BG-2020-76

[Link to manuscript](#)

[Answers to Reviews](#)

Link to manuscript

- https://www.biogeosciences-discuss.net/bg-2020-76/

Answers to Reviews

We thank the editors and reviewer team for their insightful comments and precious advice.
Since several comments coincide between reviewers we decided to address their comments in this common document, which is addressed to all reviewers. The reviewer's comments are thus reproduced here in their integrality (in blue), with specific answers including when relevant references to comments from other reviewers. Some of the most critical remarks result from imprecise wording, leading to misunderstandings. Below, we propose some rephrasing of the respective paragraphs, so that these misunderstandings can be avoided. Finally, we'd like to express here our general gratitude to the reviewers and appreciation for the pertinence of their remarks.

In summary, to address the main reviewer's comments we propose to:
- Extend the description of past Black Sea oxygenation research in the introduction, for a better positioning of the novelties of our research ->see *2.3, 2.4, 2.9, 2.11, 3.12, 3.13.*
- Enforce the links between the recent (Argo) period and the longer-term period (1955-present) *3.12, 3.13.*
- Revise the description of the statistical approach, and streamline appendixes -> see *4.6, 4.7*
- Adopt oxygen <u>saturation</u> concentration -> see *1.7*.
- Clarify technical points, regarding
  - AIC derivation -> *4.8*
  - Statistical model from atmospheric conditions -> *1.13*
  - Independence between distinct time-series -> *1.2*
- Alleviate miscomprehension due to lack of clarity -> *2.8, 2.9, 2.10, 4.11*
- Restate clearly the new insights of this study, accounting for the additions mentioned above -> *1.3*
- Change the title *-> 1.1*

Reviewer #1

The manuscript by Capet and co-authors presents a statistical analysis of cold intermediate layer (CIL) content and formation and its impact on oxygen levels in the Black Sea from the 1960s to 2019 combining different data sources (incl. observations and models). The two key findings are: (1) temporal changes in CIL water can only be described well by a model when taking into account regime shifts; (2) CIL water formation has entered a new warm regime (i.e. low formation) around 2008, which affects oxygen levels (through reduced ventilation) and is likely to also affect the biogeochemistry in the Black Sea. The manuscript is concise, well written and generally easy to follow. The presented statistical analysis seems sensible and appropriate to address the posed research questions.

__However, I need to state clearly that I am not very experienced in complex statistical analysis and it would be good if someone more knowledgeable in that field could evaluate this aspect of the study.__
-> We believe Reviewer 4 fulfills this role and we thus refer to our answer to Reviewer 4 for statistical considerations.

Most of my comments are only minor, primarily ask for some clarification, and should be easy to address by the authors. I have only a few somewhat bigger points:

1.1 *The study does not isolate the climate change impact (as suggested by the title) on the change in Black Sea ventilation. In my understanding, this cannot be achieved by the applied method, hence, I suggest adapting the title as a simple fix.*
-> This point has been underlined by several reviewers. The title will thus be modified, following the suggestion given by Reviewer1, ie. "A new intermittent regime of convective ventilation threatens the Black Sea oxygenation status".

1.2 *It is not clear from the descriptions of the datasets whether the models are completely independent from the observational data, i.e. whether the same observations have been used for model calibration. If that's the case, this needs to be discussed.*
-> The table below describes the direct and indirect dependencies that may result from the construction of the 4 times series.

	Atm	*Model3D*	*Argo*

Diva	Statistical model using Atmospheric predictors is built on the basis of the Diva time-series. So, even if Atm is more homogeneous and complete than the Diva series, those can not be considered as strictly independent.	The 3D model simulations involve no data assimilation. The model has been calibrated ("hand" screening of Atm. fluxes bulk formulation params.) using T and S in situ data from the same set that has been used to proceed with the long term Diva Analysis, but not using directly the C time series resulting from the DIVA analysis. So it's not strictly independent, in the sense that both times series are both influenced by the same initial data sets. But we consider there is no direct dependence.	Strictly independent.
Atm		Atmospheric conditions from which are extracted the predictors for the statistical model (average winter air temperature) are issued from the same data sets (ECMWF) that are used to force the model. Again, both approaches are influenced by a common data-set, but through drastically different processing pathways. We consider no direct dependence in this case.	Strictly independent.
Model3D			Strictly independent.

We propose :
- to extend Section 2.1 by adding at L145, a succinct description of direct and indirect dependence between times series, based on the information reported in the table below.
- To refer explicitly to this discussion in the sentence L151-153.

1.3 In the discussion, the authors should highlight more explicitly what the new insights of this study are compared to previous work on the topic (I understand it is the regime shift).
First, following comments from Rev2, we will further develop the introduction with a paragraph dedicated to past research on Black Sea oxygenation and ventilation dynamics, in order to better position our findings with respect to those of previous studies.

Considering the points addressed in this review, we can reformulate our main findings as follows :
- We issue an unified, synoptic metric to characterize inter-annual variations in CIL formation. The consistency between independent sources demonstrates the accuracy of this metric. Note that some dependencies exist among certain sources, this is discussed explicitly.
- Analysis of the time-series reveals a current restricted ventilation regime that is unprecedented over the last 64 years. Interannual variations appear to be better described by considering abrupt change models than smooth dynamics.

- In particular the actual regime includes years during which **no** new CIL waters are formed, which appears to be a new phenomenon. We highlighted this in conference proceedings (EGU GA, May 2018, Kiel conference on marine deoxygenation, October 2018) . Stanev et al. 2019 later discussed this issue on the basis of Argo observations. Here, we propose a long term multi-source description of the CIL dynamics, to insist on the anomaly of this feature in regards to an extended period.
- The general relationship between CIL formation and oxygenation of intermediate waters has been previously described in the literature. Here, we exploit the dense Argo data sets (7 years) to document in more detail and at a finer scale the response of oxygen conditions to CIL formation at different pycnal levels. In particular, we illustrate the fast decrease in oxygen conditions for years where no CIL water was formed.

The whole study stresses the current prevalence of a restricted ventilation regime, its impact on the Black Sea oxygenation status, and the urgency of research efforts dedicated to alleviate or at least foresee potential environmental and economic consequences.

1.4 The appendix seems to be incomplete as there are references to non-existent figures.
-> Figures A1 and D1 have been erroneously displaced at the very end of the pdf, which is probably why the reviewer missed those (also, as the figures are quite heavy. I experienced some lags while scrolling down to that part of the pdf on the Copernicus website). It will be ensured that all figures are distributed coherently in the revised manuscript. However, following the advice of reviewer 4, Appendix D will be removed, and replaced by corresponding addition in the text. The specific issue of Appendix D is further addressed in *1.7*

Specific comments
1.5 Title: I suggest changing it to "A new intermittent regime of convective ventilation threatens the Black Sea oxygenation status" as the study cannot really isolate the climate change impact from that of regional atmospheric oscillations.
-> Agreed. see *1.1.*

1.6 L58: This should also be mentioned when describing the statistical model using atmospheric forcing in the Methods
-> Several comments pointed towards a more explicit description of this statistical model. We centralized our answer in *1.13.*

1.7 L71/72: I am not fully convinced that your analysis actually separates the convective ventilation from the biogeochemical processes (BGC). The observed oxygen from Argo is affected by both physics and BGC. In order to isolate the physical (ventilation) component, would it make more sense to analyze oxygen saturation concentration (and AOU for changes in BGC)?
-> For the sake of simplicity, and according to the suggestion, all oxygen considerations will be re-issued on the basis of saturation concentration (ie. in situ concentration as a percentage of oxygen solubility at in situ conditions). We attempted to indicate with Appendix D and Fig D1 that, for the layer below sigma 14.3 kg/m³ there is a quasi-strict linear relationship between molar concentration and saturation concentration, which is due to the fact that waters below these layers have a very narrow range of thermal and haline variability, so that the oxygen solubility only vary in a very narrow range, and the consideration of molar concentration or

saturation concentration bears quasi-identical results in our analyses. However Rev4 suggested removing this part and we agree it can be simplified.

L71/72 will be rephrased. Indeed, we're unsure how ventilation effects can strictly be separated from biological effects when working with in-situ observations. What we meant in L71/72 *"Our analysis thus aims to isolate annual convective ventilation as a particular component of the complex Black Sea deoxygenation dynamics (Konovalov and Murray, 2001; Capet et al., 2016) that typically involves intricate biogeochemical and physical processes."* is only related to the protocol of the analysis. By evidencing a direct correlation between CIL ventilation and oxygenation at depth, we aimed to isolate the effect of this specific driver (ie. ventilation), although of course 1) correlation is not automatically causal, and 2) indirect impacts (eg. CIL influencing BGC processes in some ways) cannot be ruled out and are in fact most probably occurring.. To our knowledge, a strict distinction of the ventilation and BGC terms should involve a dedicated model study which is left for further work. Please consider also related issue *2.4.*

1.8 *L73-76: Maybe this "section list" is not needed? If you keep it, the last sentence seems to be incomplete.*
-> We'd like to maintain the section list to support partial reading. It will be completed.

1.9 *L85: Assuming that density increases with depth, this density criterion only defines the upper limit of the CIL. How is the lower limit defined?*
-> On the basis of the temperature threshold. 'CIL points' from a profile are required to meet both temperature and density criterion. While the density criterion is needed to define the upper boundary, the temperature criterion is sufficient to identify the lower boundary, due to the stable higher temperature of deep waters.

In our processing, if a profile does contain 'CIL points', it is only considered for analysis if the CIL's upper and lower boundaries are identifiable (ie. 'non-CIL points' are present above and below CIL points). If a profile does not contain any CIL point, it is considered (with cold content=0), if it contains more than 6 points, and the deepest observations are deeper than 80m. This excludes the possibility of missing an existing CIL layer at the sampling location. The above criterium will be detailed upon revision of the manuscript.

1.10 *L111: Include the overall number of Argo profiles used and also include the minimum and maximum number for individual years (and state corresponding years) in order to give an idea about sampling error.*
-> The numbers of used Argo profiles will be characterized accordingly.

1.11 *Given the time series of CIL content rate of change (Fig. 6), I am also wondering whether it would make more sense to only use summer/fall profiles (as rate of change is very small during that period; analogous for the other datasets)?*
-> The answer is technical, and we would like to avoid overweighting the manuscript. We'll answer the reviewer's question here, but our position would be not to include it for the sake of readability, considering that it is sufficient to state clearly that an annual averaging perspective is considered for all sources.

The most difficult time series to compose is probably the one from ship-borne profiles (ie. "Diva"), since it has to compose with (strongly) uneven sampling (both seasonally and spatially). This time series is extensively described in a dedicated publication (doi.org/10.1007/s10236-013-0683-4). In the present manuscript, we consider strictly the time series described in this publication (no temporal extension or revision of the methodology), in order to be able to simply refer to this previous work.

The time series is obtained as annual "trends" which is a by-product of the DIVA detrending methodology. To make it short, it corresponds to annual "additive anomalies" associated with observation points in order to consider these measurements in the construction of a long-term climatology. In particular, those annual trends are identified conjointly with seasonal trends, in the frame of an iterative process. So the notion of "annual average trend" is intimately embedded in the procedure. From thereon, the same "averaging-window" was adopted for each data source, as a requisite for comparability.

1.12 *Table 1: I suggest listing the different datasets in the same order as in the text.*
-> We will reorder items in the table in order to have consistent ordering in the table and text.

1.13 *I understand from the text description that the statistical model, based on atmospheric predictors, uses more than one (winter air temperature anomalies) predictors, right? If so, please state all of them either in the text, in the table or both.*
The statistical model in question does only consider winter air temperature (AT) anomalies, but the AT time series is considered several times with different time lags. So practically, it's $C \sim a\ T_{w} + b\ T_{w-1} + c\ T_{w-2} + d\ T_{w-3}$, where C is annual anomaly of CIL cold content, T_w is AT anomaly of the preceding winter, T_{w-1} is AT anomaly of preceding winter + 1year, and so on.. (eg. for C anomaly of 2001, T_{w} is from Dec2000 to Mar2001, and T_{w-1} is from Dec1999 to Mar2000, etc..). All this is explicitly stated in the reference given, as well as the fact that other predictors (e.g. winds, summer temperature) were considered but discarded by the stepwise procedure. This issue will be clarified in the text.

1.14 *For convenience, the time periods covered by the different datasets could be added to the table.*
-> Time periods will be added to Table 1

1.15 *L150-151: What is the smallest number of overlapping years between datasets? Is it appropriate to calculate correlations in the cases of least overlap?*
-> Those numbers are given explicitly in figure A1 that was erroneously transferred at the very end of the pdf. Smallest number is 8 years between "Argo" and "Atmos" time series, providing a p-value of <0.002 for the correlation.

1.16 *Regarding the atmospheric predictor model and the 3D hydrodynamical model, have any of the observational datasets (Argo and ship casts) been used to calibrate/develop either model? If so, those datasets would not be fully independent (during overlapping periods), which would affect the composite C time series (especially, its uncertainty as it may in-/decrease if they were in/dependent) and should be addressed in the discussion.*
-> This point of inter-dependencies is answered in **1.2.** The question of the composite time-series construction is addressed in **4.5** (note that model time-series that may be affected by dependencies are given a smaller weight in this composition). We did not attempt to estimate the uncertainty associated with the composite time series, as deriving uncertainty estimates for individual sources in a way that is consistent and allows a combined estimation of uncertainty for the composite time-series would be too complex.

1.17 *L156: I think Appendix A is incomplete as I couldn't find the information referred to here*
-> Indeed, the information was to be found on FigA1 (at the end of the manuscript).

1.18 *Section 2.3: It would be nice to show a time series similar to Fig. 1a also for oxygen in order to put the more recent development in CIL oxygen into historical context. I understand that this is not possible for the two models but perhaps it could be done for the ship casts? This would be particularly useful with respect to climate change.*

-> The long term relationship between oxygen and CIL cold content extracted from ship-casts has been explicitly re-illustrated in Capet et al 2016, Biogeosciences, including some of the present Argo data. We believe it is inappropriate to reproduce this here. Yet, following suggestions from Rev3, we now propose to address the CIL-oxygen in large scale historical context (based on World Ocean Database data), using a TS diagram -> see. **3.13**.

1.19 *L174/175: I am wondering whether the recently submitted work by Gordon et al. (https://doi.org/10.5194/bg-2020-119), who suggest a correction of BGC-Argo oxygen observations based on the sensor response time, could help to make use of both descending and ascending profiles?*

-> Thanks for this very interesting reference. In addition to the fact that complication arises to use both ascending and descending profiles, and although we understand that these complications can be technically alleviated on the basis of the proposed reference, we considered the following additional arguments : 1) descending and ascending profiles are very close to each other since they pertain to the same sampling cycle. While those might be relevant when considering very short time scales, they are quite redundant at the time scales of interests, and 2) in the dataset obtained from the coriolis center with the given selection criteria, we obtain 1800 descending O2 profiles and about 250 ascending profiles. We thus consider that including ascending profiles is not worth the burden of technical description in this precise case.

1.20 *L203: It's probably my personal taste but I don't really like the term "routine" regime. Maybe "standard", "normal" or simply "average"?*

-> "Routine Regime" would be replaced by "average regime". Thanks for the suggestion.

1.21 *L213-215: Would it make sense to show annual (or winter) average surface air temperature as a small panel in Fig. 5 to better demonstrate this link?*

-> In L213-215, we referred to 'warm regime' in terms of CIL cold content. Besides, the "Atmospheric" time series is explicitly based on winter temperature anomaly, in a cumulative way which is more directly relevant to the CIL cold content (cf. **1.13**). Also, we decline the reviewer's suggestion. We will, however, refer here to Oguz et al 2006 (https://doi.org/10.1016/j.jmarsys.2005.11.011) as regards long term variations in the winter air temperature.

1.22 *L218: More out of interest: would it be possible to also get intra-annual resolution from the statistical model? E.g. by not using winter time averages of the descriptor variables but monthly averages or so?*

-> We believe this level of detail is beyond the scope given to the paper. Detailed works on CIL formation and seasonal development have been done and are referenced within the manuscript. See in particular Capet et al, 2014, Akpinar et al 2017, Miladinova et al 2018. Also in L218 and paragraph, the idea is to quantify CIL formation and destruction rate from the difference between consecutive stocks. This can only be done on the basis of complete inventory such as can only be obtained for the synoptic 3D model product.

1.23 *L225: I don't understand what you mean with "before, during and at the end the thermocline setting". Mainly because I am not sure what you mean with "thermocline setting"; is it the thermocline formation or the period during which a near-surface thermocline exists?*

-> Poor wording indeed. "thermocline setting" will be replaced by "thermocline season".

1.24 *L231: Oxygen is from Argo. Would it make sense to do the intra-annual comparison of CIL formation rates also based on Argo then (or a combined C using 3D model and Argo)?*

-> This relates also to the answer given in **1.22**. CIL formation is not achieved at once, nor in a single location. Rather it proceeds from multiple small scale events in the open part, and in the northwestern shelf (see Akpinar 2017 and Miladinova 2018 for respective importance of both formation pathways). Argo sampling of the CIL is incomplete (important spatial variability and subsampling). Although we assumed to use it to build an annual metric of CIL content (with the caution notice of Sect. 2.1), we wouldn't use Argo to derive CIL formation rates at smaller temporal scales which demands differentiation between complete inventory estimates.

1.25 *Discussion: As stated in the very beginning, it would be good if differences to earlier studies/novelty would be highlighted more explicitly. In my understanding, the main differences in terms of methodology are the longer period and the basin-scale integrated approach, which is needed to detect the regime shifts. This needs to be made more clear and it should also be discussed what the advantages (for the purpose of this study) and possible limitations are.*

-> see **1.3.**

1.26 *L286/287: If one of the main conclusions focuses on the combined linear-periodic model then this combined model should be included in Figs. 3 and 4.*

-> It will be added to Figure 3 and fig 4 will be removed. Please consider **4.9** and **4.10**

1.27 *Appendix: There seem to be two figures missing: Fig. A1 (L307/308) and Fig. D1 (L379).*

-> Both figures A1 and D1 appear on the very last page of the pdf, available at the BG discussion website. As D1 is relatively heavy, scrolling down that page is somehow slow, which may explain why it has been overlooked. A lighter version will be embedded upon revision and figures will be placed consistently.

1.28 *The appendix contains a few aspects that would fit into the discussion in most journals (incl. BG), e.g. the discussion of the suitability of the statistical methods. I assume this is owed to the previous submission to GRL. The authors may consider moving parts of it into the discussion (or even into the Methods, e.g. the statement on oxygen data on L378-380). However, it also works the way it is. Maybe the editor can have a say on that.*

-> Appendix A will be maintained as is. Appendix D will be suppressed, with essential parts included in the Methods. As concerns appendix B and C, please refer to **4.6** and **4.7** (essential part of the regime shift analysis method will be embedded in the text, details will be maintained in a streamlined appendix).

Technical corrections: […]

-> These technical corrections will be considered in the revised manuscript, but are not addressed in detail in this response.

Reviewer #2

2.1 The title sounds interesting and attractive. However, for me, the result sounds as follows: There were two cold winters in the period of 2012-2019, and in each of them was not only the amount of cold waters larger, but the concentrations of oxygen were also higher. There is nothing new, except that this has been recorded by BioArgo.

This appears as a reductive description of the manuscript results.

Result section can be divided in :

- Construction of composite long term (1955-2019) time series for CIL cold content using four sources of different nature (models, observations), and assessment of the coherences and discrepancies between these sources.
- Analysis of the above time-series by comparing different descriptive models, incl. identification of regime shifts and detailing of the various regimes.
- Detailing the contribution of Cold water formation to oxygen ventilation (over Argo period), including vertical (pycnal) distribution of the ventilating action of CIL formation.

All those steps are presented with quantitative considerations that we believe to be relevant for further works addressing 1) Black Sea specific dynamics and 2) interlinks between ocean warming and deoxygenation.

2.2 Neither, the scientific content of the manuscript supports this title.

The title will be modified, following the suggestion given by Reviewer1, ie. "A new intermittent regime of convective ventilation threatens the Black Sea oxygenation status".

2.3 We read in the Abstract: "Oxygen records from the last decade indicate a clear relationship between cold water formation events and oxygenation status at different isopycnal levels, suggesting a leading role of convective ventilation in the oxygen budget of the upper intermediate layers." This just repeats what has been known for many years. This finding is just a confirmation of previous knowledge (see several papers of Konovalov et al).

We acknowledge that the term "suggesting" is inappropriate, and apologize for this. The fact that CIL formation contributes to ventilation is indeed not a novel discovery and wasn't considered as such. As underlined by reviewer2 this relationship has been considered several times in the past Black Sea litterature, including in the works of S. Konovalov.

However,

- considering this dynamics from the standing point of Argo profilers provided an insight at an intra-annual scale (ie. weekly), and enabled an detailed appreciation of the depth (pycnal) penetration of the CIL ventilation effect at an intra-annual time-scale, which is new.
- Here, we used Argo data to illustrate not only the ventilation effect, but mostly the absence of ventilation for the years where CIL ventilation does not take place, ie. in the context of the regime shift highlighted in the first part of the manuscript. This was done not to reveal the general ventilation role of CIL formation (as this was done before) but to highlight the potential consequences of a regime shift in CIL formation by revealing the absence of ventilation during the years where no new CIL waters are formed. This will be highlighted more explicitly upon revision of the Discussion Sect. 5 and abstract.

We will also further develop the introduction with a paragraph dedicated to past research on Black Sea oxygenation dynamics, in order to better position our findings with respect to those of previous studies.

__2.4__ Possible shifts of temperature-oxygen relationships in different periods have also been broadly addressed in these studies. Authors say nothing about that.

For the sake of discussion we consider that the reviewer refer to in-situ conditions, and to the long-term dynamics described and commented for instance by Konovalov et al. 2001 (in particular the well known Fig. 10, reproduced below), a dynamics that has been related to shifts in the biogeochemical terms of the oxygen budget operating over large time scales.
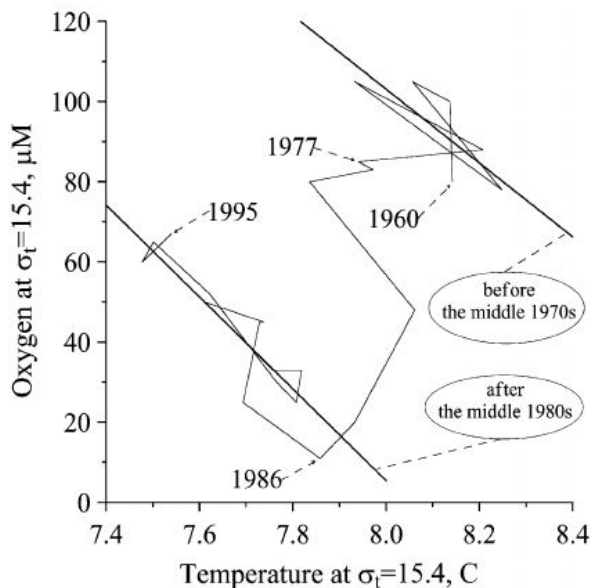

Fig. 10. Variations in oxygen concentration versus temperature in the layer of the main pycnocline ($\sigma_t = 15.4$). (Average values for individual cruises listed in Table 1 and liner fit for the periods before 1975 and after 1985).

Our manuscript does not consider the material needed to address biogeochemical terms (eg. nutrients and production). Instead it focuses on highlighting and documenting the CIL ventilation term of the oxygen budget equation and more precisely the absence of this term for some years of the recent decade. The data considered to this aim cover a relatively short period [2011-2019], for which major shifts in biogeochemical regimes, such as the historical Black Sea eutrophication and recovery, can be conservatively ruled out.

Finally, shifts in the long-term oxygen and CIL cold content relationships have been explicitly illustrated in Capet et al 2016, Biogeosciences, and we did not wish to repeat these in the present manuscript. See however request from Rev3 (*__3.12__* & *__3.13__*), to put the finding of the Argo Period, in relation with historical perspectives. .

__2.5__ This brings me to the major criticism. Authors have to make clear what the new knowledge is, which can be gained from their study compared to older ones.

-> Please refer to *__1.3__*

2.6 One big problem is that there is a lack of balance in the manuscript. Much attention has been given to the long-term variability. On this subject, I cannot find anything new.

Regarding the long term variability. To our knowledge this is the first time that CIL dynamics is documented using an unified, integrative metric over such an extended period, that multiple sources of information are compared, and that the regime shift analysis paradigm is compared to other descriptive models, which is important 1) to highlight the existence of thresholds in the response of CIL formation to atmospheric temperature and 2) to highlight the exceptional nature of the absence of CIL formation for some years within the currently prevailing regime.

2.7 The intimate link between analyses of 65-year and 7-year time series is not clear. The good part of the research is the analysis of data after 2012. However, its relation to the previous periods is completely unclear, and neither is it well articulated in the manuscript.

The last seven year period falls within the identified last regime, and is used to illustrate the impact of this regime shift on oxygenation conditions. The model time-series allows to interlink the long-term framework to the recent period, since it extend over both period while proceeding from a unique model setup (ie. forcings and physical assumptions). A dedicated effort will be deployed upon revision in order to better express this in the discussion section, which will be completed in agreement with *3.12* and *3.13*

2.8 In this part [ie. analysis of data after 2012.]*, authors should clearly describe which floats they use, and how these floats capture the temporal and spatial variability. Important to know is whether what is observed is a clear signal or just noise. The statistics of data, and how representative for the Black Sea state they are, need deeper consideration. Fig. 1 shows perhaps that data used in the analyses are not homogeneous. If the data is not homogenous, the subsequent interpretations of long-term changes are not credible.*

The spatial and seasonal distribution of oxygen Argo floats is easily accessible on numerous platforms including the ones referred to in the manuscript (see for instance the display tool of the Coriolis website). Also, adding maps of floats positions was not considered as essential.

This, of course, does not dismiss the need to properly consider the representativity of the data.

Here the reviewer comments on data analysis after 2012, hence the part on oxygen although the Fig. 1, which is referred to, concerns the CIL cold content.

- Regarding the CIL cold content series extracted from Argo, there is an expected spread between individual floats that is indeed illustrated on Fig 1 and reflects the well known spatial variability in CIL cold content. In the sake of setting this data set on the same level as other data sources annual means were extracted from the Argo dataset (see. L105-115). The resulting annual averages were afterwards compared to other data sources (see. Appendix A. and Fig A1) to ensure consistency of the analysis.
- Regarding oxygen, all available oxygen Argo data were used considering data selection criterion that are described in Sect 3. L173-179. The question of representativeness is considered and explicitly discussed on L. 180-185, and Fig. 2. In particular, it illustrates the spread of oxygen conditions derived from individual floats and shows that the use of pycnal scale provides a sufficient spatial homogeneity to characterize a basin wide condition. Finally, Fig. 7 also provides the interquartile ranges of monthly-binned Argo-derived oxygen concentrations at different pycnal levels. Considering that Argos profilers are spread across the basin, we would expect much more inconsistency and overlap between the inter-quartile ranges areas, if as the reviewer suggests, spatial variability overpasses the temporal

signal embedded in this dataset. This will be expressed explicitly in the revision of paragraph L231-235.

2.9 *The statement ". . . suggests that the CIL renewal, that was taking place systematically each year in precedent regimes, has now become occasional." contradicts what is known from earlier studies. They have to at least refer to Lee et al. (2002. Anthropogenic chlorofluorocarbons . . .) who claim that the residence time is ~5yr at 80-120m. I would recommend that they explain what the problem in this earlier estimate is, if any.*
We apologize here for a clear misunderstanding, due to inappropriate wording. Instead of "*complete renewal*" of the CIL layer as seems to have been understood here, we wanted to refer to a "partial renewal", ie. the formation of a certain volume of new cold water, adding to the remaining CIL content of the previous year and contributing to the gradual renewal of the CIL layer. We totally agree with the fact that the residence time of the CIL layer is larger than a year, this is indeed well known and was referred to (maybe imperfectly) on L.48-59.:
"If a well-formed CIL was present during the previous year, subsurface waters exposed to atmospheric cooling are already cold, which increases the amount of newly formed CIL waters (Stanev et al., 2003). Due to this positive feed-back and to the accumulation of CIL waters formed during successive years, the inter-annual CIL dynamics is better described when winter air temperature anomalies are accumulated over 3 to 4 years, rather than considered on a year-to-year basis."
        We will clarify the fact that CIL renewal time is larger than one year in the revised introduction, including references to the works advised by the reviewer.

2.10 *The issue of regime shift is not convincing. The question is: can oxygen data over 7 years only identify regime shift? What was the previous oxygen regime?*
There seems to be another misunderstanding here.
- The regime shift paradigm is used to characterize the dynamics of CIL formation, and is based on the analysis of 64 years long time series.
- Oxygen data are used to illustrate the impact that this regime shift in CIL formation may bear on the oxygen dynamics.

2.11 *What I see in the oxygen data are just two ventilation events, not a shift. Authors ignore referring to important works about regime shift (e.g., the review article of Oguz in Front. Mar. Sci., 25 April 2017). They have to study the references in this review. Based on my comments above, I am very sorry that I cannot recommend publishing.*
See **2.10** as concerns the remark on regime shift and oxygen. We will carefully reconsider *Oguz in Front. Mar. Sci., 25 April 2017),* and extend the introduction by considering environmental regime shifts in the Black Sea, however, note that we still want to consider ventilation aspects apart from other drivers. Again, the question of a shift in the biogeochemical terms of the Black Sea oxygenation balance is not directly addressed in the manuscript. We really want to focus on Cold water formation as the main ventilation mechanism, and to highlight a strong, sudden, non-linear reduction in this process, and to illustrate the impact it's likely to bear on Black Sea oxygenation status, regardless of other environmental shifts that may be happening in parallel (ie. at first order).

Reviewer #3

This is a very interesting paper. Basically the authors:
> *1. Assemble a time series data set of temperature data from the Black Sea from about 1955 to the present combining real observations and model results.*
> *2. They integrate cold temperature anomalies in the Cold Intermediate Layer (CIL), relative to a reference temperature of 8.35°C which defines the upper and lower boundary of the CIL, which they call "the cold content".*
> *3. They analyze the (extensive) variability in this "cold content", using several different approaches (linear and sine functions), but settle on a technique they call the "regime shift" paradigm or hypothesis.*
> *4. Using regime shift they identify four periods that characterize different amounts of "cold content"*
> *5. They argue that these periods reflect variable degrees of ventilation of the CIL*
> *6. The last 11 years have been a period with unusually low ventilation and they argue that this is due to ocean warming.*

I have a few specific comments,

3.1 The paper is a little hard to read because of the advanced data analysis techniques used and (though generally well written in English) some awkward word choices. I'm not sure who can fix that.

-> We refer to exchanges with Rev4 as concerns the technical justification part that needs to be retained as essential to the main text, or that can be directed to the appendixes. A certain level of technical detail can not be avoided. As concerns the english, we will ensure to have the revised manuscript checked by native english speaker.

3.2 L1 Abstract – a ~100 m ventilated surface layer is referred to but does that mean 0-100m No, it means the Cold Intermediate Layer which is more like 50 to 100m I think you should be more specific.

-> No, in this sentence we indeed refer to the part that is not anoxic so 0-100m, and not 50 to 100m. But we understand the concern as the ventilation of the upper part (ie. 0-50m) is not related to CIL formation. The abstract will be revised to avoid this confusion.

3.3 L20 The early literature (e.g., Tolmazin 1985, Progress in Oceanography 15, 217) argued that as it appeared that the sea surface in the central gyres never got cold enough for replenishment of the CIL by winter convective, that the main source of water to the CIL on an annual basis was from the NW shelf where the key density surface was cooled. I agree that we have much more data now and starting with Gregg and Yakushev (2005), who observed a ventilation event (with real data), we now know that ventilation can occur from the central gyre regions. But the NW shelf hypothesis has been totally left out of all papers since the 1990s, such as those by Akpinar, Ivanov, Oguz and others. I looked back at those papers and they don't even mention the Tolmazin argument, much less argue why it would not play a role. So as far as I can tell, the NW shelf is a possible source of ventilated water for the CIL. If the Tolmazin hypothesis has been disproved, I missed that. I think Capet et al should take that into account. It may not show up in their model, depending on how it is parameterized.

-> This important point is discussed extensively in 3.7 below.

3.4 *L29 Murray et al (1989) discovered the suboxic zone. Stanev et al (2018) is a nice paper but used model results to argue for what causes its origin.*
-> We will update the reference accordingly. We apologize for this mistake.

3.5 *L104 Why not describe the data sources in the same order as presented in Table 1?*
-> We will reorder items in the table in order to have consistent ordering in the table and text (cf. *1.12*)

3.6 *L126 How were the Atmospheric Predictors converted into C and CIL temperature variability in the water column?? I don't think anything is said.*
-> The entire procedure, and the resulting statistical model is explicitly detailed in the reference given (Capet et al 2014). We did not restate the procedure and results of this study, because we want to focus the present debate on new investigations and results. However, seeing this comment is coherent with request *1.13,* the statistical model will be given more explicitly.

3.7 *L136 Does that 3D hydrodynamic model include source water from the NW shelf?*
-> The 5 km 3Dmodel, whose setup is described in the given references, does indeed resolve a shelf source of CIL waters (as well as a source from central gyres), but this is a result of internal dynamics, not a specifically imposed behavior. Because the reviewer seems interested in this question, I may refer to Fig 4 and 5, and Section 6.3 of A. Capet et al. / Deep-Sea Research II 77-80 (2012) 128–142, that uses the same model. It depicts a second EOF of CIL Cold content which could correspond to the relative importance between the two sources of CIL water formation, which relates well to a longitudinal gradient in sea surface temperature.

The more recent works of Miladinova et al 2018, further extends on detailing the relative contribution of the two sources of CIW. I understand and share the reviewer's interest. It would be relevant to investigate whether a trend exists in the respective importance of the two CIL formation mechanisms, and if these two mechanisms imply distinct biogeochemical consequences (they certainly do, as the biogeochemical signature of those water's origin are drastically different). However, we prefer to maintain the focus and the general 'integrated' point of view. Besides, making a distinction between CIW origins would only be possible for the 3D model. In short, this is a related and relevant question, but we consider it extends beyond the scope of the present manuscript. We will however state in Sect 2.1 that the CIL dynamics as resolved by the 3D model is explicitly described in the above reference.

3.8 *Figure 5 with the intervals obtained from regime shift analysis is compelling. But I think back to the geological axiom "If I hadn't believed it, I wouldn't have seen it!". Visually (without the vertical lines) it looks like there is more variability than shows up as the std. deviations. But this is not my area of expertise so I hope other reviewers can evaluate that.*
-> See exchanges with reviewer4.

3.9 *L173 Section 2.3 The absolute values of O2 are uncertain unless the sensors are carefully calibrated. I suspect they were not. The relative changes are probably OK.*
-> Agreed. We only consider relative changes here. Also, we used monthly median values of multiple buoys with different ages we may assume an acceptable overall bias.

**3.10** *L181 Why 2018 and not 2020?*

-> This is a mistake. 2010-2018 should be replaced by 2014-2018, ie. the period where multiple buoys are present at the same time, which better illustrates spatial variability.

3.11 *L289 I think Konovalov and Murray (2001) showed this in a figure.*

-> That is totally true. Please consider our answer in **2.3**

3.12 *I really prefer the real data over the model results, which just reflect what equations and parameters were put into the model. For this reason I would really like to see a T-S plot (with real data) maybe averaged (with std. dev.) for the 3 regimes, blown up to highlight the CIL region. The intervals in Figure 5 should show up clearly.*

-> Here follows a T-S diagram, focused on the CIL part of the plot, based on in-situ data from the World Ocean Database 2018 including CTD, Bottle, and profilers data over the 4 periods (ie. 1955-2020). The four panels on the left indicate individual measurements, as well as contour surrounding 75 % of observations (obtained by two-dimensional kernel density estimation, see cf. function kde2 from the R package MASS). Panel on the right, repeats the same contours for each period. The four periods indeed show up quite clearly. Even ignoring the contours computations, it clearly appears from the scatterplots that records of temperature higher or equal to 8.35°C in the potential density anomaly range [14.5-15.0], are rare exceptions during P1-P2-P3, and appears as the main rule in P4. The figure will be included and discussed in the discussion section.
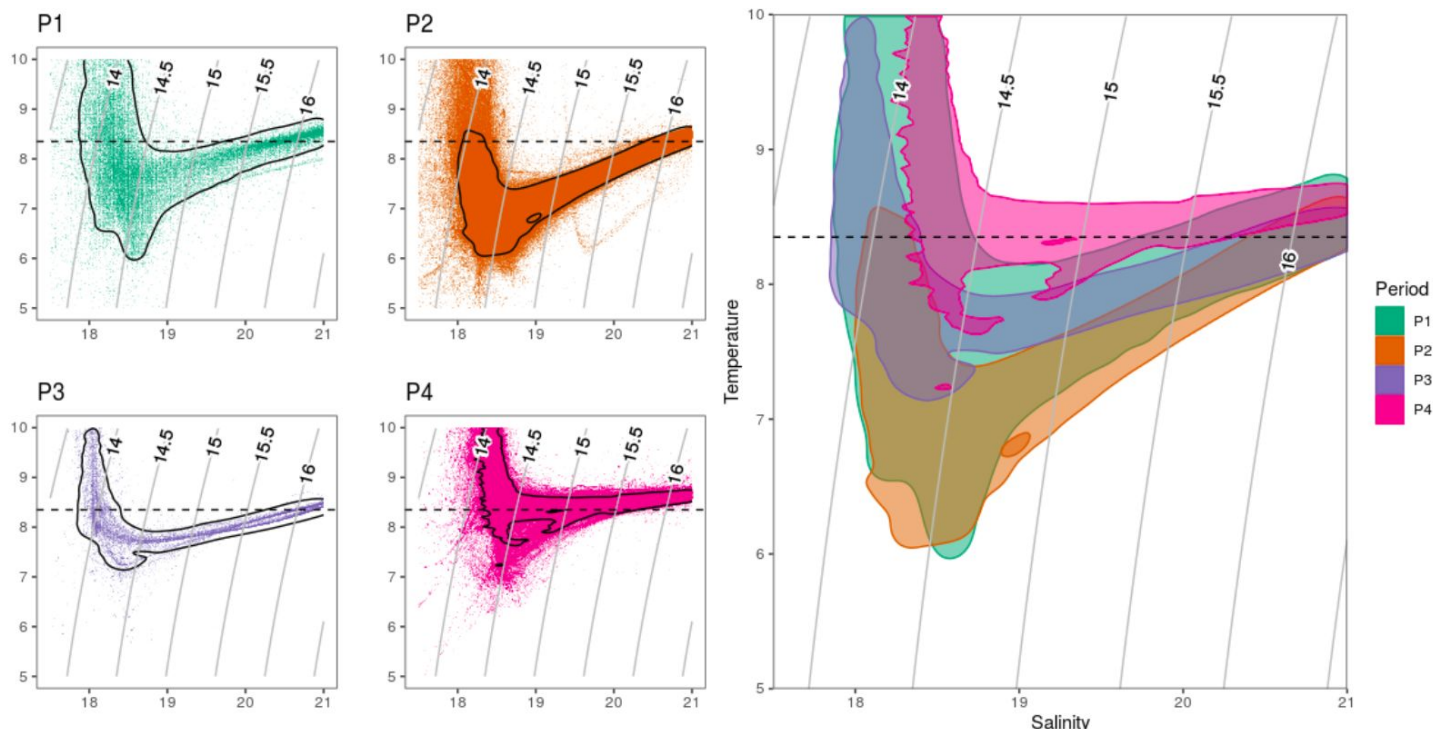
3.13 I think it would be useful to explain another reason why the CIL is important. In my view (see Murray et al., 1991) it is because it plays an important role in the formation of all deep-water in the Black Sea. To a first approximation, all deep water in the Black Sea forms by linear 2-end member mixing between the Bosporus outflow and the CIL. This must be because salinity increases all the way to the bottom and the only source of salinity is the Bosporus. See Figure 12 of Murray et al (1991). This mixing occurs on the SW shelf (Latif et al., 1991). Any curvature in the T-S plots is due to temporal variability in the signature of the CIL endmember. This mixing can be seen in the T and S sections from the Bosporus to the shelf break from Gregg et al (1999). See Murray et al (1991) for more discussion. If the CIL is warmer, and less dense, how will that impact deep water ventilation? I think the deepest layers will be ventilated less frequently.


Temperature and Salinity along the Bosporus
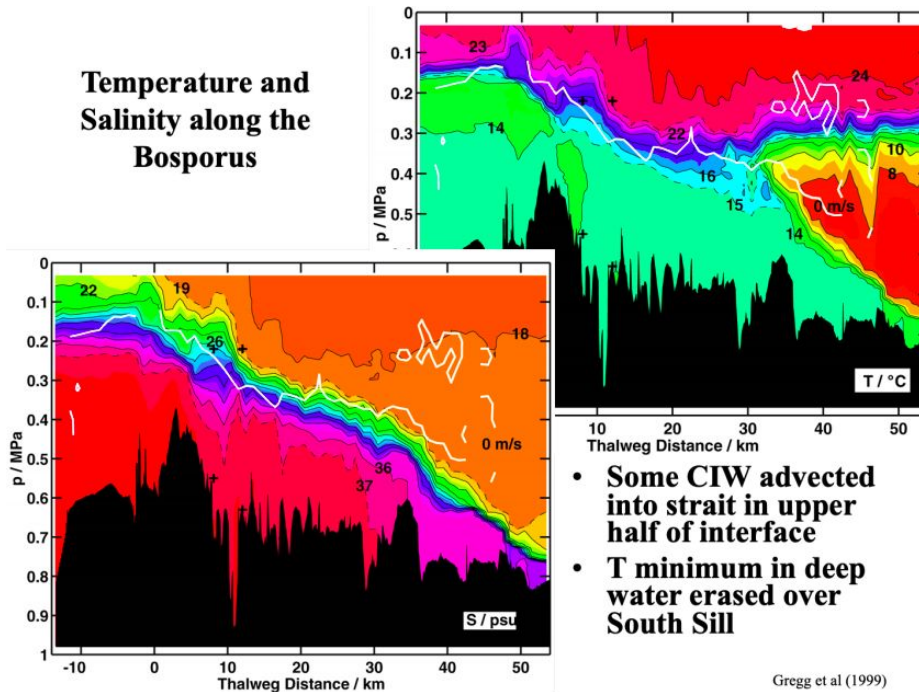
- **Some CIW advected into strait in upper half of interface**
- **T minimum in deep water erased over South Sill**

Gregg et al (1999)

-> We thank the reviewer for this very important comment. Keeping the focus on oxygen, we issued the following figure (left panel below) which represents oxygen saturation concentration on the T-S diagram (WOD data, all periods, bottles,CTD and profilers data selected with "good" quality flags)). Taking, for instance, the 14.5 kg/m³ isopycnal as the CIL end-member for mixing of deep waters, it clearly appears that this layer is less oxygenated in the case of low CIL content, which relates to the fact that the CIL layer is composed of a mixture between CIL from the previous years and annual added now oxygenated CIL waters. The right panel evidences this for specific sigma layer, In particular, it provides a link between the recent (Argo) period (fig 7 & 8), and the longer time scale as is required by Reviewer 2 (see. *2.7*). These figures will be included to enrich the discussion regarding the implication of a new CIL regime on the ventilation of deep waters.
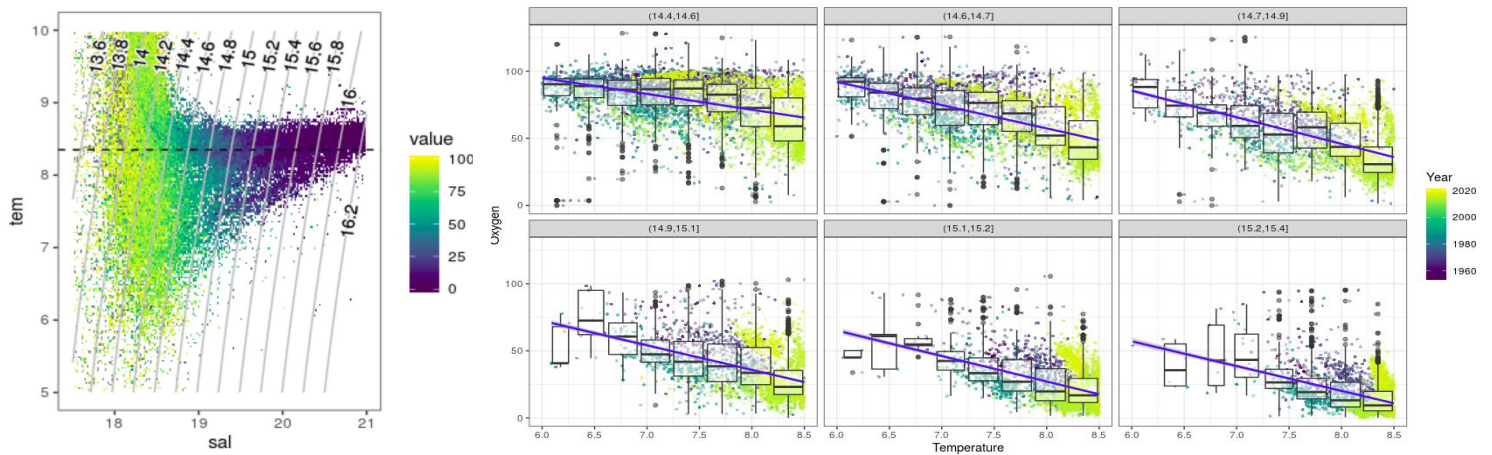
Reviewer #4

I was asked by the editors to review this manuscript with an eye towards an assessment of the statistical methods that were used. Oceanographically, the notion of different convective regimes and their link to Black Sea oxygen dynamics is established using (mainly) observational data. A causal relation, of course, cannot be inferred from this, but the physical linkages are plausible ones and great effort has gone into making sensible metrics for the cold intermediate layer, with a solid understanding of the limitations of each of the data types used. I don't know enough about the literature pertaining to convective processes, nor regional Black Sea oceanography, to comment on significance of this contribution and its novelty within oceanography. For the remainder of this this review, I will focus on the statistics and data analysis aspects.

The paper quantifies the ventilation of the Black Sea intermediate layers through a heat (cold) content metric, i.e. the CIL defined in eqn (1). Estimates for the values for the CIL over time are computed from 4 sources: the first two are observational data (ARGO floats and CTD casts); the next is output from a statistical model that relates air temperature to CIL; and the last is numerical ocean model output. The different CIL estimates have very good correspondence to one another, and a composite CIL time series is produced for further analysis. The central feature of the paper is a changepoint analysis of the CIL series was undertaken to isolate different convective regimes. This is then followed by correlation analysis to relate CIL changes to oxygen dynamics.

4.1 From an visual examination of the data, I doubt that the 4 regimes found exist as distinctive equilibrium states (if that is how one defines a regime), but are rather part of a continuously varying process with cycles and trends like we see in all climate data. These CIL data do clearly show underlying long period cycles and a decreasing trend since 1980. As for the existence of regimes, looking at the graphs without the regimes (Fig 3), and with the regimes superimposed (Fig 5) – one's eye is drawn to the regimes in Fig 5, but these are subtle at best. A statistical change point analysis will always pick up regimes, and in my experience the AIC criteria used here tends to choose overly complex model (here, implying more changepoints and regimes). But the above is a subjective assessment. Quantitatively, the analysis undertaken is clearly spelled out, its assumptions addressed, its application done properly, with the result that 4 regimes are statistically identified.

-> We will stick with the last sentence. We acknowledge the fact that the regime shift perception is one among others, and the manuscript will be screened to ensure that this acknowledgement is clearly expressed. In particular, what appears most important to us is that the recent change in CIL dynamics, as depicted by the regime change approach, is not a slow trending one, but includes a step change towards a new phenology of Black Sea ventilation. This is why we insist on this regime shift description, and on highlighting its potential consequences. In order to underline the difference between the actual and former regimes, besides statistical considerations, we followed the advice of Rev3, and produced the more direct illustration proposed in *3.12*.

4.2 These regimes here are changes in the level (mean) of theCIL over time; there appears to be no changes in the higher order statistical moments. Note that there are many different ways to do changepoint analyses, and each need not be only based on the mean level, but could use other metrics (variance, autocorrelation, skewness) to break up the series.

-> It appeared to us that considering only the first order moment would already appear as a complex approach (see other reviewer comments in that sense), yet it is sufficient to make our point. This is why we discarded the option to consider other higher order metrics.

4.3 I feel there is value in putting forth this regime analysis and the results for discussion to the wider community, and hence support publication of the paper. I would, however, downplay the ambitious claims of the title. This is an analysis that brings out the recent decline in the oxygenation state, its relation the ventilation and the CIL, and suggests the possibility of CIL regimes. But these regimes are subtle at best and not easily separable from the general downward trend in CIL. The last regime might be different from those in the past (less cold), but further work would be needed to verify/validate this as a new regime.

-> The main idea is to oppose non-linear regime dynamics to smooth linear and sine trends models. Yes, this is gladly open to criticism, and future observations will tell more about the relevance of this paradigm. Yet, it appears worthy of presentation to the interested audience.

*4.5 My main criticism is that I found it difficult to follow the methodology and how it was applied. The only reason that I was able to do so was since I have used most all these techniques before, and so could 'read between the lines' as to what was being done. The latter comment is important since it means that a**dequate and understandable descriptions of the statistical approaches used need to be included in the main body of the manuscript (not just buried the appendix details)**. Important points are glossed over (the composite time series, the atmospheric predictors model for CIL, and in particular the regime shift model). To compound this, basic statistics are discussed at length (like overfitting, AIC, and model selection, and fitting the curves of Fig 3), but in a way that is not clearly linked to the problem. Below I provide specific suggestions.*

Overall, however, it should be straightforward to insert the necessary methodological detail into the text, and streamline the appendices, so that the results would be understandable and reproducible by an educated reader. I view these as relatively minor changes (since everything is there, if you read carefully with an adequate background).

-> See specific answers below.

SPECIFIC:

4.6 *Title: a bit over-reaching. There is no explicit link made to climate. But the paper clearly demonstrates the CIL decline and de-oxygenation since 1980 whether or not these are really new regimes.*

-> Title will be changed: see **1.1.**

4.7 *Describe the "Atmospheric Predictors" model in the text (line 126). It looks like a lagged regression model. Important to be specific. Why not a basic equation?*

-> Specific description will be given in addition to reference, incl. equation. See similar comments in **1.13** and **3.6.**

4.5 *Describe the composite CIL time series (line 145) in more detail. It is a central quantity for the paper.*

- The construction of the time series was kept very simple. It simply consists of a weighted average of the 4 time-series, restricted to available sources for years during which all sources weren't available. In order to emphasize the value of direct observations, we used relative weights of 1 for observation and 0.5 for model time-series. We will ensure any ambiguity is alleviated in the text.
- Describing the time series itself is the object of Sect. 3.1. The text will be complemented with more general description (beyond the current statement of descriptive model parameters).
- The composite time series is depicted in Fig3 (stated in caption but missing in the color legend) -> **4.9**. It will be more clearly depicted in Fig5 by removing other individual sources time series (**-> 4.10**)

4.6 *Describe the regime shift model in text. It is a change-point model. These are easy to explain, if tricky to code. Be clear in the text the R-package you used, what it is based on, and that there are lots of different kinds of changepoint analysis and algorithms. Because of this lack of description, it is unclear why the authors then talk about model selection in the next paragraph (because different numbers of changepoints imply distinct statistical models).*

-> We answer here several comments related to this issue.

- We will enforce the description of the regime shift analysis and change-point model in the main text (Sect. 2.2 first paragraph). This paragraph will aim for accessibility and to allow non-specialized readers to get the basic notions of change-point models.
- The appendix will be streamlined, and organized as follows :
  - Appendix A will be maintained as is : Comparison of C time series from multiple sources and definitions of metrics used, including Fig A1, that provides important information (see for instance request **1.15**).
  - Appendix B will give technical details for the change point model and regime shift analysis including : R-package, underlying principles and options used in the R functions, precisions regarding the issue of **4.8** ("are time locations counted as parameters for AIC computation?"), and validity of underlying assumptions. This will be compiled from the actual appendixes ("

Regime shift analysis", "Statistics", "Normality","Autocorrelation" and "Biases between components of the composite time series".

○ Current appendix D will be removed and the discussion on oxygen absolute or saturation concentration will be simplified and embedded in the text.

4.7 *I think this paragraph on model selection and fit/complexity metrics to be too general and elementary – best put in the appendix.*

-> We would like to maintain the discussion about the regime shift approach, and it's relevance to describe the C time series as compared to other paradigms (linear,etc..). Mostly because this discussion highlights different conceptions of the response of CIL ventilation to atmospheric trends, and potential future evolution of this ventilation (see L244-248). We acknowledge that the statistical arguments in favor of the regime shift paradigm, although objective, are subtle at best. Our aim is mostly to propose this paradigm as a plausible point of view, this will be stressed further in the conclusions. Because we want to maintain this discussion, it appears essential to keep a few lines on model selection in the Methods section. In our opinion the fact that it is too elementary (and we are unsure that it will appear so to all readers) does not justify relegate it to appendix. So, we propose to maintain a streamlined version of this paragraph in the Method section.

4.8 *Regime shift analysis and number of changepoints. I am unclear on whether the Rpackage you used here computes the number of breakpoints as well as their time locations. The reason I bring this up is since you have used an AIC criterion to choose the number of breakpoints – is this your addition to the analysis, or is it part of the R-package? State clearly as this is the central step that determines the 4 regimes (and hence underpins your conclusions).*

-> The R package strucchange fits change-point models by computing their time location and section-specific averages (homoscedastic approach, see **4.8**) for a selected number of breakpoints (1 to 5 in our case). It provides RSS and BIC for each of those change points models. Here, we decided to use AIC for selection, following advice from colleagues, although both BIC and AIC provide the same ranking between the different change point models (ie. different by the number of change points) and between the optimal change point model and the linear, periodic and linear+periodic models. Of course, AIC computations for change point models accounts for the fact that mean for each period AND time locations of the change points are evaluated. **There is a correction here,** although it does not impact much on the results. Upon data manipulation (successive re-computation upon resubmission of this work) the time series used for AIC computation of the change point model was limited to 2017, while it is now extended to include 2019. The complete time-series has well been used for the design of the model and to get the means associated with segments, however, the AIC were derived in another version of the same scripts where the time series was erroneously limited to 2017 (and therefore shorter than the series used for the other models obviously leading to smaller AIC). So the only consequence is that AIC for the 4-segments model is corrected to 752, instead 730. 752 remains smaller than other models, so that no major changes in the conclusions arise, besides that the subtlety of this statistical support of the regime shift paradigm further needs to be stressed (cf. **4.1**).

4.9 *Figure 3 and supporting text is not needed. Fitting a linear trend and a sinusoid to the CIL doesn't add anything to the paper (and the data is repeated in Fig 5 for the regime analysis).*

Following arguments given in 4.7, we propose to maintain Fig3, in its present form, while including also the linear+periodic model (as requested in **1.26**). To avoid repetition, Fig5 will include only the composite time-series

4.10 F*igure 4. Similarly, all the model selection stuff like AIC versus changepoints is over-explained and too generic. It is enough to say in the text that AIC identified 3 breakpoints as optimal, then refer to appendix.*
-> Along successive submissions of this work, these parts have been displaced between supplementary materials, appendix and main body, several times, way and back. Here again, reviewer 1 asks instead to add a model to Figure 3 (cf. **1.26**), while Rev4 proposes to minimize this section and send it back to appendixes. We conclude that this is a matter of personal appreciation. What matters to us here is to stress the fundamental difference between slow-trending/periodic variability, as opposed to the "abrupt" change in ventilation regimes which is a consequence of the non-linear dynamics of CIL formation. To us, the regime shift analysis better corresponds to this type of dynamics. Of course, as stressed above, this will always remain a matter of personal appreciation, even if statistical considerations objectivate the approach (see **4.1**). For this reason, and because linear trends and periodic descriptive models have often been applied to describe Black Sea time series, we'd like to maintain this section in the main body. Yet, in agreement with the reviewer, the section will be streamlined and the importance of model selection minimized. It only serves the purpose to oppose two discussion points of view : the slow-trending or periodic routine to the new unprecedented regime, as will be stated more clearly. **We thus agree that the Fig4 gives unnecessary details, and propose to remove it, simply.** AIC of the models (optimal 3 change point models, linear, periodic, linear + periodic) will be given in the text as is presently done.

4.11 *Figure 6. Define dC/dt (difference between median oxygen concentrations between successive years). Is this the annual oxygenation index you refer to in line 239?*
  ● Caption will be completed : "dC/dt are basin wide CIL cold content formation and destruction rates, obtained as differences between weekly integrated CIL cold content provided by the 3D model". So dc/dt on Fig 6 is a **Weekly CIL** formation/destruction index.
  ● The **Annual oxygenation** index referred to in L239 is defined on Lines 235-236.
  ● On Fig 8, we assess the correlation between this Annual Oxygenation index, and a corresponding **Annual CIL formation** index, obtained as the differences between annual values of the CIL composite time series.
-> We apologize for the lack of clarity. Upon revision we will pay a close attention to define all those metrics in the relevant parts of the section methods (ie. 2.1 for CIL related indexes and 2.3 for oxygen related indexes) and ensure consistent references to these metrics (ie. always using the exact same words).

4.12 *Figure 8. No need for both significance (colours) and p-values (size) since they give the same information. Significance values are thresholds for p-values.*
-> We totally agree this is redundant information. We saw no harm in this, but agree now that this may be confusing. Figure 8 will be redrawn with equal size for each point, removing the point size legend, and thus considering only the significance color scale.

4.13 *Rationalize the material in appendices. It is rather a strange grab bag of material, and perhaps is a consequence of a previous review process?*

-> See answer in *4.6*

4.14 Appendix A: Put a basic understandable description of the regime analysis the main body of text (as noted). Details here. Note also that there are many changepoint analysis of various levels of statistical sophistication. You've picked one of many.
-> See answer in **4.6** (proposition of appendix structure) and **4.2** (selected of changepoint analysis)

4.15 Appendix B. I think this is a discussion of how the statistical assumptions of this regime analysis – which is based on linear regression - could be violated, and their possible consequences. Not sure you need most of this, and the important parts should be folded into Appendix A. Overall, I'll agree that within each regime, using average annual values, the normal iid assumption of OLS regression is OK (and hence the inference underlying the changepoint detection).
-> We agree with the proposition and will integrate only the most important aspects in the dedicated Appendix (see **4.6**).

4.16 Appendix D. This appendix could be omitted and a brief statement in the text made as to how and why oxygen concentration vs saturation were used.
-> We will suppress Appendix D and insert this important issue in the method. Please refer to **1.7** for a full answer on the question of the selected oxygen variable.