

1 Emergence of multiple ocean ecosystem **drivers** in a large 2 ensemble suite with an Earth System Model

3
4 **K. B. Rodgers¹, J. Lin² and T. L. Frölicher³**

5 [1]{Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, USA}

6 [2]{Princeton University, Princeton, USA}

7 [3]{Environmental Physics, Institute of Biogeochemistry and for Pollutant Dynamics, ETH
8 Zürich, Switzerland}

9
10 Correspondence to: K. B. Rodgers (krodgers@princeton.edu)

11 12 **Abstract**

13 Marine ecosystems are increasingly stressed by human-induced changes. These ocean **drivers** –
14 including warming, acidification, deoxygenation and perturbations to biological productivity -
15 can co-occur in space and time, but detecting their trends is complicated by the presence of noise
16 associated with natural variability in the climate system. Here we use Large Initial-Condition
17 Ensemble Simulations with a comprehensive Earth System Model under a historical/RCP8.5
18 pathway over 1950-2100 to consider emergence characteristics for the four individual and
19 combined drivers. Using a one-standard deviation (67% confidence) threshold of signal-to-noise
20 to define emergence with a 30 yr trend window, we show that ocean acidification emerges much
21 earlier than other drivers, namely during the 20th century over most of the global ocean. For
22 biological productivity, the anthropogenic signal does not emerge from the noise over most of the
23 global ocean before the end of the 21st century. The early emergence pattern for sea surface
24 temperature in low latitudes is reversed from that of subsurface oxygen inventories, where
25 emergence occurs earlier in the Southern Ocean. For the combined multiple-driver field, 41% of
26 the global ocean exhibits emergence for the 2005-2014 period, and 63% for the 2075-2084
27 period. The combined multiple-driver field reveals emergence patterns by the end of this century
28 that are relatively high over much of the Southern Ocean, North Pacific, and Atlantic, but

1 relatively low over the tropics and the South Pacific. For the case of two drivers, the tropics
2 including habitats of coral reefs emerges earliest, with this driven by the joint effects of
3 acidification and warming. It is precisely in the regions with pronounced emergence
4 characteristics where marine ecosystems are expected to be pushed outside of their comfort
5 zone determined by the degree of natural background variability to which they are adapted. The
6 results here also have implications for optimization of the ocean observing system.

7

8 **1 Introduction**

9 An important priority in climate research is to understand the potential vulnerabilities of marine
10 ecosystems in the face of anthropogenic climate change (e.g. Doney et al., 2012). Over the last
11 decade, multiple drivers of marine ecosystems such as ocean warming, ocean acidification,
12 nutrient stress and low oxygen levels have been identified to be among those of greatest concern
13 (e.g. Gruber, 2011; Hall et al., 2013). We have chosen to use “drivers” rather than “stressors” as
14 some drivers (for example temperature) can be beneficial to some organisms or processes. On a
15 global scale, the development of these drivers is largely a consequence of the increase in
16 atmospheric CO₂ and the associated climate change. The oceanic response to these changes,
17 namely the oceanic uptake of excess heat and anthropogenic CO₂ causes ocean warming and
18 ocean acidification, i.e. a decrease in both oceanic pH and in the saturation state of seawater with
19 regard to mineral calcium carbonate (Doney et al. 2009). The warming of the ocean tends to
20 stratify the upper ocean (Sarmiento et al. 1998), leading to a reduced supply of nutrients to the
21 euphotic zone (Bopp et al. 2001, Steinacher et al. 2010), but also to a reduced resupply of oxygen
22 to the ocean’s interior (Frölicher et al., 2009; Keeling et al., 2010), causing a loss of oxygen
23 there. The magnitude of these global drivers will likely continue to grow, given current trends in
24 fossil fuel CO₂ emissions and the strong inertia within the global community with respect to
25 efforts to decarbonize (Friedlingstein et al., 2014).

26

27 The detection of secular trends in driver fields on regional- to global-scales is complicated by the
28 presence of natural variability in the climate system, as has been shown for dissolved oxygen by
29 Frölicher et al. (2009). The presence of background natural variability motivates the introduction

1 of the concept of emergence to identify where the signal (the secular trend) becomes larger than
2 the noise (the background natural variability). Identifying and understanding when and where the
3 secular trends in ocean drivers emerge above the noise is important for two reasons. The first
4 reason is that emergence characterizes when the secular trend becomes evident or perceptible for
5 the local marine species relative to the background variability to which they have adapted. Here
6 emergence becomes a measure of perceptible changes for the ecosystems. The second reason is
7 that understanding of the emergence of multiple drivers will be important for optimizing the
8 design of the ocean observing system. Inferring trends in drivers from Repeat Hydrography is
9 complicated by natural variability in the ocean (Rodgers et al., 2009), and natural variability can
10 also complicate trend detection using time series data (Henson et al., 2014).

11
12 Previous studies exploring the concept of emergence have largely focused on physical state
13 variables of the atmosphere and ocean, such as temperature, precipitation and sea level (e.g.
14 Diffenbaugh and Scherer, 2011; Hawkins and Sutton, 2012; Mahlstein et al., 2012; Mora et al.,
15 2013; Lyu et al., 2014). Mora et al. (2013), for example emphasized that the tropics, which hold
16 the worlds greatest diversity of marine species, will exhibit emergence in ocean warming ten
17 years earlier than any of the other global ocean regions. Far less attention has been devoted to
18 date to signal-to-noise ratios in ocean biogeochemistry, a notable exception being the study of
19 Keller et al. (2014).

20
21 Here we introduce a new suite of Large Initial-Condition Ensemble Simulations using a
22 comprehensive Earth System Model to understand the local emergence characteristics of the
23 ocean ecosystem drivers sea surface temperature (SST), sea water saturation state with respect to
24 aragonite (Ω_{arag}), a mineral phase of calcium carbonate, oxygen levels (O_2), and net primary
25 productivity (NPP) over an interdecadal (30 yr) timescale. The ocean state as expressed in SST
26 exhibits pronounced decadal-to-interdecadal variability (Zhang et al., 1997), and variations on
27 this timescale are well documented for oxygen (Emerson et al., 2004; Mecking et al., 2008;
28 Kouketsu et al., 2010; Takatani et al., 2012). This has also been considered for the case of
29 phytoplankton in the study of Martinez et al. (2009).

1 In opting to use a suite of Large Initial-Condition Ensemble Simulations, we emphasize in
2 particular the uncertainty in estimates of emergence due to natural variability inherent in the
3 climate system. Deconvolving the signature of the forced response from the background natural
4 variability with one coupled model requires Large Initial-Condition Ensemble Simulations. Only
5 with a sufficiently large number of ensemble members can the effects of natural variability be
6 removed by averaging over the ensemble members. With a single model run of a coupled climate
7 model one is forced to estimate noise either from simulated pre-industrial variability or through
8 high-pass filtering of a scenario run (e.g. Deser et al., 2014). The problem is that the forced
9 response is in general imbedded in a stochastic dynamical system that exhibits variability on all
10 timescales. Additionally, the amplitude of major modes of variability such as El Nino-Southern
11 Oscillation (ENSO) is not stationary in their amplitude over climate change timescales
12 (Timmermann et al., 1999). The ensemble approach to coupled modeling thereby offers an
13 important opportunity when applied to the case of ecosystem drivers (Frölicher et al., 2009).

14

15 **2 Methods**

16 **2.1 Model and simulations**

17 We conducted 30 ensemble simulations over the 1950-2100 period following historical and
18 RCP8.5 concentration pathways (van Vuuren et al., 2011). All 30 ensemble members are run
19 with the same coupled Earth System Model developed at the Geophysical Fluid Dynamics
20 Laboratory (Dunne et al., 2012; Dunne et al., 2013): GFDL's ESM2M. The physical state model
21 underlying ESM2M is the updated version of the coupled model CM2.1 (Delworth et al., 2006),
22 consisting of the 1-degree version of the MOM4p1 ocean model (Griffies, 2009) coupled to an
23 approximately 2-degree configuration of the AM2 atmospheric model (Anderson et al., 2004).
24 The ocean biogeochemical model is Tracers of Ocean Phytoplankton and Allometric
25 Zooplankton code version 2 including 30 tracers to represent cycles of carbon, oxygen, and the
26 major macronutrients and iron (Dunne et al., 2010).

27

1 The initial conditions for the 30 ensemble members for January 1st 1950 differed in the initial
2 state of the atmosphere/land/ocean/sea ice components of the Earth System Model. This was
3 accomplished by using model state snapshots for the ends of days 1-29 in January 1950 as the
4 initial model states for January 1st 1950 for each of the ensemble members 2-30. As has been
5 shown by Wittenberg et al. (2014) using significantly smaller initial perturbations to the ocean
6 component only with nearly the same underlying physical coupled model, our initial condition
7 perturbations lead to a randomization of the ENSO state amongst the individual ensemble
8 members within five years. Given that decadal modulations of ENSO are the most pronounced
9 driver of decadal physical variability in this coupled model, decadal variability will be
10 randomized amongst the individual ensemble members.

11
12 The four drivers considered in this study are (i) surface Ω_{arag} , (ii) SST, (iii) subsurface O_2
13 vertically integrated from 100m to 600m, and (iv) NPP vertically integrated over the top 100m.
14 Our focus on subsurface O_2 concentrations is intended to characterize regimes ranging from
15 oxygen minimum zones to the main thermocline of polar and circumpolar regions.

16

17 **2.2 Confidence intervals and detection of time of emergence**

18 In order to quantify emergence characteristics, it is necessary to specify a timescale over which
19 trends are calculated. We use decadal trends considered over 30 yr intervals on a gridpoint-by-
20 gridpoint basis to quantify signal-to-noise ratios for each of the four drivers. The signal is the
21 trend obtained using the ensemble-mean, and the noise is the standard deviation of the 30 yr
22 trends for the individual ensemble members. The signal-to-noise ratio is thereby calculated as the
23 ratio of these two terms, and thereby associated with this specific timescale. The choice of a 30
24 yr trend window is motivated by the approximate length of relatively continuous elements of the
25 global observing system and by the timescale of important natural variability events such as the
26 Pacific Decadal Oscillation. Given that our model runs span 1950-2100, and our choice of 30 yr
27 trends, signal-to-noise can effectively be calculated over each year spanning the period 1965-
28 2085.

1

2 Central to our analysis is the model-derived quantification of confidence intervals for trend
3 detection. We assume that the 30 yr trends for the 30 individual ensemble members are normally
4 distributed. For a given time-window (30 yr), the signal-to-noise ratio (SNR) is calculated using
5 the relationship between the ensemble mean TREND and the standard deviation (σ_{noise}) of the
6 trends of the various ensemble members, which we denote as NOISE:

7

$$8 \text{ SNR} = \text{TREND} / \text{NOISE}. \quad (1)$$

9

10 Here we focus on a threshold of one for SNR, representing a confidence interval of 67% for
11 emergence. In other words, the threshold of one is used to characterize when the “signal” of
12 anthropogenic climate change rises above the “noise” of natural background variability.
13 However, the sensitivity to a choice of two (95% **confidence**) for the threshold will also be
14 considered. We choose two time intervals over which we consider average confidence intervals:
15 the first is for the most recent decade 2005-2014, and the second is the decade 2075-2084
16 towards the end of the 21st century. The confidence intervals are calculated using 30 yr trends
17 year-by-year over each of the 10 yr intervals before considering 10 yr averages over the
18 respective intervals. An averaging interval of 10 yr was sufficient to remove noise present in
19 analysis for individual years in the confidence intervals, presumably due to the fact that even with
20 30 ensemble members the modes of variability aren’t sufficiently randomized.

21

22 **Additionally, we also consider the sensitivity of the confidence intervals to the choice of the**
23 **width of the trend window. To that end, confidence intervals are also considered for the case of a**
24 **10 yr window. From the spectral SST characteristic of the underlying coupled**
25 **(atmosphere/ocean) model, it has been shown earlier studies that SST variability is more**
26 **pronounced over 10 yr timescales than over 30 yr timescales (Fig. 2 in Wittenberg (2009) and**
27 **Fig. 7 in Dunne et al. (2012)). Thus one may well expect that the signal-to-noise ratio**
28 **characteristics for the ecosystem drivers reflect these underlying dynamical drivers of variability,**
29 **at least in the equatorial Pacific. Thus our sensitivity analysis is intended to offer insight into**

1 both of our primary interests described in the Introduction, namely identification of perceptible
2 changes for ecosystems and optimization of the observing system.

3
4 Our analysis also includes diagnostics of Time of Emergence (ToE). In contrast to the
5 confidence interval diagnostic for fixed time intervals, ToE requires that we first specify a fixed
6 threshold (here we choose 67%) for confidence intervals and then calculate the time (year) at
7 which that is satisfied on a gridpoint-by-gridpoint basis for each driver. For consistency, a 30 yr
8 trend window is used here as well in the calculation of ToE.

9 10 **3 Results**

11 **3.1 Temporal hierarchy of global and regional emergence of individual drivers**

12 We start the analysis with the temporal hierarchy of emergence for the globe as well as for
13 various regions, such as the Southern Ocean (90-45°S), the southern subtropics (45-15°S), the
14 equatorial band (15°S-15°N), the northern subtropics (15-45°N), and the subpolar and polar
15 Northern Hemisphere (45-90°N) (**Fig. 1**). For each driver, the signal-to-noise ratio was calculated
16 gridpoint-by-gridpoint, and then the area-weighted mean of this quantity were considered by
17 region year-by-year (solid lines in **Fig. 1**). There is a distinct hierarchy in the emergence of the
18 drivers on a global scale (**Fig. 1a**), with Ω_{arag} (red line in **Fig. 1a**) already having risen above the
19 one-standard deviation (67%, black solid vertical axis in **Fig.1a**) level by the beginning of the
20 period considered (by 1965). This is followed by SST (green line in **Fig. 1a**), which globally
21 emerges from the 67% confidence level by the year 2000. O₂ inventories rising above the 67%
22 confidence interval by approximately 2060 (blue line in **Fig. 1a**), whereas NPP remains below
23 the 67% confidence level over the entire time period (purple line in **Fig. 1a**).

24
25 The regional behavior of the four drivers is shown in **Fig. 1b-f**. Overall, the hierarchy found in
26 the global analysis tends to be reproduced on a regional scale but with a few notable exceptions.
27 First, the Southern Ocean differs from the other regions in that the O₂ inventories (blue solid line
28 in **Fig. 1b**) are more detectable than SST (green solid line in **Fig. 1b**) over the duration.

1 Additionally, the subtropical SST tends to emerge from the 67% confidence interval by 1990-
2 2010 (green solid lines in **Fig. 1c-e**), with this not being the case for the Southern Ocean and the
3 Northern subpolar/Arctic regions (green solid lines in **Fig. 1b,f**). The commonality is that Ω_{arag} is
4 above the 67% confidence interval in all regions for the duration of the analysis period, whereas
5 the NPP signal remains below the 67% confidence level for the **duration of the analysis period** in
6 all regions, with the exception of the southern subtropics for the period post-2070.

7

8 **3.2 Local emergence confidence intervals for individual drivers**

9 We next consider the spatial distributions of the confidence level that the signal (ensemble mean)
10 in the four drivers has emerged from the natural variability (standard deviation among the 30
11 ensemble members) for two different time periods: 2005-2014 (**Fig. 2**) and 2075-2084 (**Fig. 3**).
12 For Ω_{arag} it is found that 99.8% of the global domain exhibits the signal emerging above the noise
13 with more than 67% certainty (**Fig. 2a**). This indicates that a **global observing platform** of 30 yr
14 duration **would be** able to detect trends nearly everywhere. The reasons for the early emergence is
15 that changes in surface Ω_{arag} are mainly dictated by the increase in surface DIC, which closely
16 follows atmospheric pCO₂ trends (e.g. Keller et al. 2014). For SST (**Fig. 2b**), 74.8% of the global
17 surface ocean has emerged with 67% certainty from the noise. Especially the lower latitudes tend
18 to have emerged with high confidence by 2005-2014, whereas the Southern Ocean south of 45°S
19 and the northern subpolar and Arctic regions reveals only few confidence that they have emerged
20 by 2005-2014. For O₂ inventories, 22.6% of the global ocean has emerged with 67% certainty
21 from the noise (**Fig. 2c**). For NPP, only 14.8% of the global ocean has emerged with 67%
22 certainty over 2005-2014 (**Fig. 2d**). Interestingly, the pattern for the O₂ inventories shown in
23 **Fig. 2c** reveals emergence over important parts of the Pacific and Atlantic sectors of the Southern
24 Ocean (defined as the region to the south of 45°S), in contrast to what is seen in SST. The lower
25 latitudes reveal relatively little emergence by 2005-2014, but there are small areas within the
26 subtropics of the North Pacific with emergence. For the case of NPP (**Fig. 2d**), there is very little
27 evidence of emergence by 2005-2014 over most of the global domain, although the Equatorial
28 Pacific Cold Tongue regions, parts of the Equatorial Atlantic, and the Agulhas regions show

1 marginal emergence. This indicates that even with complete 30 yr time series for global NPP, it
2 would not be possible in most regions to identify a significant secular trend.

3
4 Next we consider in **Fig. 3** the confidence intervals calculated over 2075-2084. The patterns
5 shown in **Fig. 3** largely reinforce what was seen in **Fig. 2**, although clearly the emergent
6 structures have expanded. Not surprisingly, Ω_{arag} (**Fig. 3a**) now stands emerged over 100% of
7 the global surface area with 67% certainty at the end of the 21st century, consistent with what
8 was shown over 2005-2014. For SST (**Fig. 3b**), 90.9% of the global ocean has emerged with
9 67% certainty, clearly having expanded beyond their limits from 2005-2014. This includes
10 expansion into the North Atlantic, and the high latitudes over the Southern Ocean. For O_2
11 inventories (**Fig. 3c**), with 42.3% of the globe emerged with 67% certainty, the Southern Ocean
12 increasingly emerges towards the end of the century, and the North Pacific subtropics have
13 emerged with high confidence. For the case of NPP (**Fig. 3d**), with 23.7% of the globe emerged
14 with 67% certainty, there continues to be only weak emergence over most of the globe, the
15 exceptions being a coherent structure spanning the southern subtropics and subtropical front
16 regions of the South Indian Ocean, the equatorial Pacific upwelling region, and then patchy
17 regions over the Southern Ocean the eastern mid-latitude and subpolar North Atlantic.

18
19 Taken together, the results here indicate that the four drivers are not advancing in unison with the
20 same patterns and rates of their detectability over the globe. Ω_{arag} emerges first, consistent with
21 what has been found in previous modeling studies (Friedrich et al., 2012), NPP emerges last, and
22 then O_2 inventories and SST have their own contrasting patterns of emergence. Stated
23 differently, random uncertainty associated with natural decadal variability poses significantly
24 more important challenges for NPP than for Ω_{arag} , with O_2 inventories and SST falling between
25 NPP and Ω_{arag} .

26

27

3.3 Local emergence confidence intervals for the multi-driver mean

We now consider an average across the four ecosystem drivers of the confidence intervals for the same time intervals 2005-2014 (**Fig. 4a**) and 2075-2084 (**Fig. 4b**). To facilitate presentation and interpretation, we calculated averages within 14 biomes as defined by Henson et al. (2010). The biome definition separates the regions where phytoplankton growth is seasonally light limited (for mid to high latitudes), regions where the ocean is gaining heat (equatorial regions) and oligotrophic regions.

For the 2005-2014 interval (**Fig. 4a**), the multi-driver confidence intervals have already risen above the threshold of one standard deviation (67% confidence interval) in the Equatorial Atlantic, the South Atlantic, and the Arabian Sea. The Indian and Pacific sectors of the Southern Ocean weigh in at 60% confidence intervals, thereby below one standard deviation. With the exception of the South Atlantic, subpolar regions are minimum regions for confidence intervals. For the case of the Equatorial Atlantic, where in fact O₂ inventories are increasing (Gnanadesikan et al., 2012; Cocco et al., 2013) (**Fig. A1e**), it is important to acknowledge that emergent O₂ inventories are thereby not to be understood as a driver. Interestingly, the hemispheric asymmetry between the northern and southern subtropics of the Pacific is the reverse of the asymmetry found in the Atlantic.

By the later period 2075-2084 (**Fig. 4b**), the confidence intervals averaged over the four drivers are higher than during the earlier period 2005-2014 except for the eastern equatorial Atlantic and the western equatorial Indian Oceans. There continues to be a hemispheric asymmetry between the subtropics of the northern and southern subtropical Pacific. Interestingly, the fact that subpolar regions have become maxima for the multi-driver mean during the later period represents a reversal of what is found during the earlier period. For both hemispheres this reflects an important contribution from the O₂ inventories, as can be seen in the subpolar time series changes in **Figs. 1b** and **1f**. A relative decrease in variability in SST can be seen over the subpolar regions (**Figs. A1d** and **A2d**), contributing to the increased confidence in emergence over these regions (the differences between **Fig. 2b** and **Fig. 3b**).

1
2 As we did for the individual drivers, we also consider here the fraction of the global area that has
3 emerged for the multi-driver mean over the periods 2005-2014 and 2075-2084. Following the
4 procedure used in **Figs. 2** and **3**, a 30 yr trend window was used to characterize signal-to-noise
5 over the respective time interval year-by-year. The result was then averaged over each of the 10
6 yr time intervals gridpoint-by-gridpoint. Subsequently, area-weighted spatial averages were
7 considered. For the earlier period (2005-2014), 40.9% of the global ocean exhibits emergence
8 with 67% confidence. By 2075-2084, 62.5% of the global ocean exhibits emergence with 67%
9 confidence.

10
11 Next we consider the emergence characteristics for the three drivers with highest emergence
12 confidence intervals (rather than all four drivers) in the subsequent two panels. This is shown for
13 the time interval 2005-2014 in **Fig. 4c**, and for the time intervals 2075-2084 in **Fig. 4d**. For each
14 case, the fields presented in **Fig. 2** and **Fig. 3** were considered gridpoint-by-gridpoint to find the
15 three maximum drivers, and these were then averaged over the same biome regions.
16 Unsurprisingly, the confidence intervals are higher for this case with three drivers than with four.
17 For the time interval 2005-2014, 87.8% of the global surface area has emerged above the 67%
18 confidence interval (**Fig. 4c**), and for 2075-2084 98.8% has emerged above 67% (**Fig. 4d**)

19
20 We then consider averages over the two top drivers for the period 2005-2014 (**Fig. 4e**) and 2075-
21 2084 (**Fig. 4f**). Clearly the emergence characteristics are significantly higher here. As is
22 revealed in the structures of locally maximum and minimum confidence intervals, the dominant
23 influences over most of the global ocean are Ω_{arag} and SST in setting the even higher confidence
24 intervals. For the time interval 2005-2014, 97.94% of the global ocean is above the 67%
25 confidence interval. For the time interval 2075-2084, this has risen to 99.92% of the global
26 ocean having risen above the 67% confidence interval.

27

1 Taken together, the six panels in **Fig. 4** reveal a strong sensitivity to the number of drivers used to
2 characterize multiple-driver emergence. For the case of emergence over 2075-2084, it is
3 interesting to note that the patterns are different. For four drivers, the Southern Ocean has the
4 strongest emergence, whereas with two drivers the tropical band has the strongest emergence.

5 **3.4 Sensitivity of Confidence Intervals to Length of Decadal Trend Window**

6 Next we consider the sensitivity of our confidence interval calculation for the period 2005-2014
7 (**Fig. 2**) to the length of the window over which trends are calculated. The sensitivity of the
8 confidence intervals over 2005-2014 to the width of the window is considered in **Fig. 5**, where a
9 window of 10 yr has been used instead of the 30 yr used in **Fig. 2**. We show that for all four
10 drivers, the choice of a 10 yr trend window results in important decreases in the confidence
11 intervals over the global domain relative to a 30 yr trend window.

12
13 For Ω_{arag} (**Fig. 5a**), the confidence intervals are much lower in the regions directly impacted by El
14 Niño variability with a 10 yr window than they are with a 30 yr window (**Fig. 2a**). Additionally,
15 smaller confidence intervals are in evidence through the North Atlantic and parts of the Arctic, as
16 well as in a well-defined band across the North Pacific. In fact, along much of both the west and
17 east coasts of North America, confidence intervals are significantly lower than with the 30 yr
18 window. This may have important implications in pointing to the need for sustained (multi-
19 decadal) observing systems for ocean biogeochemistry in these regions.

20
21 For SST as well, the pattern obtained with a 10 yr window (**Fig. 5b**) reveals large differences
22 from the pattern obtained with a 30 yr window (**Fig. 2b**). The relatively elevated values
23 throughout much of the tropics and the subtropics have now disappeared, revealing relatively
24 weak confidence intervals over global domain. A similar loss of confidence is found in for O_2
25 inventories (**Fig. 5c**) and for NPP (**Fig. 5d**). In fact for SST, O_2 inventories, and NPP, using a 10
26 yr window results in very weak confidence nearly everywhere over the global domain.

27

3.5 Time of emergence in individual ecosystem drivers

Next we consider Time of Emergence (ToE) in **Fig. 6**. Here, we use a one standard deviation threshold to define ToE, and our intention is to represent another dimension to the emergence question, namely that of time. **A 10 yr tolerance window with a robust Loess filter has been applied to the fields in Fig. 6, as described in Appendix B.** First we consider Ω_{arag} (**Fig. 6a**), with this revealing that most of the emergence has already occurred by the start of the analysis period (before 1965), consistent with what was seen in the confidence interval analysis. The only exceptions are relatively high dynamical variability zone within the equatorial Pacific, and certain high variability structures associated with western boundary current regions. For SST (**Fig. 6b**), the ToE is within the 20th century for much of the tropics, with a tendency for this ToE to shift to the 21st century for much of the subtropical and subpolar regions. For much of the high-latitude Southern Ocean region, there is no emergence during the analysis period (before 2085). This is also true for important sectors of the northern North Atlantic. For O₂ inventories (**Fig. 6c**) an early ToE is evident for the Southern Ocean as well as the eastern equatorial Atlantic, largely consistent with what was seen in **Fig. 2**. Alternating zonal structures of early and late ToE are seen in the Equatorial Pacific, and a patchwork of structures with very different ToE is in evidence over much of the ocean away from the Southern Ocean. In fact, the patchwork-like structures are in even stronger evidence for NPP (**Fig. 6d**). Consistent with what was seen in **Fig. 2d**, relatively early ToE is in evidence for both the Agulhas and the Equatorial Pacific Cold Tongue regions. Otherwise there are alternating bands of saturation for both early (pre-1965) and late (post-2085) ToE. The adjacent early and late saturation regions are the consequence of choosing a threshold for ToE. Thus the boundaries between such regions may not in general be reflecting real biome structures in the model domain. This is addressed in more detail in the Appendix (**Fig. A3** and **Fig. A4**).

4 Discussion

4.1 Interpretation of Main Results

We set out to evaluate the emergence characteristics of four ecosystem drivers (surface Ω_{arag} , SST, subsurface O_2 , and NPP) of marine ecosystems, with two questions driving this investigation. The first motivation stems from an interest in identifying when the secular trend in drivers becomes evident or perceptible for local marine ecosystems relative to the natural background decadal variability to which the organisms have adapted. The second pertains to the optimization of the ocean biogeochemical observing system, and the application of models to advance this optimization through iterative communication with the community of researchers evaluating network design. Building on the previous work of Frölicher et al. (2009), this was pursued using a suite of Large Initial-Condition Ensemble Simulations, as it is only with this approach that one can infer the secular trend (ensemble mean) for a model by filtering natural variability through an averaging procedure. The averaging procedure operates on not only patterns of climate modes such as ENSO (Wittenberg et al., 2014), but also on natural variability on smaller scales associated with variations in gyre boundaries that are not correlated to climate modes. Importantly, with this large ensemble approach, one does not need to assume that variability in the system is stationary in time.

Our main result is that there is a temporal hierarchy in the emergence of the four ocean ecosystem drivers above the level of background natural variability. This is strongly evident in **Fig. 1**. Ω_{arag} emerges earliest, NPP emerges latest, and both O_2 inventories and SST fall between the two. Additionally, three (SST, O_2 inventories, and NPP) of the four drivers considered here exhibit large regions where detection of secular trends is significantly complicated by the presence of natural decadal variability in the climate system. Our results also revealed very pronounced differences in the patterns for the confidence intervals for the emergence of SST and O_2 inventories (**Fig. 2**). For SST, the Southern Ocean emerges relatively late (post-2014) and the tropics emerge rather early (pre-2014), in line with earlier studies (Mora et al., 2013). For O_2 inventories, on the other hand, the Pacific and Atlantic sectors of the Southern Ocean exhibit regions of relatively early emergence (pre-2014) and the tropics emerge rather late (post-2014).

1 The coalescence of the different global drivers in certain regions is already creating a number of
2 hot-spots (**Fig. 4**), with the Southern Ocean and more generally the high latitudes projected to
3 increase in importance by 2075-2084.

4
5 Consistent results regarding the temporal hierarchy of ecosystem driver emergence were found
6 through the analysis of confidence intervals (**Fig. 2** and **Fig. 3**) and time of emergence (ToE)
7 (**Fig. 6**) as diagnostics. However, we prefer the confidence interval analysis over the ToE analysis
8 for two reasons. First, for the four ecosystem drivers the saturation characteristics of the ToE
9 analysis (emergence before 1965 or after 2085) are widespread, complicating interpretation.
10 Second, and perhaps more importantly, ToE diagnostics require a specification of a threshold of
11 signal-to-noise that is somewhat arbitrary (here we have considered both 1 and 2 standard
12 deviations, but have chosen to emphasize the less conservative value of 1 in **Fig. 6**).

13
14 The quantification of signal-to-noise at the center of our analysis relies on joint use of a suite of
15 Large Initial-Condition Ensemble Simulations using an individual Earth System Model (GFDL's
16 ESM2M). However, previously published analyses indicate that the collection of Earth System
17 Models developed by different modeling centers exhibit disparate amplitudes for the secular
18 trends in individual drivers (Bopp et al., 2013). For example, different ESMs are likely to differ
19 more in their projected changes in NPP than they are in Ω_{arag} or SST, since at least with Ω_{arag} or
20 SST the ESMs are largely consistent in the sign of their response (Bopp et al. 2013, Steinacher et
21 al. 2010, Cocco et al. 2013). In high latitudes, ESMs are generally consistent in simulating
22 decreasing ocean O₂ inventories under 21st century climate change (Cocco et al. 2013). However,
23 it is worth noting that in GFDL ESM2M global NPP changes by only 2% under the
24 historical/RCP8.5 scenario, whereas global NPP tends to decrease by approximately 10% for a
25 multi-model of CMIP5 ESMs under the same scenario (Laufkötter et al., 2015). Additionally,
26 different Earth System Models exhibit different noise or underlying natural variability
27 characteristics for individual drivers (Keller et al., 2014). Such inter-model differences strongly
28 suggest that the temporal and spatial characteristics of emergence should be model-dependent, as
29 has been shown for the case of surface air temperature by Hawkins and Sutton (2011).

1 Alternative emissions scenarios may also lead to changes in both the signal and the noise.
2 Investigations of the sensitivity of our results to alternative scenarios for anthropogenic emissions
3 and other model projections may be subject to further studies.

4

5 **4.2 Mechanistic Interpretation**

6

7 Although our analysis has been focused on statistical questions (namely confidence intervals and
8 time of emergence diagnostics), it is also important to consider the mechanisms that control
9 emergence timescales. The most important contrast seen in our results is in evidence in **Fig. 2**,
10 namely the early (late) emergence of SST in the tropics (Southern Ocean), and the late (early)
11 emergence of O₂ inventories in the tropics (Southern Ocean) with the 30 yr window. For SST,
12 the contrast between the tropics and the Southern Ocean in **Fig. 2** with a 30 yr window is largely
13 reflecting the weakness of the SST trend over the Southern Ocean relative to the tropics. In fact,
14 the contrast between the tropics and the Southern Ocean is more generally representative of most
15 of the rest of the global surface ocean (except in the northern North Atlantic) relative to the
16 Southern Ocean (**Fig. A1c**). The lack of SST warming reflects large-scale interhemispheric
17 asymmetries in the mean ocean circulation. The strong upwelling in the Southern Ocean nearly
18 anchors sea surface temperature at pre-industrial level (Stouffer et al. 1989; Marshall and Speer,
19 2012; Frölicher et al., 2015).

20

21 For the tropics, the secular trend is sufficiently large over a 30 yr window to be more important
22 than the natural decadal variability, but consistent with the spectral characteristics of ENSO for
23 the underlying physical model (Wittenberg, 2009). This is no longer true for the case of a 10 yr
24 window in the tropics. For the case of O₂ inventories, the reverse holds. In the Southern Ocean,
25 de-oxygenation is much larger than natural variability due to the stratification-induced reduced
26 supply of O₂ from the surface into the thermocline (Frölicher et al, 2009; Gnanadesikan et al.,
27 2012). In contrast, almost no O₂ changes are projected to occur in the low O₂ regions of the
28 tropical and subtropical thermocline owing to a reduced O₂ demand because of the lower

1 biological production and export of organic matter in the overlying near-surface waters
2 (Gnanadesikan et al., 2012; Steinacher et al., 2010). These biological drivers are expected to be
3 modulated by perturbations to the rates of ocean interior and thermocline ventilation. However,
4 the confidence of the O₂ projections in the low latitudes is low, as the GFDL ESM2M has biases
5 in its representation of today's observed O₂ distribution (Gnanadesikan et al., 2012), a feature
6 common to the current generation of Earth System models (Bopp et al., 2013).

8 **4.3 Perceptible changes in ocean drivers**

9 We have previously defined perceptible changes in drivers of ocean ecosystems as anthropogenic
10 changes that are above the noise level of the natural background decadal variability to which
11 organisms are adapted. For the case where two of the four ecosystem drivers are used (**Fig. 4e**
12 and **Fig. 4f**), our analysis has revealed that Ω_{arag} and SST are the dominant drivers with early
13 emergence in the tropics. In fact, it should be emphasized that this is a result of our two-driver
14 analysis, rather than an assumption or an imposed constraint. In particular, the two-driver
15 analysis presented in **Fig. 4e** for the recent past indicates that tropical coral reef habitats may be
16 the primary regions currently experiencing perceptible changes relative to the background natural
17 variability (Pelejero et al., 2005). Thus our results are consistent with previous studies that argue
18 that coral reefs are the marine ecosystems that are threatened most by environmental changes (see
19 the cross-chapter box on coral reefs in IPCC AR5, Gattuso et al, 2014a; also see Gattuso et al.,
20 2014b). Although the results of the two-driver analysis are seen to hold through the tropics, the
21 results may warrant particular attention in the Coral Triangle biodiversity hotspot region,
22 spanning Indonesia, the Philippines, Malaysia, Papua New Guinea, and the Solomon Islands
23 (Allen, 2008).

24
25 It is important however to emphasize that the analyses for multiple drivers seen in **Fig. 4**
26 consisted of averaging of confidence intervals obtained for the individual drivers. In interpreting
27 these results for SST and Ω_{arag} in the Coral Triangle as the confidence intervals for the impacts of
28 multiple drivers, it remains a scientific challenge to determine whether they are in fact acting

1 additively, synergistically, or antagonistically in their impact (Boldt et al., 2014, and references
2 therein). To date resource management strategies have tended to focus on the impact of
3 individual drivers, with little consideration or attention to potential relationships and feedbacks
4 between the drivers. Addressing these questions is beyond the scope of our study, but it is our
5 hope that the analysis considered here will contribute to motivating future work in this direction.
6 What is clear, however, is that anthropogenic CO₂ emissions to the atmosphere are the common
7 driver of the perturbations considered here, and this underscores the necessary of a single policy
8 response (reduction in emissions). Substantial mitigation efforts are required if ocean ecosystems
9 are to be spared from the “quadruple whammy” of the drivers considered here.

11 **4.4 Implications for Observing System Design**

12 It is also important to consider the implications of our study for optimization of the global ocean
13 observing system. With this goal in mind, our study can be considered as an Observing System
14 Simulation Experiment (OSSE). With an OSSE, one considers a model to be an analog for the
15 real ocean, for which one has the fully resolved state evolution to round-off error. Earlier OSSEs
16 (Christian et al., 2008; Park et al., 2010; Plancherel et al., 2013; Majkut et al., 2014; Cassar et al.,
17 2014) have tended to focus on one realization of the evolution of the Earth system, and focused
18 on the skill with which different observing strategies can reproduce variability in the Earth
19 System through selective sub-sampling of the model output. The target is to test the available
20 skill in reproducing the real-world trends and variability with an incomplete observing system,
21 without any claim to separating the signal associate with the secular trend and natural variability.

22
23 For our experimental configuration considered as an OSSE, we address a different but
24 complementary question. We consider the case where the observing system has perfect skill in
25 reproducing the trends and variability of the system of interest, but where the target is to identify
26 the secular trend. It is precisely this deconvolution that we address with the Large Initial-
27 Condition Simulations with the Earth System Model, thereby building on the previous analyses
28 considered with fewer ensemble members (Frölicher et al, 2009; Christian, 2014). The question
29 is then as follows: Given an observing system with perfect skill that allows one to perfectly

1 monitor the evolution of the system, how many years of continuous measurements are needed to
2 identify the secular trend above the noise of background variability? Our main result is that
3 sustained decadal measurements will be needed even for the idealized case of a perfect observing
4 system.

5
6 Viewed in this way, our main results point to the importance of maintaining a sustained multi-
7 decadal observing system for ocean biogeochemistry and ecosystem drivers. For the four drivers
8 considered here, the confidence intervals found with a 30 yr window for calculating trends (**Fig.**
9 **2**) are significantly higher than those found with a 10 yr window (**Fig. 5**). For the case with a 10
10 yr window, even Ω_{arag} reveals broad expanses of non-emergence over the decade 2005-2014.
11 This is in evidence, for example, over important parts of the Coral Triangle biodiversity hot spot
12 spanning the Indo-Pacific Warm Pool region, as well as for the North Atlantic. This underscores
13 the potential importance of sustained multi-decadal continuous measurements in order to identify
14 the rate of acidification associated with the secular trend in these regions.

15
16 More generally, our analysis of confidence intervals for emergence for two versus four drivers
17 (**Fig. 4e**) largely highlight the combined effects of Ω_{arag} and SST in the tropics. This implies that
18 even with high resolution of temporal and spatial scales, a sustained multi-decadal (30 yr)
19 observing system of the type considered by Ishii et al. (2009) in western Equatorial Pacific
20 surface waters is needed to detect rate of the secular trend in acidification against the background
21 noise of natural variability with confidence.

22

23 **5 Conclusions**

24 Here we have considered a suite of Large Initial-Condition Ensemble Simulations with GFDL's
25 Earth System Model ESM2M to evaluate the emergence characteristics of four drivers of ocean
26 ecosystems under anthropogenic climate change. The drivers chosen were Ω_{arag} , SST, upper-
27 ocean inventories of O_2 , and net primary productivity (NPP). There were two questions
28

1 underlying the analysis: First, when and where will marine ecosystems experience perceptible
2 changes in the ocean drivers, changes that lie outside of the range of natural variability to which
3 they are adapted or accustomed? Second, what are the implications for optimal design of a global
4 observing system for the four drivers?

5
6 The main result of our study is that among our four ecosystem drivers, there is a pronounced
7 temporal and spatial hierarchy to emergence over global scales. Using a one-standard deviation
8 (67%) confidence interval threshold for signal-to-noise, and a 30 yr trend window, we find that
9 the acidification driver (Ω_{arag}) emerges earliest and NPP emerges latest over global scales.
10 Between these two outliers, SST and O_2 inventories have intermediate timescales of emergence,
11 but opposing patterns between the two of them. SST emerges earlier in the low latitudes and
12 later in the high latitudes, while O_2 inventories exhibit earlier emergence over high latitude
13 regions than in the tropics. We also considered a multiple-driver analysis where we combined
14 the four individual drivers. There we found that whereas 41% of the global ocean area exhibits
15 emergence over 2005-2014, 63% has exhibited emergence by 2075-2084. This four-driver
16 analysis reveals a more pronounced emergence pattern over the extra-tropics than over the
17 tropics. We also considered the multiple-driver case where we included only two drivers by
18 biome region, with very different results. There the tropics emerge earliest, with the dominant
19 drivers being Ω_{arag} and SST. Given that Ω_{arag} and SST are the two most important drivers for
20 coral reef ecosystems, this analysis identifies coral reefs as being especially vulnerable under 21st
21 century climate change.

22
23 Considered as an Observing System Simulation Experiment (OSSE), our results emphasize the
24 need for a sustained global observing system for multiple decades for the task of identifying
25 anthropogenic trends in ecosystem drivers. This is true even for the case of a global bio-Argo
26 array within a broad multi-platform observing system, as is revealed in the contrast between the
27 confidence intervals for emergence of the secular trend between a 10 yr sustained observing
28 system (**Fig. 5**) and a 30 yr sustained observing system (**Fig. 2**).

29

1 **Appendix A: Secular trends and natural variability underlying signal-to-noise** 2 **analysis**

3 We characterize here the secular trends (left columns in **Figs. A1,A2**; referred below as TREND)
4 and standard deviations of the secular trends (right columns in **Figs. A1,A2**; referred below as
5 NOISE) separately. Recall that the signal-to-noise ratio is defined as the ratio of these two fields.
6 Ω_{arag} decreases everywhere over the global domain (**Fig. A1a**), with minimum relative rates of
7 decrease in the equatorial regions, and a general tendency towards stronger relative rates of
8 decrease at high latitudes. Largest NOISE of Ω_{arag} is simulated in a number of dynamically
9 active regions, including the margins of the subtropical gyres and the equatorial Pacific.
10 However, with the exception of a few isolated regions, the TREND is everywhere significantly
11 larger than the NOISE for Ω_{arag} . For SST, the TREND over 2005-2014 is positive (warming) over
12 most of the globe, with the notable exception of the western subpolar North Atlantic and large
13 expanses of the Southern Ocean (cooling). This stands in contrast to Ω_{arag} , where the trend had
14 the same sign over the entire domain.

15
16 The NOISE for SST finds largest expression in the subpolar regions of the Northern Hemisphere
17 and over parts of the Southern Ocean. In fact it can be seen in the **Fig. 1c,d** that the extratropical
18 regions of weak or negative TREND in SST are associated with enhanced variability. The
19 tropics, on the other hand, reveal only modest NOISE amplitude relative to the TREND. Taken
20 together, this helps to account for the fact that the confidence interval map in **Fig. 2b** reveals high
21 confidence in the tropics relative to the subpolar regions. As has been stated in the main text, it
22 may seem somewhat surprising that the signal-to-noise-ratio is relatively elevated in the
23 Equatorial Pacific, given that this is the region of largest natural variability in the climate system.
24 It is important to emphasize here that we are considering trends over a 30 yr interval in our
25 quantification of NOISE rather than considering the standard deviation associated with
26 interannual variability for each of the drivers.

27
28 For the case of O_2 inventories, a decreasing TREND can be seen in the well-ventilated
29 thermocline of the high latitudes. Within the tropics and subtropics, structures of positive trend

1 do occur. As with SST but in contrast with Ω_{arag} , the sign of the ensemble-mean response of O_2
2 inventories is not of the same sign everywhere. However, there are also pronounced structures of
3 larger NOISE in O_2 inventories, which over many regions are associated with decadal variations
4 in gyre boundaries and frontal regions. The most prominent extended region where the TREND
5 is larger than the NOISE is over a broad expanse of the Pacific and Atlantic sectors of the
6 Southern Ocean (as seen in **Fig. 2c**).

7
8 For the case of Net Primary Productivity (NPP), both the TREND and the NOISE patterns show
9 relatively narrowly-defined but large-amplitude structures. There is a trend in the zonal gradient
10 in NPP across the equatorial Pacific, as well as a trend towards enhanced NPP along the
11 poleward flanks of the Southern Hemisphere subtropical gyres. For the NOISE, a series of
12 relatively narrow structures of high amplitude are found winding through the tropics. Over most
13 regions of the globe, the NOISE is of sufficient amplitude relative to TREND to give the
14 consistently lowest Confidence Interval distribution of the four drivers considered here (**Fig. 2d**).

15
16 It is worth noting in **Fig. A1** that the unforced components (right column) of the four drivers
17 exhibit large-scale spatial coherence rather than small grid-scale noise. However, these structures
18 are distinct for each of the drivers.

19
20 The same fields for the time interval 2075-2084 are considered in **Fig. A2**. The amplitudes in
21 TRENDS have increased in general, but important elements of the TREND structures are similar
22 as for the earlier period. The structures of the NOISE are quite similar to those found for the
23 equivalent drivers during the earlier period. However, the amplitude of the variations is in many
24 cases different for 2075-2084 than for 2005-2014. This indicates that the amplitude of the
25 background natural variability is not stationary. For the case of SST and O_2 , the standard
26 deviation of natural decadal variations decreases over the Southern Ocean.

27

28 **Appendix B: Temporal filtering and trend detection**

1 The task of calculating Time of Emergence (ToE) in **Fig. 6** is complicated by the fact that the
2 signal-to-noise ratio (SNR) in time series for individual grid points in the ocean model does not
3 in general tend to be monotonically increasing over the period 1950-2100. Rather, the evolution
4 of the SNR can reflect that for this particular diagnostic, the 30 ensemble members may not be
5 sufficient to eliminate noise when averaged.

6
7 As an illustration of this problem, we consider in **Fig. A3** the evolution of the SNR for sea
8 surface temperature a region in the Pacific sector of the Southern Ocean bounded by 130°W-
9 100°W and 45°S-60°S. The time series of the SNR for the various points in this region are
10 superposed in **Fig. A3a**. The non-monotonic nature of the increase in SNR through time is
11 evident. With this unfiltered SNR calculated directly from the annual mean model output, there
12 are relatively short-timescale excursions above the two standard deviation threshold that precede
13 by a number of decades the more permanent crossing of the one standard deviation threshold.
14 However, this does not occur for all of the grid points in the domain of interest. As a
15 consequence of these early excursions above the one standard deviation threshold, the spatial
16 pattern of ToE using a strict definition of first crossing (**Fig. A3b**) reveals a spatial pattern that
17 has a ToE before 2014 (present time) over more than 50% of the region.

18
19 It is important to understand the degree to which the ToE structure in **Fig. A3b** reflects short-
20 term versus longer-term or permanent transitions of the SNR about the one standard deviation
21 threshold, rather than short-term excursions. To evaluate this, we apply to the full suite of time
22 series shown in **Fig. A3a** (gridpoint-by-gridpoint SNR for SST) a robust Loess filter. We have
23 chosen to use a 10 yr tolerance window with the robust Loess filter, with this serving effectively
24 as a low-pass filter that is obtained using 10 yr local regressions over the entire time period for
25 each individual gridpoint. The result of applying the robust Loess filter is shown in **Fig. A4a**,
26 where the filtered time series are shown as red lines overlaying the full time series shown in blue.
27 The filtered time series are seen to behave as low-pass filters that effectively remove the higher
28 frequency components. When these smoothed time series are used to define ToE with the same
29 two standard deviation threshold, the resulting pattern (**Fig. A4b**) reveals important differences

1 relative to the unfiltered time series. The structures with post-2014 ToE now occupy most of the
2 domain. This indicates the strong sensitivity of the spatial pattern of ToE to the time-filtering of
3 the SNR time series.

4
5 The sensitivity of the Time of Emergence (ToE) to the width of the window used for calculating
6 trends is considered in **Fig. A5**. Here the window is chosen to be 10 yr, and this result is to be
7 contrasted to what was found for ToE using a 30 yr window as considered in **Fig. 6**. Clearly the
8 narrower trend interval of 10-years results in saturation (post-2095 emergence for this case of a
9 10-year trend window) on nearly global scales for all of the drivers except for Ω_{arag} . This
10 indicates a strong sensitivity of ToE to the timescale chosen for the analysis, consistent with what
11 was shown in **Fig. 4** for the confidence interval sensitivity analysis. **Fig. A5** is valuable in that
12 the strong saturation characteristics for three of the drivers (post-2085 emergence) are much more
13 difficult to interpret than the parallel and favored analysis with confidence intervals.

14
15 We also consider the sensitivity of the Time of Emergence (ToE) to the width of the tolerance
16 window used for the robust Loess filter in **Fig. A6a**. For each case, the sensitivity of the ToE to
17 the width of the tolerance window (described above) is considered for each of the four drivers.
18 For each case, the area-weighted global mean sensitivity is considered. The sensitivity is weak
19 for Ω_{arag} , with the dynamic range of the sensitivity being only a few years as the tolerance
20 window is modified from 5 yr to 25 yr. The sensitivity is largest for O₂ inventories, with a
21 decrease of the ToE of 80 years as one transitions from a 5 yr to a 25 yr tolerance window. This
22 strong sensitivity is likely the expression of the red spectrum of modes of variability in
23 thermocline depth impacting O₂ inventories in their temporal variability. The sensitivities of SST
24 and NPP are quite similar, both being approximately 50% of the amplitude of the sensitivity seen
25 for O₂ inventories. In fact, the sensitivity of NPP should be expected to be larger than that of
26 SST, given that the sensitivity shown here may be obscured by the saturation characteristics of
27 NPP.

28

1 We also consider in **Fig. A6b** the sensitivity of the ToE to the choice of a SNR ratio of one in
2 **Fig. 6**. Here as well, the sensitivity considers the global area-weighted mean. The sensitivity is
3 strongest for SST, and then the second strongest is found for O₂ inventories. The sensitivity is
4 weakest for Ω_{arag} and for NPP, but as a caveat it needs to be emphasized that both of these fields
5 exhibit saturation behavior in **Fig. 6**. If the suite of 30 ensemble runs with ESM2M had been
6 runs for a significantly longer time intervals, say from 1860-2300, then one would expect that the
7 sensitivity for these two fields to be more pronounced.

8

9 **Acknowledgements**

10 The contribution of K.B. Rodgers came through awards NA17RJ2612 and NA08OAR4320752,
11 which includes support through the NOAA Office of Climate Observations (OCO). The
12 statements, findings, conclusions, and recommendations are those of the authors and do not
13 necessarily reflect the views of NOAA or the U.S. Department of Commerce. T.L. Frölicher
14 acknowledges financial support from the SNSF (Ambizione grant PZ00P2_142573). J. Lin's
15 work in a summer internship was supported through the Princeton Environmental Institute (PEI).
16 The authors would like to thank John Dunne at GFDL for his contributions with model
17 development and for access to computing resources. **We would also like to thank Brendan Carter,**
18 **Niki Gruber, Jorge Sarmiento, Richard Slater, Paul Yi, and Sarah Schlunegger for fruitful**
19 **discussions.**

20

1 **References**

- 2 **Allen, G.R.:** Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef
3 **fishes, *Aquatic Conserv. Mar. Freshw. Ecosyst.* 18, 541-556, 2008.**
- 4 Anderson, J.L., Balaji, V., Broccoli, A.J., Cooke W.F., Delworth, T.L. Dixon, K.W., Donner L.J.,
5 Dunne, K.A. Freidenreich, S.M., Garner, S.T., Gudgel, R.G., Gordon, C.T. Held, I.M. Hemler,
6 R.S. Horowitz, L.W., Klein, S.A., Knutson, T.R., Kushner P.J. Langenhost A.R., Lau, N.-C.,
7 Liang, Z., Malyshev, S.L., Milly P.C.D., Nath, M.J., Ploshay, J.J., Ramaswamy V., Schwarzkopf,
8 M.D., Shevliakova, E., Sirutis, J.J., Soden, B.J., Stern, W.F., Thompson, L.A., Wilson, R.J.,
9 Wittenberg, A.T. Wyman, B.L.: The new GFDL global atmosphere and land model AM2-LM2:
10 Evaluation with prescribed SST simulations, *J. Climate*, 17, 4641-4673, 2004.
- 11 **Boldt, J.L., Martone, R., Samhuri, J., Perry, R.I., Itoh, S., Chung, I.K., Takahashi, M., and**
12 **Yoshie, N.:** Developing ecosystem indicators for responses to multiple stressors, *Oceanography*
13 **27(4), 116-133, <http://dx.doi.org/10.5670/oceanog.2014.91>, 2014.**
- 14 Bopp, L., Monfray, P., Aumont, O., Dufresne, J.L., Le Treut, H., Madec, G., Terray, L., and Orr,
15 J.C.: Potential impact of climate change on marine export production, *Global Biogeochem.*
16 *Cycles*, 15, 81-99, 2001.
- 17 Bopp, L., Resplandy, L., Orr, J.C., Doney, S.C., Dunne, J.P., Gehlen, M., Halloran, P., Heinze,
18 C., Ilyina, T., Séférian, R., Tjiputra, J., and Vichi, M.: Multiple stressors of ocean ecosystems in
19 the 21st century: projections with CMIP5 models, *Biogeosciences*, 10, 6225-6245, 2013.
- 20 **Cassar, N., Nevison, C.D., and Manizza, M.:** Correcting oceanic O₂/AR-net community
21 **production estimates for vertical mixing using N₂O observations, *Geophys. Res. Lett.*, 41, 8961-**
22 **8970, doi:10.1002/2014GL062040, 2014.**
- 23 **Christian, J.R., Feely, R.A., Ishii, M., Murtugudde, R., and Wang, X.:** Testing an ocean carbon
24 **model with observed sea surface pCO₂ and dissolved inorganic carbon in the tropical Pacific**
25 **Ocean, *J. Geophys. Res.*, 113, C07047, doi:10.1029/2007JC004428, 2008.**
- 26 **Christian, J.R.:** Timing of the Departure of Ocean Biogeochemical Cycles from the Preindustrial
27 **State, *PloS ONE*, 9(11): e109820. doi:10.1371/journal.pone.0109820, 2014.**

1 Cocco, V., Joos, F., Steinacher, M., Frölicher, T.L., Bopp, L., Dunne, J., Gehlen, M., Heinze, C.,
2 Orr, J., Oschlies, A., Schneider, B., Setschneider, J., and Tjiputra, J.: Oxygen and indicators of
3 stress for marine life in multi-model global warming predictions, *Biogeosciences*, 10, 1849-1868,
4 2013.

5 Delworth, T.L., Broccoli, A.J., Rosati, A., Stouffer, R.J., Balaji, V., Beesley, J.A., Cooke, W.F.,
6 Dixon, K.W., Dunne, J., Dunne, K.A., Durachta, J.W., Findell, K.L., Ginoux, P., Gnanadesikan,
7 A., Gordon, C.T., Griffies, S.M., Gudgel, R., Harrison, M.J., Held, I.m., Hemler, R.S., Horowitz,
8 L.W., Klein, S.A., Knutson, T.R., Kushner, P.J., Langenhorst, A.R., Lee, H.-C., Lin, S.-J., Lu, J.,
9 Malyshev, S.L., Milly, P.C.D., Ramaswamy, V., Russell, J., Schwarzkopf, M.D., Shevliakova,
10 E., Sirutis, J.J., Spelman, M.J., Stern, W.F., Winton, M., Wittenberg, A.T., Wyman, B., Zeng, F.,
11 and Zhang, R.: GFDL's CM2 Global Coupled Climate Models. Part I: Formulation and
12 Simulation Characteristics, *Journal of Climate*, 19, 643-674, 2006.

13 Deser, C., Phillips, A.S., Alexander, M.A., and Smoliak, B.V.: Projecting North American
14 Climate over the Next 50 Years: Uncertainty due to Internal Variability, *Journal of Climate*, 27,
15 2271-2296, 2014.

16 Diffenbaugh, N. S., and Scherer, M.: Observational and model evidence of global emergence of
17 permanent, unprecedented heat in the 20th and 21st centuries. *Clim. Change*, 107, 615-624, 2011.

18 Doney, S.C., Fabry, V.J., Feely, R.A., and Kleypas, J.A.: Ocean Acidification: The Other CO₂
19 Problem, *Annu. Rev. Mar. Sci.*, 1, 169-192, 2009.

20 Doney, S.C., Ruckelshaus, M., Duffy, J.E., Barry, J.P., Chan, F., English, C.A., Galindo, H.M.,
21 Grebmeier, J.M., Hollowed, A.B., Knowlton, N., Polovina, J., Rabalais, N.N. Sydeman, W.J.,
22 and Talley, L.D.: Climate change impacts on marine ecosystems, *Annual Review of Marine*
23 *Science*, 4, 11-37, 2012.

24 Dunne, J.P., Gnanadesikan, A., Sarmiento, J.L., and Slater, R.D.: Technical description of the
25 prototype version (v0) of Tracers of Phytoplankton with Allometric Zooplankton (TOPAZ) ocean
26 biogeochemical model as used in the Princeton IFMIP model, *Biogeosciences*, 7 (Suppl.), 3593,
27 doi:10.4195/bg-7-3593-2010, 2010.

28 Dunne, J.P., John, J.G., Adcroft, A.J., Griffies, S.M., Hallberg, R.W., Shevliakova, E., Stouffer,
29 R.J., Cooke, W., Dunne, K.A., Harrison, M.J., Krasting, J.P., Halyshev, S.L., Milly, P.C.D.,

1 Phillipps, P.J., Sentman, L.T, Samuels, B.L., Spelman, M.J., Winton, M., Wittenberg, A.t., and
2 Zadeh, N.: GFDL's ESM2 Global Coupled Climate-Carbon Earth System Models. Part I:
3 Physical Formulation and Baseline Simulation Characteristics, *Journal of Climate*, 25, 6646-
4 6665, 2012.

5 Dunne, J.P., John, J.G., Shevliakova, E., Stouffer, R.J., Krasting, J.P., Malyshev, S.L., Milly,
6 P.C.D., Sentman, L.T., Adcroft, A.J., Cooke, W., Dunne, K.A., Griffies, S.M., Hallberg, R.W.,
7 Harrison, M.J., Levy H., Wittenberg, A.T., Phillips, P.J., and Zadeh, N.: GFDL's ESM2 Global
8 Coupled Climate-Carbon Earth System Models. Part II: Carbon System Formulation and
9 Baseline Simulation Characteristics, 26, 2247-2267, 2013.

10 Emerson, S., Watanabe, Y., Ono, T., and Mecking, S.: Temporal trends in apparent oxygen
11 utilization in the upper pycnocline of the North Pacific: 1980-2000, *J. Oceanogr.*, 60, 139-147,
12 2004.

13 Friedlingstein, P., Andrew, R.M., Rogelj, J., Peters, G.P., Canadell, J.G., Knutti, R., Luderer, G.,
14 Raupach, M.R., Schaeffer, M., van Vuuren, D.P., Le Quéré, C.: Persistent growth of CO₂
15 emissions and implications for reaching climate targets, *Nature Geosciences*, 7, 709-715, 2014.

16 Friedrich, T., Timmermann, A., Abe-Ouchi, A., Bates, N.R., Chikamoto, M.O., Church, M.J.,
17 Dore, J.E., Gledhill, D.K., Gonzalez-Davila, M., Henemann, M., Ilyina, T., Jungclaus, J.H.,
18 McLeod, E., Mouchet, A., and Santana-Casiano, J.M.: Detecting regional anthropogenic trends in
19 ocean acidification against natural variability, *Nature Climate Change*, 2, 10.1038/nclimate1372,
20 2012.

21 Frölicher, T.L, Joos, F., Plattner, G.K., Steinacher, M., and Doney, S.C.: Natural variability and
22 anthropogenic trends in oceanic oxygen in a coupled carbon cycle-climate model ensemble,
23 *Global Biogeochem. Cy.*, 23, GB1003, doi:10.1029/2008GB003316, 2009.

24 Frölicher, T.L., Sarmiento, J.L., Payntner, D.J., Dunne, J.P., Krasting, J.P., and Winton, M.:
25 Dominance of the Southern Ocean in Anthropogenic Carbon and Heat Uptake in CMIP5 Models,
26 *Journal of Climate*, 28, 862-886, 2015.

27 Gattuso, J.-P., Hoegh-Guldberg, O., and Pörtner, H.-O.: Coral Reefs. In Field C. et al. (Eds.),
28 *Climate Change 2014: Impacts, Adaptation and Vulnerability. Contribution of Working Group*

1 II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.
2 Cambridge: Cambridge University Press, 2014a.

3 Gattuso, J.-P., Brewer, P., Hoegh-Guldberg, O., Kleypas, J.A., Pörtner, H.-O., and Schmidt, D.:
4 Ocean acidification. In: Field C. et al. (Eds.), *Climate Change 2014: Impacts, Adaptation and*
5 *Vulnerability. Contribution of Working Group II to the Fifth Assessment report of the*
6 *Intergovernmental Panel on Climate Change.* Cambridge: Cambridge University Press, 2014b.

7 Gnanadesikan, A., Dunne, J.P., and John, J.: Understanding why the volume of suboxic waters
8 does not increase over centuries of global warming in an Earth System Model, *Biogeosciences*, 9,
9 1159-1172, 2012.

10 Griffies, S.M.: Elements of MOM4p1. GFDL Ocean Group Tech. Rep. 6, 37 pp [Available
11 online at http://data1.gfdl.noaa.gov/~arl/pubrel/o/old/doc/mom4p1_guide.pdf] (last access 2010),
12 2009.

13 Gruber, N.: Warming up, turning sour, losing breath: ocean biogeochemistry under global
14 change, *Phil. Trans. R. Soc. A*, 369, 1980-1996, 2011.

15 Hall, J., R. J. Diaz, N. Gruber, and D. Wilhelmsson.: MANAGING OCEAN ENVIRONMENTS IN A
16 CHANGING Sustainability and Economic, in *Managing ocean environments in a changing climate*,
17 edited by K. J. Noone, U. S. Sumaila, and R. J. Diaz, Elsevier Inc., **Burlington, MA, 193-222**, 2013.

18 Hawkins E. and Sutton, R.: The potential to narrow uncertainty in projections of regional precipitation
19 change, *Clim. Dyn.*, 37, 407-418, 2011.

20 Hawkins, E., and Sutton, R.: Time of emergence of climate signals, *Geophys. Res. Lett.*, 39,
21 L01702, [doi:10.1029/2011GL050087](https://doi.org/10.1029/2011GL050087), 2012.

22 Henson, S.A., Sarmiento, J.L., Dunne, J.P., Bopp, L., Lima, I., Doney, S.C., John, J., and
23 Beaulieu, C.: Detection of anthropogenic climate change in satellite records of ocean chlorophyll
24 and productivity, *Biogeosciences*, 7, 621-640, 2010.

25 Henson, S.A.: Slow science: the value of long ocean biogeochemistry records, *Phil. Trans. R.*
26 *Soc. A*, 372, 20130334, 2014.

27 Ishii, M., Inoue, H.Y., Midorikawa, T., Saito, S., Tokieda, T., Sasano, D., Nakadate, A., Nemoto,
28 K., Metzl, N., Wong, C.S. and Feely, R.A.: Spatial variability and decadal trend of the oceanic

1 CO₂ in the western equatorial Pacific warm/fresh water, *Deep-Sea Research II*, 56, 591-606,
2 2009.

3 Keeling, R.F., Koertzing, A., and Gruber, N.: Ocean Deoxygenation in a Warming World,
4 *Annu. Rev. Mar. Sci.*, 2, 199-229, 2010.

5 Keller, K.M., Joos, F., and Raible, C.C.: Time of emergence of trends in ocean biogeochemistry,
6 *Biogeosciences*, 11, 3647-3659, 2014.

7 Kouketsu, S., Fukasawa, M., Sasano, D., Kumamoto, Y., Kawano, T., Uchida, H., and Doi., T.:
8 Changes in water properties around North Pacific intermediate water between the 1980s, 1990s
9 and 2000s. *Deep-Sea Research II*, 57, 1177-1187, 2010.

10 Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., Buitenhuis, E.,
11 Doney, S.C., Dunne, J. Hashioka, T., Hauck, J., Hirata, T., John, J., Le Quéré, C., Lima, I.D.,
12 Nakano, H., Seferian, R., Totterdell, I., Vichi, M., and Völker, C.: Drivers and uncertainties of
13 future global marine primary production in marine ecosystem models, *Biogeosciences*
14 *Discussions*, 12, 3731-3824, 2015.

15 Lyu, K., Zhang, X., Church, J.A., Slangen, A.B.A., and Hu, J.: Time of emergence for regional
16 sea-level change, *Nature Climate Change*, 4, 1006-1010, doi:10.1038/nclimate2397, 2014.

17 Mahlstein, I., Hegerl, G., and Solomon, S.: Emerging local warming signals in observational
18 data. *Geophys. Res. Lett.*, 39, L21711, doi:1029/2012GL053952, 2012.

19 Majkut, J.D., Carter, B.R., Frölicher, T.L., Dufour, C.O., Rodgers, K.B., and Sarmiento, J.L.: An
20 observing system simulation for Southern Ocean carbon dioxide uptake, *Philosophical*
21 *Transactions of the Royal Society A: Mathematical, Physical, and Engineering Sciences* 372:
22 2019, 20130046-20130046, 2014.

23 Marshall, J., and Speer, K.: Closure of the meridional overturning circulation through Southern
24 Ocean upwelling, *Nature Geoscience*, 5, 171-180, 2012.

25 Martinez, E., Antoine, D., d'Ortenzio, F., and Gentili, B.: Climate-Driven Basin-Scale Decadal
26 Oscillations of Oceanic Phytoplankton, *Science*, 326, 1253-1256, 2009.

27 Mecking, S., Langdon, C., Feely, R.A., Sabine, C.L., Deutsch, C.A., and Min, D.-H.: Climate
28 variability in the North Pacific thermocline diagnosed from oxygen measurements: An update

1 based on the U.S. CLIVAR/CO₂ Repeat Hydrography cruises, *Global Biogeochem. Cycles*, 22,
2 GB3015, doi:10.1029/2007GB003101, 2008.

3 Mora, C., Frazier, A.G., Longman, R.J., Dacks, R.S., Walton, M.M., Tong, E.J., Sanchez, J.J.,
4 Kaiser, L.R., Stender, Y.O, Anderson, J.M., Ambrosino, C.M., Fernandez-Silva, I. Giuseffi, L.F.,
5 and Giambelluca, T.W.: The projected timing of climate departure from recent variability,
6 *Nature*, 502, 183-187, 2013.

7 Park, G.-H., Wanninkhof, R., Doney, S.C., Takahashi, T., Lee, K., Feely, R.A., Sabine, R.A.,
8 Trinanes, C.L, and Lima, I.D.: Variability of global net sea-air CO₂ fluxes over the last three
9 decades using empirical relationships, *Tellus, B*, 62, 362-368, doi:10.1111/j.1600-
10 0889.2010.00498.x, 2010.

11 Pelejero, C., Calvo, E., McCulloch, M.T. Marshall, J.F., Gagan, M.K., Lough, J.M., and Opdyke,
12 B.N.: Preindustrial to modern interdecadal variability in coral reef pH, *Science*, 309, 2204-2207,
13 2005.

14 Pierrehumbert, R.T.: Thermostats, radiator fins, and the local runaway greenhouse, *Journal of*
15 *the Atmospheric Sciences*, 52, 1784-1806, 1995.

16 Plancherel, Y., Rodgers, K.B., Key, R.M., Jacobson, A.R., and Sarmiento, J.L.: Role of
17 regression model selection and station distribution on the estimation of oceanic anthropogenic
18 carbon change by eMLR, *Biogeosciences*, 10, 4801-4831, 2013.

19 Rodgers, K.B., Key, R.M., Gnanadesikan, A., Sarmiento, J.L., Aumont, O., Bopp, L., Doney,
20 S.C., Dunne, J.P., Glover, D.M., Ishida, A., Ishii, M., Jacobson, A.R., Lo Monaco, C., Maier-
21 Reimer, E., Mercier, H., Metzl, N., Pérez, F.F., Rios, A.F., Wanninkhof, R., Wetzels, P., Winn,
22 C.D., and Yamanaka, Y.: Using altimetry to help explain patchy changes in hydrographic carbon
23 measurements, *J. Geophys. Res.*, 114, C09013, doi:10.1029/2008JC005183, 2009.

24 Sarmiento, J. L., Hughes, T.M.C., Stouffer, R.J., Manabe, S.: Simulated response of the ocean
25 carbon cycle to anthropogenic climate warming, *Nature*, 393, 245-249, 1998.

26 Steinacher, M., Joos, F., Frölicher, T., Bopp, L., Cadule, P., Cocco, V., Doney, S.C., Gehlen, M.,
27 Lindsay, K., Moore, J.K., Schneider, B., and Segschneider, J.: Projected 21st century decrease in

1 marine productivity: a multi-model analysis, *Biogeosciences*, 7, 979-1005, doi:10.4195/bg-7-
2 979-2010, 2010.

3 Stouffer, R.J., Manabe, S., and Bryan, K.: Interhemispheric asymmetry in climate response to a
4 gradual increase of atmospheric CO₂, *Nature*, 432, 660-662, 1989.

5 Takatani, Y., Sasano, D., Nakano, T., Midorikawa, T., and Ishii, M.: Decrease of dissolved
6 oxygen after the mid-1980s in the western North Pacific subtropical gyre along the 137°E repeat
7 section, *Global Biogeochem. Cy.*, 26, GB2013, doi:10.10292011GB004227, 2012.

8 Timmermann, A., Oberhuber, J., Bacher, A., Esch, M., Latif, M., and Roeckner, E.: Increased El
9 Niño frequency in a climate model forced by future greenhouse warming, *Nature*, 298, 694-697,
10 1999.

11 van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.,
12 Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinhausen, M., Nakicenovic, N., Smith, S.J.,
13 and Rose, S.K.: The representative concentration pathways: an overview, *Clim. Change*, 109,
14 5031, doi:10.1007/s10584-011-0148-z, 2011.

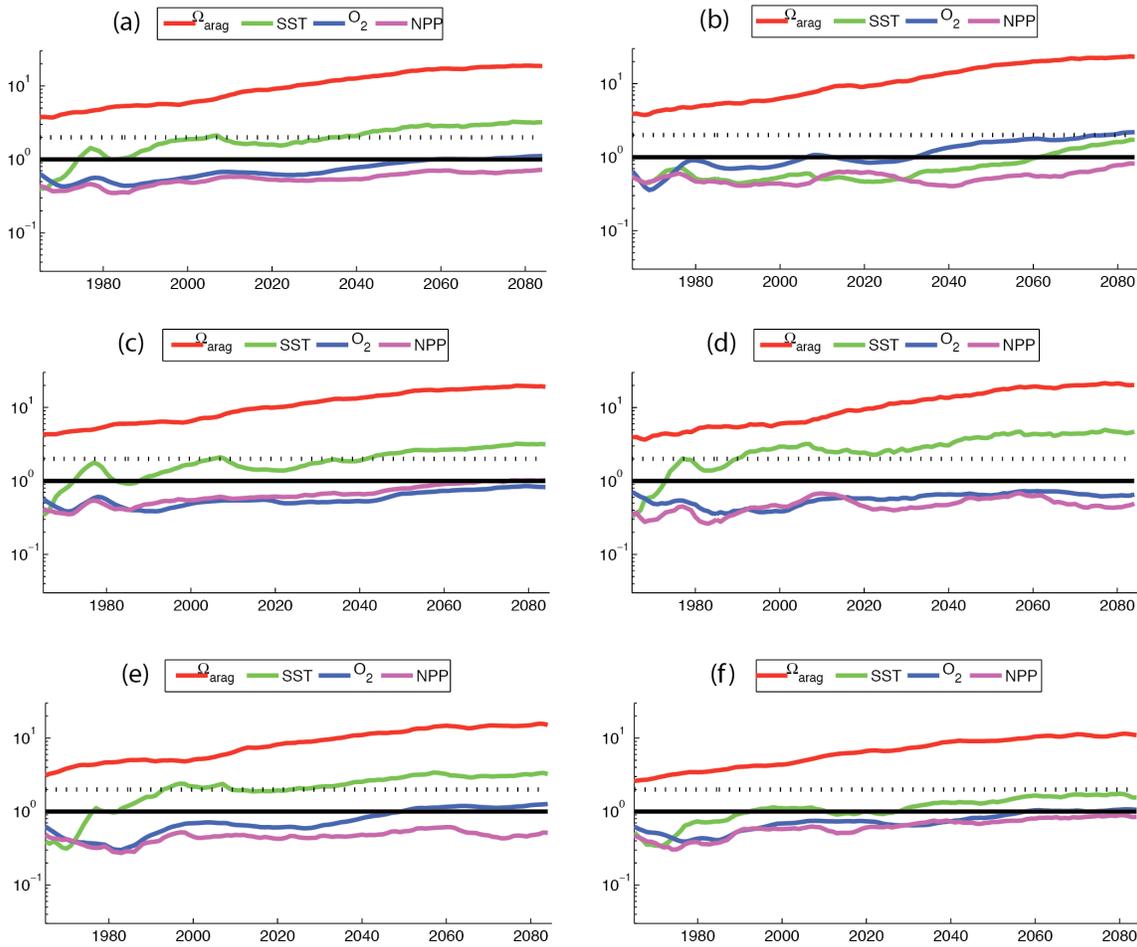
15 **Wittenberg, A.T.: Are historical records sufficient to constrain ENSO simulations *Geophys.*
16 *Res. Lett.*, 36, L12702, doi:10.1029/2009GL038710, 2009.**

17 Wittenberg, A.T., Rosati, A., Delworth, T.L., Vecchi, G.A., and Zeng, F.: ENSO Modulations:
18 Is It Decadally Predictable? *Journal of Climate*, 27, 2667-2681, 2014.

19 Zhang, Y., Wallace, J.M., and Battisti, D.S.: ENSO-like Interdecadal Variability: 1900-93,
20 *Journal of Climate*, 10, 1004-1020, 1997.

21

1



2

3

4

5

6

7

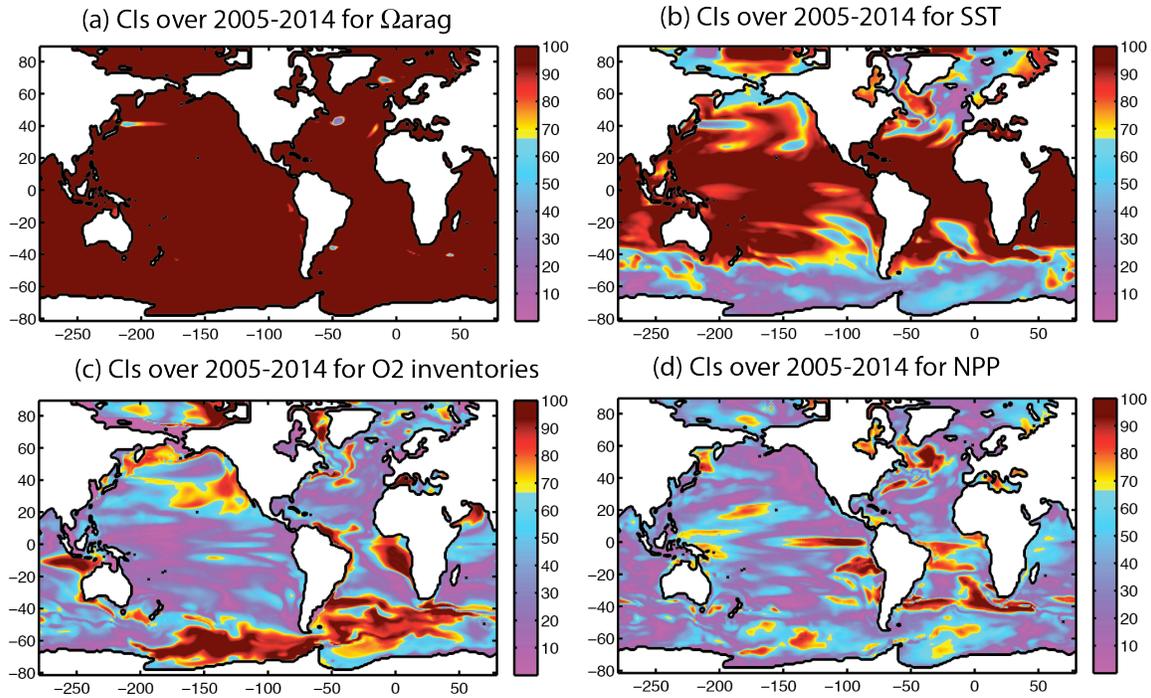
8

9

10

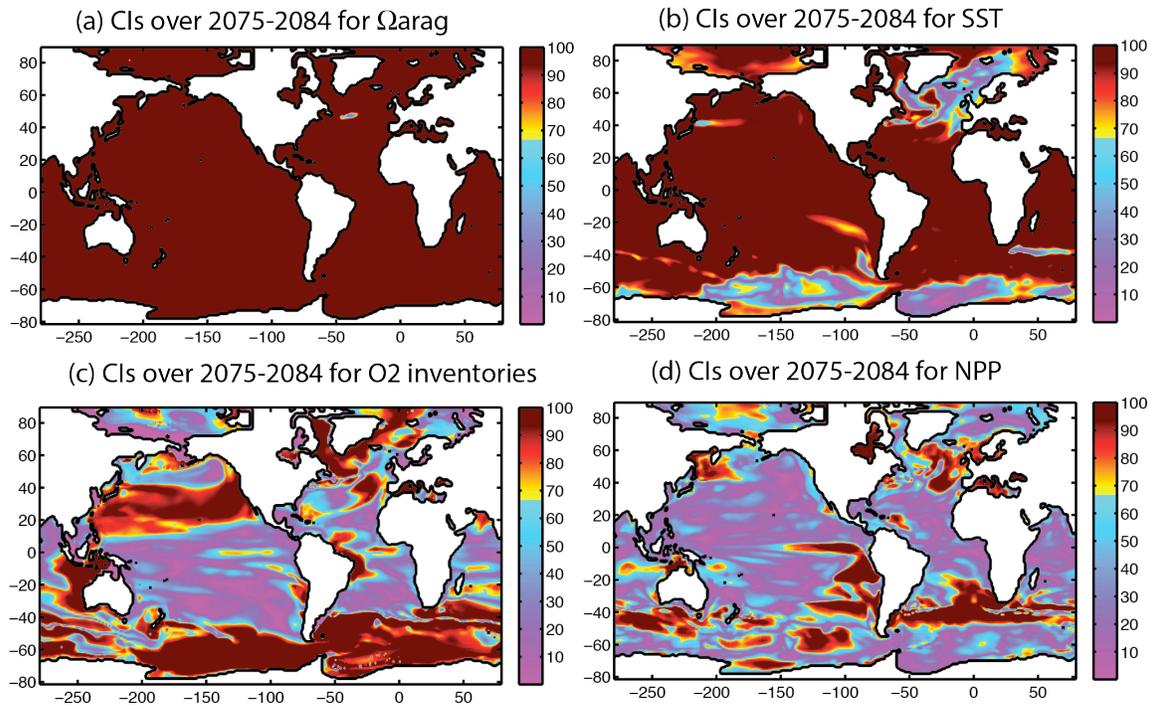
11

Figure 1. Time series of (area weighted) averages of signal-to-noise ratio (SNR) for four marine ecosystem drivers, considered over a number of ocean regions: (a) Global, (b) 90°S-45°S, (c) 45°S-15°S, (d) 15°S-15°N, (e) 15°N-45°N, and (f) 45°N-90°N. The four drivers are Ω_{arag} , SST, O_2 inventories, and NPP. For each driver, trends have been calculated individually on a gridpoint-by-gridpoint basis using a 30 yr trend window, and the SNR was subsequently calculated. The vertical axis has a logarithmic scale (non-dimensional) representing the SNR.. The 1-standard deviation threshold is shown as a solid horizontal line, and the 2-standard deviation threshold is shown as a dashed line in each panel.



1
2 **Figure 2.** Confidence interval maps averaged over 2005-2014 for (a) Ω_{arag} , (b) SST, (c) O_2
3 inventories, and (d) NPP. For the case of O_2 inventories in panel (c), shelf regions where the
4 ocean depth is less than 600m deep are not included in the analysis. For each case, a 30-year
5 window has been used to calculate trends gridpoint-by-gridpoint for each year between 2005
6 and 2014. An average over 10 years was considered to remove shorter timescale fluctuations in
7 the signal-to-noise ratio. Note that the color scheme here is chosen such that saturation occurs
8 (maroon color) above the 67% confidence interval. Warm colors indicate confidence intervals
9 ranging from 67-95% (one to two standard deviations), and cool colors span the range 0%-67%
10 (less than one standard deviation).

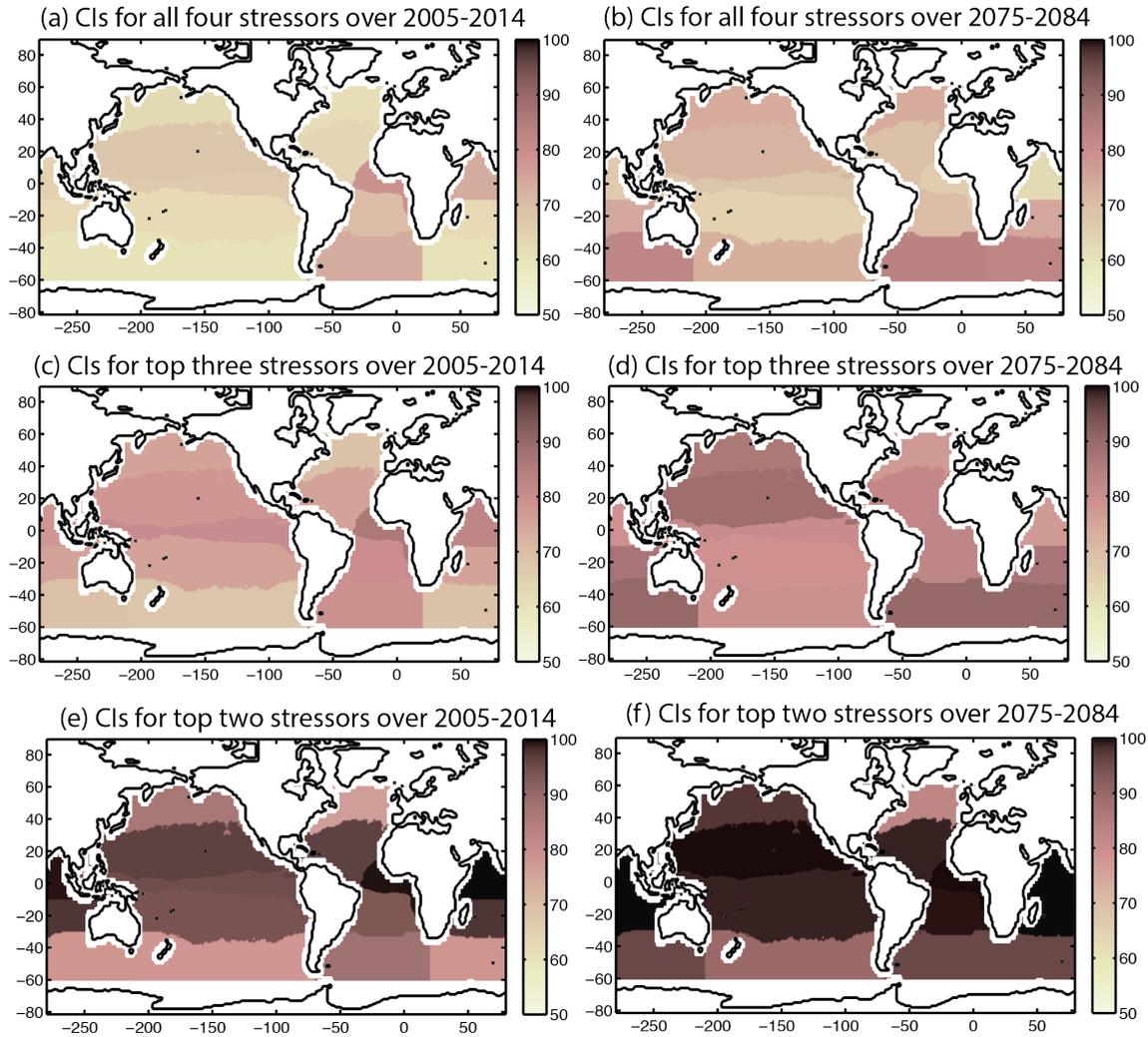
11



1

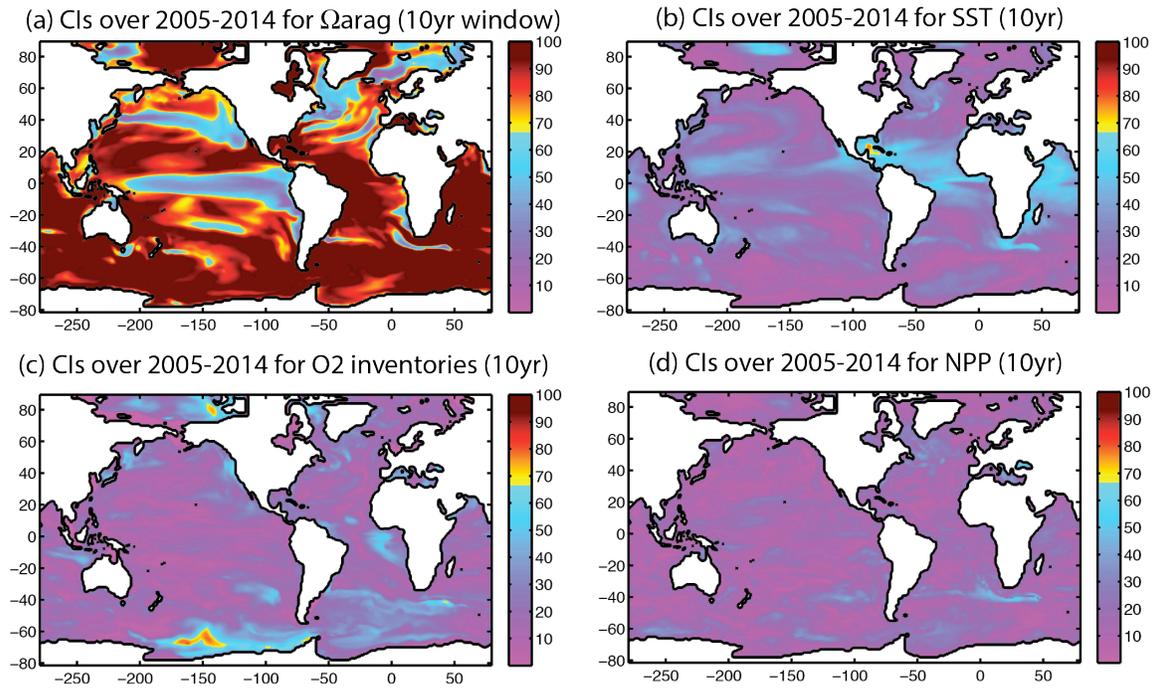
2 **Figure 3.** Same as **Fig. 2**, but averaged over the 2075-2084 period.

3



1
2 **Figure 4.** Confidence intervals for two ten-year intervals, namely 2005-2014 (left column), and
3 2075-2084 (right column). The confidence intervals for (a) the four drivers (Ω_{arag} , SST, O_2
4 inventories, and NPP) over 2005-2014 are taken as the average over the fields shown in **Fig. 2**,
5 and (b) the confidence intervals shown for the four drivers over 2075-2084 are taken as the
6 average of the fields shown in **Fig. 3**. This is then considered for the case of the dominant three
7 drivers in (c) for the period 2005-2014, and in (d) for the period 2075-2084. Finally, the
8 confidence intervals averaged for the dominant two drivers are shown in (e) over 2005-2014 and
9 (f) over 2075-2084.

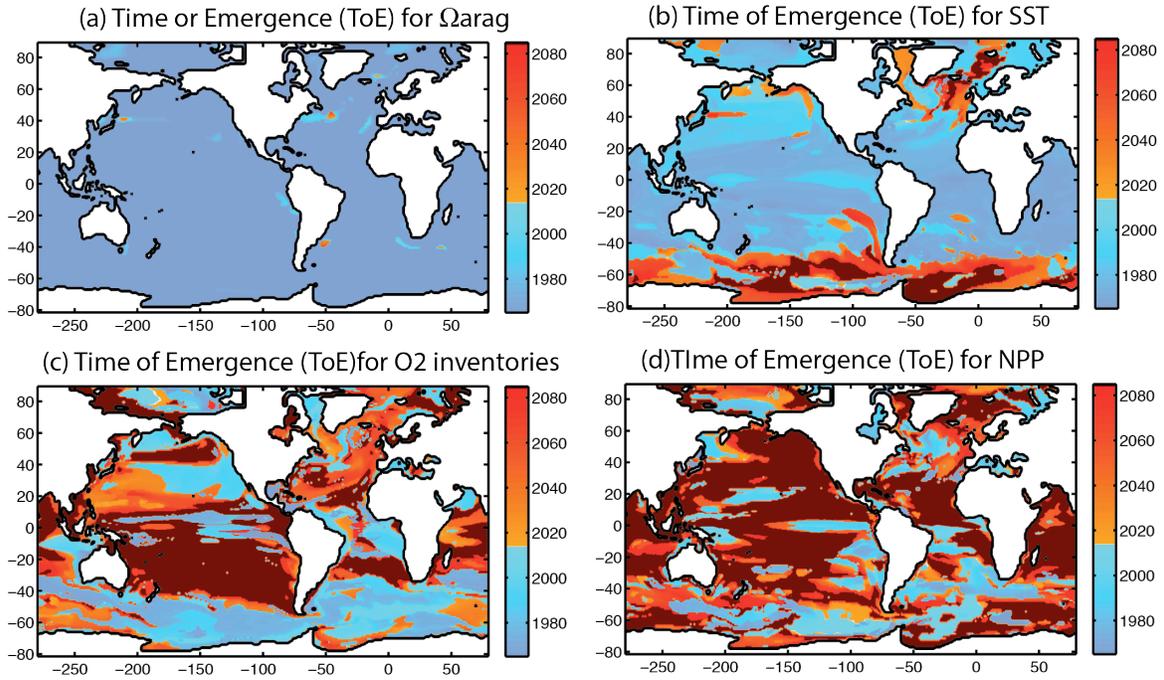
1



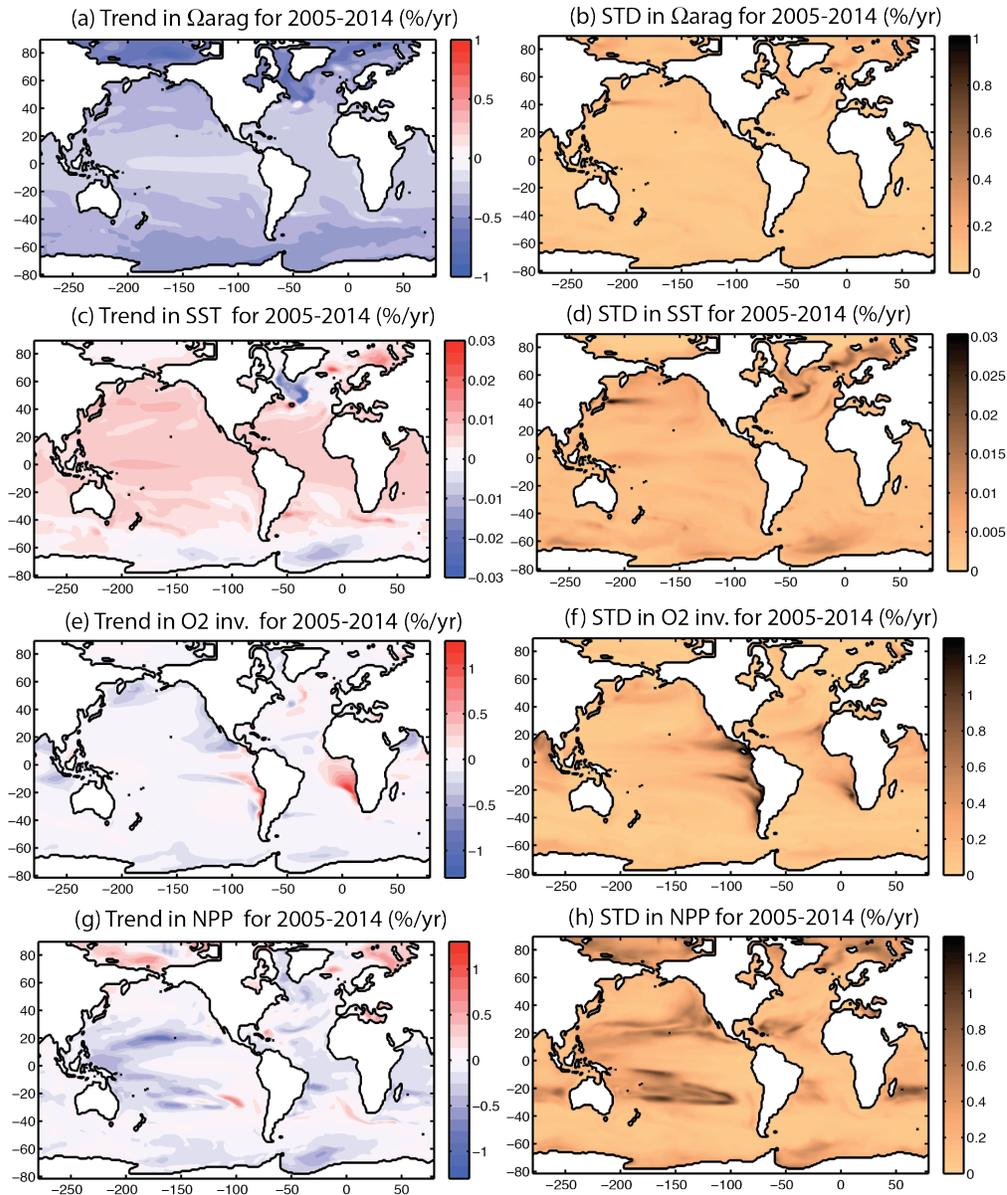
2

3 **Figure 5.** The confidence intervals for emergence of each of the drivers for the time period 2005-
4 2014, using a 10 yr window for calculating trends. The analysis here is otherwise identical to
5 that shown in **Fig. 2**, except that there a 30 yr window was chosen. The panels show the
6 distributions for (a) Ω_{arag} , (b) SST, (c) O₂ inventories, and (d) NPP. This analysis reveals
7 significantly lower confidence intervals for the 10 yr window than for the 30 yr window for each
8 of the four drivers.

9

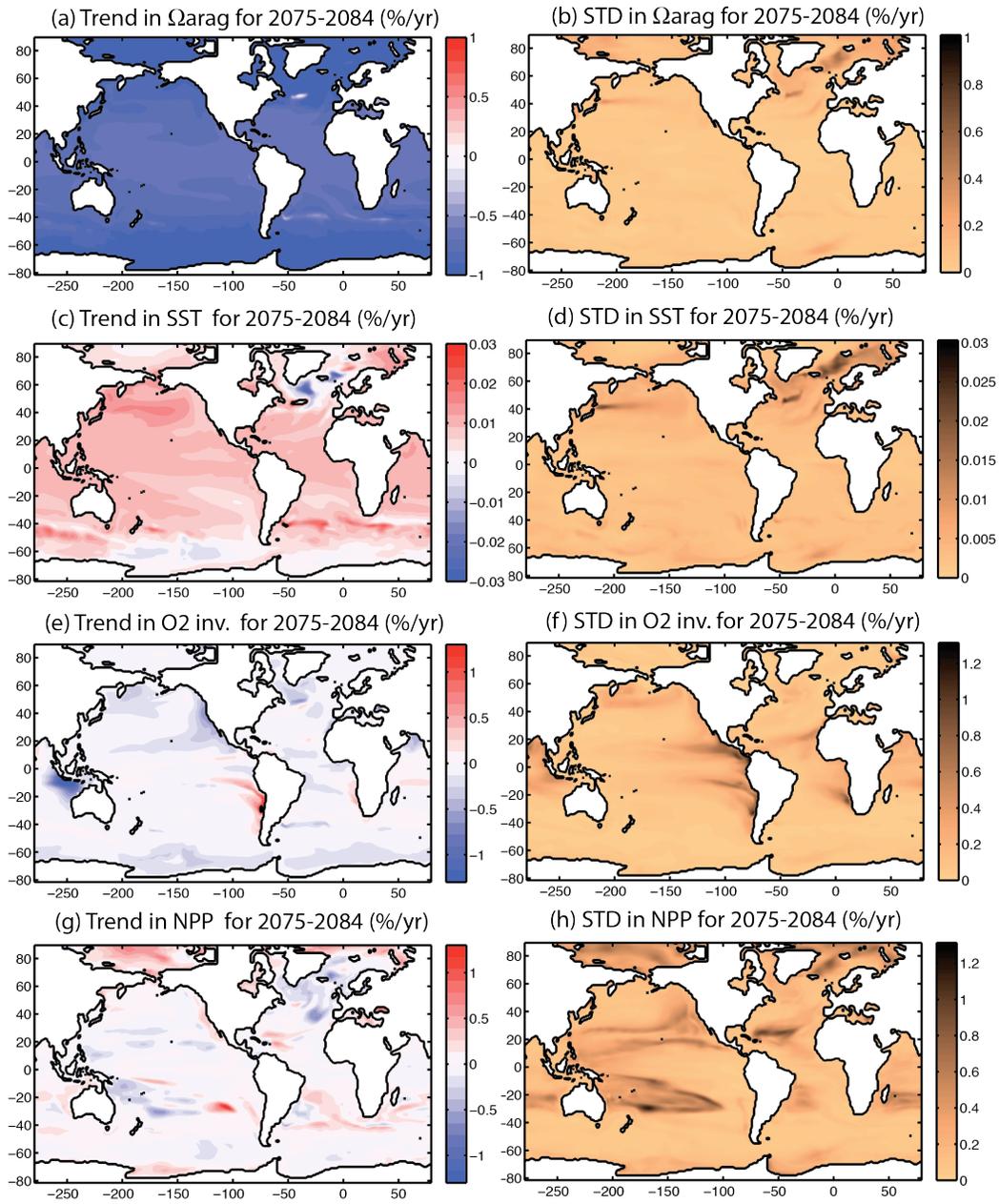


1
 2 **Figure 6.** Time of Emergence (ToE), calculated using a threshold of 1 standard deviation (67%
 3 confidence) for (a) Ω_{arag} , (b) SST, (c) O₂ inventories, and (d) NPP. The color scheme has been
 4 chosen to distinguish between relative to the present (2014) with warm colors indicating a ToE
 5 post-2014 and cold colors indicating a ToE pre-2014.
 6



1
 2 **Figure A1.** Linear trends (left) and standard deviations of the linear trends (right) for (a-b) Ω_{arag} ,
 3 (c-d) SST, (e-f) O₂ inventories from 100m-600m, and (g-h) net primary production averaged over
 4 2005-2014. All fields are calculated using 30-year trend windows, and the trends are shown in
 5 units of %/year.

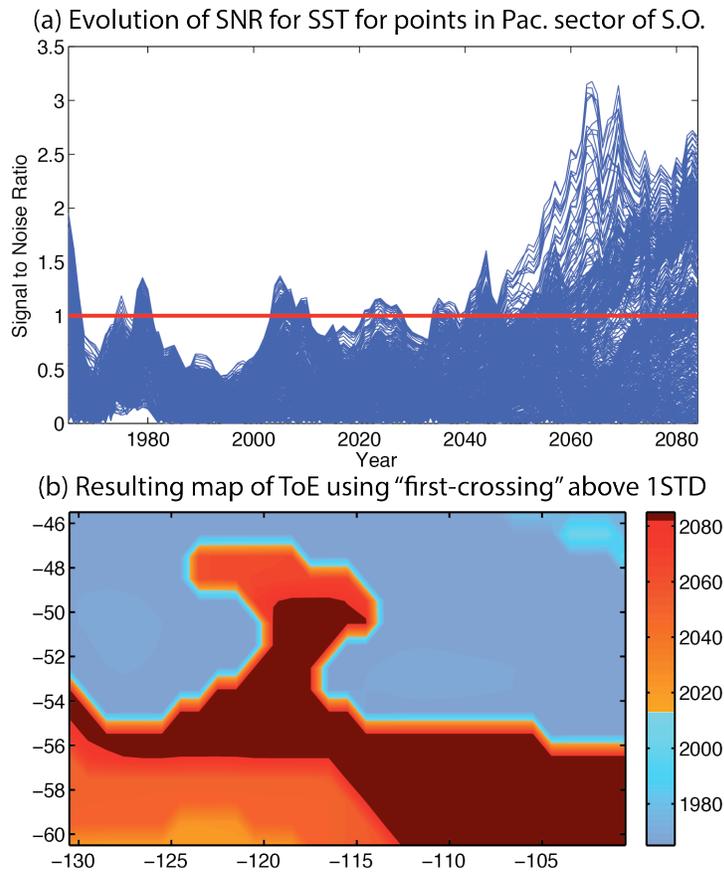
6



1

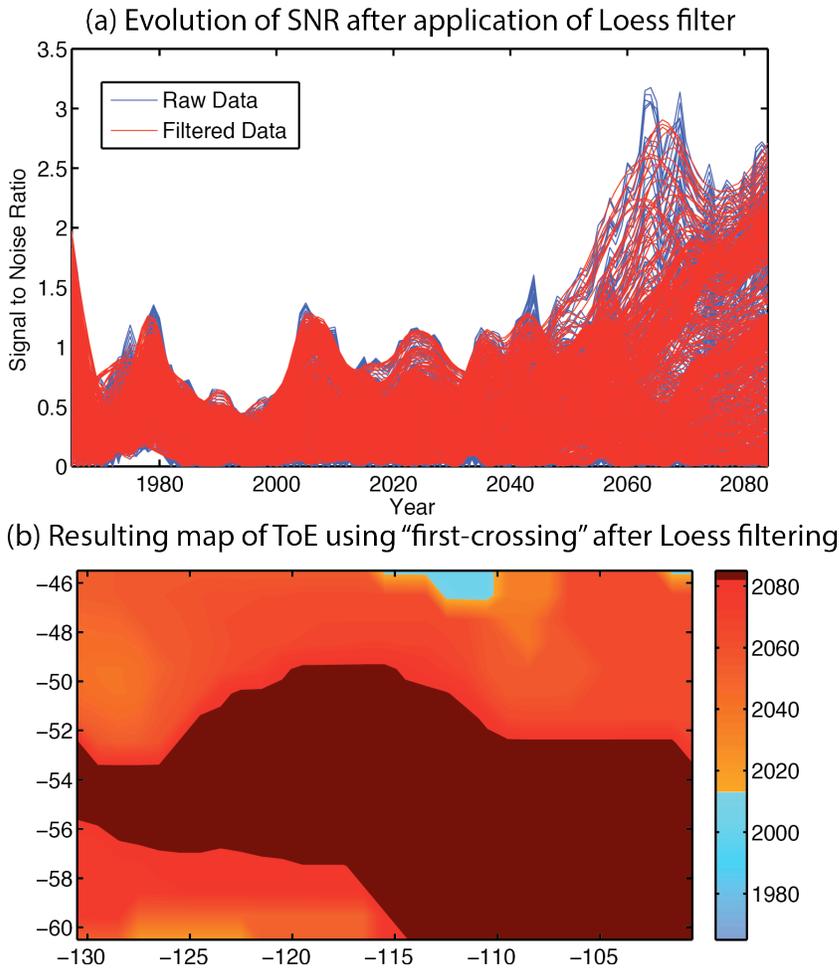
2 **Figure A2.** Same as **Fig. A1**, but averaged over 2075-2084.

3



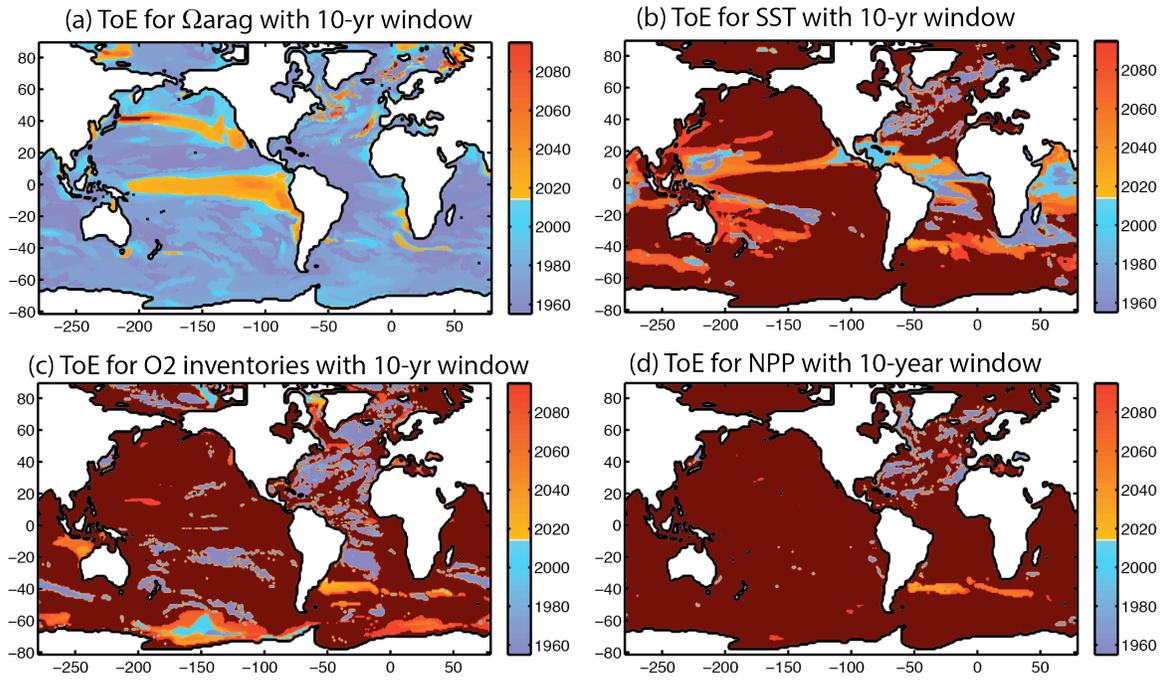
1
 2 **Figure A3.** Time of emergence of SST considered for a boxed region of the Pacific sector of the
 3 Southern Ocean (130°W-100°W and 45°S-60°S). The SNR calculated individually using annual
 4 mean SST for each gridpoint in the domain is shown in panel (a), where a 30-year window is
 5 used to calculate TREND and NOISE. Short timescale excursions of less than ten years are in
 6 evidence in modulations of maximum SNR. The spatial pattern of ToE for SST over this region
 7 is shown in panel (b). A ToE before the present time (pre-2014) is indicated by cool (blue)
 8 colors, while a later ToE is indicated with warm (orange) colors. Saturation (ToE post-2085) is
 9 represented with maroon color.

10
 11



1
 2 **Figure A4.** The ToE is considered for the same region, but this time with application of a robust
 3 Loess filter with a tolerance window of ten years applied to the SNR of SST calculated for the
 4 individual surface grid points in the domain. The SNR after application of the robust Loess filter
 5 is shown in panel (a) in red, superposed over the same blue time series considered in **Fig. A3a**.
 6 The filtering effect on the short timescale maximum excursions of the SNR is evident in the Fig.
 7 The net effect over the full domain of the robust Loess filter is shown in panel (b), revealing a
 8 later ToE over significant portions of the region of interest.
 9

1

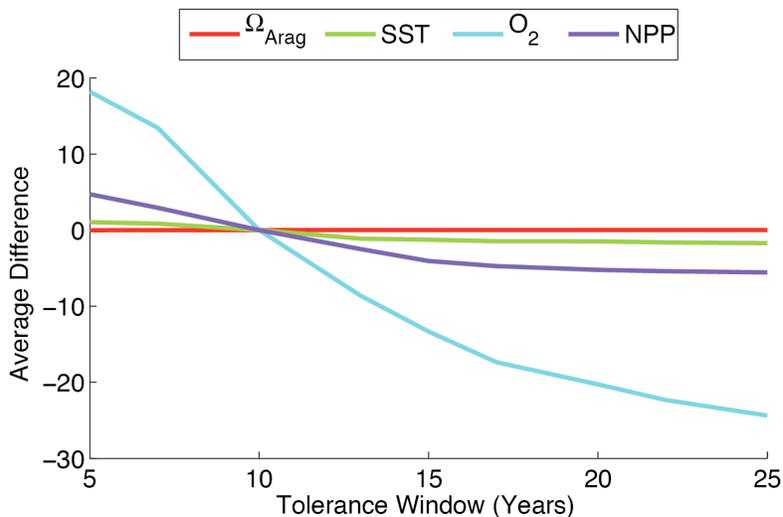


2

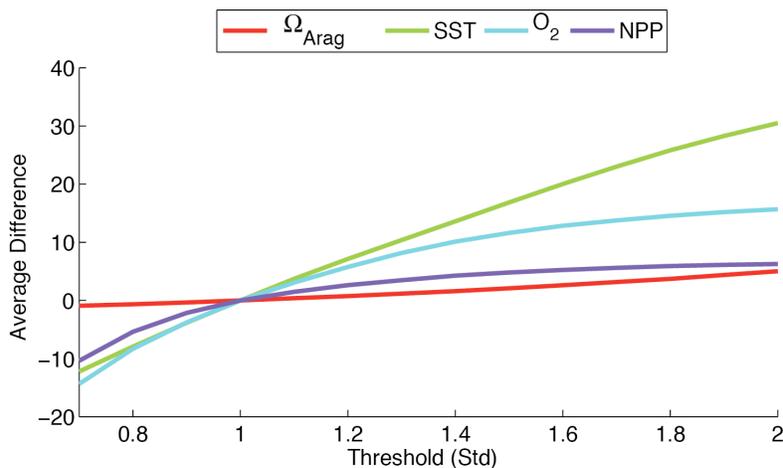
3 **Figure A5.** The Time of Emergence (ToE) for each of the four drivers has been calculated using
4 a 10-year window for the calculation of trends, considering the time interval 1955-2095. This is
5 the complement to the results with ToE for a 30 yr window, shown in **Fig. 6**.

6

(a) Sensitivity of TOE to tolerance window width using 1 STD threshold



(b) Sensitivity of ToE to choice of SNR threshold



1
2 **Figure A6.** Sensitivity analysis for tolerance window width in years (left) -- using a one standard
3 deviation threshold and revolved around a 10 yr window, as well as threshold level (right) --
4 using a 10 yr tolerance window and revolved around a standard deviation of one. Caveats
5 regarding averaging over fields that experience saturation are discussed in the text of the
6 Supplementary Materials.

7