

1 **Sensitivity of the simulated Oxygen Minimum Zone to biogeochemical processes at an**
2 **oligotrophic site in the Arabian Sea**

3
4 Sankar S.^{1,2}, Polimene L.², Marin L.^{2,3}, Menon. N. N.¹, Samuelsen A.⁴, Pastres R.³, Ciavatta S.^{2,5,*}

5
6 ¹Nansen Environmental Research Centre, India

7 ²Plymouth Marine Laboratory, United Kingdom

8 ³Ca' Foscari University of Venice, Italy

9 ⁴Nansen Environmental Remote Sensing Center and Bjerknes Centre for Climate Research, Norway

10 ⁵National Centre for Earth Observation, Plymouth Marine Laboratory, United Kingdom

11

12 * Corresponding author: Stefano Ciavatta, s.ciavatta@pml.ac.uk

13

14 **Abstract**

15

16 Oxygen minimum zones (OMZs) are large, low-oxygen areas in the global oceans. Although
17 OMZs represent a serious threat to ecosystem functioning and services, our capability of modelling
18 the main biogeochemical processes driving OMZ dynamic are still limited. Here we performed a
19 full sensitivity analysis of a complex ecosystem model to rank the most important biogeochemical
20 parameters influencing the simulation of the OMZ at an oligotrophic site in the open Arabian Sea.
21 We applied a one-dimensional configuration of the European Regional Seas Ecosystem Model
22 (ERSEM) - here advanced by including denitrification - coupled with the General Ocean
23 Turbulence Model (GOTM). The coupled model was skilled in simulating the vertical gradients of
24 climatological data of oxygen and nutrients. The sensitivity analysis of the model was carried out in
25 two steps: *i*) a preliminary Morris screening analysis of 207 ERSEM parameters, which selected the
26 three most influential groups of parameters; and *ii*) a subsequent Monte Carlo sampling-based

27 analysis for ranking the importance of the 38 parameters within the three selected groups. Overall,
28 the four most important parameters for the simulation of the minimum oxygen concentration were
29 found to be: 1) the cubic half saturation constant for oxygenic control of denitrification; 2) the
30 parameter regulating the fraction of ingested matter excreted by heterotrophic nanoflagellates; 3)
31 the bacterial efficiency at low oxygen levels; and 4) the specific rate of bacterial release of capsular
32 material. Based on these findings, and assuming that the ranking of the model parameters reflects
33 the relevance of the process they characterize, we present a conceptual model describing the most
34 important biogeochemical processes affecting the OMZ at the study site. Our results suggest that
35 including bacteria explicitly in ecosystem models is useful to simulate and predict OMZs, provided
36 that efforts are invested in estimating parameters characterizing the microbial loop in marine
37 ecosystems.

38

39 **Key words:** marine ecosystem model; ERSEM; sensitivity analysis; oxygen minimum zone;
40 bacteria; Arabian Sea

41

42 **1. Introduction**

43

44 Oxygen minimum zones (OMZs) are areas of the oceans characterized by low dissolved
45 oxygen concentrations at intermediate depths (50-1000 m). Paulmier and Ruiz-Pino (2009)
46 defined OMZs as regions where dissolved oxygen (DO) concentrations are less than 20 $\mu\text{mol L}^{-1}$,
47 decreasing to 1 $\mu\text{mol L}^{-1}$ in the core of the OMZ. In the present ocean, OMZs are expanding as a
48 consequence of eutrophication and climate change, representing a serious threat for ecosystem
49 functioning and services such as fisheries (Oschlies et al., 2008; Stramma et al., 2008; Diaz and
50 Rosenberg, 2008; Gilbert et al., 2010; Rabalais et al., 2014; Duarte et al., 2015, Breitburg et al.,
51 2018).

52 The formation, maintenance and intensification of the OMZs are governed by the interaction

53 of physical processes (oxygen solubility driven by temperature and salinity, presence of regions
54 of low ventilation and subsurface currents of poorly oxygenated water) with biological processes
55 (primary production, heterotrophic activities, bacterial respiration and remineralization of
56 organic matter).

57 Physical processes influencing OMZs are linked to global patterns of temperature, salinity and
58 circulation. For example, Bopp et al. (2002) used a coupled climate-ocean biogeochemistry
59 model to predict the decrease in DO with climate change and the net outgassing of DO from the
60 ocean. They argued that the physical processes driving the reduction in DO were: *i*) changes in
61 surface water solubility due to temperature increase; and *ii*) changes in the ocean circulation
62 pattern. Matear and Hirst (2003) used a climate model coupled with an oceanic biogeochemical
63 model to investigate the multi-century impact of protracted global warming on the ocean
64 biogeochemical cycles. Their model predicted a decline in the DO concentration through most of
65 the subsurface ocean in the future years.

66 Marine biogeochemical processes are also crucial drivers of OMZs, and OMZs strongly
67 impact global biogeochemical cycles. As a basic conceptual scheme of biogeochemical drivers of
68 OMZs (see, e.g., Sarmiento and Gruber, 2006), waters at intermediate depth receive organic
69 matter produced and sinking from the upper euphotic layers; aerobic bacteria feeding on this
70 organic matter and respiration by zooplankton consume oxygen and lower its concentration
71 within the OMZ. Diaz and Rosenberg (2008) showed that hypoxic areas in the coastal oceans
72 increased since the 1960s because of the increase in primary production fueled by riverine runoff
73 and eutrophication. Oschlies et al. (2008) showed that OMZs are particularly sensitive to changes
74 in the marine biology, by predicting a 50% increase in the global suboxic water volume by 2100
75 in response to the respiration of excess organic carbon formed at higher atmospheric CO₂ levels.
76 Increase in primary production leads to increase in accumulation of particulate organic matter
77 that, in turn, increases microbial activity and consumption of oxygen in the waters below.
78 However, other processes complicate this basic scheme of OMZs, such as the possible switch of

79 the microbial community towards anaerobic bacteria, which can reduce nitrates to N₂ gas through
80 denitrification, and can reduce sulfate to hydrogen sulfide when the OMZ reaches anoxic
81 conditions (Richards, 1965; Sarmiento and Gruber, 2006). Large amounts of biologically reactive
82 nitrogen are removed from the oceans by anaerobic denitrifying bacteria in OMZs, with crucial
83 impact on the global cycle of nitrogen (Paulmier and Ruiz-Pino, 2009). While the above-
84 mentioned studies have identified the different biogeochemical processes influencing the OMZ,
85 an understanding of their comparative impacts has not yet been achieved.

86 The overall objective of this study is to contribute to fill this gap, by ranking the importance
87 of the biogeochemical processes which need to be carefully described to understand, simulate
88 and predict OMZ formation and evolution in the oceans. This was done by ranking the
89 importance of biogeochemical parameters of a complex marine ecosystem model. This model is
90 the European Regional Seas Ecosystem Model (ERSEM) (Butenschön et al., 2016), which
91 includes most of the biogeochemical processes driving OMZ dynamics. New for this study is that
92 we included denitrification in ERSEM, since this process is relevant in OMZ systems, but it was
93 not represented in the pelagic component of the model (Butenschön et al., 2016). We ranked the
94 importance of ERSEM parameters for OMZ simulation, by using in sequence the Morris
95 screening technique, followed by a Monte Carlo sampling-based ranking. These techniques are
96 already proven to be useful with other marine biogeochemical models (Pastres and Ciavatta,
97 2005; Cossarini and Solidoro, 2008) and a marine food-web model (Morris et al., 2014). This is
98 the first systematic sensitivity analysis of ERSEM.

99 In the present study, the analysis was performed using a one-dimensional (1-D)
100 implementation of ERSEM for an oligotrophic site in the open Arabian Sea (Figure 1),
101 advancing a comparable model configuration in this region by Blackford and Burkill (2002) and
102 Blackford et al. (2004). The Arabian Sea is characterized by a vast OMZ with DO concentrations
103 below 0.05 ml L⁻¹, at depths between 150 and 1250m (Van Bennekom and Hiehle, 1994), and it
104 is one of the three major denitrification sites in world oceans (Codispoti, 1989; Naqvi et al.,

105 2005) with an annual denitrification rate of 10-30 Tg N yr⁻¹ (Fauzi et al., 1993). At the present
106 time, there is no consensus on which physical and biological processes maintain the spatial and
107 seasonal pattern of the OMZ in the Arabian Sea (McCreary et al., 2013; Roullier et al., 2014), and
108 hypotheses include high respiration related to monsoon-driven primary productivity, slow
109 advection of intermediate waters, and influx of low oxygen waters from the South Indian Ocean
110 (Swallow, 1984; Naqvi, 1987; Jayakumar et al., 2004; Wiggert et al., 2005). Results from both a
111 box model (Sarma et al., 2002) and an eddy-resolving model (Resplandy et al., 2012) showed
112 that horizontal oxygen transport is important for maintaining the OMZ. At the same time,
113 ecosystem models of varying complexity, including the seminal NPZD model of McCreary et al.
114 (1996) and the model by Ryabchenco et al. (1998) that resolved also the microbial-loop, were
115 used to investigate the contribution of biological processes to the OMZ. The three-dimensional
116 (3-D) coupled model by Anderson et al. (2007) confirmed the relevance of modelling bacteria to
117 simulate the biogeochemical fluxes and demonstrated that vertical sinking of organic particles (in
118 contrast to their horizontal transport) was a major driver of denitrification in the regional OMZ.
119 For the first time, Blackford and Burkill (2002) and Blackford et al. (2004) applied the more
120 complex ERSEM to this region, in a 1-D configuration with the physical model GOTM, and they
121 found that vertical processes and microbial trophic dynamics were important drivers of
122 biogeochemical variability in the Arabian Sea. This model added the representation of size
123 classes of detritus and variable elemental ratios in the simulation of the Arabian Sea ecosystem –
124 and these features were later recognized as essential ones to simulate OMZ in this region, as well
125 as in the global ocean (Oschlies et al., 2008; McCreary et al., 2013).

126 However, the relative contribution of the different biogeochemical processes to the formation
127 and evolution of the OMZ in the Arabian Sea remains uncertain, and further research has been
128 invoked to improve their representation in mathematical models of this ecosystem (McCreary et
129 al., 2013; Roullier et al., 2014). Therefore, we tested the sensitivity analysis methods in a case
130 study that aims to rank the biogeochemical processes that determine the annual minimum value

131 of oxygen concentration at the site in the open Arabian Sea in Figure (1). Here, spatial-temporal
132 biogeochemical variability is lower, and the OMZ thinner, than in the Northern Arabian Sea
133 (Kao et al., 2015), arguably making acceptable the application of a 1-D model configuration to
134 study the formation of the and maintenance of the upper oxycline at the study site.

135 The paper is structured as follows. Section 2 describes the coupled physical-biogeochemical
136 model, the sensitivity methods and the set-up of the analysis. In Section 3, the results are
137 presented: first the skill of the OMZ simulation is evaluated by comparing the results to
138 climatological data, and then the results of the screening and Monte Carlo-based sensitivity
139 analyses are synthesized. In Section 4 we discuss the results by presenting a conceptual model of
140 the OMZ formation, and concluding remarks are pointed out in Section 5.

141

142 **2. Methods**

143

144 **2.1 Model description**

145

146 The vertical dynamics of the water column were represented by coupling ERSEM with the 1-D
147 hydrodynamic model GOTM (Figure 2) (Butenschön et al., 2016).

148 The general equation for the coupled GOTM-ERSEM model can be written as,

149

$$150 \frac{\partial c_i(\mathbf{c}, \mathbf{p}_{bio}, \mathbf{p}_{phys}, t)}{\partial t} = \left. \frac{\partial c_i}{\partial t} \right|^{bio} + \left. \frac{\partial c_i}{\partial t} \right|^{phys} \quad (1)$$

151

152 In equation (1) the first term on the right represents the biogeochemical equations of ERSEM
153 and the second term represents the physical equations of GOTM; \mathbf{c} represents the state vector
154 collecting the model variables c_i , \mathbf{p}_{bio} the vector collecting the parameters of ERSEM, and \mathbf{p}_{phys}
155 the vector of the parameters of GOTM.

156

157 The general equation of the scalar model output y can be written as,

$$158 \quad y = g_i(\mathbf{c}, \mathbf{p}_{bio}, \mathbf{p}_{phys}, z, t) \quad (2)$$

159

160 Where g_i is a function of the model variables \mathbf{c} , of the model parameters \mathbf{p} , of the depth z in
161 the water column, and of time t . We focused this work on the sensitivity of the annual average of
162 the minimum value of dissolved oxygen simulated in the water column ($c_i=O_2$), with respect to
163 the biogeochemical parameters of ERSEM (\mathbf{p}_{bio}):

164

$$165 \quad y = \langle \min_z [O_2(\mathbf{p}_{bio})] \rangle_a \quad (3)$$

166

167 Where \min_z represents the minimum in the water column, and $\langle \rangle_a$ the annual average. This
168 minimum value was assumed to approximate the “intensity” of the OMZ (i.e. the degree of oxygen
169 depletion, e.g. McCay et al., 2005) that is sensible to the biogeochemical processes investigated in
170 this work.

171

172 *2.1.1 The biogeochemical model ERSEM*

173

174 ERSEM (Baretta et al., 1995; Blackford et al., 2004; Butenschön et al., 2016) is a biomass and
175 functional group-based biogeochemical model describing the nutrient and carbon cycle within the
176 low trophic levels of the marine ecosystem. Model state variables include living organisms,
177 dissolved nutrients, organic detritus, oxygen and CO_2 . Pelagic living organisms are subdivided in
178 three functional groups describing the planktonic trophic chain: primary producers (phytoplankton),
179 consumers (zooplankton) and decomposers (bacteria). Primary producers and consumers are
180 subdivided into 4 and 3 size-based functional types, respectively. The phytoplankton community is
181 composed of picophytoplankton, nanoflagellates, dinoflagellates and diatoms, while the
182 zooplankton community is composed of mesozooplankton, microzooplankton and heterotrophic

183 nanoflagellates (HNAN). Decomposers are modeled by one type of heterotrophic bacteria.
184 Functional types belonging to the same group share common process descriptions but different
185 parameterizations.

186 A key feature of ERSEM is the decoupling between carbon and nutrient dynamics allowing the
187 simulation of variable stoichiometry within the modeled organisms. Chlorophyll is also treated as
188 an independent state variable following the formulation by Geider et al. (1997). Consequently each
189 plankton functional type is modeled with up to five state variables describing the cellular content of
190 carbon, nitrogen, phosphorus, silicon, and chlorophyll-a. Dissolved organic matter (DOM) is
191 produced by different processes involving phytoplankton, bacteria and zooplankton while its
192 consumption is exclusively regulated by bacteria uptake. DOM is subdivided into labile, semi-labile
193 and semi-refractory components (Polimene et al., 2006), in order to provide a representation of the
194 range of organic compounds present in the marine DOM and their different levels of degradability.
195 Particulate organic matter (POM) is produced by phytoplankton and zooplankton and it is divided
196 into three size-based categories corresponding to different sedimentation rates. In this way it is
197 possible to simulate the carbon export from the surface to the intermediate OMZ layers. In the
198 version of ERSEM applied here, the decomposition of particulate organic matter is directly
199 mediated by bacteria, and the partition between labile and semi-labile organic matter occurs in
200 relation of the nutritional status of phytoplankton and bacteria (Polimene et al., 2006, 2007;
201 Butenschön et al., 2016).

202 All the ERSEM equations are detailed in Butenschön et al. (2016) and we refer the reader to
203 that paper for a comprehensive description of the mathematical formulations used in the model.
204 Here we limit our description to the ERSEM representation of oxygen dynamics, which are the
205 focus of the paper, and to the formulation describing denitrification, which was newly developed
206 in this work. The pelagic net production of oxygen is modeled through the balance between gross
207 primary production (gpp) and the whole community respiration (resp). The latter is computed as
208 sum of the contributions of bacteria (carbon biomass B_C), of NZ=3 zooplankton groups (Z_C), and

209 NP=4 phytoplankton groups (P_C) (Butenschön et al., 2016):

210

$$211 \quad \left. \frac{\partial O_2}{\partial t} \right|_{bgc} = -p_O^{resp} \left. \frac{\partial B_C}{\partial t} \right|_{resp} - p_O^{resp} \sum_{i=1}^{NZ} \left. \frac{\partial Z_C^i}{\partial t} \right|_{resp} - p_O^{resp} \sum_{i=1}^{NP} \left. \frac{\partial P_C^i}{\partial t} \right|_{resp} + p_O^{syn} \sum_{i=1}^{NP} \left. \frac{\partial P_C^i}{\partial t} \right|_{gpp} \quad (4)$$

212

213 Where p_O^{resp} is a stoichiometric factor converting the amount of carbon respired into oxygen
214 consumption, and p_O^{syn} is a stoichiometric factor converting the amount of carbon assimilated
215 through photosynthesis into oxygen production.

216 Denitrification was represented here by modifying the equation in Vichi et al (2007). This
217 process was represented as the minimum of a potential denitrification ($Denit_{pot}$) and the bacterial
218 nitrate demand (BND):

219

$$220 \quad Denit = \min(Denit_{pot}, BND) \quad (5)$$

221 Where:

$$222 \quad Denit_{pot} = Denix \cdot NO_3 \quad (6)$$

223 And

$$224 \quad BND = BOD \cdot \delta \cdot (1 - O2_{lim}) \quad (7)$$

225

226 In equation (6), NO_3 is the available concentration of nitrate and $Denix$ is the maximum specific
227 denitrification rate. In equation (7), BOD is the bacterial oxygen demand, which is a function of the
228 available organic carbon (POC+DOC), δ is a stoichiometric factor converting O₂ to NO₃, and $O2_{lim}$
229 is a cubic Michaelis-Menten function describing oxygen limitation (Vichi et al., 2007):

230

$$231 \quad O2_{lim} = \frac{(O_2)^3}{(O_2)^3 + (chN3oX)^3} \quad (8)$$

232

233 Where $chN3oX$ is the half-saturation constant for oxygenic control of nitrogen transformation.

234 Equations 7 and 8 imply that under well oxygenated conditions the BND is close to zero while it
235 increases under low oxygen condition. If the environmental nitrate concentration is insufficient to
236 satisfy the BND, then the production of reduction equivalents (HS) occurs. This latter process
237 mimics the formation of reduced sulfur (HS⁻) as observed when both oxygen and nitrate are
238 depleted:

239

$$240 \quad HS = BOD \cdot \delta' \cdot (1 - O2_{lim}) - Denit \cdot \varepsilon' \quad (9)$$

241

242 Where δ' and ε' are stoichiometric factors converting O₂ to HS⁻ and nitrate to HS⁻, respectively.

243 The above equations 5-9 represent aerobic and anaerobic (e.g. denitrifying) bacteria through a
244 single functional group. The model, in fact, describes a bulk bacteria biomass able to switch
245 between different kinds of metabolism (aerobic, anaerobic) depending on environmental conditions,
246 i.e., on O₂ and nitrate concentrations in the water column. We note that the model does not include
247 chemolithotrophic bacteria, though they can be important for the nitrogen cycle within OMZs (Lam
248 et al., 2009). However, Ward et al. (2009) found that heterotrophic denitrification is the main
249 process responsible for N loss in the OMZ of the Arabian Sea, largely exceeding the
250 chemolithotrophic anaerobic ammonia oxidation (anammox), making the absence of
251 chemolithotrophy acceptable in our application.

252

253 *2.1.2 The hydrodynamic model GOTM*

254

255 GOTM (General Ocean Turbulence Model; Burchard et al., 1999) is a 1-D water column model
256 used for the computation of hydrodynamic and thermodynamic processes related to vertical mixing.
257 The model calculates velocities, turbulence, temperature and salinity, as well as heat, momentum
258 and freshwater fluxes between the ocean and the atmosphere, when forced with local
259 meteorological inputs. Routines for nudging observations also exist in GOTM and they were
260 applied here for the relaxation of the model simulation towards salinity and water temperature

261 profiles (see also the model set up in Section 2.4.1). Such relaxation has been widely used in
 262 previous GOTM-ERSEM applications in both shelf-sea and open ocean sites (e.g. Blackford et al.,
 263 2004; Torres et al., 2006; Polimene et al., 2012; 2014; 2015, Butenschön et al., 2016).

264

265 **2.2. The screening Morris method**

266

267 The GOTM-ERSEM model was subjected initially to a screening sensitivity analysis. This aimed
 268 to identify the subset of ERSEM parameters that are most important for the simulation of the
 269 minimum oxygen concentration at the study site. The screening sensitivity analysis used the Morris
 270 method, as proposed in Morris (1991), modified by Campolongo et al. (2007), and applied with
 271 marine models by Cossarini and Solidoro (2008) and Morris et al. (2014). The Morris technique,
 272 described thoroughly in Saltelli et al. (2008), is a qualitative sensitive analysis based on the concept
 273 of Elementary Effect (EE), which is an approximation of the first order partial derivative of the
 274 model output y with respect to an input factor X_i , i.e. a model parameter. If a model has k number
 275 of independent input factors X_i , where $i=1, 2, \dots, k$, the elementary effect of the parameter X_i is
 276 given by:

277

$$278 \quad EE_i^j = \frac{y(X_1^j, X_2^j, \dots, X_{i-1}^j, X_i^j + \Delta, X_{i+1}^j, \dots, X_k^j) - y(X_1^j, X_2^j, \dots, X_k^j)}{\Delta} \quad (10)$$

279

280 Where j represents an initial point in the space of the parameters, y represents the model output,
 281 and the increment $\Delta \in [0,1]$ is a pre-defined proportion of the range of variation of the parameters,
 282 which, being constant, allows the sensitivity index to account for the statistical distribution of the
 283 input factors (see the explicative example at page 120 in Saltelli et al., 2008). All the input
 284 parameters in (10) are incremented, in random order, leading the input vector X to cover a
 285 “trajectory” j in the space of the parameters. The trajectory has $(k+1)$ nodes, that are sets of
 286 parameter values used to run the model $(k+1)$ times and compute k elementary factors EE^j . A

287 number $j=1, \dots, r$ of trajectories is built by selecting randomly their $j=1, \dots, r$ initial points in the
288 space of the parameters. The initial points of the trajectories and the increments of the input factors
289 are computed within ranges of variability that need to be defined.

290 Following Campolongo et al. (2007), we computed the sensitivity index for the input parameter
291 X_i by averaging the absolute values of the elementary effects of that parameter across all the
292 trajectories:

293

$$294 \mu_i^* = \frac{1}{r} \sum_{j=1}^r |EE_j^i| \quad (11)$$

295

296 The sensitivity index μ_i^* is computed by averaging $|EE_j^i|$ computed at points sampled within the
297 whole space of the parameters. Therefore, this technique can be considered as a global screening
298 technique, though each single elementary effect is a first order derivative, i.e. a local sensitivity
299 (Campolongo et al., 2007). Importantly, the index μ^* allows one to reduce the computational cost of
300 the screening analysis by computing the sensitivity of input parameters pooled in groups. In fact,
301 Campolongo et al., 2007 and Saltelli et al. (2008) showed that using the absolute values of the
302 Elementary Effects preserves the reliability of the sensitivity index in eq. 11 also when the
303 parameters within a group are changed by the same Δ proportion simultaneously, but in opposite
304 direction (i.e. different signs of Δ). Exploiting this property, we subdivided the parameters of
305 ERSEM into groups that refer to different ecosystem processes, and the sensitivity index μ^* was
306 calculated for each group (see Ciavatta et al., 2009 for an analogous approach). The drawback of
307 grouping input factors is the loss of information regarding the relative importance of factors
308 belonging to the same group. This was addressed by performing a Monte Carlo-based sensitivity
309 analysis of the parameters within the groups.

310

311

312

313 2.3. Monte Carlo simulations and ranking method

314

315 A Monte Carlo sampling-based sensitivity analysis was applied to rank the importance of the m
316 parameters $\mathbf{X}=(X_1, X_2, \dots, X_i \dots, X_m)$, $i=1, 2, \dots, m$, within the groups identified as most important
317 in the Morris screening analysis (Saltelli et al., 2008). A crude Monte Carlo sampling scheme was
318 used to generate a number n of realizations of the input factor vector \mathbf{X} . These realizations were
319 input to n model simulations that computed the target model output y in equation (3). The input-
320 output relationship was represented by means of a multiple-regression model:

321

$$322 \quad y = b_0 + \sum_{i=1}^m b_i X_i + residuals \quad (12)$$

323

324 and the standardized regression coefficients β_i were used as global sensitivity indices of the input
325 factors (Saltelli et al., 2008):

326

$$327 \quad \beta_i = \frac{b_i \sigma_{X_i}}{\sigma_y} \quad (13)$$

328

329 Where σ_{X_i} and σ_y are the standard deviations of the realizations of the input factor X_i and of the
330 model output y , respectively. The regression coefficients in eq. (13) provide meaningful parameter
331 rankings only when the linear regression explains a relatively large fraction of the model output
332 variability (Saltelli et al., 2000). We assessed the linear regression by computing the fraction of
333 explained variance (R^2), the regression significance (F-statistic of the null hypothesis of constant
334 model, $p<0.01$), as well as the significance of the standardized regression coefficients (t-statistic,
335 $p<0.05$).

336

337

338

339 **2.4 Set up of the analysis**

340

341 *2.4.1 Set up of the model*

342

343 The 1-D GOTM-ERSEM model was implemented for an oligotrophic site in the central Arabian
344 Sea (65°E, 13°N), which falls within the OMZ in this region (Paulmier and Ruiz-Pino, 2009; Naqvi,
345 1991). The actual depth of the central Arabian Sea is close to 4500 m, but we have simulated the
346 water column up to a depth of 500 m only, using 100 vertical levels. The selection of this maximum
347 depth was based on previous studies, indicating that the upper 500 meters include the upper
348 oxycline and the absolute minimum of oxygen (McCreary et al., 2013, Resplandy et al., 2012), as
349 confirmed here by test simulations extending till the depth of 1500 meters (not shown). A deep-
350 water remineralization closure scheme was applied to the lower boundary of the model (Figure 2).
351 The closure describes the recycling of organic matter (producing inorganic nutrients and CO₂) as a
352 linear function of the sinking biomass, at rates specified by the ERSEM remineralization parameters
353 (Blackford et al., 2004; Butenschön et al., 2016).

354 In our application, GOTM simulation was relaxed to daily profiles of salinity and temperature
355 derived from the output of the 3-D model HYCOM configured for the Indian Ocean (George et al.,
356 2010). GOTM-ERSEM was forced with daily meteorological data and cloud cover data from
357 NCEP/NCAR reanalysis (Kalnay et al., 1996) and precipitation data from GPCP (Adler et al.,
358 2003). The profiles of the initial conditions of nutrients (nitrate, phosphate and silicate) were
359 obtained from the World Ocean Atlas 2009 (Garcia et al., 2010a; Garcia et al., 2010b). The initial
360 condition of oxygen was set equal to a constant value throughout the water column (20 mmol m⁻³,
361 consistent with climatological data at depth), to avoid pre-setting the position of the OMZ in the
362 water column and letting the model simulation setting it. The model simulation was carried out for a
363 period of four years (2002 to 2005) after a spin-up time of five years, which has been shown to be
364 sufficient to achieve stable solutions of the 1D GOTM-ERSEM integration (e.g. Blackford et al

365 2004; Polimene et al., 2014) . The output of the four-year simulation was used to assess the skill of
366 the model in simulating oxygen and nutrients (nitrate, phosphate and silicate), through comparison
367 with World Ocean Atlas 2009 climatology (Garcia et al., 2010a; Garcia et al., 2010b), in the
368 absence of in-situ observations of these variables at the study site. The sensitivity analysis and
369 parameter ranking was carried out on the model output for the year 2002.

370

371 *2.4.2 Set up of the Morris screening analysis*

372

373 The configuration of ERSEM applied here has 342 pelagic parameters. However, parameters
374 defining biogeochemical constants (e.g. the inverse of the Redfield ratio of phosphorous to carbon)
375 were not object of this investigation, thus the number of parameters included in the screening
376 analysis was 207. These parameters were categorized and divided into $k=21$ groups (Table 1). The
377 increment of the input factors was set $\Delta=2/3$, following the recommendation in Saltelli et al., 2008.

378 Groups 1 to 7 comprised of parameters characterizing primary production, whereas groups 8 to
379 12 were bacteria-related parameters. The remaining groups included zooplankton parameters, food
380 matrix parameters, deep-water remineralization closure parameters, sedimentation parameters and
381 light extinction parameters.

382 The analysis was carried out with the range of variability of the uniform distribution of the
383 parameters kept within -30% to +30% of the reference value of the parameters. The 30% variation
384 with respect to the reference values of the parameters is often assumed in sensitivity analyses of
385 environmental models when the real ranges are unknown (see, e.g., Ciavatta et al., 2009; Polimene
386 et al., 2015; Pinna et al., 2015). In our application of the Morris method, we set a number $r=10$
387 trajectories for the $k=21$ groups of parameters. Thus the computational cost of the screening
388 sensitivity analysis was $(k+1) \cdot r = 220$ model runs.

389

390 *2.4.3 Set-up of the Monte Carlo sampling-based sensitivity analysis*

391

392 The Monte Carlo sampling-based sensitivity analysis was performed by selecting $n=1000$
393 random values for each of the m independent input parameters found to be the most relevant in the
394 screening analysis. In choosing this number n of model simulations, we considered the rule of
395 thumb of at least 20 realizations for each input factor desirable for multiple regression analysis
396 (Hair et al., 2006). As in the Morris application, we have used uniform distribution to generate
397 random values within the range -30% to +30% of the reference value of the input parameters (see,
398 e.g., Ciavatta et al., 2009; Polimene et al., 2015; Pinna et al., 2015). Each realization of the vector of
399 input parameters was used to run a model simulation. The multiple regression analysis of the input-
400 output relationship was performed using the software Origin, and the regression coefficients
401 defining the sensitivity index for the parameter ranking were estimated using the least-squares
402 method proposed by Draper and Smith (1981).

403

404 **3 Results**

405

406 **3.1 Skill of the reference simulation**

407

408 The model had significant skill in simulating the climatology of oxygen and nutrient
409 observations at the study site. This is illustrated in Figures 3 and 4, which show comparable
410 climatologies from the model output and the World Ocean Dataset 2009, and it is demonstrated
411 quantitatively by the Taylor diagram in Figure 5, where all the variables are close to the optimal
412 skill point.

413 The model represented the magnitude and range of all the variables, though it tended to
414 underestimate nitrate and to overestimate phosphate (Figure 3). At surface, both the climatological
415 data and the model confirm stable oligotrophic conditions, with nitrate concentrations $< 2 \text{ mmol m}^{-3}$
416 and phosphate $< 0.5 \text{ mmol m}^{-3}$ in all the months. The vertical gradients were reproduced adequately

417 by the model: oxygen decreased from the aerated surface layer downwards, nutrients were higher in
418 the ocean interior, and changes occurred more sharply between 50-200 m. In particular, in Figure 4,
419 the model represented well the average vertical profile of oxygen ($\rho>0.99$, $p<0.01$, and $\text{RMSD}=19$
420 mmol m^{-3}), the depth of its absolute minimum (200 m in both the data and model climatologies),
421 though the model tended to underestimate the oxygen climatological data on average (bias = -17
422 mmol m^{-3}).

423 The seasonal variability of the data is less well represented in the simulation. The model
424 simulated the deepening of mixed layer and associated variability in DO concentration during the
425 monsoon season from June to September (Figure 3). Some fluctuations of nutrient profiles are also
426 reproduced in the monsoon season in the upper layer, but to a much lower extent. In particular, the
427 simulation did not capture temporal variability below the oxycline and nutricline, such as the
428 increase of silicate concentration at depth ~ 400 in May-June in the climatology dataset. Biases
429 below the nutricline are likely due to the lack of lateral advection, as mentioned in Discussion.

430 The general good agreement between model and data is confirmed by the results in Figure 5,
431 since the correlation coefficients are high ($\rho>0.9$, $p<0.01$), the variability of the climatology and the
432 simulation are comparable ($\sigma/\sigma_o = 0.9$), and the biases are relatively low, with the exceptions of
433 phosphate that was overestimated (bias/ $\sigma_o=0.4$) and nitrate that was underestimated (bias/ $\sigma_o= -0.5$).
434 High scores for the skill metrics are driven primarily by the model ability in simulating the average
435 vertical gradients of the climatological data, rather than their seasonal variability.

436

437 **3.2 Screening sensitivity analysis**

438

439 The relevance of the groups of parameters resulting from the Morris analysis is shown in Figure
440 6. The 14th group (zooplankton loss parameters), 9th group (bacterial loss parameters) and 11th group
441 (additional nutrient remineralization parameters) were found to be the three most relevant groups
442 for the simulation of the OMZ, in order. The parameters included in the groups 14 and 9

443 characterize the biological processes of oxygen consumption by the zooplankton functional groups
444 and by bacterial functional group, respectively. The parameters included in group 11 are associated
445 to first order remineralization processes converting organic nutrients onto inorganic forms
446 (phosphate and ammonium). These processes are assumed to complement the biologically mediated
447 remineralization activity which is described in the model (Blackford et al., 2004; Butenschön et al.,
448 2016). A complete list of the 38 parameters included in the three above groups is given in Table 2.

449

450 **3.3 Ranking of the parameters**

451

452 The 38 model parameters that emerged collectively as the most important in the screening
453 analysis (Table 2) were the input factors to the Monte Carlo sampling-based sensitivity and ranking
454 analysis. The results are presented in Table 3, which ranks the parameters in descending order of
455 importance based on the magnitude of the standardized regression coefficient $|\beta|$ (eq. 13).

456 The four most important parameters, with $|\beta|$ higher than 0.3, were found to be: 1) the cubic half
457 saturation constant for oxygenic control of denitrification (chN3oX); 2) the parameter regulating the
458 fraction of ingested matter excreted (i.e. not assimilated) by the heterotrophic nanoflagellates
459 (pu_eaZ6X); 3) the bacterial efficiency at low oxygen levels (puB1oX); and 4) the specific rate of
460 bacterial release of capsular material (frB1R3). The first 21 parameters in Table 3 were associated
461 to significant regression coefficients (t-test, $p < 0.05$), while the remaining 17 parameters did not
462 significantly influence the simulated minimum of oxygen. Importantly, the ranking provided by the
463 sensitivity analysis was trustworthy, since the linear regression explained most of the model output
464 variability ($R^2 = 0.94$), and it was highly significant (F-test, $p < 0.01$) (see Table 3).

465 To assess further the robustness of the ranking, we performed a supplementary regression
466 analysis including only the first eleven independent variables in Table 3 (i.e. those with $|\beta| > 0.1$,
467 arbitrarily); the results confirmed the overall dominance of those parameters in explaining the
468 dissolved oxygen variability ($R^2 = 0.93$, F-test $p < 0.01$), and reproduced their ranking in Table 3. This

469 suggests that the results in Table 3 were not affected remarkably by redundancy among the many
470 model parameters included in the analysis. We note also that the parameters object of the regression
471 analysis were sampled randomly from independent uniform distributions, thus multicollinearity
472 among regressors is not an issue in our application.

473

474 **4. Discussion**

475

476 The results indicated that model parameters regulating the metabolism of aerobic and anaerobic
477 (denitrifying) bacteria and the loss terms of zooplankton (heterotrophic nanoflagellates, HNAN)
478 play a prime role in our simulation of the OMZ at the study site. Assuming that the ranking of the
479 model parameters reflects the relevance of the processes they characterize, we have inferred a
480 conceptual model describing the most important biogeochemical processes affecting the OMZ in
481 the oligotrophic site of the Arabian Sea area studied here (Figure 7).

482 At the surface, where light is sufficient to allow net growth of primary producers, oxygen is
483 produced by phytoplankton and is consumed by both autotrophic and heterotrophic (zooplankton
484 and bacteria) respiration, besides being exchanged with the atmosphere (Figure 7). Net
485 photosynthesis fades at a depth of ~100 m marking the threshold between euphotic and twilight
486 zone.

487 In the upper twilight zone (~100-200 m), heterotrophic prokaryotes are the most active
488 organisms, while grazers' biomass (mesozooplankton) is close to zero because of the negligible
489 concentration of phytoplankton. Here oxygen is consumed by remineralization of the sinking
490 detritus and therefore DO decreases drastically. Anaerobic respiration of POC via denitrification
491 becomes relevant, though the level of oxygen remains sufficient to also allow some aerobic
492 respiration (Figure 7). As a consequence, the Michaelis-Menten constant "chN3oX" emerges as the
493 most important parameter in our analysis, because it regulates the magnitude of denitrification and
494 therefore the amount of organic matter which is respired without consuming oxygen ("chN3oX" in

495 equation 8 has rank=1 in Table 3). In other words, this parameter impacts the OMZ simulation
496 because it sets the threshold at which bacteria either do or do not consume oxygen. The efficiency
497 of bacteria in using POC to grow is also crucial in determining the intensity of the OMZ, here
498 approximated by the absolute minimum value of the oxygen profile. Low efficiency implies that a
499 large portion of the carbon taken up by bacteria is respired, with a consequent high oxygen
500 consumption (or nitrate consumption, in case of anaerobic metabolism) and low net bacterial
501 production. On the contrary, high efficiency implies that a lower portion of carbon is respired,
502 resulting in a higher bacteria biomass production. This explains the high rank scored by the bacteria
503 efficiency parameter (“puB1oX” ranked 3rd in Table 3).

504 At ~200 m, both the simulation and the climatology show the absolute minimum of oxygen
505 (Figure 4). Here there is a zone where the sinking detritus is still sufficient to allow some bacterial
506 respiration. At this depth POC concentration is low and becomes a limiting factor for bacteria
507 growth and respiration. This is illustrated in Figure 8, which shows that the simulated bacteria
508 biomass follows (with a lag) the seasonal cycle of the detritus sinking from the euphotic zone. POC
509 limitation implies that the oxygen minimum is linked to the ability of heterotrophs (bacteria and
510 zooplankton predators) to survive in starvation conditions, i.e. it is linked to heterotrophs’ “basal
511 metabolism”. This explains why parameters defining the basal metabolism of heterotrophs (i.e.
512 bacteria rest respiration, and mortality of bacteria, HNAN and microzooplankton) are all within the
513 ten most important parameters in Table 3. In other words, these parameters are important because
514 they determine how much the heterotrophs are suited to survive and consume oxygen at the depth
515 where POC is a limiting food.

516 Below the depth of 200 meters and till the depth of 500 m simulated here, oxygen increases
517 slightly with depth because bacteria biomass is small (due to the reduced export of POC) and the
518 consumption of oxygen is negligible. At that depth, the lack of heterotrophic biomass and low
519 vertical transport maintain the DO concentration at ~5 mmol L⁻¹, which is a reminiscence of the
520 initial condition, slightly lower than the climatological data at the study site (Figure 4).

521 The conceptual model in Figure 7 is consistent with previous experimental and modelling
522 works, which showed that the dynamic POC-bacteria (both anaerobic and denitrifiers) is a relevant
523 biological driver of the OMZ in the global oceans, as well as in the Arabian Sea (e.g. Ulloa et al.,
524 2012; Roullier et al., 2014). Bacteria are the principal contributors of the community respiration in
525 the pelagic ecosystems (Carlson et al., 2007, Cole. et al., 1988) and diverse microbial community
526 act simultaneously both at the transition zones and within global OMZs (Beman and Carolan,
527 2013). This clearly holds for the Arabian Sea, where, for example, Gonsalves et al. (2011) observed
528 aerobic and denitrifying bacteria coexisting in both a coastal site and an off-shore site, though
529 denitrifiers were dominating the community at the coastal site. The importance of bacterial
530 degradation of detritus, rather than dissolved organic carbon, was stressed also in previous
531 modelling studies of the Arabian Sea (Anderson et al., 2007), and Roullier et al. (2014) argued that
532 the anaerobic microbial respiration enhances production and accumulation of observed particles (of
533 size $< 100 \mu\text{m}$) in the upper part of the OMZ in this region. Figure 8 shows that the absolute values
534 of POC and bacteria biomass simulated by the model are low at the study site. This can be due to
535 the oligotrophic nature of the central part of the Arabian Sea, which is a permanently stratified area
536 and has lower biomass and bacterial activity with respect to the coastal regions. In fact, Gonsalves
537 et al. (2011) measured lower concentrations of total organic carbon TOC in an offshore site with
538 respect to a coastal one, and Campbell et al. (1998) reported lower bacteria and phytoplankton
539 biomass in the central part of the basin than in in-shore waters.

540 Besides the POC-bacteria dynamics, the “complexity” of our ecosystem model pointed out the
541 relevance of two processes less extensively investigated in previous modelling analysis of the OMZ
542 in the Arabian Sea, i.e. the grazing on bacteria and the bacteria release of recalcitrant organic
543 carbon (Figure 7).

544 In the simulated trophic web, bacterial biomass (thus its overall respiration and oxygen control)
545 is top-down controlled by zooplankton grazing. Therefore, the parameter defining the efficiency of
546 heterotrophic nanoflagellates (HNAN), which are the main grazers of bacteria, became an important

547 factor for the formation of the OMZ, explaining rank=2 of “pu_eaZ6X” in Table 3. The primary
548 role of HNAN in controlling bacterial biomass in the Arabian Sea was observed previously by
549 Weisse (1999).

550 Our application suggested for the first time that bacteria production of recalcitrant organic
551 carbon can have a relevant influence on the maintenance of an OMZ. The parameter regulating the
552 release of capsular material by bacteria ranked in 4th position (frB1R3 in Table 3). This release
553 produces recalcitrant dissolved organic carbon (RDOC), which is regarded as an important element
554 for the global carbon cycle and potentially for climate regulation (Jiao et al., 2010; 2014). Our
555 results suggest that the bacterially-mediated production of RDOC influences also the maintenance
556 of the OMZs, because it reduces the bulk biomass of bacteria and therefore their oxygen
557 consumption through respiration.

558 The sensitivity analysis suggested that processes related to primary production have a less direct
559 impact on the oxygen minimum at the site investigated here, though these processes are of
560 importance in OMZ formation and evolution in general (Diaz and Rosenberg, 2008). In our study,
561 the group of parameters related to primary production had relatively low importance and were not
562 selected by the Morris screening analysis (groups 1-7 in Figure 6). This can be due to the fact that
563 the model was implemented in a relatively oligotrophic area with a fully stratified water column
564 (Figure 3) and relatively low primary production. Therefore, the mass of organic matter exported
565 from the surface to the OMZ is low in absolute value (Figure 8). The weak connection between the
566 euphotic (productive) zone and the twilight zone (where the OMZ is observed), makes the OMZ
567 weakly dependent on primary production in our simulation.

568 Though our study site was chosen in a relatively stable oligotrophic area of the open Arabian
569 Sea, where vertical 1-D processes were found to be dominant in driving both particle transport
570 (Roullier et al., 2014) and denitrification (Anderson et al., 2007), the use of a 1-D model is certainly
571 a limitation of our work. The assimilation of temperature and salinity profiles integrates the effects
572 of 3-D hydrodynamics to a certain extent (Section 2.4.1), but lateral fluxes of oxygen and other

573 biogeochemical components potentially relevant to the OMZ were not simulated. Therefore, we
574 focused our simulation on the first 500 m of the water column, where the absolute minimum oxygen
575 value is observed (Figure 4). This zone is above the deep oxycline at ~1000 m, typically observed
576 in the Arabian Sea due to the influx of deep oxygen-rich waters (Swallow, 1984; Ulloa et al., 2012;
577 Roullier et al., 2014) that clearly cannot be represented by a 1D model configuration. Furthermore,
578 the model cannot account for the episodic intrusion of oxygen within the OMZ (Ulloa et al., 2012),
579 which might contribute in sustaining aerobic activity. Our model does simulate aerobic activity
580 within the OMZ (sustained by the residual initial conditions) however it does not reproduce the
581 presence of mesozooplankton at depth, which are reported in previous works (e.g. Banse et al.,
582 2014; Roullier et al., 2014). Similarly, the lack of lateral supply of POC might also contribute to the
583 low concentration of detritus simulated at depth (Figure 8). Finally, the absence of lateral circulation
584 might explain also the discrepancies between simulated and climatological seasonal cycles of
585 nutrients and oxygen at depths below the absolute minimum (Figure 3). In particular, the model
586 could not simulate the increase in silicate concentration observed typically in June at ~400 m, which
587 extends upwards, and the decrease of nitrate concentration observed in June and October between
588 200 and 400 m. The relative increase of oxygen observed in climatological summer and autumn
589 between 200 and 400m was not captured by the model either. However, the average annual vertical
590 gradients were well represented by the model, as demonstrated by the skill metrics in Figure 5, and
591 this supports further the use of the oxygen minimum value as target metric of the sensitivity
592 analysis. The choice of this OMZ indicator is coherent with the objective of this paper, which
593 focuses on the effects of biogeochemical processes on the intensity of the minimum oxygen, rather
594 than on the extension of the OMZ (e.g. water volume; Cabré et al., 2015), which cannot be
595 represented by a one-dimensional model.

596 The 1-D ERSEM-GOTM applied here resulted adequate also in previous studies in the Arabian
597 Sea (Blackford and Burkill, 2002; Blackford et al., 2004), as well as in other shelf and open ocean
598 locations (e.g. Butenschön et al., 2016; Torres et al., 2006). We argue that our 1-D implementation

599 in a relative stable OMZ site is particularly suitable for the objective of our biology-focused
600 sensitivity analysis. On the one hand, a comprehensive sensitivity analysis can have a prohibitive
601 computational cost with 3-D implementations of complex ecosystem models (Pastres and Ciavatta,
602 2005). On the other hand, the 1-D implementation allowed us to focus on the biogeochemical
603 processes in “isolation”, i.e. without the need to disentangle them from physical-driven mechanisms
604 that could influence the simulation in 3-D implementations. In particular, the use of the ERSEM
605 model (Butenschön et al., 2016) which embeds a fully resolved microbial loop (Polimene et al.,
606 2006 and 2007) allowed us to focus on bacterial processes with a level of details not resolved in
607 most of the marine ecosystem models applied previously (e.g. Anderson et al., 2007, Resplandy et
608 al., 2012; McCreary et al., 2013), including previous versions of ERSEM as well (Blackford and
609 Burkill, 2002; Blackford et al., 2004).

610 Finally, we recognize that our approach is based on the assumption that the ranking of the
611 parameters reflects directly the ranking of the processes (i.e., if a parameter is important, the
612 equation/process that includes that parameter is important) and that this assumption could not be
613 always true. Indeed a parameter can result important (or not) because the equation/process in which
614 is included is not well represented in the model. For example, a specific process could be poorly
615 represented in the model even if it is crucial for the functioning of the real ecosystem, leading the
616 parameter to be neglected by the sensitivity analysis. These weaknesses, which are implicit in any
617 modelling study, need to be kept into account and inevitably add some degree of uncertainty to the
618 results presented here.

619

620 **5. Conclusions**

621

622 This paper identifies the most relevant biogeochemical processes involved in the ERSEM
623 simulation of the OMZ in a central oligotrophic site of the Arabian Sea. We found that processes
624 related to both aerobic and denitrifying bacteria along with the loss term of bacteria and

625 heterotrophic flagellates (HNAN) are the most important. This outcome also highlights the
626 relevance of our new representation of denitrification in ERSEM. Other processes, like primary
627 production, were found to be less relevant. These findings are consistent with previous studies,
628 which suggested that the impact of bacteria on the OMZ is important in the Arabian Sea as well as
629 in other parts of the global oceans (e.g. Ulloa et al., 2012; Roullier et al., 2014). Presently, only few
630 marine ecosystem models include an explicit description of the microbial loop, but our study
631 strongly indicates that OMZ models should explicitly include heterotrophic bacteria and their
632 production of recalcitrant carbon.

633 Despite the clear limits of our 1-D model configuration, our application provided an objective
634 list of the most important biogeochemical parameters that need to be quantified for future
635 applications of a global configuration of ERSEM (Kwiatkowski et al., 2014) aiming to simulate the
636 biogeochemical and physical dynamic underpinning OMZs and their predicted expansions. To this
637 regard, we note that the sensitivity methods proposed here are in principle applicable to OMZ scalar
638 metrics alternative to the absolute minimum applied here (e.g. OMZ area and volume below pre-set
639 oxygen thresholds, Cabré et al., 2015) more suitable for three-dimensional applications accounting
640 for horizontal transport processes.

641 Finally, we note that the analysis presented here is the first systematic sensitivity study of the
642 ERSEM model with respect to its full set of parameters. The tools developed here are not limited to
643 the study of the OMZs but can be applied straightforwardly to the study of different aspects of
644 ocean biogeochemistry (e.g. carbon fluxes in the subtropical gyres), and to prioritize the
645 parameters to be estimated in data assimilative simulations, as we are investigating in the
646 framework of ongoing work.

647

648

649

650

651 **Acknowledgements**

652

653 The research leading to these results has received funding from the European Union's Seventh
654 Framework Programme (FP7/2012-2015) under grant agreement no. 295092, INDO-MARECLIM
655 coordinated by Nansen Environmental Research Centre-India. SC was partially funded by the UK
656 NERC through the National Centre for Earth Observation (NCEO) and the Atlantic Biogeochemical
657 Fluxes Project of the RAPID-AMOC Programme, by the Copernicus Marine Environment
658 Monitoring Service (CMEMS) TOSCA and NOWMAPS projects (CMEMS is implemented by
659 Mercator Ocean in the framework of a delegation agreement with the European Union). This
660 project has received funding from the European Union's Horizon 2020 research and innovation
661 programme under grant agreement No 678396 (project TAPAS). LP was funded by the UKNERC
662 national capability in sustained observation and modelling and through the UK NERC-Defra Shelf
663 Seas Biogeochemistry program (NE/K001876/1).

664

665 **References**

666

667 Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.P., Janowiak, J., Rudolf, B.,
668 Schneider, U., Curtis, S., Bolvin, D., Gruber, A., 2003. The version-2 global precipitation
669 climatology project (GPCP) monthly precipitation analysis (1979-present). *J. Hydrometeor.* 4(6),
670 1147-1167.

671 Anderson, T.R., Ryabchenko, V.A., Fasham, M.J. and Gorchakov, V.A., 2007. Denitrification in
672 the Arabian Sea: A 3D ecosystem modelling study. *Deep Sea Research Part I: Oceanographic*
673 *Research Papers*, 54(12), pp.2082-2119.

674 Banse, K., Naqvi, S.W.A., Narvekar, P. V., Postel, J.R., Jayakumar, D.A., 2014. Oxygen
675 minimum zone of the open Arabian Sea: variability of oxygen and nitrite from daily to decadal
676 timescales. *Biogeosciences* 11, 2237–2261. doi:10.5194/bg-11-2237-2014

677 Baretta, J.W., Ebenhoh, W., Ruardij, P., 1995. The European regional Seas Ecosystem Model, a
678 complex marine ecosystem model. *Neth. J. Sea Res.* 33, 233 – 246.

679 Beman, J.M., Carolan, M.T., 2013. Deoxygenation alters bacterial diversity and community
680 composition in the ocean's largest oxygen minimum zone. *Nature Communications*, 4.

681 Blackford, J.C., Allen, J.I., Gilbert, F.J., 2004. Ecosystem dynamics at six contrasting sites: a
682 generic modelling study. *J. Marine Syst.* 52(1), 191-215.

683 Blackford, J.C. and Burkill, P.H., 2002. Planktonic community structure and carbon cycling in
684 the Arabian Sea as a result of monsoonal forcing: the application of a generic model. *Journal of*
685 *Marine Systems*, 36(3), pp.239-267.

686 Bopp, L., Le Quéré, C., Heimann, M., Manning, A.C., Monfray, P., 2002. Climate induced
687 oceanic oxygen fluxes: Implications for the contemporary carbon budget. *Global Biogeochem. Cy.*
688 16(2). doi:10.1029/2001GB001445.

689 Breitburg, D., Levin, L.A., Oschlies, A., Grégoire, M., Chavez, F.P., Conley, D.J., Garçon, V.,
690 Gilbert, D., Gutiérrez, D., Isensee, K. and Jacinto, G.S., 2018. Declining oxygen in the global ocean
691 and coastal waters. *Science*, 359(6371), p.eaam7240.

692 Burchard, H., Bolding, K., Villarreal, M.R., 1999. GOTM-A General Ocean Turbulence Model,
693 Theory, Implementation and Test cases. Technical Report EUR 18745 EN, European Commission.

694 Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., Bruggeman, J.,
695 Cazenave, P., Ciavatta, S., Kay, S., Lessin, G., van Leeuwen, S., van der Molen, J., de Mora, L.,
696 Polimene, L., Saille, S., Stephens, N., Torres, R., 2016. ERSEM 15.06: a generic model for marine
697 biogeochemistry and the ecosystem dynamics of the lower trophic levels, *Geosci. Model Dev.* 9(4),
698 1293-1339, doi:10.5194/gmd-9-1293-2016.

699 Cabré, A., Marinov, I., Bernardello, R., & Bianchi, D. (2015). Oxygen minimum zones in the
700 tropical Pacific across CMIP5 models: mean state differences and climate change trends.
701 *Biogeosciences*, 12(18), 5429-5454.

702 Campbell, L., Landry, M.R., Constantinou, J., Nolla, H.A., Brown, S.L., Liu, H. and Caron,

703 D.A., 1998. Response of microbial community structure to environmental forcing in the Arabian
704 Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, 45(10), pp.2301-2325.

705 Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity
706 analysis of large models. *Environ. Modell. Softw.* 22(10), 1509-1518.

707 Carlson C.A., Del Giorgio P.A., Herndl G.J. (2007). Microbes and the dissipation of energy and
708 respiration : from cells to ecosystems. *Oceanography-Washington DC-Oceanographic Society*, 20,
709 89–100, doi: 10.5670/oceanog.2007.52.

710 Ciavatta, S., Lovato, T., Ratto, M., Pastres, R., 2009. Global uncertainty and sensitivity analysis
711 of a food-web bioaccumulation model. *Environ. Toxicol. Chem.* 28, No. 4, 718–732.

712 Cole, J.J., Findlay, S., Pace, M.L., 1988. Bacterial production in fresh and saltwater ecosystems:
713 a cross-system overview. *Mar. Ecol. Prog. Ser.* Oldendorf, 43(1), 1-10.

714 Codispoti, L.A., 1989. Phosphorus vs nitrogen limitation of new (export) production. In: Berger,
715 W., Smetacek, V., Wefer, B. (Eds.), *Productivity of the Ocean: Present and Past*. Wiley, New York,
716 377-394.

717 Cossarini, G., Solidoro, C., 2008. Global sensitivity analysis of a trophodynamic model of the
718 Gulf of Trieste. *Ecol. Model.* 212(1), 16-27.

719 Diaz, R.J., Rosenberg, R., 2008. Spreading dead zones and consequences for marine ecosystems.
720 *Science*. 321(5891), 926-929.

721 Draper, N.R., Smith, H., 1981. *Applied Regression Analysis*, John Wiley & Sons, New York,
722 407.

723 Duarte, C.M., Fulweiler, R.W., Lovelock, C.E., Martinetto, P., Saunders, M.I., Pandolfi, J.M.,
724 Gelcich, S. Nixon, S.W., 2015. Reconsidering ocean calamities. *BioScience*. 65(2), 130-139.

725 Fauzi, R., Mantoura, C., Law, C.S., Owens, N.J., Burkill, P.H., Woodward, E.M.S., Howland,
726 R.J., Llewellyn, C.A., 1993. Nitrogen biogeochemical cycling in the northwestern Indian Ocean.
727 *Deep-Sea Res. Pt II.* 40(3), 651-671.

728 Garcia, H.E., Locarnini, R.A., Boyer, T.P, Antonov, J.I, Baranova, O.K., Zweng, M.M., Johnson,

729 D.R., 2010a. World Ocean Atlas 2009, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization,
730 and Oxygen Saturation. S. Levitus, Ed. NOAA Atlas NESDIS 70, U.S. Government Printing
731 Office, Washington, D.C., 344 pp.

732 Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Zweng, M.M., Baranova, O.K.,
733 Johnson, D.R., 2010b. World Ocean Atlas 2009, Volume 4: Nutrients (phosphate, nitrate, silicate).
734 S. Levitus, Ed. NOAA Atlas NESDIS 71, U.S. Government Printing Office, Washington, D.C., 398
735 pp.

736 Geider, R.J., H.L. MacIntyre, Kana, T.M., 1997. Dynamic model of phytoplankton growth and
737 acclimation: Responses of the balanced growth rate and the chlorophyll a: carbon ratio to light,
738 nutrient-limitation and temperature. *Mar. Ecol. Prog. Ser.* 148(1-3), 187-200.

739 George, M.S., Bertino L., Johannessen O.M., Samuelsen A., 2010: Validation of a hybrid
740 coordinate ocean model for the Indian Ocean. *J. Oper. Oceanogr.* 3(2), 25-38.

741 Gilbert, D., N.N. Rabalais, R.J. Diaz, Zhang, J., 2010. Evidence for greater oxygen decline rates
742 in the coastal ocean than in the open ocean. *Biogeosciences.* 7(7), 2283-2296.

743 Gonsalves, M.-J., Paropkari, A.L., Fernandes, C.E.G., Loka Bharathi, P.A., Krishnakumari, L.,
744 Fernando, V., Nampoothiri, G.E., 2011. Predominance of anaerobic bacterial community over
745 aerobic community contribute to intensify “oxygen minimum zone” in the eastern Arabian Sea.
746 *Cont. Shelf Res.* 31, 1224–1235. doi:10.1016/j.csr.2011.04.011

747 Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L., 2006. Multivariate data
748 analysis (Vol. 6). Upper Saddle River, NJ: Pearson Prentice Hall.

749 Jayakumar, D.A., Francis, C.A., Naqvi, S.W.A. and Ward, B.B., 2004. Diversity of nitrite
750 reductase genes (*nirS*) in the denitrifying water column of the coastal Arabian Sea. *Aquatic*
751 *Microbial Ecology*, 34(1), pp.69-78.

752 Jiao, N., Herndl, G.J., Hansell, D.A., Benner, R., Kattner, G., Wilhelm, S.W., Kirchman, D.L.,
753 Weinbauer, M.G., Luo, T., Chen, F., Azam, F., 2010. Microbial production of recalcitrant dissolved
754 organic matter: long-term carbon storage in the global ocean. *Nat. Rev. Microbiol.* 8(8), 593-599.

755 Jiao, N., Robinson, C., Azam, F., Thomas, H., Baltar, F., Dang, H., Hardman-Mountford, N.J.,
756 Johnson, M., Kirchman, D.L., Koch, B., Legendre, L., 2014. Mechanisms of microbial carbon
757 sequestration in the ocean—future research directions. *Biogeosciences*. 11(19), 5285-5306.

758 Kalnay, E.M., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha,
759 S., White, G., Woollen, J., Zhu, Y., Leetma, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins,
760 W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The
761 NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, 77(3), 437-471.

762 Kao, S-J, Wang, B. Y, Zheng, L.-W. , Selvaraj, K, Hsu, S-C, Sean Wan, X. H., Xu, M. and
763 Arthur Chen, C. T., 2015. Spatio-temporal variations of nitrogen isotopic records in the Arabian
764 Sea. *Biogeosciences*, 12, 1–14.

765 Kumar, D., A. Rajendran. A., Somasundar, K., Haake, B., Jenisch. A., Shuo, Z., Ittekkot, V.,
766 and Desai. B. N. 1990. Dynamics of dissolved organic carbon in the northwestern Indian Ocean.
767 *Marine Chemistry*, 31, 299 - 316.

768 Kwiatkowski, L., Yool, A., Allen, J. I., Anderson, T. R., Barciela, R., Buitenhuis, E. T.,
769 Butenschön, M., Enright, C., Halloran, P. R., Le Quéré, C., de Mora, L., Racault, M.-F., Sinha, B.,
770 Totterdell, I. J., and Cox, P. M., 2014. iMarNet: an ocean biogeochemistry model intercomparison
771 project within a common physical ocean modelling framework. *Biogeosciences*, 11, 7291–7304,
772 doi:10.5194/bg-11-7291-2014.

773 Lam, P. et al., 2009. Revising the nitrogen cycle in the Peruvian oxygen minimum zone, *Proc.*
774 *Natl Acad. Sci. USA* 106, 4752–4757.

775 Matear, R. J., Hirst, A.C., 2003. Long-term changes in dissolved oxygen concentrations in the
776 ocean caused by protracted global warming, *Global Biogeochem. Cy.* 17(4), 1125,
777 doi:10.1029/2002GB001997.

778 McKay, J. L., T. F. Pedersen, and J. Southon (2005), Intensification of the oxygen minimum
779 zone in the northeast Pacific off Vancouver Island during the last deglaciation: Ventilation and/or
780 export production?, *Paleoceanography*, 20, PA4002, doi:10.1029/2003PA000979.

781 McCreary, J.P., Kohler, K.E., Hood, R.R. and Olson, D.B., 1996. A four-component ecosystem
782 model of biological activity in the Arabian Sea. *Progress in Oceanography*, 37(3), pp.193-240.

783 McCreary, J.P., Yu, Z., Hood, R.R., Vinayachandran, P.N., Furue, R., Ishida, A., Richards, K.J.,
784 2013. Dynamics of the Indian-Ocean oxygen minimum zones. *Prog. Oceanogr.* 112, 15-37.

785 Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments.
786 *Technometrics*, 33, 161–174.

787 Morris, D.J., Speirs, D.C., Cameron, A.I., Heath, M.R., 2014. Global sensitivity analysis of an
788 end-to-end marine ecosystem model of the North Sea: Factors affecting the biomass of fish and
789 benthos. *Ecol. Model.* 273, 251-263.

790 Naqvi, S. W. A., 1987. Some aspects of the oxygen-deficient conditions and denitrification in the
791 Arabian Sea. *Journal of Marine research*, 45(4), 1049-1072.

792 Naqvi, W.A., 1991. Geographical extent of denitrification in the Arabian Sea in relation to some
793 physical processes. *Oceanol. Acta.* 14(3), 281-290.

794 Oschlies, A., K. G. Shulz, U. Riebesell, and A. Schmittner (2008), Simulated 21st century's
795 increase in oceanic suboxia by CO₂-enhanced biotic carbon export, *Global Biogeochem. Cycles*,
796 22, GB4008, doi:10.1029/2007GB003147.

797 Pastres, R., Ciavatta, S. 2005. A comparison between the uncertainties in model parameters and
798 in forcing functions: its application to a 3D water-quality model. *Environ. Modell. Softw.* 20(8),
799 981-989.

800 Paulmier, A., Ruiz-Pino, D., 2009. Oxygen minimum zones (OMZs) in the modern ocean. *Prog.*
801 *Oceanogr.* 80, 113-128.

802 Pinna, A., Pezolesi, L., Pistocchi, R., Vanucci, S., Ciavatta, S., Polimene, L., 2015. Modelling
803 the Stoichiometric Regulation of C-Rich Toxins in Marine Dinoflagellates. *PLoS One*, 10(9),
804 e0139046.

805 Polimene, L., Allen. J.I., Zavatarelli, M., 2006. Model of interaction between dissolved organic
806 carbon and bacteria in marine systems. *Aquat. Microb. Ecol.*, 43, 127-138.

807 Polimene, L., Pinardi, N., Zavatarelli, M., Allen, J.I., Giani, M., Vichi, M., 2007. A numerical
808 simulation study of dissolved organic carbon accumulation in the northern Adriatic Sea. *J. Geophys.*
809 *Res-Oceans*. 112(C3).

810 Polimene, L., Mitra, A., Saille, S.F., Ciavatta, S., Widdicombe, C.E., Atkinson, A., Allen, J.I.,
811 2015. Decrease in diatom palatability contributes to bloom formation in the Western English
812 Channel. *Prog. Oceanogr.* 137, 484-497.

813 Polimene, L., Brunet, C., Butenschon, M., Martinez-Vicente, V., Widdicombe, C., Torres, R.,
814 Allen, J.I., 2014. Modelling a light-driven phytoplankton succession. *Journal of Plankton Research*
815 36, 214–229.

816 Polimene, L., Archer, S.D., Butenschon, M., Allen, J.I., 2012. A mechanistic explanation of the
817 Sargasso Sea DMS ‘summer paradox’. *Biogeochemistry* 110, 243–255.
818 <http://dx.doi.org/10.1007/s10533-011-9674-z>.

819 Rabalais, N.N., Cai, W.J., Carstensen, J., Conley, D.J., Fry, B., Hu, X., Quinones-Rivera, Z.,
820 Rosenberg, R., Slomp, C.P., Turner, R.E., Voss, M., 2014. Eutrophication-driven deoxygenation in
821 the coastal ocean. *Oceanography*. 27(1), 172-183.

822 Resplandy, L., Lévy, M., Bopp, L., Echevin, V., Pous, S., Sarma, V.V.S.S., Kumar, M.D., 2012.
823 Controlling factors of the oxygen balance in the Arabian Sea's OMZ, *Biogeosciences*, 9, 5095-5109,
824 doi:10.5194/bg-9-5095-2012.

825 Richards, F.A., 1965. Anoxic basins and fjords. In: Riley, J.P., Skirrow, G. (Eds.), *Chemical*
826 *Oceanography*, vol. 1. Academic Press, New York, 611-645.

827 Roullier, F., Berline, L., Guidi, L., Durrieu De Madron, X., Picheral, M., Sciandra, A., Pesant, S.
828 and Stemann, L., 2014. Particle size distribution and estimated carbon flux across the Arabian Sea
829 oxygen minimum zone. *Biogeosciences*, 11(16), pp.4541-4557.

830 Ryabchenko, V.A., Gorchakov, V.A. and Fasham, M.J.R., 1998. Seasonal dynamics and
831 biological productivity in the Arabian Sea Euphotic Zone as simulated by a three-dimensional
832 ecosystem model. *Global Biogeochemical Cycles*, 12(3), pp.501-530.

833 Saltelli, A., Chan, K., Scott, M., 2000. Handbook of Sensitivity Analysis. Probability and
834 Statistics Series. John Wiley & Sons Publishers, New York, USA.

835 Saltelli, A., Ratto, M., Andres, T., Campolongo, M., Cariboni, J., Gatelli, D., Saisana, M.,
836 Tarntola, S., 2008. Global Sensitivity Analysis. The Primer. John Wiley & Sons Ltd, Chichester,
837 UK.

838 Sarma, V.V.S.S., 2002. An evaluation of physical and biogeochemical processes regulating
839 perennial suboxic conditions in the water column of the Arabian Sea. *Global Biogeochem. Cy.*
840 16(4).

841 Sarmiento, J.L., Gruber, N., 2006. Ocean biogeochemical dynamics. Princeton University Press.

842 Stramma, L., Johnson, G.C., Sprintall, J., Mohrholz, V., 2008. Expanding oxygen minimum
843 zones in the tropical oceans. *Science*. 320, 655–658.

844 Swallow, J.C., 1984. Some aspects of the physical oceanography of the Indian Ocean. *Deep Sea*
845 *Research Part A. Oceanographic Research Papers*, 31(6), pp.639-650.

846 Torres, R., J. I. Allen, Figueiras, F.G., 2006. Sequential data assimilation in an upwelling
847 influenced estuary. *J. Mar. Syst.* 60(3-4), 317-329, doi:10.1016/j.jmarsys.2006.02.001.

848 Ulloa, O., Canfield, D. E., DeLong, E. F., Letelier, R. M., & Stewart, F. J. (2012). Microbial
849 oceanography of anoxic oxygen minimum zones. *Proceedings of the National Academy of*
850 *Sciences*, 109(40), 15996-16003.

851 Van Bennokom, A, J., Hiehle, M.A., 1994. CTD operations and calibrations during legs D1, D2
852 and D3 of the Netherlands Indian Ocean Programme, in *Geological Study of the Arabian Sea*,
853 edited by W. J. M. van der Linden and C. H. van der Weijden, 37-66, Neth. Geosci. Found., The
854 Hague, Netherlands.

855 Vichi, M., Pinardi, N., Masina, S., 2007. A generalized model of pelagic biogeochemistry for the
856 global ocean ecosystem. Part I: theory. *J. Mar. Syst.* 64, 89-109.

857 Ward, B.B., Devol, A.H., Rich, J.J., Chang, B.X., Bulow, S.E., Naik, H., Pratihary, A.,
858 Jayakumar, A., 2009. Denitrification as the dominant nitrogen loss process in the Arabian Sea.

859 Nature. 461(7260), 78-81.

860 Weisse, T., 1999. Bacterivory in the northwestern Indian Ocean during the intermonsoon–
861 northeast monsoon period. *Deep Sea Research Part II: Topical Studies in Oceanography*, 46(3),
862 pp.795-814.

863 Wiggert, J.D., Hood, R.R., Banse, K. and Kindle, J.C., 2005. Monsoon-driven biogeochemical
864 processes in the Arabian Sea. *Progress in Oceanography*, 65(2), pp.176-213.

865

Table 1. List of the 21 groups of pelagic parameters investigated in the screening Morris analysis.

Group	Description	Number of parameters
1	Photosynthetic parameters	20
2	Metabolic carbon lost parameters (respiration)	5
3	Lost carbon by lysis parameters	4
4	Nutrient parameters	38
5	Q ₁₀ parameters : regulating temperature factors	4
6	Photosynthetically available fraction of irradiation	1
7	Other primary production parameters	17
8	Maximum specific gross uptake of bacteria	6
9	Bacterial loss parameters	6
10	Nutrient uptake / remineralization	4
11	Additional nutrient remineralization	11
12	Other bacteria parameters	2
13	Maximum zooplankton uptake	13
14	Zooplankton loss parameters	21
15	Q ₁₀ of zooplankton	7
16	Zooplankton nutrient quotas	6
17	Food matrix parameters	20
18	Deep-water remineralization closure parameters	4
19	Sedimentation parameters	7
20	Cellular structural parameters	4
21	Light extinction parameters	7

866
867

Table 2. Parameters included in the groups 9, 11 and 14 of the Morris screening analysis, which were investigated in the Monte Carlo-based sensitivity analysis.

No.	Notation	Description and units
Group 9: Bacterial loss parameters		
1	frB1R3	Fraction of activity respiration of bacterial uptake converted to semi-refractory DOC [unit]
2	puB1oX	Bacterial growth efficiency at low oxygen levels [unitless]
3	puB1X	Bacterial growth efficiency at high oxygen levels [unitless]
4	sdB1X	Specific mortality of bacteria at reference temperature [day^{-1}]
5	srsB1X	Specific rest respiration at reference temperature [day^{-1}]
6	DeniX	Maximum specific denitrification rate [day^{-1}]
Group 11: Additional nutrient remineralization parameters		
7	chN3oX	Michaelis-Menten constant for oxygenic control of denitrification [$\text{mmol O}_2 \text{ m}^{-3}$]
8	puR4_B1X	Fraction of small size POM available for bacteria uptake [unitless]
9	puR6_B1X	Fraction of medium size POM available for bacteria [unitless]
10	puR8_B1X	Fraction of large size POM available for bacteria [unitless]
11	redfieldX	Carbon to Nitrogen Redfield ratio [unitless]
12	rR2R1X	Specific rate for breakdown of semi-labile to labile DOC [unitless]
13	sN4N3X	Specific nitrification rate at reference temperature and silt concentration [day^{-1}]
14	sR1N1X	Specific dissolution of labile DOP to phosphate [day^{-1}]
15	sR1N4X	Specific dissolution of labile DON to ammonium [day^{-1}]
16	reoX	Specific reoxidation rate of reduction equivalents [day^{-1}]
17	R1R2X	Labile fraction of DOM production [unitless]
Group 14: Zooplankton loss parameters		
18	pe_R1Z4X	DOM-fraction of uptake excreted by mesozooplankton [unitless]
19	pu_eaRZ4X	Fraction of POM-uptake excreted by mesozooplankton [unitless]
20	pu_eaZ4X	Fraction of prey-uptake excreted by mesozooplankton [unitless]
21	puZ4X	Mesozooplankton assimilation efficiency [unitless]
22	sdZ4oX	Specific mortality of mesozooplankton due to oxygen limitation [day^{-1}]
23	sdZ4X	Specific basal mortality of mesozooplankton [day^{-1}]
24	srsZ4X	Specific rest respiration of mesozooplankton at reference temperature [day^{-1}]
25	Z4mortX	Specific overwintering mortality of mesozooplankton [day^{-1}]
26	Z4repwX	Specific overwintering respiration of mesozooplankton [day^{-1}]
27	pe_R1Z5X	DOM-fraction of uptake excreted by microzooplankton [unitless]
28	pu_eaZ5X	Fraction of prey-uptake excreted by microzooplankton [unitless]
29	puZ5X	Microplankton assimilation efficiency [unitless]
30	sdZ5oX	Specific mortality of microzooplankton due to oxygen limitation [day^{-1}]
31	sdZ5X	Specific basal mortality of microzooplankton [day^{-1}]
32	srsZ5X	Specific rest respiration of microzooplankton at reference temperature [day^{-1}]
33	pe_R1Z6X	DOM-fraction of uptake excreted by heteroflagellates [unitless]
34	pu_eaZ6X	Fraction of prey-uptake excreted by heteroflagellates [unitless]
35	puZ6X	Heteroflagellates assimilation efficiency [unitless]
36	sdZ6oX	Specific mortality of heteroflagellates due to oxygen limitation [day^{-1}]
37	sdZ6X	Specific basal mortality of heteroflagellates [day^{-1}]
38	srsZ6X	Specific rest respiration of heteroflagellates at reference temperature [day^{-1}]

Table 3. Ranking of the parameters from the regression analysis of the output of the Monte Carlo simulations. $|\beta|$ is the absolute value of the standardized regression coefficients, which are reported with their standard errors, and p-value is the level of significance of the t-test on the parameters. See Table 2 for description of the parameter notations. N.S. indicates model parameters that were not associated to significant regression coefficients (t-statistic).

Rank	Parameter	$ \beta $	Standard error	p-value	Group
1	chN3oX	0.462	0.008	p<0.05	11
2	pu_eaZ6X	0.422	0.008	p<0.05	14
3	puB1oX	0.373	0.008	p<0.05	9
4	frB1R3	0.314	0.008	p<0.05	9
5	pu_eaZ5X	0.300	0.008	p<0.05	14
6	srsB1X	0.251	0.008	p<0.05	9
7	sdZ6oX	0.185	0.008	p<0.05	14
8	sdB1X	0.170	0.008	p<0.05	9
9	puZ5X	0.139	0.008	p<0.05	14
10	sdZ5oX	0.110	0.008	p<0.05	14
11	puZ6X	0.109	0.008	p<0.05	14
12	srsZ5X	0.080	0.008	p<0.05	14
13	puZ4X	0.044	0.008	p<0.05	14
14	srsZ6X	0.037	0.008	p<0.05	14
15	puB1X	0.036	0.008	p<0.05	9
16	sdZ6X	0.035	0.008	p<0.05	14
17	pe_R1Z4X	0.031	0.008	p<0.05	14
18	pu_eaZ4X	0.031	0.008	p<0.05	14
19	sdZ5X	0.021	0.008	p<0.05	14
20	srsZ4X	0.020	0.008	p<0.05	14
21	rR2R1X	0.018	0.008	p<0.05	11
22	sR1N1X	0.010	0.008	N.S.	11
23	DeniX	0.009	0.008	N.S.	9
24	Z4mortX	0.009	0.008	N.S.	14
25	pe_R1Z5X	0.008	0.008	N.S.	14
26	sdZ4X	0.008	0.008	N.S.	14
27	sR1N4X	0.006	0.008	N.S.	11
28	pe_R1Z6X	0.005	0.008	N.S.	14
29	redfieldX	0.005	0.008	N.S.	11
30	sdZ4oX	0.005	0.008	N.S.	14
31	puR4_B1X	0.004	0.008	N.S.	11
32	reoX	0.004	0.008	N.S.	11
33	Z4repwX	0.004	0.008	N.S.	14
34	sN4N3X	0.004	0.008	N.S.	11
35	R1R2X	0.003	0.008	N.S.	11
36	pu_eaRZ4X	0.002	0.008	N.S.	14
37	puR6_B1X	0.001	0.008	N.S.	11
38	puR8_B1X	0.001	0.008	N.S.	11

Regression statistics:

Number of cases = 1000; Coefficient of determination: $R^2=0.94$; F-value= 414.2

870 **Figure Captions**

871

872 **Figure 1.** Location of the study site in the Arabian Sea (65°E, 13°N).

873

874 **Figure 2.** Schematic of the coupled GOTM-ERSEM model configuration used in this study.

875 ERSEM describes the biogeochemical and trophic processes that drive the evolution of inorganic

876 and organic variables in the simulated pelagic environment, and the exchanges of oxygen and

877 carbon dioxide with the atmosphere. Remineralization closure equations represent the fluxes at the

878 deep water boundary. GOTM describes the physical vertical mixing in the water column, taking

879 account of the meteorological forcing. The black arrows represent ecosystem processes described

880 by Butenschön et al. (2016), and the white arrow represents denitrification, which was included in

881 ERSEM in this work.

882

883 **Figure 3.** Model climatology computed from a simulation of years 2002-2005 (left), versus

884 climatology data derived from the World Ocean Dataset 2009 (right), for oxygen and nutrients.

885

886 **Figure 4.** Annual average profile of oxygen in the climatologies derived from the model simulation

887 of years 2002-2005 (“Model”) and from the World Ocean Dataset 2009 (“Data”).

888

889 **Figure 5.** Taylor diagram summarizing the model skill in reproducing the climatological data of

890 oxygen (O₂), nitrate (NO₃), phosphate (PO₄) and silicate (SiO). The diagram represents the

891 Pearson correlation coefficient (ρ), the standard deviations of model and data (σ and σ_0 ,

892 respectively) and the model bias. The optimal skill point is represented by the black dot with

893 coordinates (1,0).

894

895

896 **Figure 6.** Result of the screening analysis based on the Morris method applied to 21 groups of

897 model parameters. The three most important groups (i.e. the ones with the highest values of the
898 sensitivity index μ^*) were, in order: I) group 14 (zooplankton loss parameters); II) group 9
899 (bacterial loss parameters); and III) group 11 (additional nutrient remineralization parameters).
900

901 **Figure 7.** Conceptual diagram of the most relevant biogeochemical processes driving the
902 Oxygen Minimum Zone (OMZ) at the oligotrophic study site in the open Arabian Sea. At surface,
903 oxygen (O₂) is exchanged with the atmosphere, produced by phytoplankton, and consumed by both
904 autotrophic and heterotrophic (zooplankton and bacteria) respiration. Net photosynthesis fades at
905 the depth of ~100 m, which parts euphotic and twilight zone. Below this depth, bacteria
906 remineralize aerobically the detritus sinking from the surface, thus consuming oxygen down to its
907 lowest value at ~200 m. At low oxygen values, bacteria respire by reducing nitrate (NO₃) via
908 denitrification. We found that the release of recalcitrant dissolved organic carbon (RDOC) and
909 grazing of heterotrophic nanoflagellates (HNAN) also contribute significantly to the formation of
910 the OMZ, by reducing the bulk biomass of bacteria, hence their overall respiration.
911

912 **Figure 8.** Simulated annual evolution of particulate organic carbon (POC, continuous blue line)
913 and bacteria biomass (dashed red line), at the depth of 200 m where the average annual minimum
914 of dissolved oxygen occurs.