Review of Capet et al.: ***"Climate change induced a new intermittent regime of convective ventilation that threatens the Black Sea oxygenation status"***

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 seem 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.

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) 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.
(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.
(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).
(4) The appendix seems to be incomplete as there are references to non-existent figures.

Overall, I recommend publication after moderate revisions. My detailed comments follow below; line numbers are given as, e.g. 'L123'.

Specific comments

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.

L58: This should also be mentioned when describing the statistical model using atmospheric forcing in the Methods

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)?

L73-76: Maybe this "section list" is not needed? If you keep it, the last sentence seems to be incomplete.

L85: Assuming that density increases with depth, this density criterion only defines the upper limit of the CIL. How is the lower limit defined?

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. 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)?

Table 1: I suggest listing the different datasets in the same order as in the text. 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. For convenience, the time periods covered by the different datasets could be added to the table.

L150-151: What is the smallest number of overlapping years between datasets? Is it appropriate to calculate correlations in the cases of least overlap? 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.

L156: I think Appendix A is incomplete as I couldn't find the information referred to here.

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.

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?

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

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

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?

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?

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)?

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.

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

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

Technical corrections

L2-3: "from the mid-1970s to the early 1990s"; specify "recent years", e.g. "post-2005"? comma after "Here"
L6: maybe "years without renewal of intermediate water"?
L7: "density levels"
L15: "While the reduction"
L20: "at the surface"
L24" "Strait"
L27: no comma after "(halocline)"; remove "the" after "prevents"
L33: "winter time"
L40: "upper ~100"
L41: "forcing"; "e.g." (check throughout the manuscript, also for "i.e.", the first "." is often missing)
L42: give time scale for alterations, e.g. "(order of days)"
L53: "Miladinova et al. (2018)"
L57: "feedback"
L62: (2005-2018)
L63: "trend leading to conditions"
L65: "with increasing trends"
L67: comma before "in particular"
L68: remove "extended"
L71: "an individual component"

L74: “regime shift analysis”
L78: use full term for CIL in section heading
L81: “provides”
L87: “water mass”
L89: “for consistency with existing literature”
L93: “i.e.”
L99: “dataset”
L100: I’d suggest “errors” instead of “artifacts” since one usually uses the term “sampling error”; space before “Each data”
L102: “netCDF”
L107: “This data was”
L110: “Good].”
L115: Could you provide the total numbers of profiles in the central and peripheral basin regions?
L123: again I’d suggest “errors”
Table 1: “guaranteed” instead of “granted” (under drawbacks for atmospheric predictors); “Three-dimensional” instead of “3D” in the rationale for GHER3D
L124: please cite the “reference study” here
L126: “consists of”
L129: remove “defined earlier”
L134: please cite the “reference study” here
L136: “Three-dimensional (3D) hydrodynamic model”
L136-144: Please include the time period for which the model was run.
L147: “metric for the”
L151: specify those series with little overlap
L154: here and throughout the manuscript: no space between “M J” in the unit; just “MJ”
L159: “inter-annual”
L161: “i.e.”
L164: “models’ ability”
L168: “i.e.”
L169: “which” instead of “that”
L171: “shift”
Fig. 1: add legend; larger panel and axes labels; “MJ” instead of “M J” in y axes units; no “-“ between quantity and units on y axes (this applies to all figures); caption: “Table” instead of “Tab.”; “time series” instead of “trend”; why does the statistical model show a range and what value (e.g. mean) of that model was used for the analysis?
L181: “Argo floats were operating”
L182/183: “Argo floats profiling”
Fig. 2: larger axes and panel labels; caption: “Argo floats”
L185: use full term for CIL
L187: “in Fig.”
L188: include use latex command “\cdot” instead of “.” In “$l1.i$”; state what “i” is
L193: “overestimate”
Fig. 3: in the legend, you use “Model3d”, while you use “C^3dModel” on L140 (I suggest to use one term consistently); axes labels and legend could be larger; suggest adding the linear and periodic functions to the legend (or the terms and put the functions into the caption)
L196: “shift”; “i.e.”

L198: “;” before “not shown”
L199: “i.e.”
Fig. 4: caption: “shift”; “point”
L200: suggest using full term for “CIL”
Fig. 5: here you use “3D model” in the legend (different to Fig. 3), be consistent; caption: “blue shaded area” looks more purple; “gray shaded area” very hard to see, perhaps make slightly less transparent; no comma after “i.e.”
L207: no comma after “e.g.”
L209: “;” after “oscillation”
L210: “20th”; “1950s; Ivanov et al., 2000)).”
L213: “prevails in the Black Sea since about ten years.”; “low cold content”
L217: suggest using full term for “CIL”
L219: “before 2008” instead of “in precedent regimes”; specify “the latest period”
L224: “period P2, P1/3 and P4, respectively.”
L226: for better readability: “about -1 MJ”
L227: “in more detail”
L228: “quasi-absence” and add reference to figure panel
L229: “during” instead of “among”; “depict”
L230: “while lower”; “simulated” or “shown” (this figure is not based on observations)
L231: unit missing for “16.0”
L232: “increases”; “in the years 2012 and 2015—2017 when CIL formation is significant”
L234: no space between “~ 14”
L241: “decreases continuously”
L242: “remain”; space before unit
L247: “regime shift model” instead of “first”
Fig. 6: the yellow dots in P4 panel are difficult to see, suggest using different color; add panel labels (a, b, c, d) for mores specific in-text referencing; add “regime name” to panel titles; caption: “time series”; “in Fig. 5”
L250: “built”
L251: “prevail” instead of “be considered as routine”? I think it’s more important to emphasize that earlier assumptions might not be valid anymore
L254: “trend”
Fig. 7: the shaded areas are hard to see, maybe make them a little less transparent? Caption: “areas”; “.” at end of caption
Fig. 8: significance is expressed by p-value, so why additional log10 of it? “Pearson Correlation Coefficient”; Labels and legend should be a bit larger
L265: “i.e.”
L269: “Sea”
L271: put “, respectively” at end of sentence
L284: “which has been”
L286: space before “Statistical”; “indicate”
L289: “i.e.”; parenthesis not closed after sigma values
L297: “feedbacks”
L301/302: all links should be included in that section, not as footnotes
L306: “denoted”
L310-317: equations should be numbered; in the RMS equation, “N^m,n” should be under the square root; in the relative bias equation, use “\cdot” instead of “.”

L319: “;” before “Fig. 1”; no “.” after “Fig. 1”
L321: add reference for “published prognostic values”
L331: “p-value” with cursive “p”
L334: “six reject”
L340: “i.e.”; use “Akaike Information Criterion (AIC)” as you refer to this appendix before introducing AIC in the main text
L341: this should be “Appendix C” (afterwards the numbering of appendix sections B1/B2/C1 seems off)
L343: “guaranteed”
L356: “Table C1; the”
L360: “Table C1”
Table C1, caption: “second column” and “third column”
L365: “coverage”; comma after “For instance”
L367: remove one “identified”
L370: “in-situ data”
L371: use “in situ” or “in-situ” consistently throughout the manuscript (preferably the former in cursive letters), here you use both
L380: no space before the “.”; don’t use “in this study” instead of “in the following”
L381: The heading for the “Author contributions” section is missing