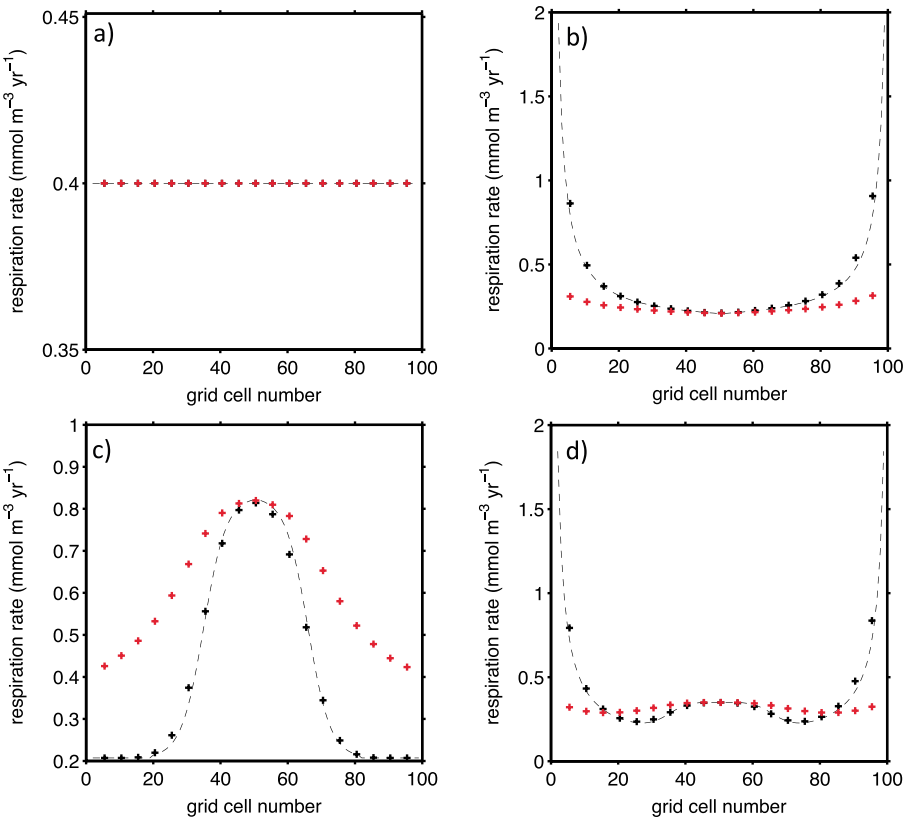


Figure 2. One-dimensional tube experiments: diffusion only. Imposed (dashed line) and diagnosed (red cross) respiration rates for $K_H = 1000\,m^2\,s^{-1}$, $A_H = 0\,cm\,s^{-1}$. Integrated respiration for cells 2 to 99 is the same in all experiments (black cross denotes imposed respiration averaged over 10 neighboring grid cells). (a) EVEN, (b) DISTANT (c) CENTRAL, and (d) COMBINATION. See text for details of the four experiments.

In a second set of experiments (Figures 2 and S2), we consider diffusive mixing between adjacent compartments as the exclusive transport mode (i.e., there is no advection in the tube). In the standard case, we set the respective horizontal eddy diffusion coefficient (K_H) to a value of 1000 m^2/s, consistent with horizontal (latitudunal) scales, which our tube model is to resolve. In the EVEN case, OUR reproduces the imposed respiration rate everywhere (Figure 2a). However, with a distribution of respiration closer to reality (DISTANT, Figure 2b), OUR agrees with the imposed values only in the middle of the tube, while imposed respiration is severely underestimated toward both outcrops. For the CENTRAL case, with maximum respiration in the tube's center, we observe an overestimate almost everywhere (Figure 2c). When combining these two cases in COMBINATION, there is both overestimation and underestimation of diagnosed OUR compared with imposed respiration (Figure 2d).

The results show that when diffusive mixing is dominant, the distribution of the respiration rate is decisive for the degree to which it is reproduced by the OUR. While aging is uniform (i.e., it proceeds with the same rate everywhere) the steady state distribution of age depends solely on the design of the tube and diffusivity both of which are kept unchanged in these four experiments. Hence, age distribution is the same in all experiments with $K_H = 1000\,m^2\,s^{-1}$ and $A_H = 0\,cm\,s^{-1}$. The steady state distributions of both age and TOU are not only a consequence of their production (by aging and respiration), but they are also modified by diffusive transports, which are not the same for age and TOU unless both respiration and aging have the same distribution. This is the case only when respiration is the same everywhere. In steady state, the net diffusion of age and TOU is always directed out of the tube. Elevated respiration (relative to a uniform rate) in the center of the tube will enhance the gradient of TOU relative to that of age: dividing them will result in a higher value almost everywhere. The reverse is true in the opposite case of elevated (again relative to the uniform rate) respiration in the distant parts of the tube, near the outcrop boxes. In this case the TOU gradient inside the tube is not influenced; i.e., the extra respiration does not affect the TOU/age gradient—and is invisible to the OUR method.

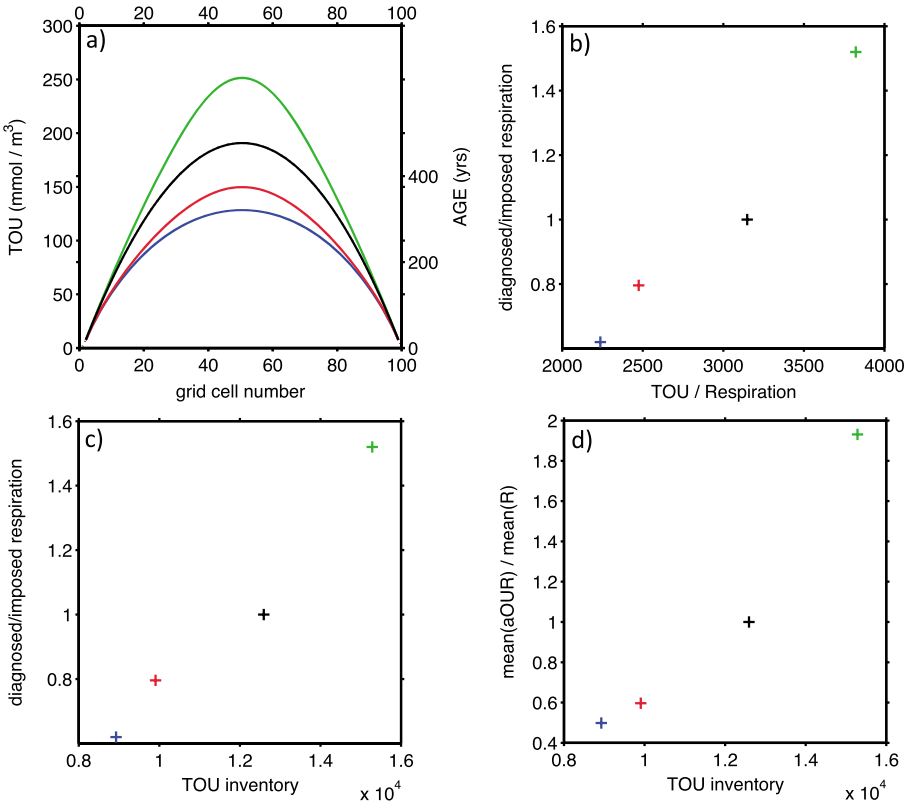
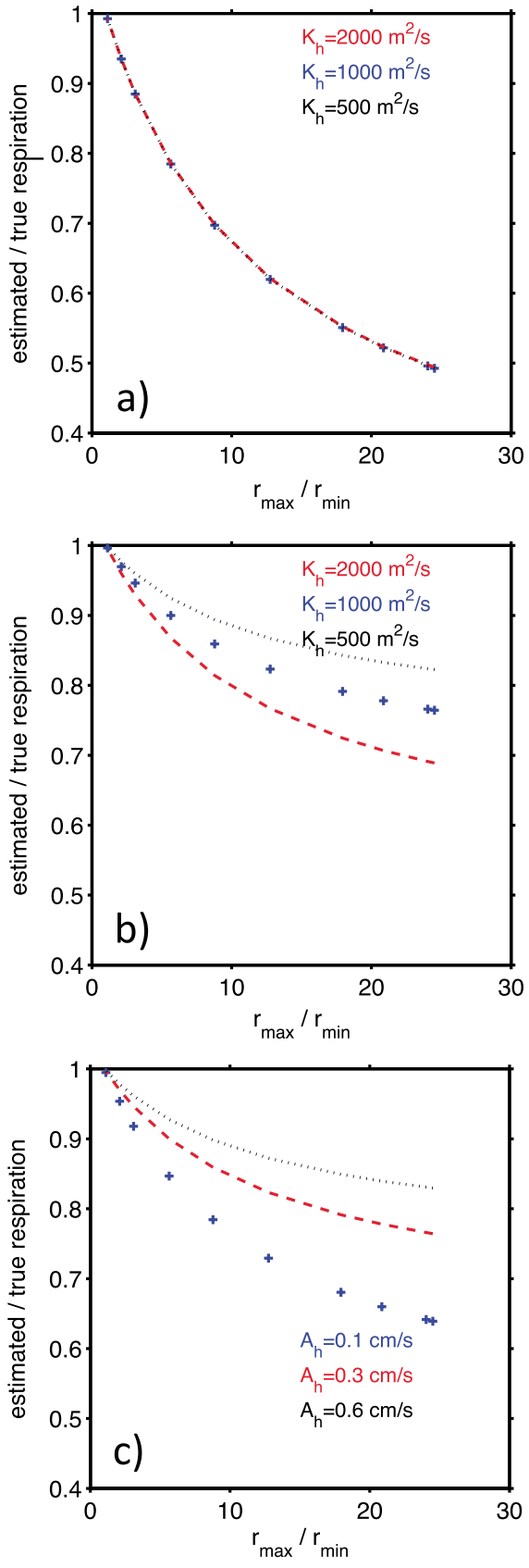


Figure 3. Analysis of several 1-D tube model experiments ($K_H = 1000\ m^2\ s^{-1}$, $A_H = 0\ cm\ s^{-1}$). See Figure 1 for details of these experiments. Color code: EVEN (black), DISTANT (blue), COMBINATION (red), and CENTRAL (green). (a) Distribution of TOU along the tube (left-hand axes) and distribution of age along the tube (right-hand axes). Note that the age distribution is the same in all four experiments. For simplicity we have therefore scaled the age axes relative to the TOU axes such that the black line represents also the age distribution of all four experiments. (b) Relationship between the integrated (cells 2–99) ratio diagnosed/imposed respiration and the integrated ratio stored TOU / imposed respiration. (c) Relationship between the integrated (cells 2–99) ratio diagnosed/imposed respiration and the TOU inventory. (d) Relationship between the ratio mean aOUR / mean imposed respiration and the TOU inventory.

Despite integrated respiration along the tube being the same in all four of experiments (Figure 2), the respective TOU inventories differ considerably (Figure 3a). Highest inventories and largest gradients of TOU are observed with imposed respiration elevated only in the center of the tube (CENTRAL). In this case the TOU inventory is larger than in the case of EVEN respiration in the tube. With elevated respiration toward the outcrops (experiment DISTANT) the TOU inventory is smaller, or the loss of TOU from the tube is larger, compared with that of the tube with constant imposed respiration (EVEN). There is a clear relationship between the ratio of TOU inventory/integrated respiration and the ratio diagnosed/imposed respiration (Figure 3b). When the TOU inventory is smaller than that of the EVEN case, we observe that OUR underestimates true respiration, where it is larger OUR overestimates true respiration (Figure 3c). Excess loss (or storage) of the TOU tracer relative to the EVEN case, and hence, implicitly, relative to that of the age tracer, is thus the key to understand the inability of the OUR approach to correctly estimate oxygen consumption with uneven distributions of respiration in the presence of diffusive mixing.

In the ocean neither uniform rates nor central maxima are expected, but rather maxima near the outcrops as these are situated in more productive regions or at least close to the productive surface. Therefore, we perform several sensitivity runs for the DISTANT case by varying the ratio of maximum imposed respiration (close to the outcrops) and minimum imposed respiration (in the tube's center). We find that the larger this ratio, i.e., the more the distribution of the respiration rate differs from an even distribution, the stronger is the underestimate of oxygen consumption by the OUR approach. This relationship is independent of the imposed horizontal diffusivity (K_h) (Figure 4a). In order to bridge the gap between the tube model results with diffusion only (Figures 2 and 4a) (i.e., Péclet number, $Pe = 0$) and advection only (Figure 1)



(i.e., Péclet number, *Pe* = infinity) we additionally perform experiments in which we combine advection and diffusion. We find that increasing the ratio of advection to diffusive mixing generally reduces the degree of the OUR underestimate (Figures 4b and 4c). In other words OUR is a poor estimate of R in regimes of low *Pe* numbers. The overall relationship between the degree of heterogeneity of imposed respiration and OUR underestimation persists.

Diffusive mixing showed to be the major process resulting in the disproportional transport and loss of age and TOU, modifying their respective gradients, and, given a distribution with distant respiration maxima, causing OUR to underestimate imposed respiration in the tube model. In the ocean there are more transport modes affecting the distributions of TOU and age than can be represented in the tube model: advection and diffusion within isopycnal layers are two-dimensional, and there is diffusion in the vertical (diapycnic diffusive mixing). Also, diffusion and advection in the different directions are not constant throughout the ocean but vary in space and time. Three-dimensional ocean circulation models are designed to reproduce all these transports. In the next section we asses the ability of OUR to estimate true respiration in the framework of such models.

Figure 4. Sensitivity of the integrated ratio of diagnosed/imposed respiration to the degree of heterogeneity of imposed respiration along the tube. All experiments are based on the distribution of imposed respiration shown in Figure 2b (DISTANT). Values on the *x* axes (r_{max}/r_{min}) reflect the ratio of the maximum imposed respiration (in cells 2 and 99, respectively) to the minimum imposed respiration (in the center of the tube). Values on the *y* axes are based on step wise (10 neighboring grid cell) diagnosed OUR and averaged *R*, integrated over the tube, cells 2 to 50. (a) Tracer transport is by diffusion only ($A_H = 0$ cm s^{-1}). Indicated color and symbol combinations represent different diffusivities of 500 m^2 s^{-1} (red dashed line), 1000 m^2 s^{-1} (blue cross), and 2000 m^2 s^{-1} (black dotted line). (b) Tracer transport is by advection ($A_H = 0.3$ cm s^{-1}) plus diffusion with K_H of 500 m^2 s^{-1} (black dotted line), 1000 m^2 s^{-1} (blue cross) and 2000 m^2 s^{-1} (red dashed line). (c) Tracer transport is due to diffusion ($K_H = 1000$ m^2 s^{-1}) plus advection with A_H of 0.1 cm s^{-1} (blue cross), 0.3 cm s^{-1} (red dashed line), and 0.6 cm s^{-1} (black dotted line). Please see captions of Figures 1 and S1 for details of the advective flow.

4. Global Biogeochemical Models

4.1. Models Used

We use three 3-dimensional (3-D) global ocean circulation and biogeochemical models (see supporting information methods section for details) in which oxygen is affected by air-sea exchange with a one-box atmosphere, photosynthetic production in surface water, and respiration in the interior of the ocean. Oxygen sinks and sources are computed in NPZD-type (nutrient, phytoplankton, zooplankton, detritus) modules with fixed elemental ratios and explicit sinking of detritus particles. In the TMM-MIT28 [*Khatiwala et al*., 2005] (see supporting information), the first 3-D model we employed, TOU agrees to within a few percent with AOU computed from O_2^{sat} and O_2 [*Duteil et al*., 2013]. We make sure to be in full seasonal cyclic-tracer equilibrium such that at any point in space and time the following tracer budget is satisfied: $O_2 = O_2^{pref} - TOU$. O_2^{pref} is a preformed-oxygen model tracer [*Duteil et al*., 2013]. This state is reached after 6000 model years. TOU and ideal age tracer are implemented like in the tube model, see section 2 for details.

4.2. Isopycnal Analysis

We diagnose regionally averaged OUR from d*TOU*/d*t* (modified equation (1)) on selected density levels for a north-south transect along 30°W in the Atlantic Ocean in the TMM-MIT28 model (Figure 5). Interpolating TOU and age to the density layer $\sigma_\Theta = 27.0$, for example, two distinct patterns are evident. In young water (younger than approximately 20 years), the TOU versus age gradient is steep. For older waters (20 to 110 years) all data points appear to line up along a straight line indicating that an average OUR can be computed. Regressing TOU versus age for $\sigma_\Theta = 27.0$ along 30°W for waters older than 20 years yields an average OUR of about 0.8 mmol m^{-3} yr^{-1} (Figure 5a). The true model respiration along this transect and density layer is both much higher and more variable (Figure 5b). Local rates range between 1.9 mmol m^{-3} yr^{-1} (in the subtropics) to 6 mmol m^{-3} yr^{-1} (close to the outcrops). Averaging true respiration over the region between 40°S and 45°N for which we computed OUR, we find a mean value of 3 mmol m^{-3} yr^{-1}; i.e., the true respiration is about 4 times the OUR estimate. This cannot be an effect of averaging since true respiration is never smaller than the OUR estimate on the selected density layer. Repeating this comparison for a suite of density layers (Figure 5c), we find larger differences between the OUR estimate and true respiration in the upper 500 m (density layers are plotted at their mean depth in Figure 5c). Around 1000 m the OUR estimates and true respiration rates agree, at greater depth respiration rates become too small to be estimated by a linear regression of TOU versus age along 30 W in our model. Overall, the multi-isopycnal analysis suggests that the underestimate of respiration by OUR is larger near the surface than in the deep.

The multi-isopycnal analysis of Figure 5c shows hemispheric (North Atlantic Ocean, South Atlantic Ocean) segments of isopycnals by excluding the initial steep part in the TOU versus age plot (see example in Figure 5a) and averaging over the rest of the segment. Thus, we average over larger distances than has been done in most field studies (e.g., those of *Jenkins* [1982, 1987]). This is dictated by the coarse resolution of the TMM-MIT28. To obtain more regional values, closer to existing field studies, we computed one OUR estimate for the very young part of the isopycnal segment (but with ideal ages > 1 year) and split the inner part of each hemispheric segment into two subsegments wherever possible. We find that all OUR estimates are smaller compared to the respective average true respiration (Figure 5d). We further find that the mismatch between both decreases with increasing age on a given isopycnal, resembling qualitatively what we observed in the tube model (Figures 2b and 2d).

Neither the tube model nor the multi-isopycnal analysis provides a global quantification of the degree of underestimation of respiration by OUR. In the following section we attempt such a quantification based on a slightly modified method to estimate OUR.

4.3. Global Integral of OUR

A simplified approach of OUR determination, in which the oxygen consumption is computed as the ratio of AOU and age (*t*) (equation (3)) provides an averaged estimate of respiration over the entire distance between a given parcel of water and its outcrop location. This approach, dubbed "aOUR" [after *Karstensen et al*., 2008] has also been used by a number of other authors [e.g., *Jenkins*, 1982; *Mecking et al*., 2006; *Stanley et al*., 2012; *Sonnerup et al*., 2013] though under different names. *Sonnerup et al*. [2013] provide a

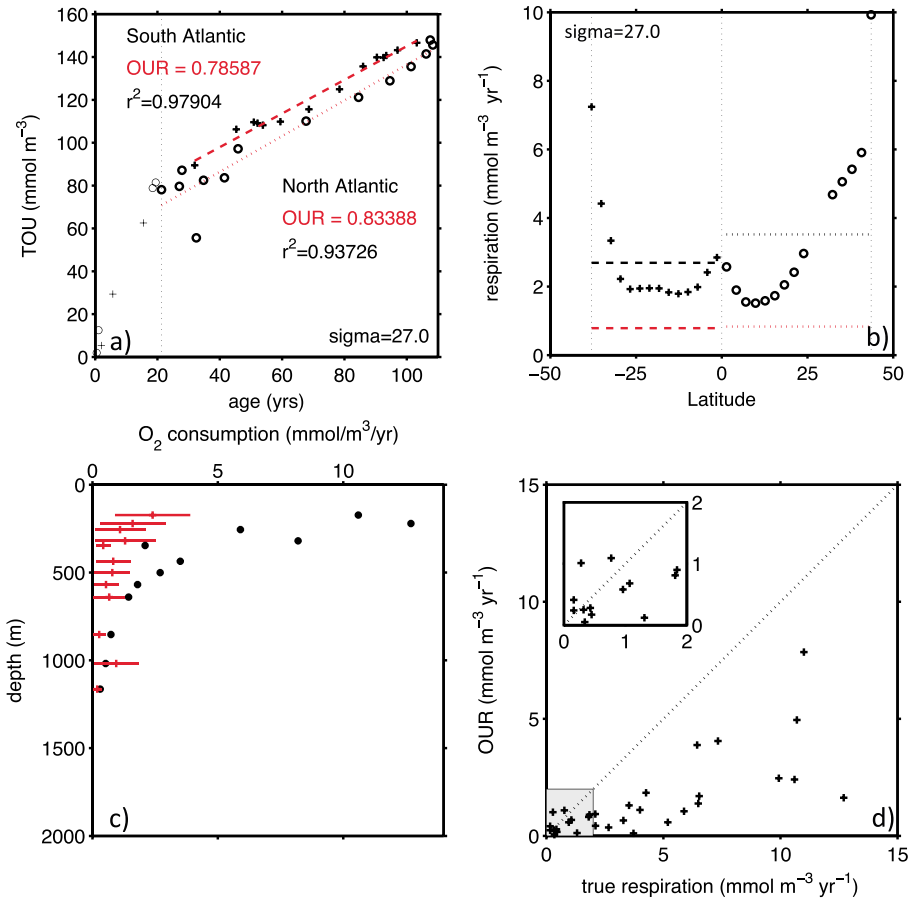


Figure 5. Oxygen consumption along 30°W in the TMM-MIT28 model after 6000 years of model runtime. TOU (mmol O_2 m^{-3}) versus ideal age (years) interpolated to the density layer $\sigma_\Theta = 27.0$ kg m^{-3}. Waters with a mean age of $>$ 20 years (bold symbols) are from 40°S to 45°N (z_{sig27} = 400–600 m). OUR is computed for a segment in the north (cross, dashed line) and couth (circle, dotted line) Atlantic Ocean, respectively. (b) True model respiration rates (mmol O_2 m^{-3} yr^{-1}) (symbols according to hemisphere, as in Figure 5a) for waters from the same density layer. Diagnosed OUR (dashed and dotted red lines), and black lines represent the mean true respiration rates and red lines the diagnosed OUR (dashed: South Atlantic, dotted: North Atlantic). (c) Vertical distribution of diagnosed OUR (red cross) and true respiration rates (dot) on different isopycnals along 20°W in TMM-MIT28. Like in Figure 5a we exclude waters close to the outcrops, which showed steeper d*TOU*/d*t* gradients but which where represented by few model grid cells only. The error bars in Figure 5c represent the 95% confidence limits of OUR estimates. Individual r^2 values ranged between 0.88 and 0.99. (d) Scatterplot of diagnosed OUR versus true respiration. For this plot, we analyzed smaller segments of different isopycnals from the upper 1000 m by splitting the segments shown in Figure 5c into two subsegments. Estimates from isopycnal segments close to outcrops are also included. The insert shows the concentration range 0–2 mmol m^{-3} yr^{-1} in more detail.

detailed comparison of path-averaged aOUR and the OUR from the regression slope of AOU over age. The aOUR approach lends itself to basin scale rather than regional studies where large-scale ratios of diagnosed to true respiration in models are aimed at.

$$aOUR = AOU/t \tag{3}$$

We compute aOUR (equation (3)) from TOU/age for each model grid point below 100 m depth. aOUR is the average rate of oxygen utilization along the trajectory of a given water mass since its last contact with the atmosphere. While the individual local values (x, y, z) of aOUR are not very informative, their horizontal integrals are. In our TMM-MIT28 model run, volume-weighted global mean rates of aOUR decrease with depth from maximum values of 4 mmol m^{-3} yr^{-1} just below 100 m depth to values below 0.5 mmol m^{-3} yr^{-1} in the deep ocean (Figure 6a). Summing up aOUR values over the model ocean, we find a volume-weighted global integral OUR of 0.62 Pmol O_2 yr^{-1}. Integrating the model's true respiration over the same ocean volume yields 1.49 Pmol O_2 yr^{-1}, i.e., a value larger by a factor of 2.4. Most of the difference between both

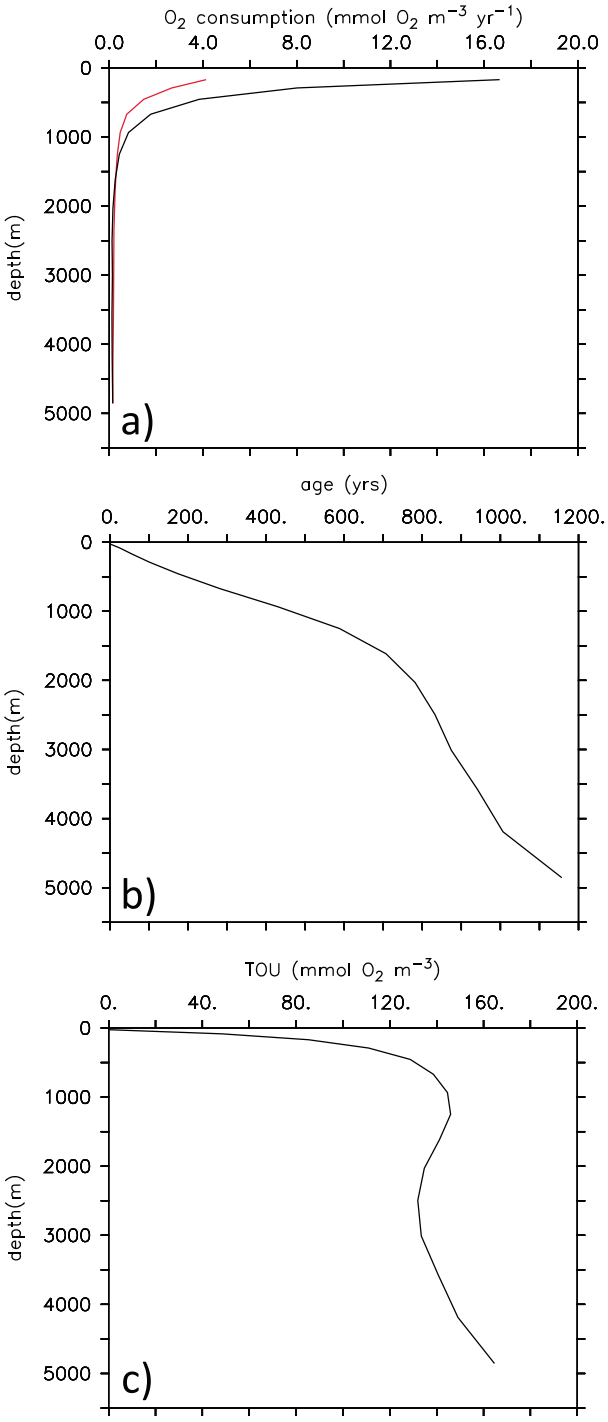


Figure 6. Global estimates. (a) Global mean profiles of oxygen consumption in a model (TMM-MIT2.8). Estimate based on path-averaged OUR (aOUR = TOU/t) (red) and the mean of true respiration (black) in the model. In the upper 1000 m true respiration exceeds mean aOUR by up to a factor of 4. (b) Global mean profile of ideal age (years). (c) Global mean profile of TOU (mmol m^{-3}).

estimates is in the upper 1000 m. The global mean profile of respiration rates from the diagnostic tracer shows about 17 mmol O_2 m^{-3} yr^{-1} immediately below 100 m, which is higher than the corresponding OUR by a factor of 4. This global finding is similar to the observation along 30°W from the same model (Figure 5c).

We briefly explore TOU, age, OUR, and true model respiration in two other models. In the UVic Earth System Model (rev. 2.9, Kiel-Version [*Keller et al.*, 2012]) we compute AOU from the model's oxygen, potential temperature, and salinity distribution and combine this with the age from an ideal age tracer to compute aOUR. The global integral of OUR is 0.55 Pmol O_2 yr^{-1} or 47% of the global integral of respiration computed from detrital nitrogen remineralization in this model (corresponding to 1.17 Pmol O_2 yr^{-1}). The ocean biogeochemical component of UVic does not include a DOM compartment, and all organic matter export in this model is due to sinking particles. In an earlier version of the UVic (2.8) [*Oschlies et al.*, 2008; A. Oschlies, personal communication, 2013] we find a higher global integral of respiration (1.46 Pmol O_2 yr^{-1}) and a lower global integral of OUR (0.55 Pmol O_2 yr^{-1}) equivalent to 38% of diagnosed respiration. Differences in the model's ocean physics between UVic 2.8 and 2.9 translate into different age structures with overall slightly higher ages (slower ventilation rates) in UVic 2.9. Differences of the ecosystem and biogeochemical model are discussed in *Keller et al.* [2012].

We compute profiles of aOUR also for runs of the 1-D tube model described earlier (section 3). First, we present the four standard runs of the diffusive tube ($K_h = 1000$ m^2/s; $A_h = 0$ cm/s) for experiments with a distribution of imposed respiration rates following the design of EVEN, DISTANT, CENTRAL, and COMBINATION (Figures 7a–7d). We compare the imposed local respiration rate (black dashed), the aOUR (red solid), and the average of the imposed respiration rate between the outcrop and the reference point (average rate, blue solid). Since in the aOUR method no local regressions of d*TOU*/d*t* are performed, it does not provide local rates. However, it should reproduce the average rate

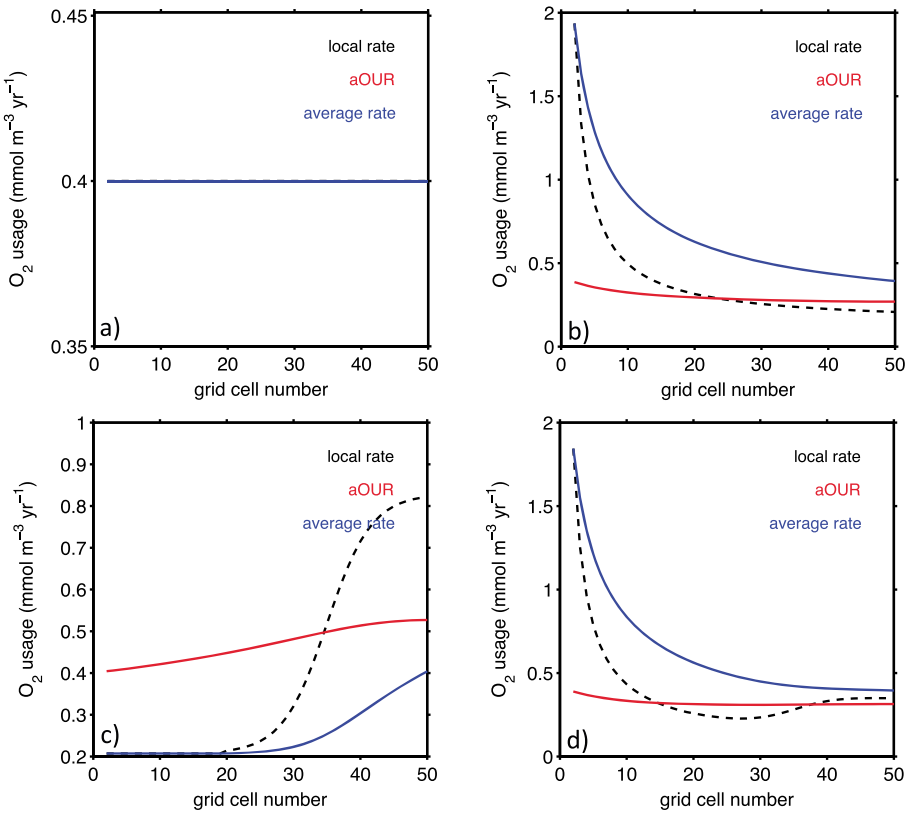


Figure 7. Imposed respiration (black dashed line), average imposed respiration since outcrop (blue solid line), and diagnosed aOUR (red solid line) for several tube model experiments ($K_H = 1000\ m^2\ s^{-1}$, $A_H = 0\ cm\ s^{-1}$). (a) EVEN, (b) DISTANT, (c) CENTRAL, and (d) COMBINATION. See text for details of experimental design of the four experiments. Please note in Figure 7a the local rate = average rate = aOUR everywhere; the blue solid line is the top most printed one and hides the other two lines.

between outcrop and reference point. Similar to the findings for the d*TOU*/d*t* estimate, aOUR is able to correctly estimate the respective rate only in the case of constant imposed respiration. Again, aOUR is an underestimate when imposed respiration is elevated toward the outcrops and an overestimate when imposed respiration is elevated in the tube's center. The average ratio of aOUR/R_{ave} along the tube shows a clear correlation with the TOU inventory (Figure 3d), as did the tube-integrated ratio of diagnosed OUR to imposed respiration (see section 3 and Figure 3c). Studying the ability of aOUR to reproduce the averaged imposed respiration in the tube under more realistic conditions (i.e., including advection and with elevated respiration toward the outcrops), we find that by increasing advection the discrepancy was reduced (Figure S3). Similar to model experiments discussed in section 3, an increase in the degree of heterogeneity of the imposed respiration rates decreases the quality of the aOUR estimate and aOUR performs particularly badly close to the outcrops (Figure 7). Collectively, we take these findings to support our conclusion that the underestimate of respiration by regional OUR and by ocean-scale aOUR is ultimately caused by the same processes: nonproportional redistribution and differential losses of TOU and age tracer.

5. Discussion

5.1. Why Does OUR Underestimate Respiration?

Currently, AOU is thought to reflect the total oxidation of biogenic carbon along specified surfaces [e.g., *Ogura*, 1970; *Carlson et al.*, 2010; *Sonnerup et al.*, 2013] and that the vertical integral of OUR is related to the amount of organic matter exported from the surface layer in particulate and dissolved form (export or new production) and subsequently respired in the interior of the ocean [e.g., *Sarmiento and Gruber*, 2006; *Emerson and Hedges*, 2008]. In contrast, we find that OUR and its vertical integral to substantially underestimate respiration and export production in the models employed.

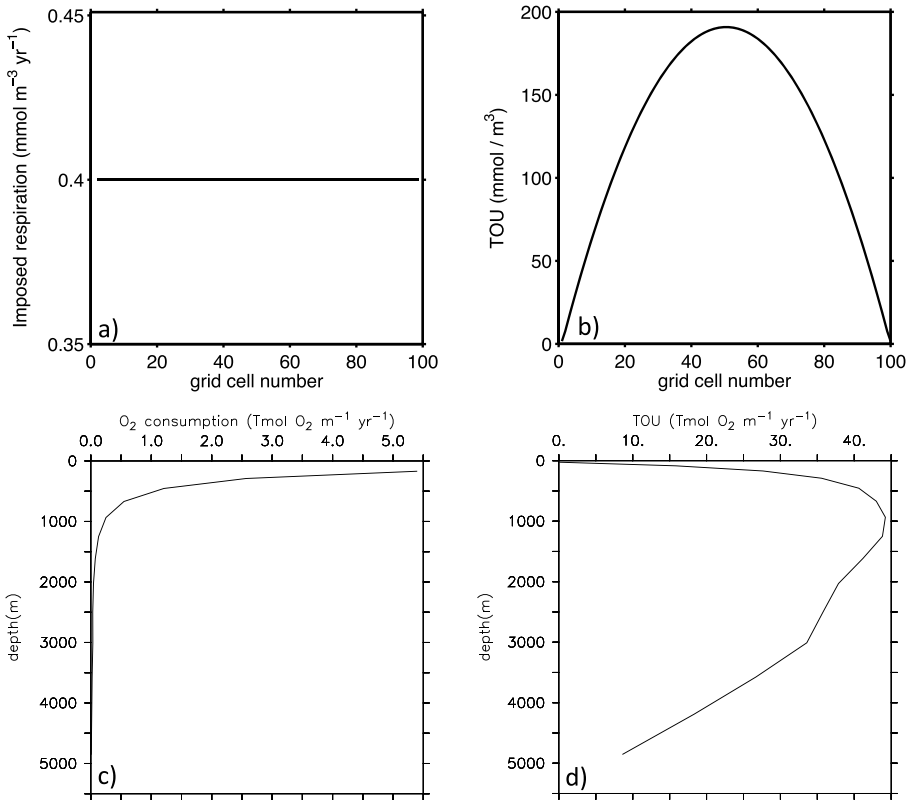


Figure 8. Distributions of respiration and TOU. An example from the tube model, with (a) the distribution of respiration from the experiment shown in Figure 2a and (b) the respective distribution of TOU. TMM-MIT28 with (c) the global profile of horizontally integrated respiration and (d) the respective profile of horizontally integrated TOU.

In the tube model, OUR is a numerically correct or acceptable representation of respiration in two cases: water transport by advection only or uniform respiration rates (Figures 1 and 2a). Both conditions are highly unlikely and met nowhere in the ocean. In particular, isopycnal diffusion is known to significantly contribute to tracer distribution in the ocean [*Iselin*, 1939; *Mecking et al*., 2004]. The relative importance of isopycnal diffusion and advective transports varies. Combining different transient tracers into transit time distributions (TTD) *Waugh et al*. [2004] showed that TTDs in the North Atlantic are usually broad, i.e., show a ratio "width" to mean age (second to first moment of TTD, respectively) on the order of one. This is consistent with a Péclet number for one-dimensional flow of about 1. The upper limit of this ratio (i.e., the actual importance of diffusion relative to advection), however, is not well constrained given the limited time span covered by transient tracers [*Waugh et al*., 2004; *Sonnerup et al*., 2013].

Our experiments with the 1-D tube model, in which mixing contributes to tracer transport using implemented diffusion coefficients consistent with observations [e.g., *Mecking et al*., 2004], show that the intensity of the OUR mismatch is related to the nonproportionality of the rates of respiration and aging. Aging proceeds monotonously and with the same rate everywhere, but respiration rates vary on a given isopycnal. OUR is calculated from the imprints of aging and respiration preserved in the ocean, TOU (AOU), and age. The more (less) a water mass is isolated from the surface, the higher (lower) is the share of the preserved effect of the respective rate (i.e., age versus aging and AOU (TOU) versus respiration). On a given isopycnal, the preserved effect of any process varies with distance from the outcrop—it is low there and high in its center. This is particularly obvious from the tube experiment where respiration is constant but TOU (the preserved effect of respiration) is higher in the tube's center (Figures 8a and 8b). In the 3-D models the same relationship is observed in the vertical. Deep model layers preserve more TOU per unit respiration compared with shallow ones (Figures 8c and 8d).

When considering isopycnal processes exclusively (Figure 4), the OUR underestimates true respiration by about 20–30% on average. Mind that this estimate is based on averaging over segments of the tube where

diagnosed OURs differ strongly from true respiration and those where it does not (see Figure 2b and caption of Figure 4 for details). Locally, i.e., close to the outcrops, the underestimate is larger. Moreover, while the analysis of the spatial distribution of the rates (respiration and aging) and their preserved effect (AOU and age) in the 1-D experiments helps to understand basic principles of the OUR analysis, these experiments are no realistic representations of conditions in the ocean. This can only be approached in complex 3-D ocean models.

Since the source of organic matter supporting oxygen consumption in the ocean interior is the ocean surface, oxygen consumption on an isopycnal in the real ocean can be expected to be higher near the outcrops but also higher on shallow isopycnals in general. There, near the surface, an overproportional loss of AOU relative to age should occur. This is confirmed in the runs with the 3-D models, the results of which resemble those of the 1-D tube model in that the degree of underestimation increases toward the surface (Figure 5c). It is noteworthy that the mismatch between true and diagnosed respiration is so much greater in the 3-D model than in the 1-D tube model, e.g., compare Figures 5c and 6 with 4. We think that diapycnal mixing, absent in the 1-D, but included in the 3-D case, is the principal cause of this difference. Vertical gradients are drivers of diapycnal mixing, and they are strongest near the surface so is the deviation of respiration versus aging from linearity (imagine a diffusive-only, vertical 1-D tube). This is yet another reason for the difference between true and diagnosed respiration to be more pronounced near the surface than deeper down.

Diffusive mixing as a potential troublemaker in OUR calculations has been treated in the literature mostly with regard to the mixing of nonideal age tracers. Mixing complicates the calculation of age from age tracer distributions [*Sarmiento et al*., 1990]. *Mecking et al*. [2004, 2006], for example, applied a two-dimensional diffusion-advection model to quantify and correct for effects of nonlinear tracer mixing on CFC-12 ages. At first sight, this is not relevant to our study since we use an ideal age tracer. At second sight, however, it is relevant to the OUR calculation. The effects described for ages of nonideal age tracers apply also to AOU; both do not scale linearly with ideal age. A potential effect of nonlinear mixing on diagnosed OUR has been suspected earlier [e.g., *Mecking et al*., 2004] but not quantified.

The concept behind the OUR determination does not include mixing in the first place. It is a "box car" concept [*Sarmiento et al*., 1990], which assumes that advective transports dominate where it has been applied. Mixing has rather been treated as a source of (minor) error in this context. *Jenkins* [1987] explores the errors introduced with respect not only to nonideal age tracers but also to AOU (or O_2) diffusive mixing in the subtropical North Atlantic in an area he dubbed the "beta triangle." He estimated the error introduced by the isopycnal mixing of AOU to be of a certain importance (between 10 and 20%) in the OUR determination and even smaller by an order of magnitude with respect to diapycnal mixing in the location he analyzed. These errors are a lot smaller than the discrepancies described in our comparison of true respiration and diagnosed OUR. In the case of the beta triangle area, the dominance of advection for oxygen transport may be used as an argument that the steady state distribution of oxygen (or AOU) may be successfully combined with a transient tracer, reflecting the advective rather than its diffusive component of transports, in order to diagnose oxygen consumption. *Jenkins* [1998] compared OUR with estimates of respiration (the J_i term in his equation (4)) derived from a least square solution to a set of 2-D advection-diffusion equations of ^{3}H-^{3}He age, salinity, and oxygen applied to data from the subtropical North Atlantic. He finds differences between the J_i term and OUR to generally increase with depth. Below 250 m depth OUR from regressions of AOU over ^{3}H-^{3}He age is larger compared to the respective J_i estimates. By explicitly considering the effects of diffusive mixing and advection, the comparison of *Jenkins* [1998] clearly points to weaknesses of the regression-based OUR approach. However, working in the real ocean, neither of the various estimates of respiration could be compared with a suitable independent reference estimate of respiration.

A direct comparison between the results of our study and applications of the OUR method in the field [e.g., *Jenkins*, 1982, 1987; *Warner et al*., 1996; *Sonnerup et al*., 1999; *Feely et al*., 2004] is complicated by the nature of the age tracers used in the respective studies. In the real ocean there is no tracer of ideal age which we use in the models. Instead, ages are often derived from transient tracers like tritium/helium [e.g., *Jenkins and Clarke*, 1976] or chlorofluorocarbons (CFCs, e.g., *Sonnerup et al*. [1999]) with a number of known limitations. For example, the mixing of tracers like tritium/helium causes the computed tracer ages to be biased low [*Jenkins*, 1987; *Khatiwala et al*., 2001; *Koeve et al*., 2015]. Also, since boundary conditions have varied with time, tracer ages are not fixed but known to have changed with the time of observation even under conditions

of constant circulation [*Sonnerup*, 2001; *Khatiwala et al*., 2001; *Waugh et al*., 2003]. Tracer ages from different transient tracers may differ [*Waugh et al*., 2003; *Tanhua et al*., 2013].

Transient ages are always biased toward younger values compared to a steady state ideal age [*Khatiwala et al*., 2001] (Figure S4). This is because the transient age tracers have been in the world only for a few decades, which is a short time compared with observation-based estimates of the mean age (the "last passage times" of *DeVries and Primeau* [2011]). Given the short presence of transient tracers in the ocean, transient ages cannot reflect all advective and diffusive pathways which fluid elements constituting a water parcel may have taken [*Khatiwala et al*., 2001]. Rather, they will predominantly reflect the fast, advective pathways. Combining several age tracers, however, allows to derive ocean age spectra and thereby compute mean ages which are principally consistent with the ideal age concept used in our study. Still, differences between such TTD-based mean ages and ideal age may arise from that part of the TTD older than the time covered by transient tracers [*Waugh et al*., 2004]. Recently, mean ages from this transit time distribution (TTD) approach have been used to compute OUR from the Pacific Ocean [*Sonnerup et al*., 2015].

OUR estimates which combine tracer ages from transient tracers [*Jenkins*, 1982, 1987; *Warner et al*., 1996; *Sonnerup et al*., 1999; *Feely et al*., 2004] with steady state oxygen data (i.e., AOU) make the implicit assumption that all AOU observed on a given isopycnal has been produced over the time span which is described by the respective transient tracer. This assumption is certainly not true. In steady state all AOU produced in the deep ocean must eventually return to the surface ocean and thereby affect the AOU distribution in more shallow waters. Whether said implicit assumption of the OUR method is a reasonable approximation, at least in some regions of the ocean, has never been shown and can only be studied in dedicated model experiments.

5.2. Real Ocean Problem or Model Deficiency?

The discrepancy of OUR and respiration observed in our models is large. While in the tube model, there are cases where there is agreement or near-agreement (constant respiration or regions of near-constant respiration in the diffusive tube; advection-only experiments), there is no such case in the 3-D models we analyzed, except for some deep-sea regions of minimal oxygen consumption (Figure 6). This points either to a severe inadequacy of the underlying concept or to a severe inadequacy of the 3-D models employed. Both possibilities are disturbing since they concern an important process fundamental to our understanding of ocean biogeochemistry or of much-used tools employed in its study and to predict potential future changes [e.g., *Oschlies et al*., 2008, 2010; *Duteil and Oschlies*, 2011; *Taucher et al*., 2014; *Keller et al*., 2014].

We find the underestimate of respiration by OUR in different 3-D models, which were independently coded, both by us and by others. This gives us confidence as to the robustness of the results. The concept of OUR as a measure of ocean respiration has been tested in model environments thought to include all components relevant to this concept. This includes aspects of model physics, like the relative roles of advection and diffusion and of the models' biogeochemistry, in particular vertical and regional gradients of respiration in the models, which are subject to model formulations and the parameters of organic matter sinking and decay, as well as the relative roles of sinking versus nonsinking organic matter [*Najjar et al*., 1992; *Kriest et al*., 2012]. None of these will be perfect in coarse-resolution global ocean models of the class we employed. In particular the coarse vertical resolution (e.g., 15 vertical layers in the TMM-MIT28) combined with steep vertical oxygen gradients in the upper ocean may give rise to vertical diffusive transports being too high in the models applied. The computation time needed to spin up a global model with significantly higher vertical and horizontal resolution into full steady state (6000 model years), however, precluded further experiments so far.

The overall performance to reproduce large-scale patterns of abiotic as well as biotic tracers is usually taken as evidence that a model reasonably represents the real ocean [e.g., *Dutay et al*., 2002; *Najjar et al*., 2007; *Duteil et al*., 2012; *Kriest et al*., 2010, 2012]. Should a concept like OUR not perform as expected in such a system, this is strong evidence for this concept to be problematic also in reality. Note that in our model experiments, OUR has been found to substantially underestimate respiration in any constellation in which the distribution of respiration resembled that in the real ocean. Still our results are from simulations, not from real-world data. Whether the degree of underestimation of respiration by OUR (or aOUR) observed in the models we used applies in the same way to the real world should depend on the ability of these respective models to strike the correct balance between (explicit and implicit) diffusive and advective transports (in the sense of Figures 4 and 8).

Another way to evaluate a method is to compare its results with those of an alternative method employed to quantify the same process. Agreement between such independent estimates has been interpreted as mutual confirmation. Electron transport system (ETS)-based respiration [*Packard and Williams*, 1981; *Aristegui et al*., 2005] measured by Packard and *Packard and Codispoti* [2007] agreed well with OUR determined in the same region by *Zheng et al*. [1997]. A general agreement between the divergence of organic matter fluxes into sediment traps and OUR at appropriate depth levels [e.g., *Feely et al*., 2004] has likewise been viewed as evidence of the suitability of either method [*Sarmiento and Gruber*, 2006] so has the comparison of new production determined by nitrate transport into the euphotic zone and by added-up OURs on stacked isopycnals [*Sarmiento et al*., 1990]. On the other hand, inconsistencies among various methods used to determine respiration and related rates have been documented. ETS-derived respiration rates at low levels rely on conversion factors determined at high levels [*Aristegui et al*., 2005]. For sediment trap flux measurements a number of severe biases have been described [*Gardner*, 2000; *Kähler and Bauerfeind*, 2001]. Each of the methods with which OUR may be compared has, apart from the necessarily different timescale, drawbacks of its own (for an overview see *Burd et al*. [2010]). The agreement of two methods may tell that both are reliable just as well as that both are unreliable.

There is at present no method representing an undisputed absolute standard against which to evaluate or calibrate the OUR method or any other method to determine respiration that can be employed in the ocean interior given the low rates prevailing there. Instead, OUR has been viewed as the method yielding the most reliable results. *Emerson and Hedges* [2008] call it "the best available method" given the availability of a suitable age tracer and the careful consideration of errors [*Jenkins*, 1987]. Reliable determination of the mean age is possible today by the combination of two transient age tracers in the TTD approach [*Hall et al*., 2002; *Waugh et al*., 2003; *Tanhua et al*., 2013] and the evaluated oxygen utilization approach of *Duteil et al*. [2013] corrects for uncertainties of AOU.

However, despite using the ideal tracers of age and true oxygen utilization, our assessment showed that OUR does not provide a reliable estimate of respiration in the ocean's interior. Our assessment takes advantage of the fact that in an ocean model an absolute standard exists, against which diagnosed OUR can be compared.

6. Conclusions and Outlook

The misrepresentation of oxygen consumption in the ocean by OUR is due to the nonproportional diffusive mixing of AOU and water age resulting in an excess loss of AOU versus age related to the vertical and regional distribution of respiration in the ocean (elevated rates close to the surface). Since the nonproportional relationship of aging and respiration is a fact in the ocean, there is no solution to improve the concept. We consider the underestimate basically not to be a model artifact, but whether the degree of underestimation found in our study is realistic will surely depend decisively on whether the models used represent the relative importance of advection and diffusion realistically.

Historical simulation as well as projections until the end of this century carried out with state-of-the-art climate models suggests that global export production is likely to decrease considerably as a consequence of surface ocean warming and increasing stratification [*Steinacher et al*., 2010; *Bopp et al*., 2013; *Cabré et al*., 2015]. Detecting such changes in the real ocean over the decades to come requires methods suitable to quantify export production on large scales of space and time. OUR and its vertical integral may be considered as a tool to monitor changes in export production in the future. This requires, however, a solution to the concerns raised in this study regarding the conceptional reliability of the OUR approach.

Acknowledgments
We thank Arne Körtzinger, Andreas Oschlies, Angela Landolfi, and Heiner Dietze (GEOMAR, Kiel, Germany) for discussion and comments. Comments and suggestions from two anonymous reviewers and the editor where most helpful. We acknowledge funding by the German Science Foundation (DFG) to P.K. via the SFB 754 and by the German BIOACID program (BMBF grant FKZ 03F0608A) to W.K. We thank Iris Kriest and Samar Khatiwala for providing the Kiel TMM-BGC model environment, which we have used and modified for our purpose and David Keller and Andreas Oschlies for providing the Kiel version of the UVic Earth System model. The data presented in this work are available from http://thredds.geomar.de/. No conflicts of interest.

References

Anderson, L. A., and J. L. Sarmiento (1994), Redfield ratios of remineralization determined by nutrient data analysis, *Global Biogeochem. Cycles*, *8*, 65–80, doi:10.1029/93GB03318.

Aristegui, J., S. Agusti, J. J. Middelburg, and C. M. Duarte (2005), Respiration in the mesopelagic and bathypelagic zones of the oceans, in *Respiration in Aquatic Ecosystems*, edited by P. A. del Giorgio and P. J. l. B. Williams, pp. 181–205, Oxford Univ. Press, Oxford, U. K.

Bopp, L., et al. (2013), Multiple stressors of ocean ecosystems in the 21st cnetury: Projections with CMIP5 models, *Biogeosciences*, *10*, 6225–6245.

Burd, A. B., et al. (2010), Assessing the apparent imbalance between geo- chemical and biochemical indicators of meso- and bathypelagic biological activity: What the @$#! is wrong with present calculations of carbon budgets?, *Deep Sea Res., Part II*, *57*(16), 1557–1571, doi:10.1016/j.dsr2.2010.02.022.

Cabré, A., I. Marinov, and S. Leung (2015), Consistent global response of marine ecosystems to future climate change across the IPCC AR5 earth system models, *Clim. Dyn.*, *45*, 1253–1280.

Carlson, G. A., D. A. Hansell, N. B. Nelson, D. A. Siegel, W. M. Smethie, S. Khatiwala, M. M. Meyers, and E. Halewood (2010), Dissolved organic carbon export and subsequent remineralisation in the mesopelagic and bathypelagic realms of the North Atlantic basin, *Deep Sea Res., Part II*, *57*, 1433–1445.

del Giorgio, P. A., and C. M. Duarte (2002), Respiration in the open ocean, *Nature*, *420*(6914), 379–384, doi:10.1038/nature01165.

DeVries, T., and F. Primeau (2011), Dynamically and observationally constrained estimates of water-mass distributions and ages in the global ocean, *J. Phys. Oceanogr.*, *41*, 2381–2401, doi:10.1175/JPO-D-10-05011.1.

Dietze, H., and A. Oschlies (2005), Modeling abiotic production of apparent oxygen utilisation in the oligotrophic subtropical North Atlantic, *Ocean Dynam.*, *55*, 28–22, doi:10.1007/s10236-005-0109-z.

Dutay, J.-C., et al. (2002), Evaluation of ocean model ventilation with CFC-11: Comparison of 13 global ocean models, *Ocean Mode.*, *4*, 89–102.

Duteil, O., and A. Oschlies (2011), Sensitivity of simulated extent and future evolution of marine suboxia to mixing intensity, *Geophys. Res. Lett.*, *38*, L06607, doi:10.1029/2011GL046877.

Duteil, O., et al. (2012), Preformed and regenerated phosphate in ocean general circulation models: Can right total concentrations be wrong?, *Biogeosciences*, *9*, doi:10.5194/bg-9-1-2012.

Duteil, O., W. Koeve, A. Oschlies, D. Bianchi, E. Galbraith, I. Kriest, and R. A. Matear (2013), A novel estimate of ocean oxygen utilisation points to a reduced rate of respiration in the ocean interior, *Biogeosciences*, *10*, 7723–7738, doi:10.5194/bg-10-7723-2013.

Emerson, S. R., and J. I. Hedges (2008), *Chemical Oceanography and the Marine Carbon Cycle*, 453 pp., Cambridge Univ. Press, Cambridge, U. K.

England, M. H. (1995), The age of water and ventilation timescales in a global ocean model, *J. Phys. Oceanogr.*, *25*, 2756–2777.

Eppley, R. W., and B. J. Peterson (1979), Particulate organic matter flux and planktonic new production in the deep ocean, *Nature*, *282*, 677–680.

Feely, R. A., C. L. Sabine, R. Schlitzer, J. L. Bullister, S. Mecking, and D. Greeley (2004), Oxygen utilization and organic carbon remineralization in the upper water column of the Pacific Ocean, *J. Oceanogr.*, *60*, 45–52.

Gardner, W. D. (2000), Sediment trap sampling in surface waters, in *The Changing Ocean Carbon Cycle: A Midterm Synthesis of the Joint Global Ocean Flux Study*, edited by R. B. Hanson, H. W. Ducklow, and J. G. Field, pp. 240–281, Cambridge Univ. Press, Cambridge, U. K.

Hall, T. M., T. W. N. Haine, and D. W. Waugh (2002), Inferring the concentration of anthropogenic carbon in the ocean from tracers, *Global Biogeochem. Cycles*, *16*(4), 1131, doi:10.1029/2001GB001835.

Iselin, C. O. (1939), The influence of vertical and lateral turbulence on the characteristics of mid-depths, *Eos Trans. AGU*, *20*, 414–417, doi:10.1029/TR020i003p00414.

Ito, T., M. J. Follows, and E. A. Boyle (2004), Is AOU a good measure of respiration in the ocean?, *Geophys. Res. Lett.*, *31*, L17305, doi:10.1029/2004GL020900.

Jenkins, W. J. (1982), Oxygen utilization rates in North Atlantic subtropical gyre and primary production in oligotrophic systems, *Nature*, *300*, 246–248.

Jenkins, W. J. (1987), ^{3}H and ^{3}He in the beta triangle: Observations of gyre ventilation and oxygen utilization rates, *J. Phys. Oceanogr.*, *17*, 763–783.

Jenkins, W. J. (1998), Studying subtropical thermocline ventilation and circulation using tritium and ^{3}He, *J. Geophys. Res.*, *103*, 15,817–15,831, doi:10.1029/98JC00141.

Jenkins, W. J., and W. B. Clarke (1976), The distribution of ^{3}He in the western Atlantic Ocean, *Deep Sea Res.*, *23*, 481–494.

Kähler, P., and E. Bauerfeind (2001), Organic particles in a sediment trap: Substantial loss to the dissolved phase, *Limnol. Oceanogr.*, *46*, 719–723.

Kähler, P., H. Dietze, W. Koeve, and A. Oschlies (2010), Oxygen, carbon, and nutrients in the oligotrophic eastern subtropical North Atlantic, *Biogeosciences*, *7*, 1143–1156.

Karstensen, J., L. Stramma, and M. Visbeck (2008), Oxygen minimum zones in the eastern tropical Atlantic and Pacific oceans, *Prog. Oceanogr.*, *77*, 331–350.

Keller, D., A. Oschlies, and M. Eby (2012), A new marine ecosystem model for the University of Victoria Earth System Climate Model, *Geosci. Model Dev.*, *5*, 1195–1220.

Keller, D., Y. Feng, and A. Oschlies (2014), Potential climate engineering effectiveness and side effects during a high CO_2-emission scenario, *Nat. Commun.*, *5*, 3304, doi:10.1038/ncomms4304.

Khatiwala, S. (2007), A computational framework for simulation of biogeochemical tracers in the ocean, *Global Biogeochem. Cycles*, *21*, GB3001, doi:10.1029/GB002923.

Khatiwala, S., M. Visbeck, and P. Schlosser (2001), Age tracers in an ocean GCM, *Deep Sea Res., Part I*, *48*, 1432–1441.

Khatiwala, S., M. Visbeck, and M. A. Cane (2005), Accelerated simulation of passive tracers in ocean circulation models, *Ocean Model.*, *9*, 51–69.

Koeve, W., H. Wagner, P. Kähler, and A. Oschlies (2015), ^{14}C-age tracers in global ocean circulation models, *Geosci. Model Dev.*, *8*, 2079–2094, doi:10.5194/gmd-8-2079-2015.

Kriest, I., S. Khatiwala, and A. Oschlies (2010), Towards and assessment of simple global marine biogeochemical models of different complexity, *Prog. Oceanogr.*, *86*, 337–360.

Kriest, I., A. Oschlies, and S. Khatiwala (2012), Sensitivity analysis of simple global marine biogeochemical models, *Global Biogeochem. Cycles*, *26*, GB2029, doi:10.1029/2011GB004072.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey (1997), A finite-volume, incompressible Navier-stokes model for studies of the ocean on parallel computers, *J. Geophys. Res.*, *102*, 5733–5752, doi:10.1029/96JC02776.

Martin, J. H., G. A. Knauer, D. M. Karl, and W. W. Broenkow (1987), VERTEX: Carbon cycling in the northeast Pacific, *Deep Sea Res. Part A*, *34*, 267–285.

Matsumoto, K. (2007), Radiocarbon-based circulation age of the world ocean, *J. Geophys. Res.*, *112*, C09004, doi:10.1029/2007JC004095.

Mecking, S., C. E. Greene, S. L. Hautala, and R. E. Sonnerup (2004), Influence of mixing on CFC uptake and CFC ages in the North Pacific thermocline, *J. Geophys. Res.*, *109*, C07014, doi:10.1029/2003JC001988.

Mecking, S., M. J. Warner, and J. L. Bullister (2006), Temporal changes in pCFC-12 ages and AOU along two hydrographic sections in the eastern subtropical North Pacific, *Deep Sea Res., Part I*, *53*, 169–187.

Najjar, R. G., J. L. Sarmiento, and J. R. Toggweiler (1992), Downward transport and fate of organic matter in the ocean: simulations with a general circulation model, *Global Biogeochem. Cycles*, *6*, 45–76, doi:10.1029/91GB02718.

Najjar, R. G., et al. (2007), Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2), *Global Biogeochem. Cycles*, *21*, GB3007, doi:10.1029/2006GB002857.

Ogura, N. (1970), The relation between dissolved organic carbon and apparent oxygen utilization in the Western North Pacific, *Deep Sea Res.*, *17*, 221–231.

Orr, J., R. Najjar, C. L. Sabine, and F. Joos (2000), Abiotic-HOWTO. Internal OCMIP Report, LSCE/CEA, Saclay, Gif-sur-Yvette, France, 25 pp.,revision: 1.16. [Available at http://ocmip5.ipsl.jussieu.fr/OCMIP/phase2/simulations/Abiotic/HOWTO-Abiotic.html, last access: 28 November 2013.]

Oschlies, A., and P. Kähler (2004), Biotic contribution to air-sea fluxes of CO_2 and O_2 and its relation to new production, export production, and net community production, *Global Biogeochem. Cycles*, *18*, GB1015, doi:10.1029/2003GB002094.

Oschlies, A., K. G. Schulz, U. Riebesell, and A. Schmittner (2008), Simulated 21th century's increase in oceanic suboxia by CO_2-enhanced biotic carbon export, *Global Biogeochem. Cycles*, *22*, GB4008, doi:10.1029/2007GB003147.

Oschlies, A., W. Koeve, W. Rickels, and K. Rehdanz (2010), Side effects and accounting aspects of hypothetical large-scale Southern Ocean iron fertilization, *Biogeosciences*, *7*, 4017–4035, doi:10.5194/bg-7-4017-2010.

Packard, T. T., and L. A. Codispoti (2007), Respiration, mineraliation, and biochemical properties of the particulate matter in the southern Nansen Basin water column in april 1981, *Deep Sea Res.*, *54*, 403–415.

Packard, T. T., and P. J. l B. Williams (1981), Rates of respiratory oxygen consumption and electron transport in surface seawater from the northwest Atlantic Ocean, *Oceanol. Acta*, *4*, 351–358.

Pytkowicz, R. M. (1971), On the apparent oxygen utilization and the preformed phosphate in the oceans, *Limnol. Oceanogr.*, *16*, 39–42.

Redfield, A. C. (1942), The processes determining the concentration of oxygen, phosphate and other organic derivatives within the depths of the Atlantic Ocean, in *Papers in Physical Oceanography and Meteorology*, vol. 9, No. 2, 22 pp., Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, Cambridge and Woods Hole, Mass.

Redfield, A. C., B. H. Ketchum, and F. A. Richards (1963), The influence of organisms on the composition of seawater, in *The Composition of Seawater: Comparative and Descriptive Oceanography. The Sea: Ideas and Observations on Progress in the Study of the Seas*, vol. 2, edited by M. N. Hill, pp. 26–77, Interscience Publishers, New York.

Sarmiento, J. L., and N. Gruber (2006), *Ocean Biogeochemical Dynamics*, 503 pp., Princeton Univ. Press, Oxford, U. K.

Sarmiento, J. L., G. Thiele, R. M. Key, and W. S. Moore (1990), Oxygen and nitrate new production and remineralization in the North Atlantic subtropical gyre, *J. Geophys. Res.*, *95*, 18,303–18,315, doi:10.1029/JC095iC10p18303.

Schmittner, A., A. Oschlies, X. Giraud, M. Eby, and H. L. Simmons (2005), A global model of the marine ecosystem for long-term simulations: Sensitivity to ocean mixing, buoyancy forcing, particle sinking, and dissolved organic matter cycling, *Global Biogeochem. Cycles*, *19*, GB3004, doi:10.1029/2004GB002283.

Schmittner, A., A. Oschlies, H. D. Matthews, and E. D. Galbraith (2008), Future changes in climate, ocean circulation, ecosystems, and biogeochemical cycling simulated for a business-as-usual CO_2 emission scenario until year 4000 AD, *Global Biogeochem. Cycles*, *22*, GB1013, doi:10.1029/2007GB002953.

Sonnerup, R. E. (2001), On the relation among CFC derived water mass ages, *Geophys. Res. Lett.*, *28*, 1739–1742, doi:10.1029/2000GL012569.

Sonnerup, R. E., P. D. Quay, and J. L. Bullister (1999), Thermocline ventilation and oxygen utilization rates in the subtropical North Pacific base on CFC distributions during WOCE, *Deep Sea Res., Part I*, *46*, 777–805.

Sonnerup, R. E., S. Mecking, and J. L. Bullister (2013), Transit time distribution and oxygen utilization rates in the Northeast Pacific Ocean from chlorofluorocarbons and sulfur hexafluoride, *Deep Sea Res., Part I*, *72*, 61–71.

Sonnerup, R. E., S. Mecking, J. L. Bullister, and M. J. Warner (2015), Transit time distribution and oxygen utilization rates from chlorofluorocarbons and sulfur hexafluoride in the Southeast Pacific Ocean, *J. Geophys. Res. Oceans*, *120*, 3761–3776, doi:10.1002/205JC010781.

Stammer, D., K. Ueyoshi, A. Kuhl, W. G. Large, S. A. Josey, and C. Wunsch, (2004) Estimating air-sea fluxes of heat, freshwater, and momentum through global ocean data assimilation, *J. Geophys. Res.*, *109*, C05023, doi:10.1029/2003JC002082.

Stanley, R. H. R., S. C. Doney, W. J. Jenkins, and D. E. I. I. I. Lott (2012), Apparent oxygen utilization rates calculated from tritium and helium-3 profiles at the Bermuda Atlantic Time-series Study site, *Biogeosciences*, *9*, 1969–1983, doi:10.5194/bg-9-1969-2012.

Steinacher, M., et al. (2010), Projected 21st century decrease in marine productivity: A multi model analysis, *Biogeosciences*, *7*(3), 979–1005, doi:10.5194/bg-7-979-2010.

Tanhua, T., D. W. Waugh, and J. L. Bullister (2013), Estimating changes in ocean ventilation from early 1990s CFC-12 and late 2000s SF6 measurements, *Geophys. Res. Lett.*, *40*, 927–932, doi:10.1002/grl.50251.

Taucher, J., L. T. Bach, U. Riebesell, and A. Oschlies (2014), The viscosity effect on marine particle flux: A climate relevant feedback mechanism, *Global Biogeochem. Cycles*, *28*, 415–422, doi:10.1002/2013GB004728.

Thiele, G., and J. L. Sarmiento (1990), Tracer dating and ocean ventilation, *J. Geophys. Res.*, *95*, 9377–9391, doi:10.1029/JC095iC06p09377.

Trenberth, K. E., W. G. Large, and J. G. Olson (1989), A global ocean wind stress climatology based on ECMWF analyses, National Center for Atmospheric Research, 93 pp.

Wanninkhof, R. (1992), Relationship between wind speed and gas excvhange over the ocean, *J. Geophys. Res.*, *97C*, 7373–7382, doi:10.1029/92JC00188.

Warner, M. J., J. L. Bullister, D. P. Wisegarver, R. H. Gammon, and R. F. Weiss (1996), Basin-wide distribution of chlorofluorocarbons CFC-11 and CFC-12 in the North Pacific: 1985–1989, *J. Geophys. Res.*, *101*, 20,525–20,542, doi:10.1029/96JC01849.

Waugh, D. W., T. M. Hall, and T. W. N. Haine (2003), Relationships among tracer ages, *J. Geophys. Res.*, *108*(C5), 3138, doi:10.1029/2002JC001325.

Waugh, D. W., T. W. N. Haine, and T. M. Hall (2004), Transport times and anthropogenic carbon in the subpolar North Atlantic Ocean, *Deep Sea Res., Part I*, *51*, 1475–1491.

Weaver, A. J., et al. (2001), The UVIC earth system climate model: Model description, climatology, and applications to past, present and future climates, *Atmos. Oceans*, *39*, 361–428.

Williams, P. J. l B., P. J. Morris, and D. M. Karl (2004), Net community production and metabolic balance at the oligotrophic ocean site, station ALOHA, *Deep Sea Res., Part I*, *51*, 1563–1578.

Zheng, Y., P. Schlosser, J. W. Swift, and E. P. Jones (1997), Oxygen utilization rates in the Nansen Basin, Arctic Ocean: Implications for new production, *Deep Sea Res., Part I*, *44*, 1923–1943.