Evolution of oxygen and stratification in the North Pacific Ocean in CMIP6 Earth System Models

Lyuba Novi[1], Annalisa Bracco[1], Takamitsu Ito[1], Yohei Takano[2,3]

[1]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA.
[2]British Antarctic Survey, Cambridge, UK.
[3]Los Alamos National Laboratory, Los Alamos, NM, USA.

Correspondence to: L. Novi (lnovi3@gatech.edu)

Abstract. This study examines the linkages between the upper ocean (0-200 m) oxygen (O_2) content and stratification in the North Pacific Ocean in four Earth system models (ESMs), an ocean hindcast simulation, and ocean reanalysis data. Trend and variability of oceanic O_2 content are driven by the imbalance between physical supply and biological demand. The physical supply is primarily controlled by ocean ventilation, which is responsible for the transport of O_2-rich surface waters into the subsurface. Isopycnic Potential Vorticity (IPV) - is a quasi-conservative tracer proportional to density stratification that can be evaluated from temperature and salinity measurements - is used as a dynamical proxy for ocean ventilation. The predictability potential of the IPV field is evaluated through its information entropy.

Results highlight a strong O_2-IPV connection and somewhat higher (than in rest of the basin) predictability potential for IPV across the tropical Pacific, where the El Niño Southern Oscillation occurs. This pattern of higher predictability and strong anticorrelation between O_2 and stratification is robust across multiple models and datasets. In contrast, IPV at mid-latitudes has low predictability potential and its center of action differs from that of O_2. In addition, the locations of extreme events or hotspots may or may not differ among the two fields, with a strong model dependency, which persists in future projections. These results, on one hand, suggest the possibility to monitor ocean O_2 through few observational sites co-located with some of the more abundant IPV measurements in the tropical Pacific, and, on the other, question the robustness of the IPV-O_2 relationship in the extra-tropics. The proposed framework helps characterizing and interpreting O_2 variability in relation to physical variability and may be especially useful in the analysis of new observationally-based data products derived from the BGC-ARGO float array in combination with the traditional but far more abundant ARGO data.

1 Introduction

Dissolved oxygen (O_2) in the oceans is crucial for biogeochemical cycling, marine ecosystem and redox chemistry of seawater. O_2 is a key element for the survival and functioning of marine organisms as fish, shellfish, marine mammals, and other aquatic life rely on O_2 to breathe and carry out essential metabolic processes. Many commercially important fish species and shellfish thrive only on well-oxygenated waters. Growth, reproduction, and overall health of marine organisms

depends on the balance between metabolic demands and O_2 supply (Deutsch et al., 2015).

Ocean deoxygenation refers to the long-term decrease in the concentration of O_2 in the Earth's oceans. At the global scale, the O_2 inventory has been declining significantly over the past decades according to historical observations (Ito et al., 2017; Schmidtko et al., 2017). Changes in O_2 concentrations can reflect the impacts of climate change, nutrient pollution,

35 eutrophication, and other human-induced stressors (e.g. Breitburg et al., 2018). Monitoring ocean O_2 levels helps scientists assessing the health and resilience of marine ecosystems and identifying areas that may be prone to O_2 depletion or hypoxic events. Monitoring and predicting O_2 levels in the oceans are especially important around and within Oxygen Minimum Zones (OMZs), which are characterized by layers in the water column with very low O_2 concentration due to biological, chemical, and physical processes. As oceans warm, OMZs are posed to increase in number and size across the globe. In the

40 North Pacific, a large OMZ exists on the eastern side of the tropical Pacific, and its variability and trends are important also for nitrogen cycling and production of N_2O, a potent greenhouse gas (Nevison et al., 2003; Yang et al., 2017).

Interpreting changes in O_2 concentrations requires understandings how ocean circulation, mixing, air-sea gas exchange, biological productivity and respiration operate. The air-sea gas exchange for O_2 is relatively efficient, and it maintains the surface water close to saturation with the overlying atmosphere for ice-free regions. Ocean circulation is the primary

45 pathway through which O_2 is supplied (or ventilated) into the thermocline and deep ocean. In the subsurface, O_2 is gradually consumed by respiration due to the decomposition of dissolved and particulate organic matter. The O_2 concentration progressively decreases as water masses age. At climatological timescale, the rates of O_2 supply and consumption are balanced to sustain a steady state. In another words, changes in O_2 concentration are caused by an imbalance between O_2 supply and O_2 consumption.

50 On the supply side, the ventilation of O_2 is essentially controlled by the ocean circulation and mixing processes. Broadly, ventilation refers to the exchange of waters between the surface layer and the ocean interior (Talley et al., 2011), and involves a wide range of physical processes such as the wind-driven shallow overturning associated with the Subtropical cells (Brandt et al., 2015; Duteil et al., 2014; Eddebbar et al., 2019), the formation of mode and intermediate waters (Claret et al., 2018; Sallee et al., 2010, 2012; Gnanadesikan et al., 2012) and the lateral (isopycnal) eddy stirring (Rudnickas et al.,

55 2019; Gnanadesikan et al., 2013, 2015). These circulation systems are ultimately driven by the atmospheric winds and air-sea buoyancy fluxes which exhibit significant interannual, decadal and multi-decadal variability.

Fluctuations in ventilation rates as well as ocean stratification are known to impact both O_2 levels (Ridder & England, 2014; Duteil et al., 2014; McKinley et al., 2003) and the distribution of isopycnal potential vorticity (IPV), a dynamical tracer which is proportional to the local stratification and the Coriolis parameter. The use of the absolute value of the Coriolis

60 parameter in the formula, indicated by *, guarantees that the relationship with stratification holds with the same sign in both hemispheres, so that higher IPV* indicates stronger stratification and vice versa. A strong winter-time convective mixing will produce weakly stratified, O_2-rich water masses (low IPV* and high O_2), and vice versa. These properties are then brought together into the ocean interior following the pathway of large-scale ocean currents.

In this study, we build upon this relationship and explore the overarching hypothesis that isopycnic potential vorticity (IPV*) may be used as a proxy for O_2 with a focus on the North Pacific basin. If this was the case, then IPV* may provide a path to monitor and predict the evolution of O_2.

In the North Pacific, the Pacific Decadal Oscillation (PDO) is the mode of climate variability that exerts the greatest control on stratification and O_2, as shown in Ito et al. (2019). Indeed, the dominant mode of oxygen variability in the North Pacific Ocean is correlated with the PDO index such that the PDO explains about 25% of its variance. In the tropics, the PDO modulates the depths of isopycnal surfaces and biological productivity/respiration together with the El Niño Southern Oscillation (ENSO), while at mid-latitudes is the dominant mode influencing the depth of the winter mixed-layer ventilation and the ventilation processes. Our specific objectives are to analyze Earth System Models (ESMs) to evaluate the hypotheses that (HYP 1) the ocean ventilation (IPV*) regulates O_2 variability in the oceans; (HYP 2) the PDO/ENSO-ventilation-O_2 linkage provides the basis for the predictability of O_2 whenever IPV* is predictable; and (HYP 3) the linkage can be exploited to identify hotspots of O_2 changes in variability, means and extremes (see Methods).

Rather than working directly with the observational data, this study applies data-mining tools to a combination of Earth System Models (ESMs) outputs. ESMs provide a mathematical representation of underlying physical and biogeochemical processes in the form of coupled partial differential equations that are discretized and computationally integrated using high performance computing infrastructures. The time-evolving, three-dimensional distribution of physical and biogeochemical variables are generated as outputs of such calculations. Here we analyze outputs from the Coupled Model Intercomparison Project Phase 6 (CMIP6, Eyring, 2016), a major international effort with the primary objective of providing a standardized framework for simulating past, present, and future climate conditions. The participating modeling groups run their climate models under specified radiative forcing scenarios, and generate a comprehensive set of output datasets freely available to the scientific community through data portals and archives provided by the Earth System Grid Federation (ESGF). Using a suite of ESMs we will address the following questions:

- How robust is the relationship between O_2 and IPV* in the North Pacific across several ESMs and how may it evolve by the end of the 21st century? ($\rightarrow$ HYP 1)

- What are the linkages between O_2 and IPV* versus large-scale modes of climate variability such as PDO and ENSO? ($\rightarrow$ HYP 2)

- Where are the hotspots of changes in IPV* and O_2, both in the historical period and in the projections, and are they co-located or differ in space and time? ($\rightarrow$ HYP 3)

To address the above questions, we apply a data-mining tool for dimensionality reduction and network analysis (δ-MAPS, Fountalis et al., 2018), and apply concepts such as information entropy (IE), and the standard Euclidean distance index

(SED, Diffenbaugh and Giorgi, 2012). Since the above approaches are relatively new to the ocean biogeochemistry community, a brief overview is provided here, followed by detailed definitions in section 2.

First, Information Entropy (IE) is a metric that measures the amount of randomness and therefore unpredictability in a dataset. For example, given a time series, IE measures how much information is contained in the data. If the time series is a random sequence, the entropy is high because the complexity is high, and the system is chaotic. On the other hand, if it approximately follows a sinusoidal curve, its predictability is high, and the IE would be low.

Second, δ-MAPS (Fountalis et al., 2018) combines feature extraction and network analysis into a single framework. The goal of δ-MAPS is to identify key features and to visualize how those features relate to one another. It establishes sets of spatially connected chunks of grid points sharing similar dynamical features, called domains. This is simpler and easier to interpret than empirical orthogonal functions (EOFs) which suffer from orthogonality constraints. It also allows to investigate the network of domains, but this step is skipped in this work. The benefits of δ-MAPS include simplicity, interpretability and overfitting prevention relative to conventional EOF-based approaches when extracting climate patterns from high-dimensional datasets.

Finally, the Standard Euclidean Distance (SED) index is a simple and flexible method used to detect *total* changes in one or more variables in a given dataset (in other words to identify regions that stand out for changes in means, extremes and variability), through measuring the distance in multi-variate space between a baseline period and any other (Diffenbaugh et al., 2008; Diffenbaugh and Giorgi, 2012, Williams et al., 2007). The SED is a non-parametric method, meaning it does not assume a specific probability distribution for the data. This flexibility makes it applicable to a wide range of datasets, regardless of their underlying distribution.

Before delving into these methodologies, we stress that this work aims not only at testing the hypothesis we put forward, but also at introducing recently developed approaches for model intercomparison and more generally data analysis to the ocean biogeochemistry community.

2 Materials and methods

In this section, we describe in more details the three tools recently developed for climate science applications and adopted in our analysis and how we calculate IPV*.

We evaluated the predictability potential using the Information Entropy (IE). IE is defined following Prado et al. (2020) and is based on the recurrence of microstates in a recurrence plot (RP). A RP (Eckmann et al 1987) is a visualization technique for trajectory recurrence of a given dynamical system described in phase space by a matrix R_{ij} such that

$$R_{ij}(\epsilon) = \Theta\left(\epsilon - |x_i - x_j|\right), \ x_i \in \mathbb{R}^d, \ i,j \ = \ 1,2, \dots, K, \tag{1}$$

where Θ is the Heaviside function, $|\ \ |$ is an Euclidean distance, in our work, x_i and x_j are states at time steps i and j, ε is a threshold distance (the maximum distance between two states to be considered mutually recurrent), d is the x_i space dimension and K is the number of states considered (the length of each analyzed time series). R_{ij} is a matrix which represents non-recurrent (as zeros) and recurrent (as ones) states in phase space respectively, and it is explicitly dependent on ε. Corso et al. (2018) introduced the *Recurrence Entropy* quantifier, for which for a given time series, the probability of occurrence of microstates in its RP is quantified without the need for a space-state reconstruction. A microstate of dimension N is a NxN matrix sampled inside the RP, with probability of occurrence $P_k = n_k/N_{tot}$, where n_k are the number of occurrences of the microstate, and N_{tot} is the total number of possible configurations of 0 and 1 of the microstate (see Ikuyajolu et al (2021) and Prado et al (2020) for more details). The information entropy IE is then defined as

$$IE(N_{tot}) = \ - \ \sum_{k=1}^{N_{tot}} P_k \ \ln P_k \tag{2}$$

where k refers to the kth microstate. When IE is normalized by the maximum entropy (corresponding to when all microstates show the same probability) then IE=0 corresponds to perfect predictability, while IE=1 represents chaos. Furthermore, the explicit dependence of the entropy quantifier on ε is removed using the maximum entropy formulation. Prado et al. (2020) have shown that a value for which IE is maximum exists, does not vary much for varying ε and is strongly correlated with the Lyapunov coefficient of the system, (we refer to Ikuyajolu et al (2021) for the details of the heuristic used to estimate the maximum entropy). In brief, Prado et al. (2020) suggests a technique to eliminate the dependence of the entropy computation on the selection of a distance threshold (ε) by finding a clearly defined maximum in the relationship between epsilon and the entropy (Smax in Fig. 4 of Prado et al. (2020)). This maximum is robust and relatively stable in a range of epsilon values. Furthermore, there is a strong correlation between the maximum entropy and the Lyapunov exponent. In our work, the code used to compute the entropy (see Data Availability section) uses the heuristic explained in detail in Ikuyajolu

145 et al (2021) for the calculation of Smax through an iterative procedure that calculates the recurrence entropy for varying ε until a maximum is found and retained. This algorithm requires three input parameters: the microstate dimension (that we set at 4, but explored other values as shown in Results), the number of random samples to compute the microstates distribution in the RP (here 10000) and a random sub-sample used to determine the ε for which entropy is max (here 1000). We compute the entropy field of the deseasonalized and detrended IPV* (full signal) using monthly data over the whole historical and

150 future periods. In each point, the entropy of the IPV* field is associated with recurring microstates in its time series (that define the system and thereby impacts its predictability). The higher the predictability of a time series the more recurrent are its temporal dynamics, i.e. the easiest will be to predict its future evolution.

δ-MAPS (Fountalis et al., 2018) is an unsupervised network analysis method that identifies spatially contiguous and possibly overlapping regions referred to as domains, and the lagged functional relationships between them. In short, domains are

155 spatially contiguous regions that share a highly correlated temporal activity between grid cells of the same domain. In this work we apply it to the sea surface temperature (SST) anomaly field (see Data) to identify the major modes of climate variability in the north Pacific in a reanalysis and in the ESMs. δ-MAPS is an alternative approach to reduce the dimensionality of spatio-temporal data to EOFs (standard or rotated). The orthogonality between EOF components complicates the interpretation of the results, especially when comparing models and observational datasets, as discussed, for

160 example, in Dommenget and Latif (2002) and Falasca et al. (2019), and δ-MAPS offers a powerful solution to this problem. Given any spatio-temporal field, its local homogeneity is hypothesized to be highest at "epicenters" or "cores". For each grid point, a local homogeneity is defined as the average pairwise cross-correlation between that grid cell and a set of K nearest neighbors (see Fountalis et al., 2018 for details). Cores are then determined as neighbors of points where the local homogeneity is a local maximum and above a threshold δ. Each core is iteratively expanded and merged using a greedy

165 algorithm to iteratively find domains as large as possible that are (i) spatially contiguous, (ii) include at least a core and (iii) have homogeneity higher than δ. δ is computed using a significance test for the unlagged cross-correlations. Given any random pair of grid points, the significance of the Pearson's correlation of their timeseries is assessed through the Bartlett's formula (Box et al., 2011) with the null hypothesis of no coupling. The significance of each correlation is tested for a user-specified significance level α, and δ is computed as the average of significant correlations. Here, we applied δ-MAPS with K

170 = 8 and α = 0.01.

The Identification of hotspots of change follows the approach Introduced by Diffenbaugh and Giorgi (2012) (which builds on Williams et al., 2007 and Diffenbaugh et al., 2008 and references therein), and applied by Turco et al. (2015) to the

analysis of global atmospheric data. Hotspots are quantified through a Standard Euclidean Distance index (SED) that aggregates the changes in means, variability and extremes of the given spatio-temporal field according to:

175 $$SED = \sqrt{\sum_{i=1}^{Nind} \sum_{j=1}^{4} \left(\frac{ind_{ij}}{p95(|ind_{ij}|)} \right)^2}$$ (3)

We compute in each grid point two SED indices, separately for O_2 and IPV*. N_{ind} is the total number of indicators per each variable and i the index identifying each indicator, j spans the seasons, so that ind_{ij} is the ith indicator in the jth season, and p95 is the 95th percentile. As the indicators and SEDs are computed point by point, i.e. each grid point has one value, the percentile is computed spatially over all the grid points. Here we consider December-January-February as (boreal) winter,

180 March-April-May as spring, and so on. We consider three indicators for each variable, evaluating changes in means, variability and extremes between two periods of equal length. *Period 1* covers 1950-1981 (1960-1986 for reanalysis and E3SM-2G ocean hindcast), and *Period 2* 1983-2014 (1988-2014 for reanalysis and hindcast) for the historical time, and 2036-2067 and 2069-2100 for the projected future. In equation (3) indicators of both periods are normalized on the 95th percentile calculated over *Period 1*, in order to fairly compare changes of hotspots intensity over time. We chose not to

185 compare 2069-2100 with 1950-1981, but with 2036-2067 instead, because we want to track changes in each period compared to the preceding timeslot in order to quantify how rapidly they occur in future projections compared to historical. For each variable, we compute three indicators at each grid point and for each season using the Climate Data Operators (Schulzweida, 2022) as follows:

- Changes in means are estimated in each season separately by $Ind1\_means = yseasm_2 - yseasm_1$, where *yseasm₁*
190 and *yseasm₂* are multi-year seasonal means in *Period 1* and *2*, respectively. Therefore, taking for example the O_2 historical simulations over 1950-2014 (but similar expressions hold for IPV* and the other periods), $Ind1_{DJF} = <O_{2DJF}>_{1983-2014} - <O_{2DJF}>_{1950-1981}$, $Ind1_{MAM} = <O_{2MAM}>_{1983-2014} - <O_{2MAM}>_{1950-1981}$, $Ind1_{JJA} = <O_{2JJA}>_{1983-2014} - <O_{2JJA}>_{1950-1981}$ and $Ind1_{SON} = <O_{2SON}>_{1983-2014} - <O_{2SON}>_{1950-1981}$, where $< ...>$ is a time mean (seasonal climatology).

- Changes in multi-year seasonal variability are evaluated by (i) detrending each variable point by point in the two periods
195 separately, (ii) computing the multi-year seasonal standard deviation of these detrended fields, yseasσ, for each period for each season, (iii) computing *Ind2_variability* as the percentual changes such that $Ind2\_variability = 100 \cdot \left(\frac{yseas\sigma_2 - yseas\sigma_1}{yseas\sigma_1} \right)$. Therefore, with the example of O_2 historical simulations over 1950-2014, $Ind2_{DJF} = 100 \, (std(O_{2DJF})_{1983-2014} - std(O_{2DJF})_{1950-1981})/std(O_{2DJF})_{1950-1981}$, where std(…) is the multi-year seasonal (winter) standard deviation over the specified periods (and equivalent formulations hold for the other seasons: MAM,
200 JJA and SON).

- Finally, changes in extremes (in our case specifically overshots of IPV* and undershots of O_2) are computed through the following steps: (i) for each season, we compute at each grid point the multi-year seasonal O_2 minimum or IPV* maximum over *Period* 1 and we consider it as a threshold (one threshold map per season); (ii) we count how many times in each corresponding season of *Period 2* $O_2 <$ threshold$_{O2}$ (IPV* > threshold$_{IPV*}$) is verified; (iii) the percentage of occurrences computed at point (ii) is taken as indicator of percentual changes in extremes and estimated by $Ind3\_extremes = 100 \cdot \left(\frac{N_{occ}}{N_T}\right)$, where N_{occ} is the number of extremes occurrences (by season) and N_T is the total number of months in all the corresponding seasons (96 for the models and 81 for reanalysis and hindcast). Building on previous works (Falasca et al. 2019; Falasca et al. 2022), we expect the topology of a given model to remain relatively stable across ensemble members, i.e. we do not expect the member choice to significantly influence the calculation of extremes and hotspots and especially their relationships, and verified that this is indeed the case in one of the models by testing four additional randomly chosen ensemble members of CanESM5 (see Data). A major advantage of this hotspot definition is that it accounts for changes in mean, variability and extremes at the same time in the identification of the hotspots. In other words, it accounts for the topology of the simulated climate fields, which can be characterized by considering all the three aspects together. This definition of extremes aims at including information on seasons exceeding corresponding baseline extremes, without choosing a priori a threshold on the current distribution, which is especially relevant for comparing changes with respect to a reference baseline. The three indicators, grouped into four seasons for each variable, are then used to compute the SED indices.

Finally, the IPV* (m^{-1} s^{-1}) is used as a proxy of stratification and is defined as the isopycnic potential vorticity (Talley et al., 2011) with the absolute value of the Coriolis parameter in its formula:

$$IPV* = \frac{|f|}{g} N^2 \qquad (4).$$

Here N^2 is the Brunt–Väisälä frequency (which is a measure of a fluid stability to vertical displacements), g is the gravitational acceleration and f is the Coriolis parameter. IPV is a conservative tracer in frictionless and adiabatic circulation. IPV* is calculated over the three-dimensional ocean volume using Eq. 4 and we consider the 0-200 m vertical weighted average. This procedure allows us to compare datasets with different vertical discretization.

3 Data

We consider four ESMs from the CMIP6 catalog, a hindcast and reanalysis data as summarized in Table 1. Whenever multiple ensemble members were available, we selected the first (r1i1p1f1). We randomly selected four additional ensemble members for CanESM5 (r5i1p1f1, r10i1p1f1, r15i1p1f1 and r20i1p1f1) to further verify the robustness of the hotspots

calculation to the member choice. All ESMs are forced by the historical radiative forcing from 1850 to 2014 and we analyze the monthly outputs from 1950 to 2014. We further discuss future ssp585 scenarios and focus on the 2036-2100 period, indicated as *future*.

The hindcast is a new ocean-ice biogeochemistry simulation (referred to as the G-Case), E3SMv2.0-BGC (hereafter, E3SM-2G, Takano et al, 2023), based on the Model for Prediction Across Scales-Ocean (MPAS-O), an ocean component of the Energy Exascale Earth System Model (E3SM) version 2 (Golaz et al., 2022). Details on ocean physics updates can be found in Golaz et al. (2022). One of the major updates is the introduction of Redi isopycnal mixing (Redi, 1982). Along with ocean physics updates, this version also incorporated a uniform background vertical diffusion specifically developed for simulations of the ocean biogeochemistry to enhance ocean carbon uptake and thermocline ventilation of dissolved inorganic carbon (DIC). Incorporating this mixing parameterization results in an improved representation of climatological O_2 distributions in the v2.0 version compared to its predecessor (Burrows et al., 2020). The Marine Biogeochemistry Library (MARBL, Long et al, 2021) is used to simulate the ecosystem dynamics and cycling of biogeochemical elements. After the spin-up period, the model is forced by a meteorological reanalysis dataset, JRA-55do version 1.4 (Tsujino et al., 2020) from 1958 onward. As ocean reanalysis, we use the ORAS4 product (Balmaseda et al.,2012; Mogensen et al.,2012) available from 1959 onward, which includes a direct surface fluxes implementation from ERA40 and ERA-Interim and multi-scales bias correction. When analyzing the E3SM-2G hindcast and the ORAS4 reanalysis, we focus on the 1960-2014 interval to avoid the spurious presence of an anticyclonic tropical cyclone in the NE Pacific in 1959 in JRA-55do v1.4.

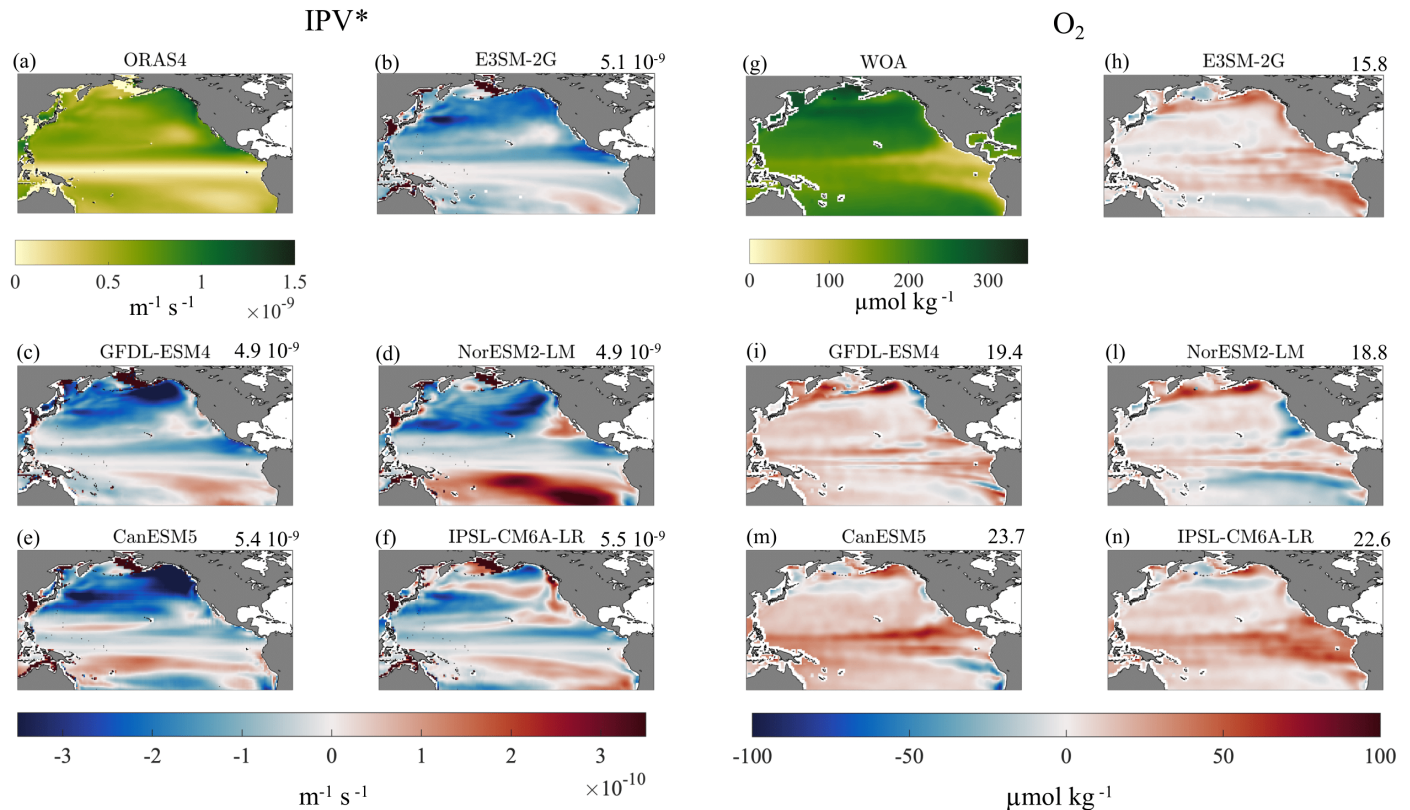
All the data are remapped at 1°x1° horizontal resolution.

Table 1. CMIP6 Earth System Models, global ocean hindcast and reanalysis used in this work.

Modeling Group/Center	Model Name	Atmospheric Component/Resolution	Oceanic Component/Resolution	Reference
National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory	GFDL-ESM4	AM4.0, ~1°, 49 levels	OM4 MOM6, 0.5°x0.5°, 75 vertical levels (hybrid pressure/isopycnal) Biogeochemical component: COBALTv2	(Dunne et al., 2020; Stock et al., 2020)

Canadian Earth System Model version 5	CanESM5	CanAM5, T63 (~2.8º), 49 levels	CanNEMO, 45 vertical levels, NEMO3.4.1, ORCA1 tripolar grid, 1° with refinement to 1/3° within 20° of the equator Biogeochemical component: CMOC	(Swart et al., 2019; Christian et al., 2022)
NorESM Climate modeling Consortium	NorESM2-LM	CAM-OSLO, 2° resolution; 32 levels.	MICOM, 1°, 70 vertical levels Biogeochemical component: iHAMOCC	(Seland et al.,2020; Tjiputra et al., 2020)
Institut Pierre-Simon Laplace	IPSL-CM6A-LR	LMDZ, NPv6, N96; 1.25°Lat x 2.5° Lon, 79 levels	NEMO-OPA (eORCA1.3, tripolar primarily 1°, 75 vertical levels. Biogeochemical component: PISCESv2	(Boucher et al., 2020)
Department of Energy, Energy Exascale Earth System Model	E3SMv2.0-BGC (E3SM-2G)	JRA55do reanalysis (55km, 3hr resolution)	MPAS-O (30 to 60km resolution) Biogeochemical component: E3SMv2.0-BGC, MARBL	(Golaz et al, 2022; Takano et al., 2023; Long et al., 2021)
ECMWF Ocean reanalysis System	ORAS4	_	Global, 1°, 42 Levels	(Balmaseda et al.,2012; Mogensen et al.,2012)

We begin our analysis with a brief evaluation of the ESM biases in the two main fields of interest, IPV* and O_2. For the IPV*, the ocean reanalysis dataset is used for validating the model outputs for the maximum possible time overlap in the historical configuration (1959-2014). For O_2, we can only contrast the annual mean O_2 climatology between the World Ocean Atlas (Garcia et al., 2019) and the ESMs (Fig. 1). We additionally compared the ORAS4 IPV* climatology over

255 1988-2014 (i.e. "period 2" for this reanalysis) with the corresponding climatology computed using SODA3.4.2, which uses a different ocean component compared to ORAS4. The differences across reanalysis that use different models but assimilate the same observations are much smaller (about one order of magnitude) than the signal (Suppl. Fig. S1), and smaller than any model bias.

260 **Figure 1**. (Left) IPV* annual mean climatology (1959-2014) weighted averaged over 0-200 m depth in the North Pacific basin. (a) ORAS4. (b-f) Model biases (model – ORAS4) difference. (Right) O_2 annual mean climatology (1950-2014) weighted averaged over 0-200 m depth. (g) World Ocean Atlas climatology. (h-n) Model biases (model – WOA) difference. The RMSE of the modelled IPV* ($m^{-1} s^{-1}$) and O2 (micro mol kg^{-1}) are shown on top of each panel.

The E3SM-2G hindcast is forced by observed atmospheric fields and not surprisingly displays the smallest bias and root
265 mean square error (RMSE). The RMSE of the modelled IPV* ($m^{-1} s^{-1}$) and O_2 (micro mol kg^{-1}) are shown atop of each panel in Figure 1. Overall, the IPV* and O_2 biases have broadly anticorrelated patterns, with the models being generally less stratified and more oxygen rich than observed in the extra-tropical North Pacific, and often too stratified and with a larger O_2 deficit than observed south of the Equator. However, maximum and minimum biases in the two fields only seldom coincide.

Regionally, the E3SM-2G is generally less stratified than observed with a relatively low O_2 bias and an overestimation of approximately 10 micro mol kg^{-1} in the subtropical thermocline of the North Pacific basin. The hindcast performs especially well in the tropical thermocline. Among the CMIP6 models, CanESM5 shows a slightly higher IPV* underestimation in the subpolar gyre and a O_2 overestimation in the subtropics compared to the other ESMs, while NorESM2-LM emerges as the most stratified south of the Equator. In O_2 larger biases (positive or negative) are found generally in the tropical thermocline and the tropical/subtropical boundaries. The sign and magnitude of the biases are model dependent. Interestingly, models generally overestimate O_2 at subpolar latitudes.

4 Results

4.1 Predictability potential (HYP 1)

We begin our analysis with the predictability potential of IPV*, quantified through the information entropy (IE, see Methods). The goal is to verify if and where IPV* has an elevated predictability skill, owing to the presence of quasi-recurrent behaviors in its time-series. We also aim to examine whether O_2 is correlated with IPV* in regions where the latter has a high predictability potential. As a reminder, IE values close to 1 indicate high complexity and unpredictability, and close to 0 perfect predictability (the signal is recurrent, for example constant or periodic). We preliminary tested the sensitivity of the entropy field to the microstate dimension, within a meaningful range according to previous literature (Ikuyajolu et al., 2021), using microstates of dimension 2, 3, 4 and 5 for GFDL-ESM4 over 1950-2014 (Suppl Fig. S2). The IE pattern, i.e. areas more (less) predictable relative to the surroundings are substantially unchanged, i.e. the geographical patterns are robust, in agreement with Ikuyajolu et al. (2021). Both microstate dimensions 4 and 5 show reasonable entropy values and we chose to use a microstate dimension of 4 to conduct all the analysis because it spans the widest range of possible values.

O_2 – IPV* relationship across ESMs and its future evolution

IE maps for IPV* are shown in Fig. 2 for both historical and future times, with superposed the contours of the areas where the (lagged) anticorrelation between IPV* and O_2 is at least -0.5 (see Suppl. Fig. S3-S4 for the anticorrelation and lag maps). Higher predictability in the historical period is found in the tropical Pacific areas close to the geographical location of ENSO (i.e. the area most impacted by ENSO being the domain identified as ENSO-related by δ-Maps, which well maps the area identified by an EOF analysis over the SST field for having the greatest variance explained by PC1). The predictability potential is generally highest along two stripes enclosing the ENSO pattern and excluding the upwelling cold tongue. The distribution of IE follows broadly that found in a much longer simulation of the IPSL model covering the past 6,000 years

and analyzed by Falasca et al. (2022) and appears to be robust. The western boundary current region and the Kuroshio-Oyashio extensions have low predictability across all datasets considered. In NorESM2-LM and CanESM5, and to a lesser degree in ORAS4 and IPSL-CM6A-LR, the higher predictability of the ENSO area extends to the north-eastern portion of the basin. In general, in both the hindcast and the models, strong anticorrelations between the time series of IPV* and O_2 (c.c $\leq$ -0.5) coincide with low IE regions and are linked to ENSO affecting concurrently stratification and O_2 in the tropics and south of the upwelling area. Very limited IPV* predictability is found in the central and western North Pacific, where the variability is dominated by the PDO signal. The PDO does not emerge as easily predictable both in the interval considered, in agreement with e.g. Gordon et al. (2021) and, at least in the IPSL model, across the whole second half of the Holocene (Falasca et al., 2022). In those areas, anticorrelations between O_2 and IPV* are relatively weak (generally > -0.4 but for NorESM2-LM). The entropy and the regions where the evolution of IPV* and O_2 are strongly anticorrelated do not change significantly in the future projections in the four models. We further explored whether oxygen solubility, (O_2sol), which is modulated by ocean warming/cooling, and the apparent oxygen utilization AOU, which is controlled mostly by the biogeochemical processes affecting oxygen demand, may be independently linked to IPV* predictability. The areas where IPV* and AOU time series are positively correlated with correlation coefficients $\geq$ 0.5 are very similar to the ones obtained by analyzing the O_2-IPV* relationship. For O_2sol, which well approximates preformed O_2 at the depths considered, the anticorrelations areas (i.e. where c.c. $\leq$ -0.5) are quite extensive, especially in the hindcast, but mostly superimposed to high-entropy/low predictability IPV* areas (Suppl. Fig. S5).

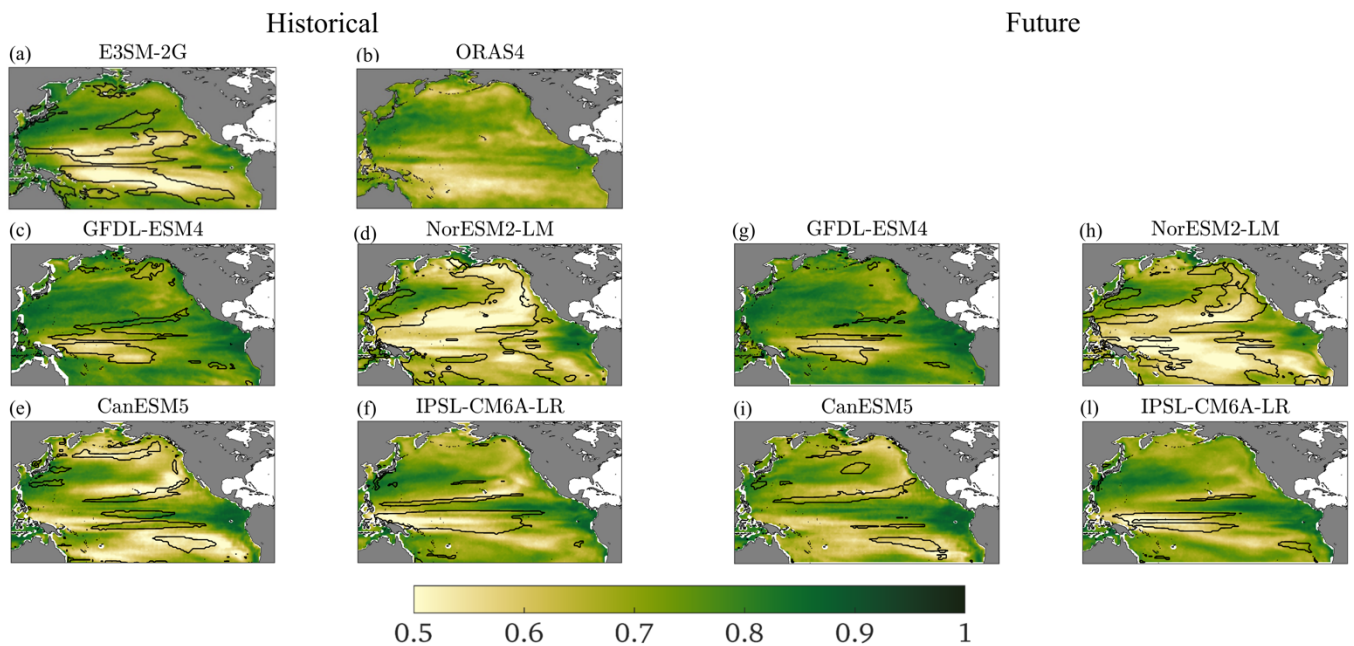



Figure 2: IPV* entropy field in the historical interval (left) and in the future (right) for the ESMs, and in the historical 1960-2014 period for the hindcast and ORAS4, with superposed the contours of the areas where IPV* and O_2 time series are anticorrelated with correlation coefficients $\leq$ -0.5.

In the next section we isolate the PDO signal to explore whether the low predictability in the northern Pacific Ocean (north of ENSO region) related to time scales, i.e. if there is a low frequency PDO modulation with a high frequency "noise" due to both atmospheric and oceanic variability. The PDO is indeed a lower-frequency mode compared to ENSO and has most loading at extra-tropical latitudes where the "noise" is greater.

4.2 Trends and PDO impact on O_2 and IPV* (HYP 2)

The limited predictability found in the North Pacific does not exclude the possibility of the PDO modulating both IPV* and O_2 inventories simultaneously. As explored in previous works (for example, Ito et al., 2019), the dominant mode of observed O_2 variability in the northern Pacific Ocean is correlated with the PDO index which explains about 25% of the variance. Observations, however, offer only sparse coverage, in both time and space. To further verify the PDO modulation in the present work, we computed the first EOF for the E3SM-2G hindcast 0-200m O_2 and IPV* anomalies over 1960-2014 over the northern Pacific (20.5°N-69.5°N;115.5°E-60.5°W) and the corresponding time series for the first principal component *(pc1)*. The first EOF explains 25% of the oxygen variance and about 12% of the IPV* variance. The computed *pc1* shows a significant and strong correlation (Pearson's R coefficient) with PDO timeseries computed using SST anomalies with $|R|$ = 0.83 (p < 0.01) for O_2, and $|R|$ = 0.44 (p <0.01) for IPV*, after applying 5-year moving means. The correlation with the PDO is higher than with ENSO, which is at most $|R|$ = 0.34 (p < 0.01) for O_2, after applying a 3-month moving mean. This is consistent with previous knowledge that oxygen variability in the North Pacific is mostly associated with the PDO. We hereby quantify the (linear) impact of the PDO on the two fields of interest, and then evaluate their residuals. If the PDO is the main predictor of IPV* and O_2 distributions, its impact on the two fields should be strongly anticorrelated and larger than the residual. As mentioned, the objective is to verify if the greater availability of IPV* in observations, reanalyses and modeled fields could be used to extrapolate information about O_2 and its evolution in time, bypassing – at least to a certain degree – the need to run full biogeochemical models or measure O_2 directly.

Estimation of climate indices in ESMs and linear regression

We use δ-MAPS (see Material and Methods) applied to the SST field to evaluate the main modes of Pacific climate variability, ENSO and PDO with a greater focus on the latter, and their time evolution in the models, the ocean hindcast and the reanalysis. While the evolution of ENSO using δ-MAPS is straightforwardly described by the timeseries of the cumulative anomalies in the ENSO-related domain (e.g. Falasca et al., 2019), for the PDO we must consider the difference between the SST cumulative anomalies of two domains. The domains are identified by the complex network algorithm, and

we applied a 5-yr running mean to produce the PDO time-series shown in Fig. 3. The domains shape and size are indicated in Fig. 4. For ORAS4 and E3SM-2G over 1960-2014, we computed the 0-lag Pearson's correlation coefficients between these timeseries and the commonly defined indices of PDO (following Mantua et al. 1997) and Nino34 (average SST anomalies over the box 5ºN-5ºS, 170ºW-120ºW) for validation. Both timeseries are retrieved from NOAA

350 (https://psl.noaa.gov/data/climateindices/list/), and a 3-month moving average is applied to all the ENSO timeseries (signals and indices) and a 5-year moving average to all the PDO timeseries (signals and indices). Correlation coefficients are 0.88 for PDO and 0.93 for ENSO in ORAS4, and are 0.89 for PDO and 0.91 for ENSO in E3SM-2G.

Moving to the models (Fig. 3), in the historical period GFDL-ESM4 slightly underestimates the PDO strength, while the opposite is verified in CanESM5 and NorESM2-LM. In the latter, the frequency of the signal is also higher than observed.

355 By the end of the 21st century, the strength of the PDO remains unaltered in GFDL-ESM4 and IPSL-CM6A, while decreases in NorESM2-LM and especially in CanESM5, following a decrease in size of the eastern domain. A decrease in amplitude and increase in frequency of the PDO was found also in several models in the CMIP5 ensemble (Li et al., 2020).

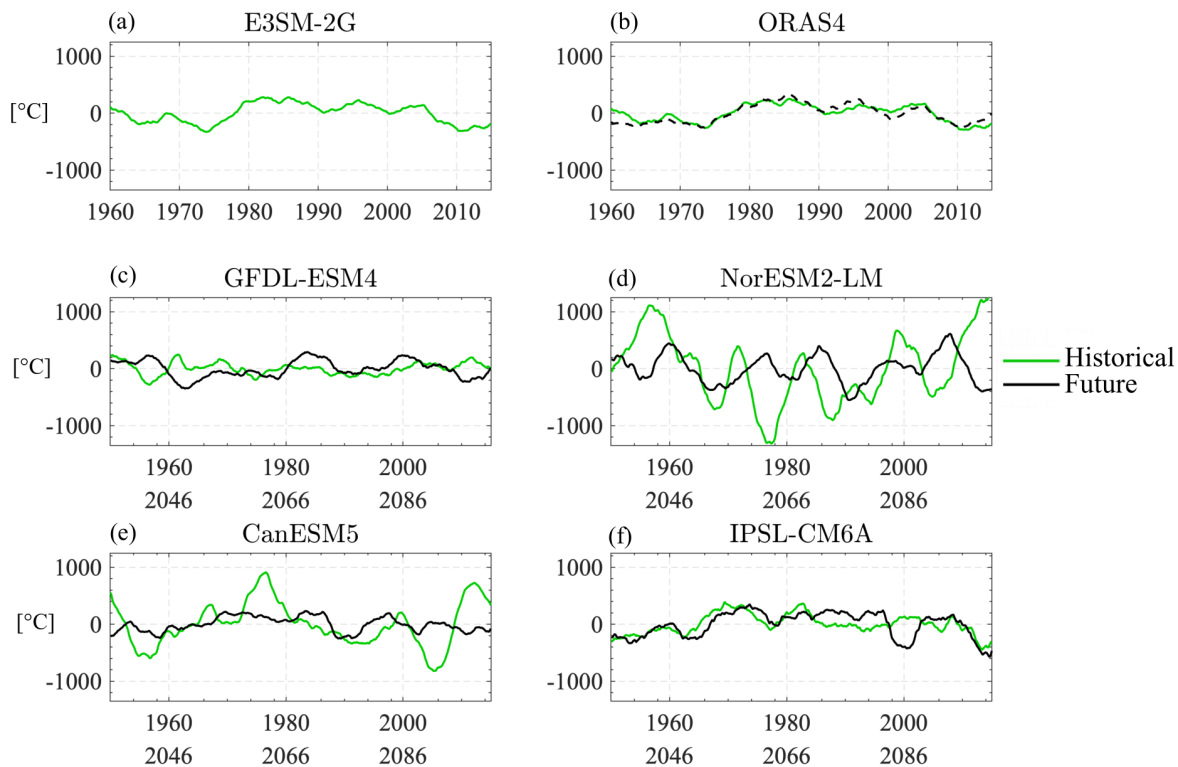



Figure 3: PDO indices (SST cumulative anomalies) calculated using δ-MAPS (see text) in the historical and future time periods. The

360 dashed line in panel (b) is the PDO index timeseries from NOAA (available at https://psl.noaa.gov/data/climateindices/list/), after a 5-year moving mean is applied and rescaled to fit in the plot (multiplied by the standard deviation of the ORAS4 timeseries in panel (b)).

Given the PDO(t) indices, the residual component of the fields of interest that is not linearly forced by the PDO can be separated as a function of time (see e.g. Kucharski et al., 2008) so that for O_2 (but the same procedure was applied to IPV*):

$$O_{2res}(x,y,t)=O_2(x,y,t)-O_{2PDO}(x,y,t) \qquad (5),$$

365 where

$$O_{2PDO}(x,y,t) = b_{O2}(x,y)*PDO(t) \qquad . \qquad (6)$$

$b_{O2}(x,y)$ is constant in time and determined by least-square fitting through a linear regression for each dataset separately. Figure 4 shows b_{IPV*} and b_{O2} for all datasets with superposed the boundaries of the domains corresponding to the ENSO mode and those contributing to the PDO, in the historical period. In most cases there is an overall anticorrelation signal

370 between the maps for the two fields, as to be expected, but also several important differences. First, the regions where b_{O2} is strongest (both positive and negative values) do not correspond to minima and maxima in b_{IPV*}. Second, the equatorial upwelling tends to have a strong positive signal in b_{O2} and only a weak one, but of the same sign, in b_{IPV*}. Third, the PDO impacts on the two fields vary substantially among models, as quantified by the correlation among the respective fields indicated atop the b_{O2} plots, with GFDL being the closest to the hindcast and, for the IPV* case, also to the reanalysis. In

375 NorESM2-LM the anticorrelation between the regression fields is too strong and the PDO has both a shape and loading in the Pacific interior which is different than observed. CanESM5 and IPSL-CM6A-LR display positive spatial correlations, with important biases with respect to the hindcast at the equator and along the eastern boundary. Furthermore, the trends of the residuals have amplitude comparable to the PDO-forced signal in both fields in the historical period in all cases (see Suppl. Fig. S6), and the linear trends of the whole signals are nearly identical to those of the residuals. We will further

380 discuss the trends shape when presenting *Ind1 means* (i.e., the indicator for changes in means).

We performed a comparable linear regression analysis also using the ENSO index, instead of the PDO, and obtained similar shapes of the b coefficients -as to be expected – but much lower absolute values (Suppl. Fig. S7). This further confirms that in the North Pacific the PDO is the dominant role of climate variability.

385

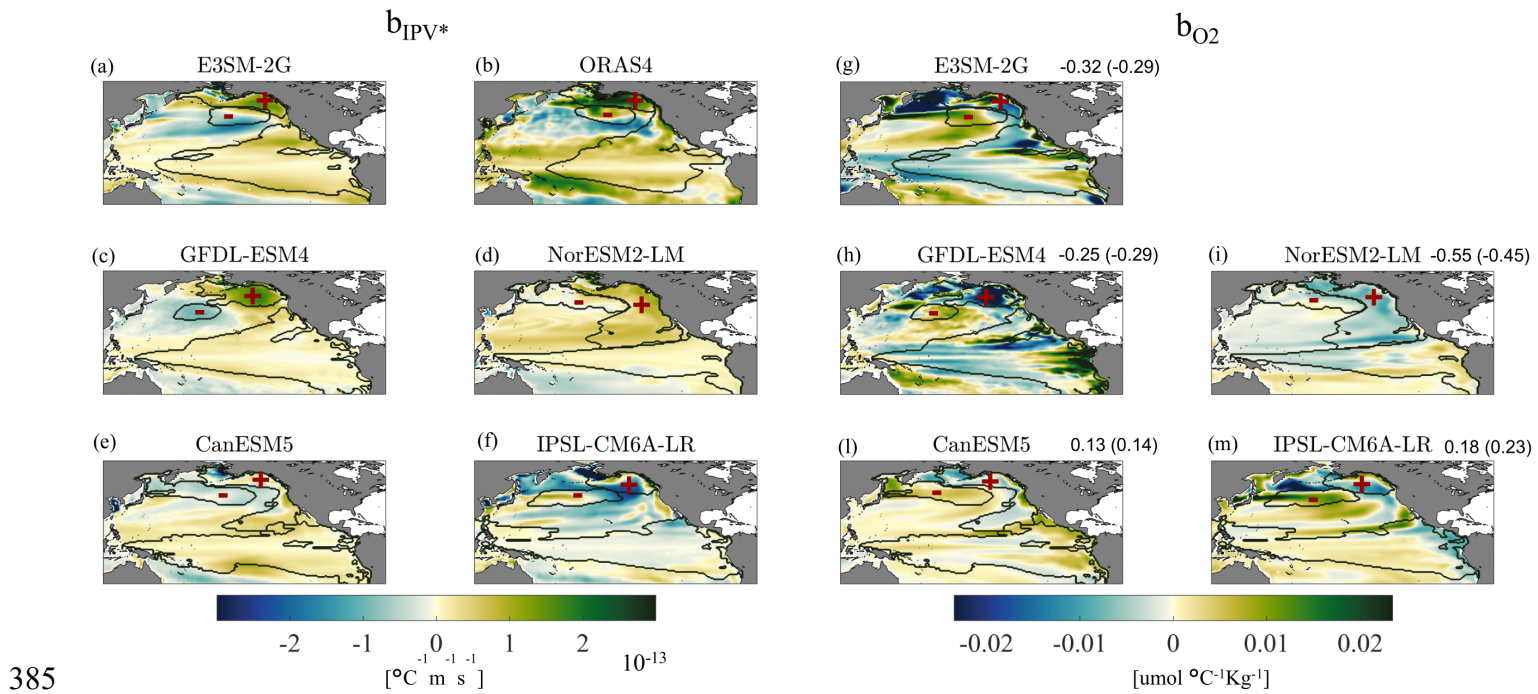



Figure 4: b_{IPV*} (left) and b_{O2} (right) regression coefficient maps with superposed contours of the ENSO and of the PDO+ and PDO-domains. B_{IPV*} represents the change in IPV* per change in SST, b_{O2*} represents the change in O_2 per change in SST. The correlation coefficients among the corresponding maps for the same model or hindcast are also indicated. Color limits are fixed as +/- 3 standard

deviations of the ensemble for each variable over the whole area (+/- 2.99 10⁻¹³ for IPV* and +/- 0.023 for O₂). Values in parentheses are

390   c.c. computed north of 20ºN. All the c.c. values passed a shuffling significance test at 5% level (see Suppl. Mat.).

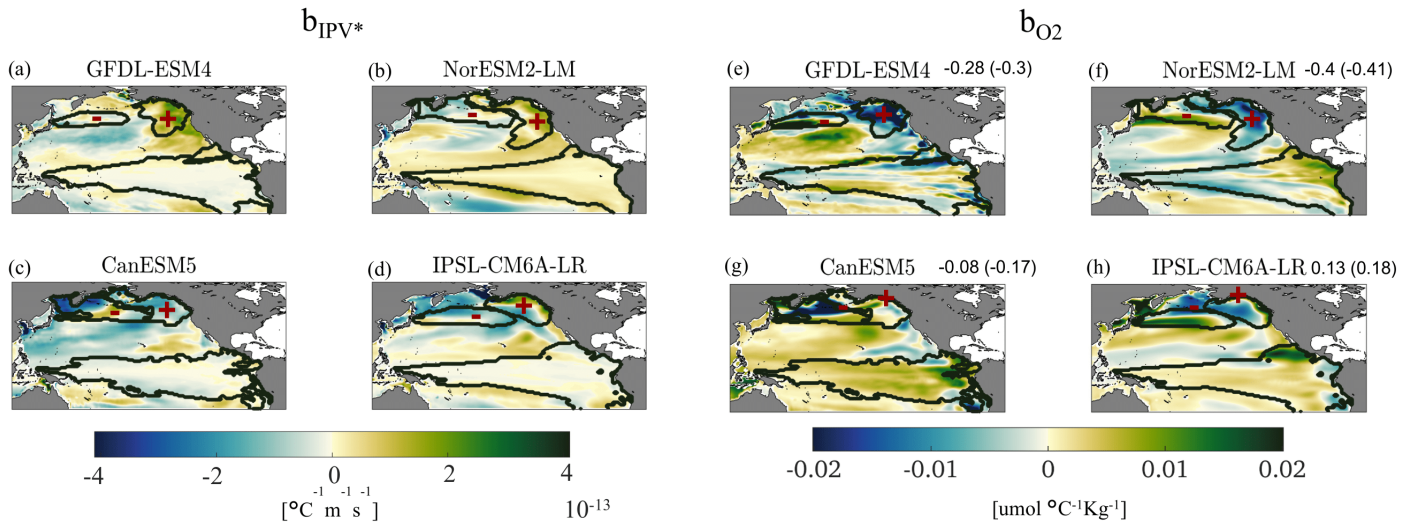



Figure 5: as in Figure 4 but for the future projections. Color limits are fixed as +/- 3 standard deviations of the ensemble for each variable over the whole area (+/- 4.1 10⁻¹³ for IPV* and +/- 0.02 for O₂). Values in parentheses are c.c. computed north of 20ºN. All correlation coefficients passed a significance test at 5% level (see Suppl. Mat.).

395

Moving to the projections, the maps of the regression coefficients do not change considerably in three of the models considered (Fig. 5). In CanESM5, on the other hand, b_{IPV*} changes sign over most of the domain. The residual trends, when compared to the regression coefficients, are stronger and dominate the evolution of both fields, especially in the subtropical and subpolar gyres of the North Pacific (Suppl. Fig. S8), superseding the PDO signal.

400   **4.3 Hotspots of change (HYP 3)**

As a last step, we evaluate changes in means, extremes and variability for both variables (considering whole signals, i.e. not just the residuals) using the indicators introduced in the Methods. For the historical time, we divide the 1950-2014 interval in two periods of equal length covering 1950-1981 and 1983-2014 (1960-1986 and 1988-2014 for E3SM-2G and ORAS4). We evaluated the indicators using all seasons separately or averaged together, and found that differences across seasons were

405   small, as measured by the standard deviation of the indicators (Suppl. Fig. S9-S11). In the following we discuss only the all-seasons averaged indicators without any loss of information.

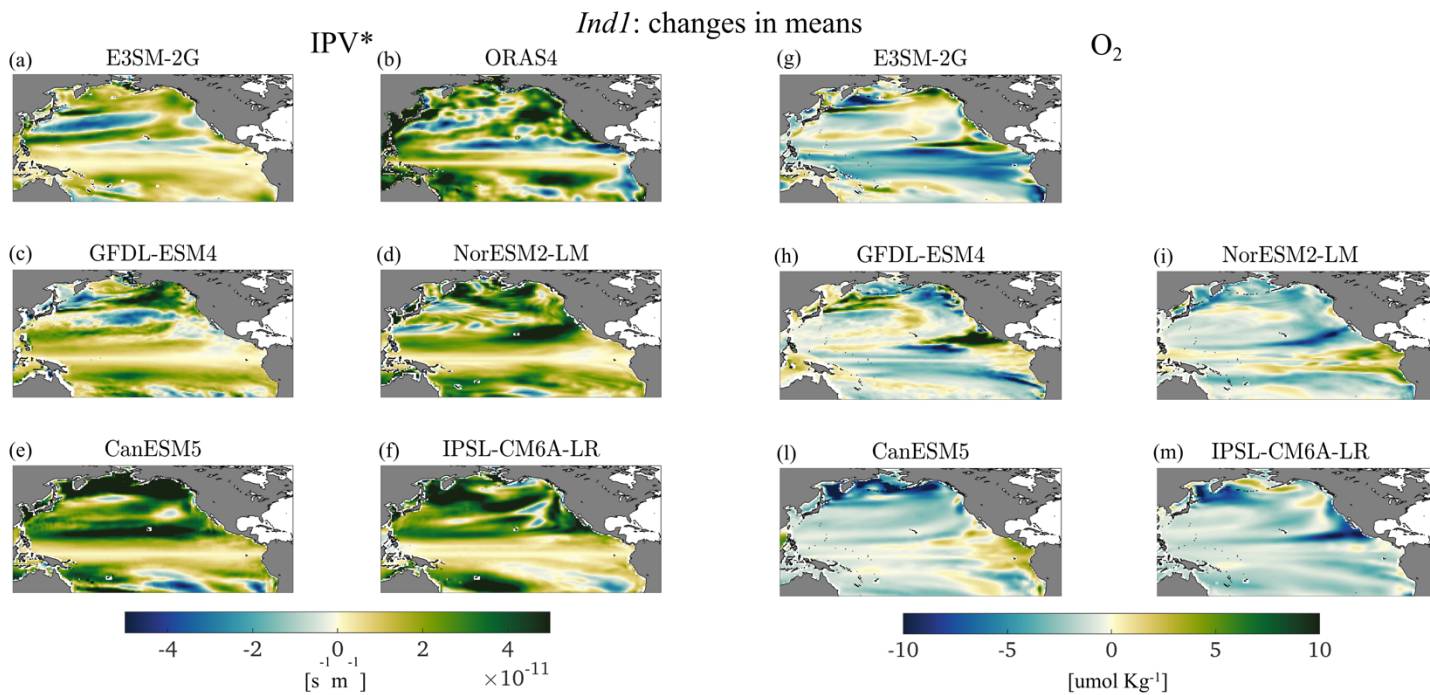



Figure 6. 1950-2014 *Ind1 means* for IPV* (left) and O_2 (right). All indicator maps are obtained by averaging the respective seasonal maps.

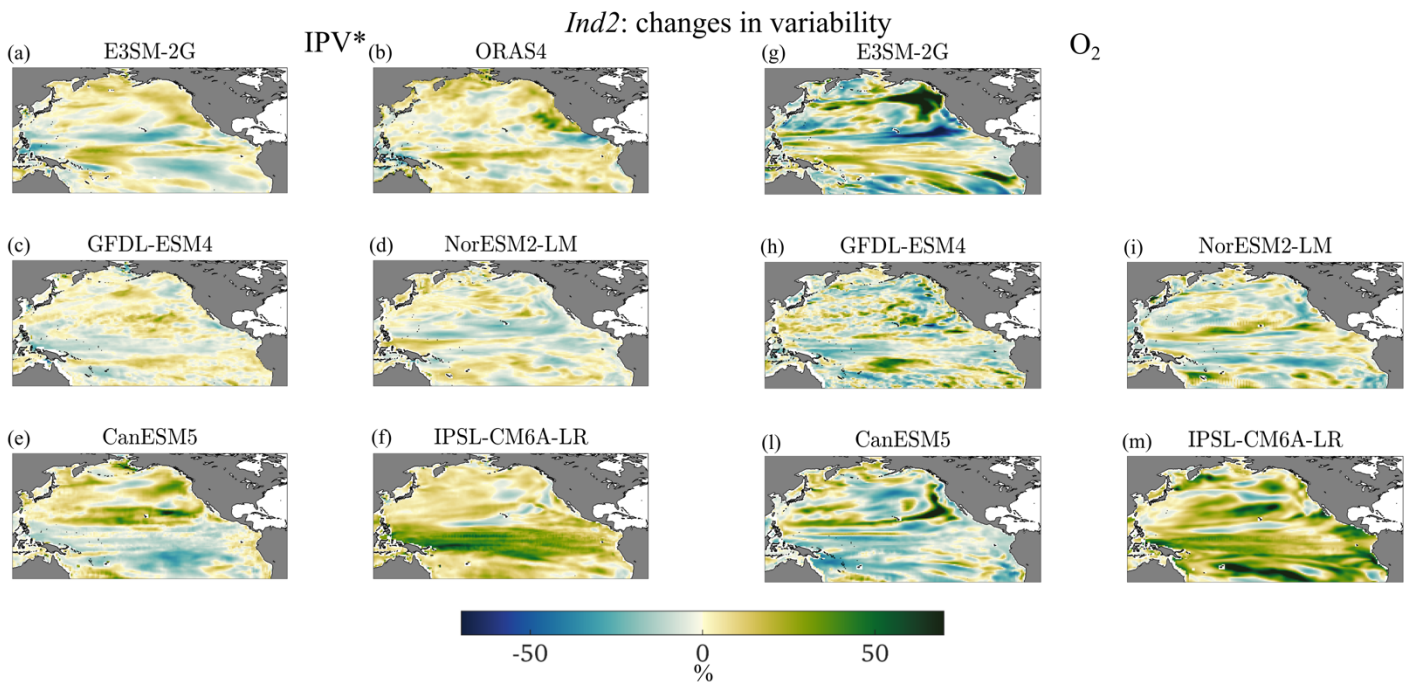
Ind1 means in Fig. 6 shows the changes in the mean fields, which have very similar patterns to the linear trend in both IPV* and O_2 (see Suppl. Fig. S6). By 2015 stratification has increased nearly everywhere in the ESMs, but for the equatorial upwelling region, where is mostly unchanged, and the Kuroshio-Oyashio extension. In ORAS4 there is also a prominent band where stratification decreases between 10° and 20°N extending from the coast of the American continent to 150°W in the second period and in the overall trend. O_2 decreases in most of the north Pacific, especially in the subpolar gyre around the Kamchatka peninsula, and increases in the upwelling areas along the coast of Peru, Central America, and the California Current System. Areas of increasing O_2 are also found in the North Equatorial Current region in the E3SM-2G hindcast, the GFDL-ESM4 and CanESM5 models, the Equatorial upwelling band in NorEMS2-LM, and portions of the subpolar gyre around Alaska in E3SM-2G and IPSL-CM6A-LR.

420

Figure 7. 1950-2014 $Ind2_{variability}$ for IPV* (left) and O_2 (right). The oxygen panel is lacking for ORAS4 as it has no biogeochemistry.

Indicators of change in (seasonal) variability ($Ind2_{variability}$, Fig. 7) show strong differences across models in patterns and, at least for O_2, intensity. Whenever corresponding maps of O_2 and stratification have the same sign and comparable amplitude at corresponding locations, they indicate that increments or decreases in IPV* variability at seasonal scales are associated

425 with corresponding increments in 0-200m O_2 variability. In the hindcast, changes are greater for residual O_2 than stratification. This is verified also in three of the models in the north-eastern extratropics. Among the models, GFDL-ESM4 and NorESM2-LM show patchy changes, both positive and negative, across the domain, with the smallest amplitudes among the datasets considered. CanEMS5 undergoes predominantly positive changes north of the equator in IPV* and negative to the south of it, while the variability in the O_2 field decreases also in the central portion of the subtropical gyre. In IPSL-

430 CM6A-LM the variability increases nearly everywhere in both fields, but especially at the equator and to the south of it in IPV* and more uniformly at all latitudes in O_2.

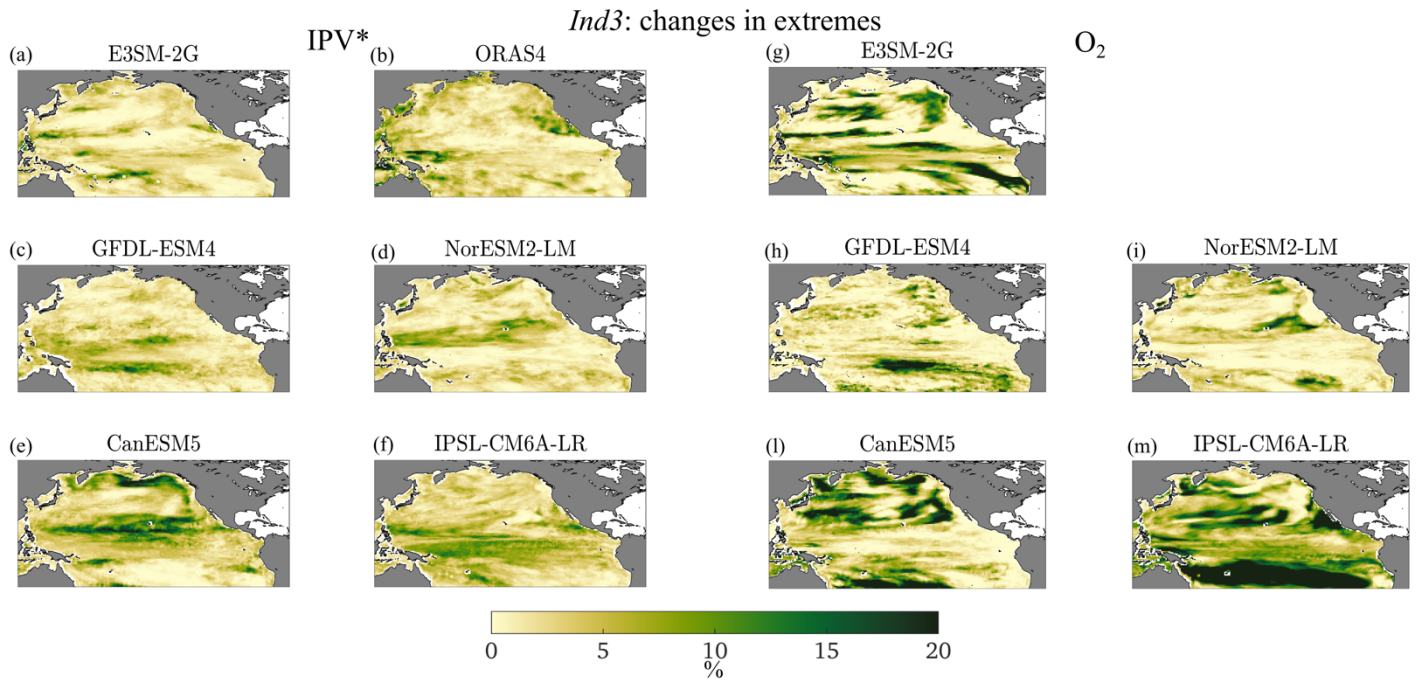



Figure 8. 1950-2014 $Ind3_{extremes}$ for IPV* (left) and O_2 (right). The oxygen panel is lacking for ORAS4 as it has no biogeochemistry.

435 Changes in extremes ($Ind3_{extremes}$) for the O_2 field are stronger than for stratification (Figure 8). Episodes of strong O_2 decrease and stratification increase are more frequent in *Period 2*. For O_2 the equatorial regions to the north and south of the upwelling band emerge as especially impacted in the E3SM-2G hindcast and GFDL-ESM4, while the subtropical gyre displays an increase in extreme events nearly everywhere in CanESM5 and IPSL-CM6A-LM, or at its boundary in E3SM-2G, or in its eastern portion in GFDL-ESM4 and NorEMS2-LM. The subpolar gyre is affected especially in CanESM5 and

440 IPSL-CM6A-LM. Changes in IPV* extremes have less clear latitudinal differences and do not display a robust (across models) intensification at extratropical latitudes. In ORAS4 maxima are found near the California Current System and in the Warm Pool area.

	E3SM-2G	GFDL-ESM4	NorESM2-LM	CanESM5	IPSL-CM6A-LR
Ind1_means (means)	-0.01 (-0.02)	-0.07 (**-0.07**)	**-0.23** (**-0.16**)	**-0.28** (**-0.16**)	**-0.12** (-0.07)
Ind2_variability (variability)	**0.32** (**0.24**)	**0.28** (**0.23**)	0.04 (**0.15**)	**0.33** (**0.21**)	**0.24** (**0.18**)
Ind3_extremes (extremes)	0.01 (-0.08)	**0.29** (0.05)	**0.35** (**0.47**)	0.09 (**0.34**)	-0.03 (**0.17**)

Table 2 1950-2014 Correlation coefficients (c.c) between the corresponding indicator maps for IPV* and O_2. Bold underlined values indicate c.c. ≥ 0.1 that passed the shuffling significance test at 5% level (see Suppl Mat.). Numbers in parentheses reflect c.c. computed north of 20ºN.

Table 2 summarizes the correlation coefficients between the maps of the three indicators for the two fields considered. Coefficients are negative for all models but small for *Ind1$_{means}$*, slightly larger in amplitude and positive for the variability indicator in all cases (*Ind2$_{variability}$*), and very small for *Ind3$_{extremes}$* in the hindcast, CanEMS5 and IPSL-CM6A-LM, while larger in amplitude and positive for GFDL-ESM4, with a strong contribution from the Equatorial region, and especially NorESM-LM, where positive values are also found north of 20ºN.

The resulting hotspot indices (SED), computed separately for the IPV* and the O_2 indicators (see Methods) are reported in Fig. 9. Except for IPSL-CM6A-LM, the hotspots are found outside the equatorial band. Those for O_2 are generally stronger along the eastern part of the subtropical gyres, in the eastern part of PDO region and along the California upwelling system, and the IPV* hotspots are more commonly found over the western parts of the basin and along the southern boundary of the subtropical gyre. This result suggests a longitudinal decoupling between hotspots in O_2 and stratification in at least three of the models and in the hindcast, with NorESM2-LM being the exception due to the simulated superposition of the changes in extremes in the two fields. We computed also the SED for the residual fields, obtaining similar results (Suppl Fig. S12).

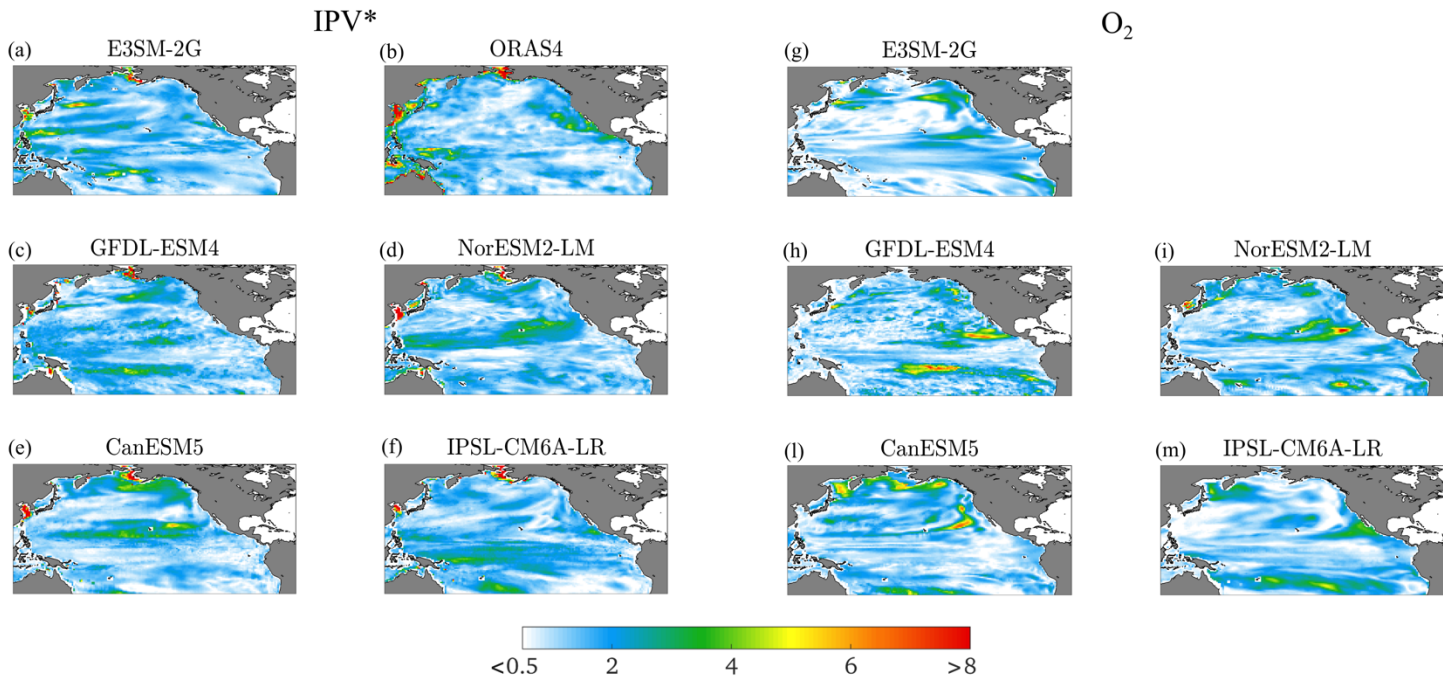



Figure 9. 1950-2014 (1960-2014 for E3SM-2G and ORAS4) SED *index* for IPV* (left) and O_2 (right). The colorscale is realized with rgbmap (Greene, 2023). The oxygen panel is lacking for ORAS4 as it has no biogeochemistry.

The maps of the indicators for the future projections follow in Fig. 10-12, again averaged over seasons for consistency with the historical ones. The associated standard deviations are reported in Suppl. Fig. S13-S15. In the projections, the seasonal differences are slightly stronger compared to the historical period for $Ind1_{means}$ (IPV*) and $Ind3_{extremes}$ (both IPV* and O_2), especially for CanESM5, NorESM2-LM and IPSL-CM6A-LR, in the northern subpolar gyres for $Ind1_{means}$ IPV* (Fig. 10) and along the subtropical and the northern subpolar gyres for $Ind3_{extremes}$ IPV* (Fig. 12). Standard deviations for $Ind3_{extremes}$ O_2 are stronger along the extratropical gyres and weaker in the tropical upwelling region (Suppl. Fig. S15). Areas of higher standard deviations in the projections are, however, associated with much stronger values of $Ind1_{means}$ and $Ind3_{extremes}$ compared to the historical period. In the projections, $Ind1_{means}$ strengthens significantly and is stronger than the actual trend shown in Supp. Fig. S8, indicating an acceleration of the changes in the last portion of the 21st century. This is especially relevant for IPV* north of the Equator. Stratification increases everywhere but for areas in the southern hemisphere which have different extension in the four models and are found in the central and eastern portions of the basin. O_2 decreases everywhere but for small areas around the equatorial upwelling band. The decrease is very strong along the northern

boundary of the Pacific Ocean and, depending on the model, at the subtropical gyre boundary (NorESM2-LM and to a lesser degree CanESM5) and south of the Equator along the coast of Central and South America (IPSL-CM6A-LR).

480

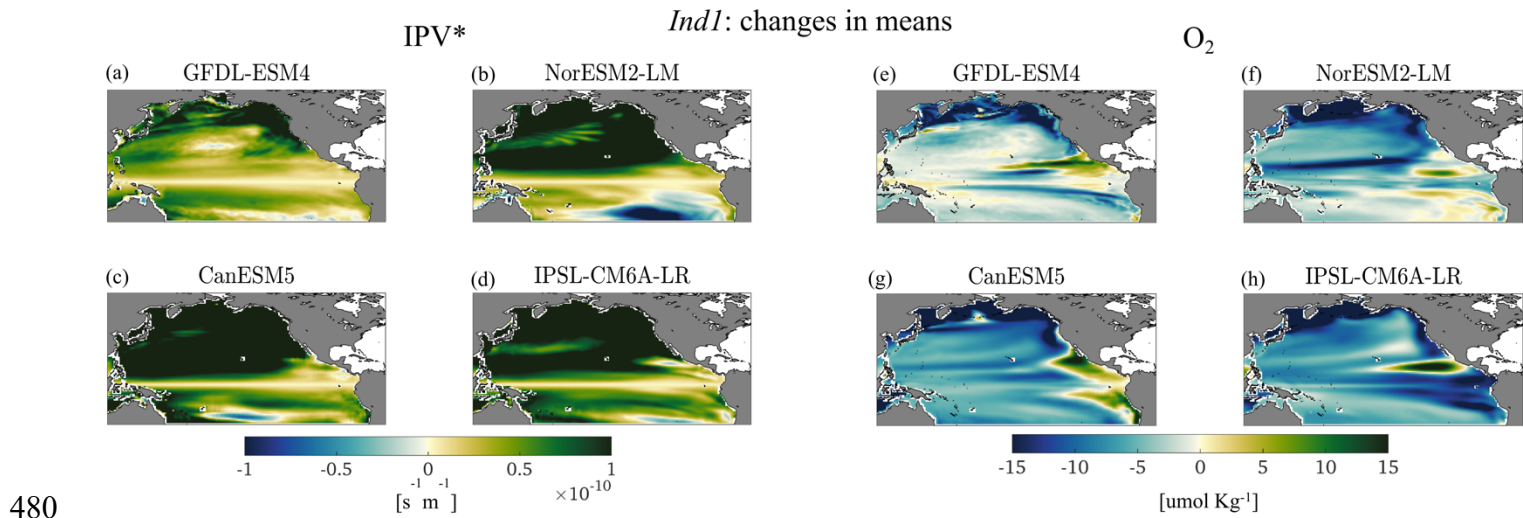



Figure 10. 2036-2100 Ind1$_{means}$ for IPV* (left) and O$_2$ (right).

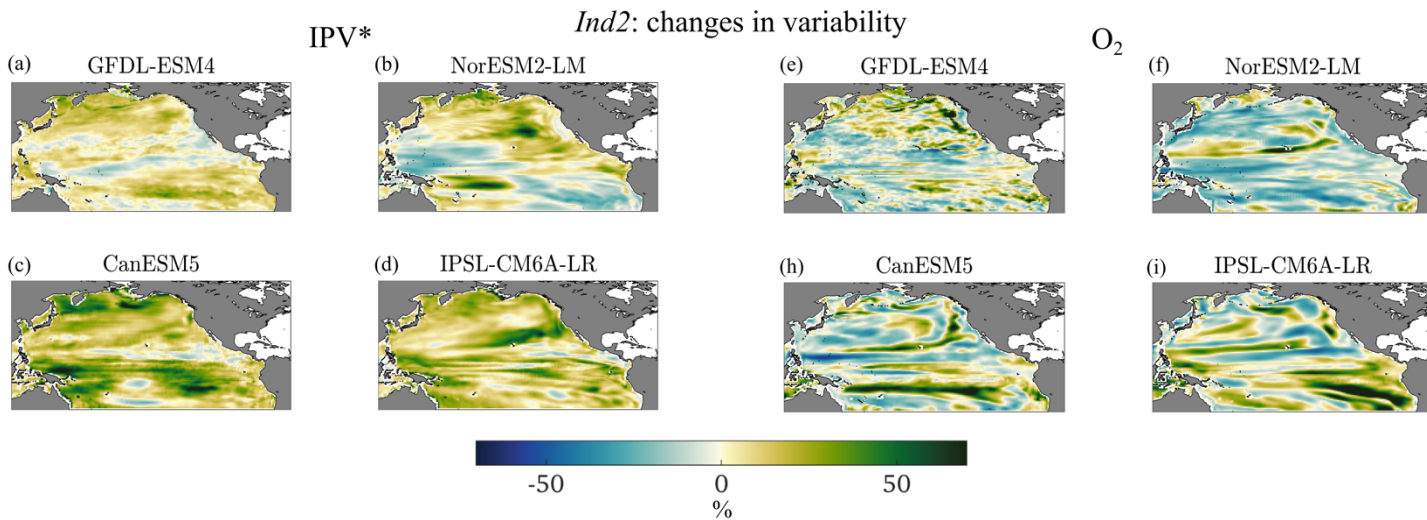



Figure 11. 2036-2100 *Ind2* $_{variability}$ for IPV* (left) and O$_2$ (right).

In terms of variability, very few areas with comparable sign and amplitude (which would indicate comparable increases or

485 decreases) can be seen in Fig. 11 when comparing the two variables. IPV* variability increases almost everywhere in three

of the models, NorESM2-LM being the exception in the Warm Pool and to the south of the Equator in the eastern portion of the basin. O_2 variability increases in patchy areas mostly in the eastern half of the basin in GFDL-ESM4, only along the southern boundary of the subtropical gyre in NorESM2-LM, roughly along the boundaries of the gyres in CanESM5 and along the northern gyre boundary and south of the Equator in IPSL-CM6A-LR. Lastly, extremes (*Ind3 extremes*, Fig. 12) are

490   found to increase nearly everywhere but for the equatorial upwelling area in both variables for CanESM5 and IPSL-CM6A-LR, in the northern hemisphere in NorESM2-LM, in the ENSO region, especially in the Warm Pool, and in GFDL-ESM4 along the northern boundary of the basin for IPV* and along the northern and southern portion of the domain for O_2. Correlations among maps of the two variables are generally very small for all indicators in the projections (Table 3) with |c.c.| < 0.4, except for *Ind1 means* in NorESM2-LM and CanESM5. Finally, we verified the robustness of our results to the

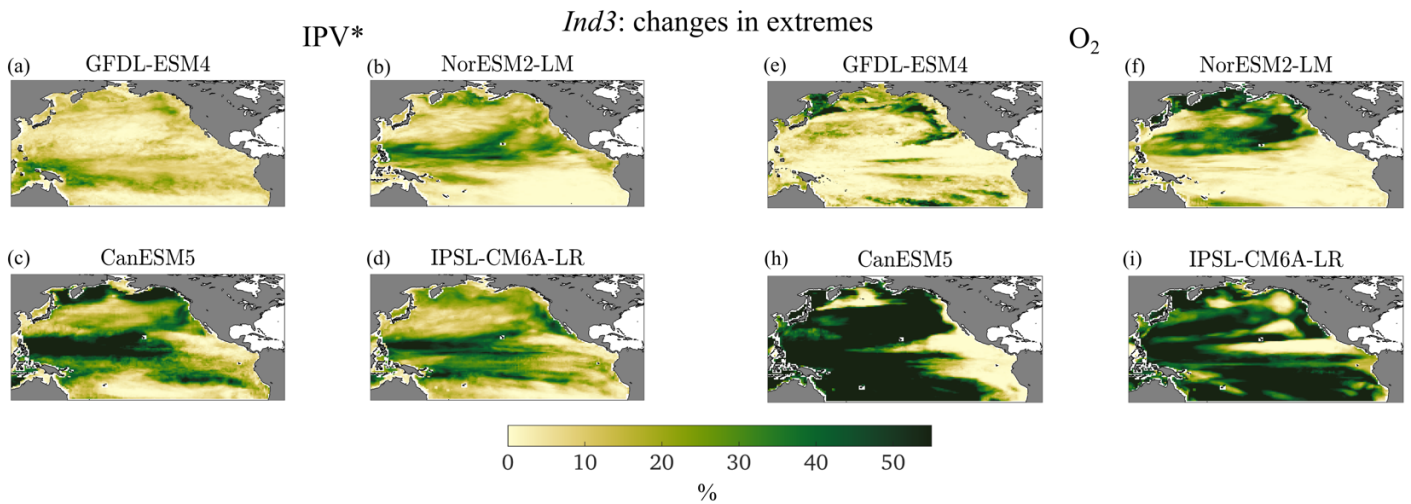
495   choice of the ensemble member, computing the extremes indicators of four randomly-chosen ensemble members of the CanESM5 model for the whole signals of IPV* and O_2 during the historical periods. We found no significant changes in extremes and SED (Suppl Fig S16-17).

500

Figure 12. 2036-2100 *Ind3 extremes* for IPV*$_{res}$ (left) and residual O_{2res} (right). The percentage shown reaches 60% (three times more than during historical, Figure 8).

	GFDL-ESM4	NorESM2-LM	CanESM5	IPSL-CM6A-LR
Ind1	**- 0.21 (-0.15)**	**-0.55 (-0.22)**	**-0.60 (-0.59)**	**-0.32 (-0.28)**
Ind2	**0.30 (0.30)**	**0.14 (0.26)**	**0.21 (0.19)**	0.02 (0.00)
Ind3	-0.01 **(0.56)**	**0.35 (0.40)**	**0.23** (0.01)	0.08 (-0.02)

Table 3. 2036-2100 Correlation coefficients (c.c) between the corresponding indicator maps for IPV* and O_2. Bold underlined values indicate c.c. $\geq$ 0.1 that passed the shuffling significance test at 5% level (see Suppl Mat.). Numbers in parentheses reflects c.c. computed north of 20°N.

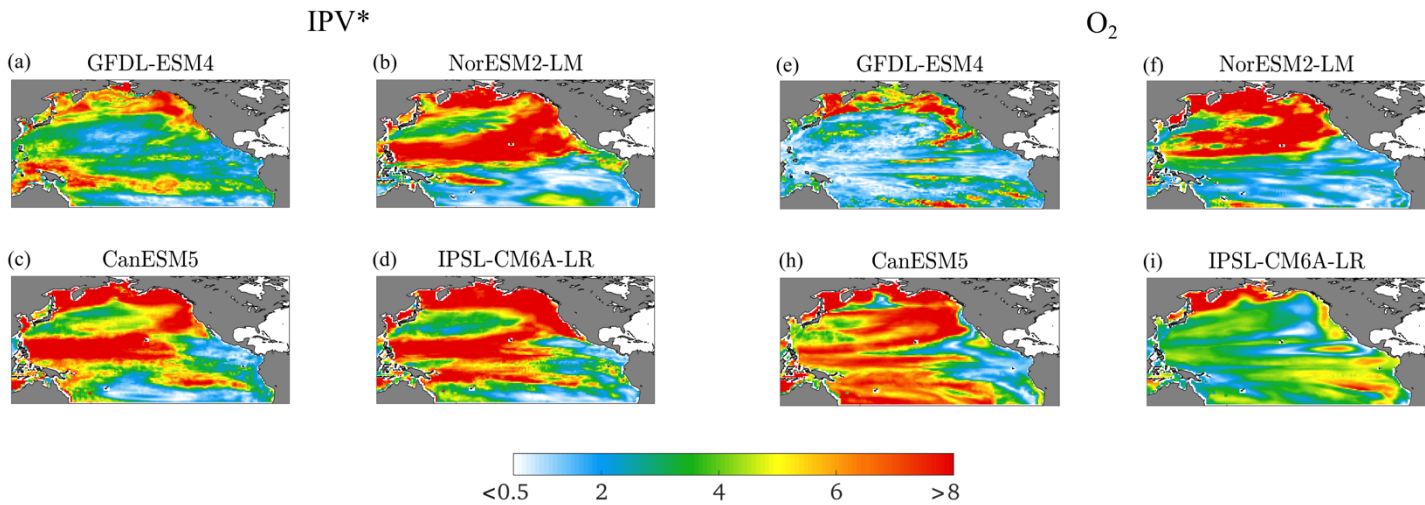



Figure 13. 2036-2100 SED *index* for IPV* (left) and O_2 (right). The colorscale is produced with rgbmap (Greene, 2023).

510

5. Discussion and Conclusions

515 Earth System Models (ESMs) can simulate the Earth's climate and biogeochemical processes with good accuracy, offering valuable insights into the future. Challenges persist, however, in representing reliably ocean biogeochemical dynamics. Biogeochemical processes can involve intricate interactions between multiple components of the earth system. These processes are often nonlinear and their representation and coupling with physical variables are complex and challenging to interpret, therefore requiring advances in diagnosis methods and interpretation. To assess model performance, continued

520 efforts to develop metrics for model evaluation and intercomparison are needed. In this study we presented a set of tools stemming from data-mining techniques that may contribute to this end. These quantitative approaches, together with advances in observation-based gridded products, can better characterize and extract information about linkages between physical and biogeochemical variables. In particular, the availability of biogeochemical data, including dissolved oxygen, remains sparse compared to that of physical data. Limited observational data hinders model validation. Exploiting linkages

525 between physical climate and oceanic O_2 can enhance understanding and predictive skills for biogeochemical tracers. Examples of recently developed tools that take advantage of these linkages can be found in Giglio et al. (2018) and Sharp et al. (2022), who applied machine learning tools to the ARGO-O_2 dataset to generate time-evolving maps of dissolved O_2 concentrations from seasonal to interannual timescales.

The overarching hypothesis in this work was that in the North Pacific the spatial-temporal variability of O_2 reflects that of

530 ocean ventilation, which can be measured by the magnitude of the isopycnic potential vorticity (IPV). A recent study (Ito et al., 2019) found that at subtropical latitudes the variability of winter-time mixed layer depths and the subduction of O_2 are linked to the PDO. Elevated O_2 levels emerge downstream of the deepened winter mixed layer during the positive phase of the PDO. According to the same study, in the equatorial Pacific, the variability of upper ocean O_2 is linked to the stratification and the depth of thermocline, which in turn are modulated by the PDO. There has been a wide range of

535 mechanisms suggested for the connection between upper ocean O_2 and ventilation, many of which can be represented in ESMs. We should note, however, that Ito et al., (2019) also showed that extra-tropical O_2 variability involves multiple types of physical-biogeochemical coupling that may compensate one another. For example, ventilation variability can have opposite imprint on O_2 than water mass shifts depending on the vertical stratification of temperature and O_2. In the subtropical thermocline, both temperature and O_2 decrease with depth, and vertical shifts of water masses generate positive

540 correlation between them. However, a negative relationship is expected between temperature and O_2 under ventilation-driven

variability, as colder conditions are typically associated with stronger ventilation (thus higher O_2). The superposition of these two processes may cause partial compensations and could amplify inter-model differences, especially in O_2.

Our goal was to challenge the overarching hypothesis using four ESMs, a hindcast and reanalysis data. We verified the simplistic view that the spatial-temporal variability of O_2 reflects that of ocean ventilation through the analysis of potential
545   predictability, of the linkages between ventilation and O_2 with the dominant climate modes of the North Pacific, and of the patterns of extreme events in ventilation and O_2. As tracer of physical ventilation, we chose Isopycnic Potential Vorticity or IPV*: a strong ventilation is assumed to generate a negative anomaly in IPV*, which then is advected and mixed through physical transport. Ventilation supplies O_2-rich surface waters into the interior ocean, implying a negative correlation between O_2 and IPV*. First, the information entropy (IE) was adopted to identify the areas where IPV* has a high
550   predictability potential. Predictability was generally high along two stripes enclosing the ENSO pattern and excluding the upwelling cold tongue regions, which were found to correspond to areas where O_2 and IPV* are strongly anticorrelated. The underlying mechanisms are relatively well understood (Ito et al., 2019) and this behavior is robust across all the analyzed datasets and does not change significantly in the future projections in the four ESMs. Therefore, around the Pacific Equator IPV*, which is easily retrievable from temperature and salinity data, has a good predictability potential (higher than in the
555   rest of the basin) and can be used as proxy for O_2. The greater availability of temperature and salinity (and therefore stratification) observations from ARGO floats, reanalyses and modeled fields could be used in conjunction to the fewer co-located observations of O_2 to validate our findings and further extrapolate information about O_2 and its time evolution in these tropical areas.

Secondly, the variability of O_2 and IPV* was examined in relation to large-scale modes of climate variability in the extra-
560   tropical North Pacific. At mid-latitudes, the regional climate variability is PDO-dominated and our analysis shows very low predictability of IPV*, unlike the ENSO-dominated equatorial regions. The low predictability extends to the western boundary current and the Kuroshio-Oyashio extension region. In the extra-tropical North Pacific, the (linear) contribution of the PDO on O_2 and IPV*, and the trends of their residuals have comparable amplitude over the historical period. This is not verified in the future projections, when the trends become increasingly dominant. Pattern correlations in the PDO regression
565   maps (b coefficients) are generally quite small across models.

Thirdly, we evaluated the hotspots of change in IPV* and O_2 in the historical period and in the future projections. Overall, the historical hotspot indices or SED, computed separately for IPV* and O_2, suggest a longitudinal decoupling across the two variables for all datasets except for the NorESM2-LM. In addition, most of hotspots are in the extratropics. O_2 SED tend to be stronger along the eastern portion of the basin, while IPV* hotspots are mostly found over the western side of the basin
570   and along the southern boundary of the subtropical gyre. The intensity of the SED increases over time from the historical period to the end of the 21st century, as to be expected. Larger changes and hotspots are found at the gyre boundaries and in the northern portion of the basin, from the Kamchatka peninsula to the Gulf of Alaska. While O_2 loss is broadly linked to the

strong increase in stratification, there are significant differences across model patterns, pointing to the need of further investigation.

575 The existing uncertainty In the CMIP6 models' representation of oxygen inventory changes limits the information that may be extracted from the projections. A detailed model intercomparison of ocean deoxygenation in CMIP6 models is beyond the scope of our work, and has been addressed by Abe and Minobe (2023). Major sources of uncertainty in the future projections reside, for example, in their ENSO-amplitude representation as detailed in Beobide-Arsuaga et al. (2021), and in uncertainties in the amount of future warming (Tokarska et al., 2020), and consequently in changes in upper-ocean

580 stratification. Compared to the CMIP5 catalog, CMIP6 models tend to warm more, and show a decline of subsurface oxygen ventilation with no consistent decrease of inter-model uncertainties (Kwiatkowski et al. 2020). Here we found that while in some models the relationship between IPV* and O_2 becomes stronger, that is not the case for all, and in particular it is not verified in GFDL-ESM4, which has the highest horizontal resolution and the best representation of the historical period.

A note of caution should be spent on the representation of regional changes and hotspots. The currently available spatial

585 resolution for CMIP6 models does not resolve the fine-scale (mesoscale and finer) physical and biogeochemical processes occurring near the coast. This is especially true in regions of elevated nutrient supply such as along the California Current System and more generally the Eastern Boundary Upwelling Systems (EBUS). Consequently, projected oxygen trends may exhibit variability even within subregions under the same scenario as shown, for EBUS, by Bograd et al. (2023). Analysis at the scales required to capture coastal dynamics, however, would require higher resolution models that will need – if

590 projected into the future – boundary conditions from CMIP6 simulations. CMIP6 models indeed remain the only tool for evaluating changes in large-scale modes of climate variability at interannual to decadal times. While resolution is an important limitation for coastal areas, our main findings remain relevant in the interpretation of the large scale forcing. In particular, the outcome of the hotspots analysis, i.e., that there is a large-scale longitudinal de-coupling between the areas of most prominent changes in IPV* and O_2 despite the PDO imprinting, is unlikely to be influenced by the models' resolution.

595 If this was the case, independently of the PDO or ENSO representation – which may differ in each model – we could reasonably conclude that IPV and O_2, which can be monitored together, for example through the accumulation of ARGO floats equipped with CTD and O_2 sensor, could be used to track the large-scale co-variability of physics and biogeochemistry. We found that the linkages between extra-tropical O_2 and PDO are model-dependent, and the new BGC-Argo array may be able to validate the performance of each model by testing relationships between temperature, IPV and O_2.

600 This could also be done through models with different resolutions, where including smaller-scale physical processes (potentially with a simple biogeochemical module) may better represent physical-biogeochemical coupling and they can be validated against new BGC-Argo observations. Indeed, the variability across the current generation of CMIP models is, for some of our hypotheses, too large to reach any conclusion and the relationship not as strong as hypothesized on the base of the available sparse observations. At the same time, at least for the historical period, our analysis allows for identifying

605 which models may be more realistic in its representation of the large-scale variability of the North Pacific.

Additionally, models and reanalyses or hindcasts such as E3SM-2G allow for testing if there may be predictability notwithstanding their biases, and for the case of the North Pacific, if there is a robust relationship across models between large-scale climate modes of variability, stratification and O_2. The predictability potential extrapolated by global ESMs at the resolution adopted in CMIP6 represents an upper bound of the actual one, but it is useful for identifying when further

610   exploration may be warranted or where such exercise may simply be futile. For example, the information entropy could be evaluated using ARGO datasets opportunely interpolated onto a regular distribution (e.g., Smith and Murphy 2007, Cheng and Zhu 2016, and for BGC-ARGO Turner et al. 2023, Keppler et al. 2023, Sharp et al., 2022). The interpolation alone will increase the uncertainty. In regions where the predictability potential is high, such an exercise is warranted, wherever the potential predictability is low, futile. In reference to the second hypothesis, we found that the PDO modulates IPV* and O_2,

615   but the signal is not robust across models, limiting the possibility to monitor the large-scale evolution of O_2 from temperature and salinity data alone. Relatively high predictability of IPV* and O_2 in the equatorial regions can be exploited as the tropics are generally under-sampled in historical O_2 datasets. While there is relatively weak predictability at mid-latitudes, the sampling density is relatively higher in historical O_2 datasets and there are more direct observations likely due to the proximity to population centers.

620   As a final remark, we carried out our analysis on a subsample of the CMIP6 catalog. However, adding more models will not challenge the main conclusions of this work: 1) There is significant inter-model differences in the representation of climate variability in the North Pacific. This is not just reflected in the representation of the patterns of physical climate variability, but also in the representation of the relationships between physical (IPV*) and biogeochemical (O_2) variables; 2) The IPV* - O_2 relationship in the North Pacific is not robust in terms of patterns and time evolution across the datasets analyzed, but is

625   nonetheless statistically significant under several metrics in the hindcast and in some models, GFDL-ESM4 being the best example and also the model that best represents matches the hindcast and the reanalysis over the historical period.

630

Data availability

The python version of δ-MAPS is available at https://github.com/FabriFalasca/py-dMaps . The code for the Information Entropy computation is available at https://github.com/FabriFalasca/NonLinear_TimeSeries_Analysis . Climate indices used

in this study are from NOAA at https://psl.noaa.gov/data/climateindices/list/). The CMIP6, Earth system model output used in this study is available via the Earth System Grid Federation (https://esgf-node.llnl.gov/search/cmip6/).

Authors contribution

LN performed all analysis, AB and TI conceived the study, TI and YT helped supervise the project, YT led the E3SM-2G development and integration. AB took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank Fabrizio Falasca and Ilias Fountalis, who developed several of the software tools for the data mining component of this project. We acknowledge the Gibbs SeaWater (GSW) Oceanographic Toolbox for Python and the Python package (eofs) for EOF analysis of spatial-temporal data, which we employed. The authors were supported by the Department of Energy, Regional and Global Model Analysis (RGMA) Program, Grant No. 0000253789. YT was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, as part of the Energy Exascale Earth System Model project, and through NERC-NSF grant (C-STREAMS, reference NE/W009579/1).

References

Abe, Y. and Minobe, S.: Comparison of ocean deoxygenation between CMIP models and an observational dataset in the North Pacific from 1958 to 2005. Frontiers in Marine Science. 10. 10.3389/fmars.2023.1161451, 2023

Balmaseda, M. A., Mogensen, K. and Weaver, A. T.: Evaluation of the ECMWF ocean reanalysis system ORAS4. Q.J.R. Meteorol. Soc., doi: 10.1002/qj.2063, 2012.

Beobide-Arsuaga, G., Bayr, T., Reintges, A., and Latif, M.: Uncertainty of ENSO-amplitude projections in CMIP5 and CMIP6 models, Clim. Dynam., 56, 3875–3888, https://doi.org/10.1007/s00382-021-05673-4, 2021.

Bograd, S. J. et al.: Climate change impacts on eastern boundary upwelling systems. *Annu. Rev. Mar. Sci*. **15**, 303–328, 2023.

Box, G. E., Jenkins, G. M. and Reinsel, G. C.: Time series analysis: forecasting and control, Wiley, 2011.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp,
665   L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F.,
Davini, P., Lavergne, C. de, Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C.,
Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B.,
Guez, L., Guilyardi, É., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M.,
Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G.,
670   Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin,
P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R.,
Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and evaluation of the
IPSL-CM6A-LR climate model, J. Adv. Model. Earth Sy., 12, e2019MS002010, https://doi.org/10.1029/2019MS002010,
2020.

675   Breitburg, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, D.,
Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S. W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose,
K. A., Seibel, B. A., Telszewski, M., Yasuhara, M., and Zhang, J.: De- clining oxygen in the global ocean and coastal
waters, Science, 359, https://doi.org/10.1126/science.aam7240, 2018

Brandt, P., Bange, H. W., Banyte, D., Dengler, M., Didwischus, S.-H., Fischer, T., Greatbatch, R. J., Hahn, J., Kanzow, T.,
680   Karstensen, J., Körtzinger, A., Krahmann, G., Schmidtko, S., Stramma, L., Tanhua, T., and Visbeck, M.: On the role of
circulation and mixing in the ventilation of oxygen minimum zones with a focus on the eastern tropical North Atlantic,
Biogeosciences, 12, 489–512, https://doi.org/10.5194/bg-12-489-2015, 2015.

Burrows, S. M., Maltrud, M., Yang, X., Zhu, Q., Jeffery, N., Shi, X., Ricciuto, D., Wang, S., Bisht,G., Tang, J., Wolfe, J.,
Harrop, B. E., Singh, B., Brent,L., Baldwin, S., Zhou,T., Cameron-Smith, P., Keen, N., Collier, N., Xu, M., Hunke, E. C.,
685   Elliott, S. M., Turner, A. K., Li, H., Wang, H., Golaz, J.-C., Bond-Lamberty, B., Hoffman, F. M., Riley, W. J., Thornton, P.
E., Calvin,K., Leung, L. R.: The DOE E3SM v1.1 biogeochemistry configuration: Description and simulated ecosystem-
climate responses to historical changes in forcing. Journal of Advances in Modeling Earth Systems, 12(9), e2019MS001766.
https://doi.org/10.1029/2019ms001766 , 2020.

Cheng, L., and J. Zhu, 2016: Benefits of CMIP5 Multimodel Ensemble in Reconstructing Historical Ocean Subsurface
690   Temperature Variations. J. Climate, 29, 5393–5416, https://doi.org/10.1175/JCLI-D-15-0730.1.

Claret, M., Galbraith, E.D., Palter, J.B., Bianchi, D., Fennel, K., Gilbert, D. and Dunne, P.J.: Rapid coastal deoxygenation
due to ocean circulation shift in the northwest Atlantic. Nature Clim Change 8, 868–872, https://doi.org/10.1038/s41558-

018-0263-1, 2018.

Corso, G., Prado, T., Lima, G., Kurths, J., and Lopes, S.: Quantifying entropy using recurrence matrix microstates. Chaos
695   28:083108. doi: 10.1063/1.5042026 , 2018.

Deutsch C., Ferrel A., Seibel B., Pörtner H.O., Huey R.B.: Ecophysiology. Climate change tightens a metabolic constraint on
marine habitats. Science, 348(6239):1132-5. doi: 10.1126/science.aaa1605. PMID: 26045435. 2015.

Diffenbaugh, N. S., and Giorgi, F.: Climate change hotspots in the CMIP5 global climate model ensemble, Clim. Change,
114 (3–4), 813–822, 2012.

700   Diffenbaugh, N.S., Giorgi, F. and Pal, J.S.: Climate change hotspots in the United States. Geophys Res Lett 35,
doi:10.1029/2008GL035075 , 2008.

Dommenget, D. and Latif, M.: A cautionary note on the interpretation of EOFs. Journal of Climate, 15, 216–225, 2002.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V.,
705   Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin,
R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R.,
Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T.,
Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng,
Y., and Zhao, M.: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and
710   simulation characteristics, J. Adv. Model. Earth Sy., 12, e2019MS002015, https://doi.org/10.1029/2019MS002015, 2020.

Duteil, O., Böning, C.W., Oschlies, A.: Variability in subtropical-tropical cells drives oxygen levels in the tropical Pacific
Ocean. Geophys Res Lett 41(24):8926–8934, 2014.

Eckmann, J. P., Kamphorst, S. O. and Ruelle, D.: Recurrence Plots of Dynmiacal Systems, Europhys. Lett. 4, 973 (1987).
715   Eddebbar, Y. A. et al.: El Niño-like physical and biogeochemical ocean response to tropical eruptions. J. Clim. 32, 2627–
2649, 2019.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled
Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958,
https://doi.org/10.5194/gmd-9-1937-2016, 2016.

720   Falasca, F., Bracco, A., Nenes, A., and Fountalis, I.: Dimensionality Reduction and Network Inference for Climate Data
Using δ-MAPS: Application to the CESM Large Ensemble Sea Surface Temperature, J. Adv. Model. Earth Sy., 11, 1479–
1515, https://doi.org/10.1029/2019MS001654, 2019.

Falasca, F., Crétat, J., Braconnot, P. and Bracco: A. Spatiotemporal complexity and time-dependent networks in sea surface
temperature from mid- to late Holocene, Eur. Phys. J. Plus 135, 392, https://doi.org/10.1140/epjp/s13360-020-00403-x ,

725   2020.

Fountalis, I., Dovrolis, C., Bracco, A., Dilkina, B. and Keilholz, S.: δ-MAPS: from spatio-temporal data to a weighted and lagged network between functional domains. Appl Netw Sci **3**, 21, https://doi.org/10.1007/s41109-018-0078-z , 2018.

Garcia, H. E., Boyer, T. P., Baranova, O. K., Locarnini, R. A., Mishonov, A. V., Grodsky, A., Paver, C.R, Weathers, K.W., Smolyar, I.V., Reagan, J.R., Seidov, D. and Zweng, M.M.: World ocean atlas 2018: Product documentation. A. Mishonov,
730   Technical Editor, 2019.

Giglio, D., Lyubchich, V. and Mazloff, M. R.: Estimating Oxygen in the Southern Ocean Using Argo Temperature and Salinity. *J. Geophys Res. Oceans, 123* (6), 4280−4297. 2018.

Gilpin, W: Deep Reconstruction of Strange Attractors From Time Series. Cambridge, MA: Harvard University, 2020.

Gnanadesikan, A., Dunne, J. P., and John, J.: Understanding why the volume of suboxic waters does not increase over
735   centuries of global warming in an Earth System Model, Biogeosciences, 9, 1159–1172, https://doi.org/10.5194/bg-9-1159-2012, 2012.

Gnanadesikan, A., Bianchi, D., and Pradal, M. A.: Critical role for mesoscale eddy diffusion in supplying oxygen to hypoxic ocean waters, Geophys. Res. Lett., 40, 5194–5198, https://doi.org/10.1002/GRL.50998, 2013.

Gnanadesikan A., Pradal M., Abernathey R.: Isopycnal mixing by mesoscale eddies significantly impacts oceanic
740   anthropogenic carbon uptake. Geophys. Res. Lett. 42, 4249–4255. doi: 10.1002/2015GL064100, 2015.

Golaz, J.C., Van Roekel, L.P., Zheng, X., Roberts, A.F., Wolfe, J.D., Lin, W., Bradley, A.M., Tang, Q.; Maltrud, M.E., Forsyth, R.M. et al.: The DOE E3SM Model Version 2: Overview of the physical model and initial model evaluation. J. Adv. Model. Earth Syst. 14, e2022MS003156, 2022.

Gordon, E. M., Barnes, E. A., and Hurrell, J. W.: Oceanic harbingers of Pacific Decadal Oscillation predictability in CESM2
745   detected by neural networks. Geophysical Research Letters, 48, e2021GL095392. https://doi.org/10.1029/2021GL095392, 2021.

Greene, C.: rgbmap color maps (https://www.mathworks.com/matlabcentral/fileexchange/46874-rgbmap-color-maps), MATLAB Central File Exchange. Retrieved June 27, 2023

Ikuyajolu, O. J., Falasca, F. and Bracco, A.: Information Entropy as Quantifier of Potential Predictability in the Tropical
750   Indo-Pacific Basin. Front. Clim. 3:675840. doi: 10.3389/fclim.2021.675840, 2021.

Ito, T., Minobe, S., Long, M. C., and Deutsch, C.: Upper ocean O_2 trend.: 1958–2015, Geophys. Res. Lett., 44, 4214–4223, https://doi.org/10.1002/2017GL073613, 2017.

Ito, T., Long, M. C., Deutsch, C., Minobe, S., and Sun, D.: Mechanisms of low-frequency oxygen variability in the North Pacific. Global Biogeochemical Cycles, 33(2), 110–124. https://doi.org/10.1029/2018GB005987, 2019.
755   Kucharski, F., Bracco, A., Yoo, J.H. and Molteni, F.: Atlantic forced component of the Indian monsoon interannual variability. Geophys. Res. Lett. 35 (4) https://doi.org/10.1029/2007gl033037 , 2008.

Keppler, L., Landschützer, P., Lauvset, S. K., & Gruber, N. (2023). Recent trends and variability in the oceanic storage of dissolved inorganic carbon. Global Biogeochemical Cycles, 37, e2022GB007677. https://doi.org/10.1029/2022GB007677

760 Li, S., Wu, L., Yang, Y., Geng, T., Cai, W., Gan, B., Chen, Z., Jing, Z., Wang, G. and Ma, X.:The Pacific Decadal Oscillation less predictable under greenhouse warming. Nat. Clim. Chang. 10, 30–34, https://doi.org/10.1038/s41558-019-0663-x, 2020.

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., Dunne, J. P., Gehlen, M., Ilyina, T., John, J. G., Lenton, A., Li, H., Lovenduski, N. S., Orr, J. C., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R.,
765 Stock, C. A., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., and Ziehn, T.: Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections, Biogeosciences, 17, 3439–3470, https://doi.org/10.5194/bg-17-3439-2020, 2020

770 Long, M. C., Moore, J. K., Lindsay, K., Levy, M., Doney, S. C., Luo, J. Y., Krumhardt, K. M., Letscher, R. T., Grover, M., and Sylvester, Z. T.: Simulations With the Marine Biogeochemistry Library (MARBL), J. Adv. Model. Earth Sy., 13, e2021MS002647, https://doi.org/10.1029/2021MS002647, 2021.

Marwan, N., Romano, M. C., Thiel, M., and Kurths, J.: Recurrence plots for the analysis of complex systems. Phys. Rep.
775 438, 237–329. doi: 10.1016/j.physrep.2006.11.001 , 2007.
`

Mantua, N. J., Hare, S. R., Zhang, Y, Wallace, J.M., Francis, R. C.: A Pacific interdecadal climate oscillation with impacts on salmon production Bull. Am. Meteorol. Soc., 78, pp. 1069-1080, 1997.
McKinley, G. A., Follows, M. J., and Marshall, J.: Mechanisms of air-sea CO2 flux variability in the
780 equatorial Pacific and the North Atlantic. Global Biogeochemical Cycles, 18(2), C07S06. https://doi.org/10.1029/2003GB002179, 2004.
Mogensen, K., Alonso Balmaseda, M., Weaver, A.: The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System4. ECMWF Technical Memorandum 668. 59 pages, 2012.
785 Nevison, C., Butler, J. H., and Elkins, J. W.: Global distribution of N2O and the ΔN2O-AOU yield in the subsurface ocean, Global Biogeochem. Cy., 17, 1119, https://doi.org/10.1029/2003GB002068, 2003.
Prado, T., Corso, G., Santos Lima, G., Budzinski, R., Boaretto, B., Ferrari, F., Macau, E.E.N. and Lopes, S.R.: Maximum entropy principle in recurrence plot analysis on stochastic and chaotic systems. Chaos 30:043123. doi: 10.1063/1.5125921, 2020.

790   Redi, M.H.: Oceanic Isopycnal Mixing by Coordinate Rotation. Journal of Physical Oceanography, 12, 1154-1158, 1982.

Ridder, N. N. and England, M. H.: Sensitivity of ocean oxygenation to variations in tropical zonal wind stress magnitude, Global Biogeochem. Cy., 28, 909–926, 2014.

Rudnickas, D. Jr., Palter, J., Hebert, D., Rossby,H.T.: Isopycnal mixing in the North Atlantic oxygen

795   minimum zone revealed by RAFOS floats J. Geophys. Res.: Oceans, 124, 10.1029/2019JC015148, 2019.

Sallée, J. B., Rintoul, S. R. and Wijffels, S. E.: Southern ocean thermocline ventilation. J. Phys. Oceanogr. 40, 509–529, 2010.

Sallée, J.B., Matear, R., Rintoul, S. et al.: Localized subduction of anthropogenic carbon dioxide in the Southern Hemisphere oceans. Nature Geosci 5, 579–584, https://doi.org/10.1038/ngeo1523, 2012.

800   Schulzweida, U.: CDO User Guide (2.1.0). Zenodo. https://doi.org/10.5281/zenodo.7112925, 2022.

Schmidtko, S., Stramma, L. And Visbeck, M.: Decline in global oceanic oxygen content during the past five decades. Nature 542, 335–339, https://doi.org/10.1038/nature21399 , 2017.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J.,

805   Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geosci. Model Dev., 13, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., and Dunne, J. P.: GOBAI-O$_2$:

810   temporally and spatially resolved fields of ocean interior dissolved oxygen over nearly two decades, Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-2022-308, in review, 2022.

Smith, D. M. and Murphy, J. M.: An objective ocean temperature and salinity analysis using covariances from a global climate model, J. Geophys. Res.-Oceans, 112, C02022, https://doi.org/10.1029/2005JC003172, 2007

815

Stock, C. A., Dunne, J. P., Fan, S., Ginoux, P., John, J., Krasting, J. P., et al.: Ocean biogeochemistry in GFDL's Earth System Model 4.1 and its response to increasing atmospheric CO2. Journal of Advances in Modeling Earth Systems 12, e2019MS002043, 2020.

820 Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geosci. Model Dev., 12,

4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

825   Takano, Y., Maltrud, M., Sinha, A., Jeffery, N., Smith, K., Conlon, L., Wolfe, J., and Petersen, M.: Global Ocean Carbon Cycle Simulations with the 2 E3SM version 2 (E3SMv2), https://doi.org/10.5281/zenodo.10093369, Zenodo, 2023

Talley, L. D., Pickard, G. L., Emery, W. J. , Swift, J. H.: Descriptive Physical Oceanography: An Introduction, Academic Press, London, 2011.

830   Tjiputra, J. F., Schwinger, J., Bentsen, M., Morée, A. L., Gao, S., Bethke, I., Heinze, C., Goris, N., Gupta, A., He, Y.-C., Olivié, D., Seland, Ø., and Schulz, M.: Ocean biogeochemistry in the Norwegian Earth System Model version 2 (NorESM2), Geosci. Model Dev., 13, 2393–2431, https://doi.org/10.5194/gmd-13-2393-2020, 2020.

Turco, M., Palazzi, E., Hardenberg, J. and Provenzale, A.: Observed climate change hotspots Geophys. Res. Lett. **42** 3521–8, 2015.

835   Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., and Knutti, R.: Past warming trend constrains future warming in CMIP6 models, Sci. Adv., 6, eaaz9549, https://doi.org/10.1126/sciadv.aaz9549, 2020.

Tsujino, H., Urakawa, L. S., Griffies, S. M., Danabasoglu, G., Adcroft, A. J., Amaral, A. E., Arsouze, T., Bentsen, M., Bernardello, R., Böning, C. W., Bozec, A., Chassignet, E. P., Danilov, S., Dussin, R., Exarchou, E., Fogli, P. G., Fox-

840   Kemper, B., Guo, C., Ilicak, M., Iovino, D., Kim, W. M., Koldunov, N., Lapin, V., Li, Y., Lin, P., Lindsay, K., Liu, H., Long, M. C., Komuro, Y., Marsland, S. J., Masina, S., Nummelin, A., Rieck, J. K., Ruprich-Robert, Y., Scheinert, M., Sicardi, V., Sidorenko, D., Suzuki, T., Tatebe, H., Wang, Q., Yeager, S. G., and Yu, Z.: Evaluation of global ocean–sea-ice model simulations based on the experimental protocols of the Ocean Model Intercomparison Project phase 2 (OMIP-2), Geosci. Model Dev., 13, 3643–3708, https://doi.org/10.5194/gmd-13-3643-2020, 2020.

845

Turner, K. E., Smith, D. M., Katavouta, A., and Williams, R. G.: Reconstructing ocean carbon storage with CMIP6 Earth system models and synthetic Argo observations, Biogeosciences, 20, 1671–1690, https://doi.org/10.5194/bg-20-1671-2023, 2023.

850   Webber, C. L., Jr. and Zbilut, J. P.: Dynamical assessment of physiological systems and states using recurrence plot strategies, J. Appl. Physiol. 76, 965, 1994.

Williams, J. W., Jackson, S.T. and Kutzbach, J.E.: Projected distributions of novel and disappearing climates by 2100AD, Proc. Natl. Acad. Sci. U. S. A., 104, 5738– 5742, 2007.

Wright, D. G. (1997). An equation of state for use in ocean models: Eckart's formula revisited. Journal of Atmospheric and

855   Oceanic Technology, 14(3), 735–740. https://doi.org/10.1175/1520-0426(1997)014<0735:AEOSFU>2.0.CO;2

Yang S., Gruber N., Long M. C., Vogt M.: ENSO-driven variability of denitrification. Global Biogeochemical Cycles, 31

(10), 1470–1487. doi: 10.1002/2016gb005596, 2017.

Zbilut, J. P. and Webber, C. L., Jr.: Embeddings and delays as derived from quantification of recurrence plots, Phys. Lett. A 171, 199, 1992.

860