The paper focuses on the Benguela upwelling system using a global model and a new biogeochemical parameterization scheme forced by 110 years of reanalysis forcing to diagnose deoxygenation and carbon trends in the region. The authors fail to qualify why they focus on the Benguela region. This seems like an odd choice for the questions posed, and they treat the NBUS and SBUS the same, but they are not the same. The authors need to cite more SBUS papers here – not a lot of representation of the SBUS system on the author team either.

The manuscript is not suitable for publication in its present state, and as such I suggest rejection. While I echo the other reviewers’ thoughts in many ways, including that there clearly has been a large amount of work put into this study, the analysis is not suitable as the model is inadequate to address the questions posed by the author team.

Their physical model is not adequate for the research questions. Physical model – agulhas leakage is weak as simulated – something that impacts one of their key points. The Indian Ocean in particular performs poorly in their streamfunction plot on Figure 4. The Agulhas Current is much too weak and too broad in their high resolution model (in fact, it looks better in the low resolution model). When running a model like this, that includes the Agulhas and the Agulhas retroflection, they need to be explicitly evaluated - the latter by looking at MADT for satellite and model - could be a mean state, and the former by using both satellite and in situ data - the ASCA and ACT data is available on the RSMAS website. While Indian Ocean features, they are highly relevant to the Benguela, especially the southern part, given the importance on water masses, transport of the Benguela, the generation of the shelf-edge jet, turbulence fields in the Cape Basin etc .. Also, there has been no effort to match their transport of the Agulhas against even published literature (e.g. Beal et al 2015: Capturing the transport variability of a western boundary Jet: results from the Agulhas Current Time-Series Experiment (ACT)). The Agulhas should be 70-80Sv, the 'mean' Benguela of the order of 20Sv.

This will impact their leakage and their assessment of the extent of south Indian Ocean
water (FiG 5).

Their use of a climatological mean in Figure 8, used for an assessment of the coastal upwelling cells, doesn't make sense given that there is a strong seasonal signal in the south and a weak one further north. The southern Benguela upwelling cells will then of course be underestimated here. Given that their focus was on the coastal upwelling here, it is unclear why the authors took vertical velocities at a depth of 100 m when something more like 30/50 m would be more appropriate. Maybe the vertical layers in their model does not allow for this?

The description of Figure 7 made out that it captured the velocities well, however, their HR model shows intense offshore flow at ~34S which is just not realistic, and their offshore meridional velocities are underestimated, while they are too strong and uniform over the shelf. Instead of separating meridional and zonal velocities, one possible way forward is to show (as a proxy for surface geostrophic flow): either MADT (from satellite) compared with mean SSH from the models and/or current speeds overlaid with vectors.

Another metric for evaluation used in this region (given the intense offshore turbulence) is surface eddy kinetic energy (model vs. satellite). This would've given them a good sense of how well the leakage is reproduced.

Minor comments:

Looking at annual mean is biased in a seasonal upwelling system like the SBUS.

Hypoxia in the NBUS is source water dependent which changes seasonally, unlike the SBUS – which seems important to the mechanism described within. See Jarre et al. 2019.
Water column denitrification is not thought to happen broadly except in pockets

Coupled denitrification with nitrification in water column – this is observed in the SBUS see Flynn et al. 2021.

Methods:

Line 112-113 – what is the ratio for oxygen and alkalinity chosen?

Sinking fluxes seem slow for shell material – which would sink out more like 80 m/d

Model evaluation is weak. Please consider comparison to glodapv2 O2? What about CLIVAR sections from the region?

There is also a regional climatology for O2 as well along the St. Helena Bay line.

Did they show us observed decadal O2 trends vs simulated? With statistics? See line 601
What impact does the bias in PP have on the other bgc discussed namely O2 and co2?

How do you calculate and show increased leakage? Given the inadequacies of the physical model, how can this conclusion be confident?

Need to show a figure with varying resolution simulations alongside observations more directly – with statistics

No trend in pH does not mean no Ocean Acidification – see Salisbury and Jonnson (2019) – temperature and S trends play a role in pH trends in particular – as do the trends in NPP and O2... but are there trends in Ω or pCO2?