

REVIEWER 3

MAJOR COMMENT

REV3:

The flawed general premise is that somehow in situ $[O_2]^{sat}$ controls in situ $[O_2]$, while other mechanisms drive the saturation state. However, in the ocean interior and particularly near the seabed, far away from the surface, changes in solubility alone (from changes in temperature or salinity) should have zero effect on in situ $[O_2]$, except in the case where the solubility is reduced below the in situ $[O_2]$. If a parcel of water with salinity S , temperature T , and oxygen concentration $[O_2]$ was artificially cooled down, its O_2 solubility would increase, its O_2 saturation state would decrease, but its O_2 content would remain unchanged. While $\Delta[O_2]^{sat}$ may correlate well with $\Delta[O_2]$ in the real ocean and in marine biogeochemistry models, there is no causation.

ANSWER:

We understand the concerns that Reviewer 3 raises and, while we agree Reviewer 3 is essentially correct, we also believe that, in the specific case of the NWES, that is the object of our study, the less rigorous approach adopted here and in Wakelin et al. (2020) still represents a good approximation that is useful to understand the processes involved. Clearly this will need additional clarification in the text. Our argument is as follows:

There is a main assumption under the method we used, that the water column equilibrates with atmospheric O_2 on timescales short enough (every winter in seasonally stratified regions, more frequently in regions that are well-mixed year-round) to establish a causal link between $[O_2]^{sat}$ and $[O_2]$. This is similar to what is assumed in other widely used metrics like AOU.

This clearly may be far from true in the ocean interior, close to sea-ice interface or in upwelling areas, when a water parcel has been separated from the atmosphere long enough to degrade such causal link.

However the North Western European Shelf has characteristics that do not preclude using the decomposition adopted here: it is a highly dynamic system, characterised by relatively shallow depths, short residence times (2-4y for the North Sea, 100d for the Norwegian Trench) and, crucially, intense mixing, both wind- and tidally-driven; many regions (Irish Sea, English Channel, Southern North Sea) are known to be well mixed year-round, and the regions that stratify do so only seasonally. This effectively resets $[O_2]$ towards equilibrium with atmospheric pO_2 every winter, i.e. towards $[O_2]^{sat}$ (and SS toward 1), hence the causal link.

See also Ito et al. 2004, where differences between simulated preformed $[O_2]$ and $[O_2]^{sat}$ are small over much of the global ocean up to depths of 500-1000m and away from the poles.

Our main (and only) conclusion that concerns causal links between $[O_2]^{sat}$ and $[O_2]$ is that there is a component of $[O_2]$ change that is solely warming driven and mediated by the effect of temperature on solubility. This is quite trivial as a result, and it is already well established (e.g. Kwiatkowski et al. 2020, section 3.3 and fig 3).

For these reasons, we believe that our approach is justified, despite some limitations that we will better highlight in the paper to respond to the concerns of reviewer 3.

In the manuscript we improved the methods section by explicitly stating the hypotheses behind our method (and other metrics such as AOU), its limitations and potential pitfalls, and the reason why the method is still valid in the case of the NWES. Then, in the discussion, we mentioned again the

method's limitations and cited some existing alternative metrics (TOU, Ito et al. 2004, EOU, Duteil et al. 2013) that overcome them.

MINOR COMMENTS

REV3:

The authors notation is sometimes hard to parse and confusing, particularly when triple subscripts are used. I would recommend using a different notation, which I hope helps clarifying the comments presented here. (Note that the authors' notation is already different from the preceding work by Wakelin et al. (2020).) I recommend using a simpler symbol, such as $f = [O_2] / [O_2]^{sat}$ for the saturation state. Below I use the "0" subscript for "at t0", i.e., the 1990–2019 average, and the "1" subscript means "at t", i.e., the 2070–2099 average, so that for any quantity X, its 21st-century change is denoted by $\Delta X = X_1 - X_0$.

ANSWER:

While we recognise that other possible notations are in current use, the notation we use is at least partially consistent with published literature (e.g. Kwiatkowski et al. 2020, " ΔO_{2sat} ", Ito et al. 2004, " $O_{2,sat}$ ", Duteil et al. 2013, " O_{2pre} ").

We believe that SS for "oxygen saturation state" may be quite immediate for readers (we changed this from SS_{O_2}), same for "t0" for "at time 0" and "t" for "at time t". As for the difference between our notation and that used in Wakelin et al. 2020, we changed it because we believe the notation in Wakelin et al. 2020 may have been confusing (e.g. DOs for "oxygen saturation state", dimensionless, and DO for "oxygen concentration", concentration).

REV3:

"comment on the product rule" see online

ANSWER:

Reviewer 3 is correct in pointing out that " $\Delta O_{2,other}$ arbitrarily combines a 1st-order term with the 2nd-order term" and that the second order contribution, $\Delta f \times \Delta [O_2]^{sat}$, should be quantified separately. We modified the methods section, and the results, were relevant, to account for this more rigorous approach. This includes computing the second order term.

However, the main body of results are not based on the ΔO_2 decomposition, but on correlations between $[O_2]^{sat}$ and f with other variables, and it can be demonstrated how, given any variable X,

$\text{corr}(X, f_0 \times \Delta [O_2]^{sat}) = \text{corr}(X, [O_2]^{sat})$, and

$\text{corr}(X, \Delta f \times [O_2]^{sat}_0) = \text{corr}(X, f)$

hence the main bulk of our results still stands. We however amended the description of the decomposition of $\Delta [O_2]$ in the methods and presented the separate contribution of the second order term in the results. The contribution of the second order term turned out to be negligible.