1. Major comments

1. The quality of English can be improved in places, particularly the choice of words.

After having been revised to match the reviewers advices, the manuscript will be reviewed by an external English adviser.

―――――――――――

2. The setup of the model is not clear enough. The individual components (the hydrodynamic model, sediment model, and biogeochemical model) are explained by giving references to previously published works. It is not clear whether those models are used unchanged and whether any site-specific parameterizations are used in this study. Equally important, the coupling between these model components is not described. One can imagine that this is not a trivial task and more information is needed to assess this work.

The physical model is the GHER 3D hydrodynamical model already described in a lot of publications (e.g. Beckers [1], Delhez [4]) and in the particular case of the Black sea in Grégoire et al. [5, 6], Stanev and Beckers [11], Beckers et al. [2]). The implementation used in the frame of these simulations is described in Capet et al. [3]

The biogeochemical model is described in details in Grégoire et al. [7]. Compared to Grégoire et al. [7], the model has been coupled (online) in 3D with the GHER3D hydrodynamic model and extended with a dynamic representation of the benthic compartment (in Grégoire et al. [7], a reflective boundary condition was used for describing benthic degradation). This representation is based on the comparative analysis performed by Soetaert et al. [9] on the representation of benthic processes that is both reliable and tractable for long term simulations in a 3D ocean model. Hence, the benthic compartment is described by a semi-empirical model whose state variables are 2D variables (vertically integrated C and N content with two degrees of liability). From these 2D stock variables, the fluxes of solutes (NO_3, NH_4, O_2, ODU) are estimated using bulk parameters (i.e. fraction of organic matter degraded by denitrification and anaerobic oxidation, oxygen consumption in nitrification). As in Soetaert et al. [9], these bulk parameters are estimated from Monte Carlo simulations performed with an existing 1D diagenetic model of the Black sea sediments developed and validated by Wijsman et al. [13]. Concerning the resuspension, we used a setting similar to that used in Stanev and Kandilarov [10].

We are on the point of finalizing a manuscript describing in details the mechanisms and results of the sediment model, and presenting the validation of bentho-pelagic fluxes against in-situ estimates. However, the detailed description and validation of such a model would be too long for inclusion in this manuscript that we prefer to keep focused on the hypoxia issue and oxygen dynamics. Therefore, here, we propose to extend the description of the benthic model and we will refer to the work of Soetaert et al. [9] and to the submitted manuscript.

These 3 model components (hydrodynamic, biochemical and benthic) are integrated online within one single 3D modeling framework, meaning that, for example, biochemical variables are influenced at every time step by the state of the physical variables (e.g. advection, diffusion,

temperature effect). Similarly the benthic state variable are updated at every time step and for every grid point from the sedimenting variable of the biochemical model and provides, at every time step, the bottom boundary conditions for the pelagic model.

From the text of the manuscript it is not even clear if the modeling results used in this study are original or were taken from the cited studies. If original results are used, a section on modeling results is needed before proceeding to assessing the model performance by comparing it to observations (section 3).

The objective here was to restrain the focus of the manuscript to the issue of hypoxia and hence on the oxygen dynamics and the ability of the model to reproduce this dynamics. We therefore present the model results for the oxygen dynamics together with the comparison with available oxygen observation.

Concerning the hydrodynamics, the interannual variability of the model hydrodynamics and physics has been explored in Capet et al. [3] and validated in details with satellite SST and Sea level elevation data. Among others this validation exercise demonstrates the ability of the model to correctly reproduce the interannual variability of the SST, and that of the mixed layer depth. Further aspects of the model results concerning the processes driving oxygen dynamic are given in Section 4.

----

3. The statistical approach for estimating the model's suitability appears good.

----

4. The criterion on the Brunt-Vaisala stability frequency is not clear. The stability frequency varies with depth. At which depth is it taken here? At maximum? Similarly, a better definition is needed for the density difference that is used in defining the mixed layer depth: deltaRho over which depth interval?

We effectively used the maximal frequency on the vertical. The reference used for the computation of the $\Delta\rho$ interval is 3 m , as defined in Kara et al. [8].

We refined the definitions by replacing the sentence

"*A criteria on the Brunt-Vaïasala frequency $\sqrt{-g\frac{\delta\rho}{\rho}} > 0.05s^{-1}$ is used to define the period of enhanced stratification, referred to as the stratification (resp. mixing) period in the following (Fig. 7). The mixed layer depth z_ρ, defined by a density difference of $\Delta\rho|_3^{z_\rho} = 0.0125 \ kg/m^3$ [8] reaches $\sim 10 \ m$ in average during the stratification period.*"

by

" The stratification and mixing periods (Fig. 7) are defined using a criteria on the maximum value of the Brunt-Vaïasala frequency along the vertical, $\max_{z=0}^{Z}(\sqrt{-g\frac{\delta\rho(z)}{\rho(z)}}) \geq$ and $< 0.05 \ s^{-1}$, for stratification and mixing period, respectively. The mixed layer depth, z_ρ, is defined by a density difference $\Delta\rho = 0.0125 \ kg/m^3$[8] between z_ρ and the 3 m depth reference. It reaches $\sim 10 \ m$ in average during the stratification period. "

---

5. A hypoxic index H is suggested as an indicator of the severity of hypoxia, which combines both the area and duration of hypoxia. The attempt to introduce such an index (presumably to be used in management practice) is laudable. It may be worth specifying, however, that some of the effects of hypoxia are not linear with the duration of the hypoxic episode, which may be important if such an index is accepted in management practice.

The damaging effect of hypoxia on the ecosystem is a complex matter that cannot be completely addressed in this study. There is a great variability, among groups or species, for the sensitivity to hypoxia which depends on the level of de-oxygenation and the time of exposure [12]. On a ecosystem scale, the issue is even more complex as one should consider also the area exposed to hypoxia, the assemblage of species, and which key ecosystem functionality is affected.

Rather, we proposed here a general H index representing the intensity of H as an environmental stressor, acknowledging the fact that the link to ecosystem health issue remains to be done in the first paragraph of Sect. 6 :

" Lethal and non-lethal effects of low oxygen concentrations on living organisms strongly depend on the tolerance of the considered species, on the level of oxygen concentration and the duration of the hypoxic event (Vaquer-Sunyer and Duarte, 2008). The choice of a sustainable level of H that meets the requirements of Good Environmental Status (GES), as defined by the European Marine Strategy Directive (Cardoso et al., 2010), is therefore a very delicate issue which requires the combination of appropriate tools, as well as a dedicated and site-specific study."

This issue is further clarified in the revised manuscript.

Also, one can imagine that in some instances normalizing by the total area, as well as by the entire time period (one year), may be beneficial, as it produces a non dimensional index that could be easier to compare among environments.

As accurately underlined by the referee, it is true that units of area are more interpretable and friendly than unit of area times unit of time and we revised the index definition accordingly (see below). However, we think that the yearly averaged hypoxic area does not bear sense since we know it is a seasonal process. Eventually, such an H expressed in unit of area could be wrongly interpreted as an area of hypoxic area effectively encountered at a given time, which should be avoided

In order to produce an index expressed in area units, that can be compared from year to year and that corresponds to some meaningful value, we thus propose to use a typical temporal scale of hypoxia to normalize the index.

$$H = \frac{1}{\overline{D}} \int_{year} A(t)dt, \tag{1}$$

where $\overline{D}$ denotes the average D over all the years.

If we define, for each year, the duration of hypoxia by

$$D = \frac{1}{\max A(t)} \int_{year} A(t)dt, \tag{2}$$

which corresponds roughly to the time during which hypoxia occurs over half of the maximal area (see Fig. 1). With this definition of the duration of hypoxia, Eq. 1 implies that H is equivalent to the maximal extension if $D = \overline{D}$, i.e. if the duration is assumed to be equal to the average (see Fig. 1). Finally, this formulation allows an easy algebraic relationship between the index H and the duration and maximal area of hypoxia, which permits a graphical representation (Fig. 1).

As it appears that this issue deserves more attention, we develop Section 5.1 in the revised version of the manuscript and discuss the difficulties of defining H and D.

In particular we introduce the equivalence in terms of ratio of the shelf area for the sake of comparison with other sites suffering from seasonal hypoxia.
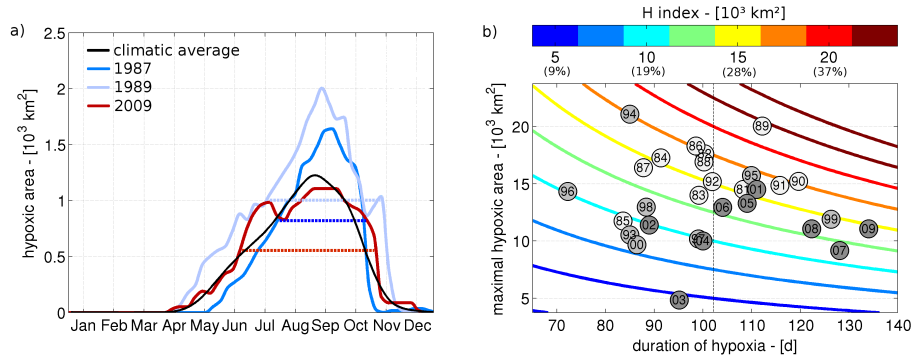

Figure 1: (a) Seasonal dynamics of the hypoxic area for selected years (color plain lines) and climatological average (black plain line). Dotted lines indicate the duration of hypoxia defined by Eq. 2. (b) The H index (colored lines) quantifies the intensity of the seasonal hypoxia event for a given year, and is computed (Eq. 1) as the integral of the hypoxic area over the year, normalized for the average duration of hypoxia (dotted line). The circles locates the hypoxic event diagnosed for each years (1981 -1991 : light gray, 1992-2000 : mid gray and 2001-2009: dark gray) which evidence the variabilities and the lack of relationship between the spatial and temporal extension of hypoxia. H is expressed in units area (if the duration is equal to the average duration, H equates the maximal extension of hypoxia). For inter-site comparability purposes the percentages given below the color bar express H in terms of ratio of the shelf area depicted in Fig. 1.

──────────────

6. The definition of "the winter sediments stock of semi-labile detritus C" needs to be better. What are the units, what is 'semi-labile', what role does sediment resuspension lay (i.e. would you define it in the same way in deeper waters)? Fig. 12 uses C in mmolC/m2, which implies a vertically integrated quantity. Over which depth is

the integration carried out? The stated values 10 mmolC/m2 are very small compared to the total pool of labile carbon in the surface decimeters of sediment.

The units of figure 12 are effectively wrong, in the sense that a mention "10^3" has been omitted for the Y-axis. This figure is corrected in the revised manuscript.

The definition of C is clarified by extending the description of the benthic part of the model in Sect. 2. The sediment layer is represented by 2D variables, i.e. integrated vertically to represent the total pool of carbon sediment involved in driving bentho-pelagic fluxes. The units is thus mmolC/m2.

We use a decay rate of 0.003 d^{-1}(~ 1 yr^{-1}) for the semi-labile stock while a value of 0.0753 d^{-1}(~ 27.5 yr^{-1}) is used for the labile part. Resuspension consists in a flux of materials from the sediment layer to the water column in particulate form. Resuspension, and its counterpart, deposition, are parameterized as in Stanev and Kandilarov [10], i.e. using critical thresholds on the bottom stress. This bottom stress is computed from current induced bottom stress and wave induced bottom stress which accounts for the bathymetry.

---

7. The regression model is defined so that it uses predictor values that are normalized by their mean and variance (Eq. 3). Does this imply that the occurrence of hypoxia is considered to depend on the deviation of nitrate loadings, temperature, and stratification strength from their average values but not on their absolute values? One would think that, for example, large percentage variations in nitrogen discharges in clean oligotrophic systems may have smaller effects than relatively small variations in a system that is already on a brink of hypoxia. This seems like an issue that can complicate transferring the developed model between environments.

Equation 3 describes a multi-linear regression using normalized response and predictors. This regression comes out of a stepwise procedure which allows to identify among a set of potential candidates the drivers of hypoxia (those that explain the largest part of hypoxia variability as assessed by R^2) and to differentiate the impacts of group of drivers (here "climate" versus "Eutrophication") on the variability of Hypoxia (thanks to the use of standardized values for the predictors and for H). The regression law also gives some indication on what could be the variability of hypoxia if some drivers are modified (provided of course that these drivers are modified in their range of values used for establishing the statistical law). The different nature of the potential drivers (i.e. different physical units) also justifies the normalization in this statistical analysis.

The use of a statistical model for prediction purposes is extremely delicate since this type of model is only valid in conditions that are similar to that used for their establishment. It means that using equation 3 for predicting hypoxia in other environments characterized by different atmospheric and environmental conditions is not valid. Similarly, to use equation 3 for predicting hypoxia in the Black sea in the future with atmospheric and river conditions totally different from those encountered during 1980-2010 is delicate. We would prefer to use a mechanistic model for that purpose. Therefore, equation 3 can not fit directly to other environments but it evidences key parameters whose role could effectively be assessed for other environments by a similar approach. Geomorphology of the different estuaries is expected to strongly constrain the sensitivity of hypoxia

to these predictors. We will clarify that point in the text. According to both reviewers comments, section 5.4 is extended in the revised manuscript and provides now a relationship between H and the predictors (using absolute values), obtained by reinserting the linear impact of climatic predictors to the power law obtained in this section.

It is also confusing that the regression model in Eqs. 3 and 7 uses the normalized H* whereas the figures use non-normalized H."

We decided to use non-normalized values for the Figures because it bears more meaning and can be connected more easily to measurements (see our discussion of the definition of the H index). That choice is clarified in the revised version of the manuscript.

2. Minor comments

8. Fig. 2: Caption for panel (a) is missing a number for the hypoxic threshold: "[O2]¡X? mmol/m3".

The caption has been completed "[O2]¡X? mmol/m3"

---

9. Fig. 7. The potential energy anomaly is badly defined as "the amount of energy needed to mix the entire water". I believe it is essential to say that it is the volume-specific difference in potential energy between stratified and mixed state and that it is averaged over the entire water column (is it?). Otherwise the standard units of J/m3 don't match the definition, which in its current form implies J/m2.

We apologize for this erroneous statement. The caption :

"*Seasonal evolution of the intensity of the vertical stratification as appraised by the potential energy anomaly (ϕ) (i.e. the amount of energy needed to mix the entire water), distinguishing the additive roles of haline (ϕ_S) and thermal (ϕ_T) stratification. Averages are computed over the investigation area presented in Fig.1. Vertical lines indicate the beginning and the end of the stratification period defined by a criterion on the Brunt-Vaïasala frequency (i.e. $\sqrt{-g\frac{\delta\rho}{\rho}} > 0.05s^{-1}$)* "

has now been changed to

"*Seasonal evolution of the vertical stratification as appraised by the potential energy anomaly, $\phi = -\frac{g}{h}.\int_Z^0 z(\rho(T,S) - \rho(\overline{T},\overline{S}))dz$ (i.e. the volume-specific difference in potential energy between stratified and mixed state of the water column). The color lines depicts the additive roles of thermal ($\phi_T = -\frac{g}{h}.\int_Z^0 z(\rho(T,\overline{S}) - \rho(\overline{T},\overline{S}))dz$) and haline ($\phi_S = \phi - \phi_T$) stratification. Averages are computed over the investigation area presented in Fig. 1. Vertical lines indicate the beginning and the end of the stratification period defined by a criterion on the Brunt-Vaïasala frequency (Sect. 4.1)* "

Bibliography

[1] Beckers, J.-M.: Application of the GHER 3D general circulation model to the Western Mediterranean, J. Marine Syst., 1, 315–332, doi:DOI:10.1016/0924-7963(91)

90001-B, URL http://www.sciencedirect.com/science/article/B6VF5-48B0D4F-2H/2/3638e58a3b0406bfbe4ba4e771dbfad2, 1991.

[2] Beckers, J.-M., Grégoire, M., Nihoul, J. C., Stanev, E., Staneva, J., and Lancelot, C.: Modelling the Danube-influenced North-western Continental Shelf of the Black Sea. I: Hydrodynamical Processes Simulated by 3-D and Box Models, Estuar. Coast. Shelf S., 54, 453–472, doi: DOI:10.1006/ecss.2000.0658, URL http://www.sciencedirect.com/science/article/pii/S0272771400906589/, 2002.

[3] Capet, A., Barth, A., Beckers, J.-M., and Grégoire, M.: Interannual variability of Black Sea's hydrodynamics and connection to atmospheric patterns, Deep-Sea Res. Pt. II, 77–80, 128–142, doi:10.1016/j.dsr2.2012.04.010, URL http://www.sciencedirect.com/science/article/pii/S0967064512000586, satellite Oceanography and Climate Change, 2012.

[4] Delhez, E.: Reconnaissance of the general circulation of the North-Western European Continental Shelf by means of a three-dimensional turbulent closure model, Earth Sci. Rev., 41, 3–29, doi:DOI:10.1016/0012-8252(96)00012-8, URL http://www.sciencedirect.com/science/article/pii/0012825296000128, 1996.

[5] Grégoire, M., Beckers, J.-M., Nihoul, J. C., and Stanev, E.: Reconnaissance of the main Black Sea's ecohydrodynamics by means of a 3D interdisciplinary model, J. Marine Syst., 16, 85–105, doi:10.1016/S0924-7963(97)00101-2, URL http://www.sciencedirect.com/science/article/pii/S0924796397001012, 1998.

[6] Grégoire, M., Soetaert, K., Nezlin, N., and Kostianoy, A. G.: Modeling the nitrogen cycling and plankton productivity in the Black Sea using a three-dimensional interdisciplinary model, Journal of Geophysical Research (Oceans), 109, C05 007, doi:10.1029/2001JC001014, 2004.

[7] Grégoire, M., Raick, C., and Soetaert, K.: Numerical modeling of the central Black Sea ecosystem functioning during the eutrophication phase, Progr. Oceanogr., 76, 286–333, doi: 10.1016/j.pocean.2008.01.002, 2008.

[8] Kara, A. B., Helber., R. W., Boyer, T. P., and Elsner, J. B.: Mixed layer depth in the Aegean, Marmara, Black and Azov Seas: Part 1: General Features, J. Marine Syst., 78, S169–S180, 2009.

[9] Soetaert, K., Middelburg, J. J., Herman, P. M., and Buis, K.: On the coupling of benthic and pelagic biogeochemical models, Earth Sci. Rev., 51, 173–201, doi:DOI: 10.1016/S0012-8252(00)00004-0, URL http://www.sciencedirect.com/science/article/pii/S0012825200000040, 2000.

[10] Stanev, E. and Kandilarov, R.: Sediment dynamics in the Black Sea: numerical modelling and remote sensing observations, OD, 62, 533–553, doi:10.1007/s10236-012-0520-1, URL http://dx.doi.org/10.1007/s10236-012-0520-1, 10.1007/s10236-012-0520-1, 2012.

[11] Stanev, E. V. and Beckers, J.-M.: Numerical simulations of seasonal and interannual variability of the Black Sea thermohaline circulation, Journal of marine systems, 22, 241–267, 1999.

[12] Vaquer-Sunyer, R. and Duarte, C. M.: Thresholds of hypoxia for marine biodiversity, Proceedings of the National Academy of Sciences, 105, 15 452–15 457, 2008.

[13] Wijsman, J. W., Herman, P. M., and Gomoiu, M.-T.: Spatial distribution in sediment characteristics and benthic activity on the northwestern Black Sea shelf, Marine Ecology Progress Series, 181, 25–39, 1999.