Increase in ocean acidity variability and extremes under increasing atmospheric CO_2

Friedrich A. Burger[1,2], Thomas L. Frölicher[1,2], and Jasmin G. John[3]

[1]Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland.
[2]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland.
[3]NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA.

Correspondence: Friedrich A. Burger (friedrich.burger@climate.unibe.ch)

Abstract. Ocean acidity extreme events are short-term periods of extremely high $[H^+]$ concentrations. The uptake of anthropogenic CO_2 emissions by the ocean is expected to lead to more frequent and intense ocean acidity extreme events, not only due to mean ocean acidification, but also due to increases in ocean acidity variability. Here, we use daily output from ensemble simulations of a comprehensive Earth system model under a low and high CO_2 emission scenario to isolate and quantify the

5   impact of changes in variability on changes in ocean acidity extremes. We show that the number of days with extreme $[H^+]$ conditions for surface waters is projected to increase by a factor of 14 by the end of the 21st century under a high CO_2 emission scenario relative to preindustrial levels. The duration of individual events is projected to triple, and the maximal intensity and the volume extent in the upper 200 m to quintuple. Similar changes are projected in the thermocline. At surface, the changes are mainly driven by increases in $[H^+]$ seasonality, whereas changes in interannual variability are also important in the ther-

10 mocline. Increases in $[H^+]$ variability and extremes arise predominantly from increases in the sensitivity of $[H^+]$ to variations in its drivers. In contrast to $[H^+]$ extremes, the occurrence of short-term extremes in low aragonite saturation state due to changes in variability is projected to decrease. An increase in $[H^+]$ variability and an associated increase in extreme events superimposed onto the long-term ocean acidification trend will enhance the risk of severe and detrimental impacts on marine organisms, especially for those that are adapted to a more stable environment.

15 # 1 Introduction

Since the beginning of the industrial revolution, the ocean has absorbed about a quarter of the carbon dioxide (CO_2) released by human activities (Friedlingstein et al., 2019). Oceanic uptake of anthropogenic CO_2 mitigates global warming by reducing atmospheric CO_2, but also leads to major changes in the chemical composition of seawater through acidification (Gattuso and Buddemeier, 2000; Caldeira and Wickett, 2003; Orr et al., 2005; Doney et al., 2009). When CO_2 dissolves in seawater, it forms

20 carbonic acid that dissociates into bicarbonate ($[HCO_3^-]$) and carbonate ions ($[CO_3^{2-}]$), releasing hydrogen ions ($[H^+]$) and thereby reducing pH (pH = -log($[H^+]$)). The rise in $[H^+]$ is partially buffered by the formation of $[HCO_3^-]$, thereby decreasing $[CO_3^{2-}]$ and the calcium carbonate saturation state (Ω) that describes whether water is supersaturated or undersaturated with respect to calcium carbonate. Over the last four decades the surface ocean pH has declined by about 0.02 pH units per decade (Bindoff et al., 2019). Continued carbon uptake by the ocean will further exacerbate ocean acidification in the near future

25 (Caldeira and Wickett, 2003; Bindoff et al., 2019) with potential major consequences for marine life (Doney et al., 2009) and ocean biogeochemical cycling (Gehlen et al., 2012).

Superimposed onto the long-term ocean acidification trend are short-term extreme events, during which ocean pH or the calcium carbonate saturation state are extremely low (Hofmann et al., 2011; Joint et al., 2011; Hauri et al., 2013). These
30 events may be driven by ocean mixing processes, biological production and remineralization, mineral dissolution, temperature changes, air-sea gas exchange or a combination thereof (Lauvset et al., 2020). In eastern boundary upwelling systems, for example, short-term upwelling events and mesoscale processes can lead to low surface pH events and to short-term shoaling of the saturation horizon (Feely et al., 2008; Leinweber and Gruber, 2013). Ocean pH also rapidly changes as a consequence of microbial activity (Joint et al., 2011). Phytoplankton blooms and accompanying respiration drastically increase the partial
35 pressure of CO_2 (pCO_2) and reduce pH in the thermocline (Sarmiento and Gruber, 2006). Such extreme events may have pH levels that are much lower than the mean pH conditions projected for the near future.

The vast majority of scientific literature on ocean acidification has focused so far on gradual changes in the mean state in ocean chemistry (e.g. Orr et al. (2005); Bopp et al. (2013); Frölicher et al. (2016)). However, to understand the full conse-
40 quences of ocean acidification for marine organisms and ecosystem services, it is necessary to understand how ocean acidity extremes change under increasing atmospheric CO_2 (Kroeker et al., 2020). The ability of marine organisms and ecosystems to adapt to ocean acidification may depend on whether these species have evolved in a chemically stable or a highly variable environment (Hofmann et al., 2011). If the frequency and intensity of short-term ocean acidity extreme events strongly increases, one might expect that some organisms may exhibit difficulties to adapt, especially if carbonate chemistry crosses some critical
45 thresholds. Key plankton species such as coccolithophores (Riebesell et al., 2000), foraminifera and pteropods (Bednaršek et al., 2012) were found to be negatively impacted by low carbonate concentrations. After only several days of undersaturation, some species such as pteropods already show reduced calcification, growth and survival rates (Bednaršek et al., 2014; Kroeker et al., 2013). Carbonate system variability also plays a role in shaping the diversity and biomass of benthic communities (Kroeker et al., 2011; Hall-Spencer et al., 2008). In laboratory experiments, some deep water corals undergo a decline
50 in calcification for low-pH conditions over a week, while recovery may be possible when the low-pH condition persists for six months, stressing the importance of high-frequency variability and short-term acidification events (Form and Riebesell, 2012). There is also evidence that the organism response to variability in ocean acidity could change with ocean acidification (Britton et al., 2016). Understanding the temporal variability of ocean carbonate chemistry and the changes therein is therefore of critical importance for understanding the impacts of ocean acidification on marine organisms and ecosystems (Hofmann et al.,
55 2011).

Changes in extremes arise from changes in the mean, variability, or shape of the probability distribution (Coles, 2001). Under ongoing long-term ocean acidification (i.e. changes in the mean), one can expect that extreme events in [H$^+$] and Ω may become more frequent and intense (Hauri et al., 2013). In addition to the changes in the mean, recent studies suggest that

60 the seasonal cycles in [H$^+$] and Ω are also strongly modulated under elevated atmospheric CO$_2$ due to non-linear carbonate chemistry processes (Kwiatkowski and Orr, 2018; Fassbender et al., 2018; Gallego et al., 2018; Landschützer et al., 2018; McNeil and Sasse, 2016; Rodgers et al., 2008; Hauck and Völker, 2015). Over the 21st century and under a high greenhouse gas emission scenario, Earth system model simulations suggest that the winter-summer difference in surface [H$^+$] is projected to increase by 81%, whereas the seasonal amplitude for aragonite saturation state (Ω_A) is projected to decrease by 9 % on global

65 average (Kwiatkowski and Orr, 2018). Recent observational-based estimates as well as theoretical arguments support these projected increases in seasonality for [H$^+$] and pCO$_2$ (Landschützer et al., 2018; Fassbender et al., 2018). We can therefore expect that changes in variability may also impose changes in the frequency and intensity of extreme events in ocean acidity.

Unlike for marine heatwaves and extreme sea level events (Frölicher et al., 2018; Collins et al., 2019; Oppenheimer et al.,

70 2019), little is known about the characteristics and changes of extreme ocean acidity events and if so, only on seasonal timescales. A global view of how extreme events in ocean chemistry will unfold in time and space and a mechanistic understanding of the relevant processes is currently missing. This knowledge gap is of particular concern as it is expected that ocean acidity extreme events are likely to become more frequent and intense under increasing atmospheric CO$_2$. Given the potential for profound impacts on marine ecosystems, quantifying trends and patterns of ocean acidity extreme events is a pressing issue.

75

In this study, we use daily output of a five-member ensemble simulation under a low and a high CO$_2$ emission scenario with a comprehensive Earth system model to investigate how changes in interannual, seasonal, and residual daily variability under rising atmospheric CO$_2$ levels imprint on the occurrence, intensity, duration and volume of [H$^+$] and Ω extreme events. The remainder of this paper is organized as follows: Section 2 presents the coupled carbon-climate Earth system model,

80 the ensemble simulations, analysis methods, and a brief model evaluation. The global and regional changes in ocean acidity extremes and their drivers are examined in Sections 3.1-3.4. The discussion and conclusions are given in Section 4.

2 Methods

2.1 Model & experimental design

The simulations used in this study are performed with the CMIP5-generation fully coupled carbon-climate Earth system model

85 developed at the NOAA Geophysical Fluid Dynamics Laboratory (GFDL ESM2M) (Dunne et al., 2012, 2013). GFDL-ESM2M consists of ocean, atmosphere, sea ice, and land models, and includes land and ocean biogeochemistry. The ocean component is the Modular Ocean Model version 4p1 (MOM4p1), with a nominal 1° horizontal resolution increasing to 1/3° meridionally at the equator, with a tripolar grid north of 65°N, and with 50 vertical depth levels (Griffies, 2009). The dynamical sea ice model uses the same tripolar grid as MOM4p1 (Winton, 2000). The Atmospheric Model version 2 (AM2) has a horizontal resolution

90 of 2° × 2.5° with 24 vertical levels (Anderson et al., 2004). The Land Model version 3 (LM3) simulates water, energy, and carbon cycles dynamically and uses the same grid as AM2 (Shevliakova et al., 2009).

The ocean biogeochemical and ecological component is version two of the Tracers of Ocean Phytoplankton with Allometric Zooplankton (TOPAZv2) module that parametrizes the cycling of carbon, nitrogen, phosphorus, silicon, iron, oxygen, alkalin-

95   ity, lithogenic material, and surface sediment calcite (see supplementary material in Dunne et al. (2013)). TOPAZv2 includes three explicit phytoplankton groups: small, large, and diazotrophs, and one implicit zooplankton group. The ocean carbonate chemistry is based on the OCMIP2 parametrizations (Najjar and Orr, 1998), with the dissociation constants for carbonic acid and bicarbonate ions from Mehrbach et al. (1973) and the carbon dioxide solubility from Weiss (1974). TOPAZv2 also simulates diurnal variability in ocean physics as well as in phytoplankton growth. While diurnal variations in open ocean pH

100 are therefore simulated to some extent, we do not expect the model to fully capture the high diurnal variability in seawater chemistry in coastal regions with large biological activity (Kwiatkowski et al., 2016; Hofmann et al., 2011).

We performed a 220 year spin-up simulation at prescribed preindustrial CO_2 concentration started from a quasi-equilibrated 1000 year preindustrial control simulation to ensure stability of the model under a new computing infrastructure. We then ran a

105 five-member ensemble simulation covering the historical 1861-2005 period, followed by a high (RCP8.5; RCP: Representative concentration pathway) and a low greenhouse gas emission scenario (RCP2.6) over the 2006-2100 period with prescribed atmospheric CO_2 concentrations. RCP8.5 is a high emission scenario without effective climate policies, leading to continued and sustained growth in greenhouse gas emissions (Riahi et al., 2011). In GFDL ESM2M, atmospheric surface temperature in the RCP8.5 ensemble is projected to increase by 3.24 (ensemble minimum: 3.17 - ensemble maximum: 3.28) °C between

110 preindustrial and 2081-2100. The RCP2.6 scenario represents a low emission and high mitigation future (van Vuuren et al., 2011). The RCP2.6 ensemble simulated here undergoes an atmospheric warming of 1.21 (1.18-1.26) °C by the end of the 21st century relative to preindustrial levels. The five ensemble members over the historical period were initialized from an extension of the spin-up simulation, that includes historical land-use over the 1700-1860 period (Sentman et al., 2011). The five ensemble members were generated by adding different very small SST disturbances of the order 10^{-5} K to a surface grid

115 cell in the Weddell sea at 70.5° S, 51.5° W on January 1st 1861 (Wittenberg et al., 2014; Palter et al., 2018). Although the ocean biogeochemistry is not perturbed directly, $[H^+]$ and Ω differences between the ensemble members spread rapidly over the globe. On average, the ensemble members can be regarded as independent climate realizations after about three years of simulation for surface waters and about eight years at 200 m depth (Frölicher et al., 2020). Neither the choice of the perturbation location nor the choice of the perturbed variable has a discernible effect on the results presented here (Wittenberg et al., 2014).

120 In addition, an accompanying 500 year long preindustrial control simulation with potential vegetation was performed that extends the 220 year long spin-up simulation.

2.2   Analysis methods

2.2.1   Extreme event definition and characterization

We analyze daily data of $[H^+]$ and Ω in the upper 200 m of the water column. In this study, we focus on the aragonite

125 saturation state Ω_A. The calcium carbonate saturation state Ω differs between different mineral forms of calcium carbonate,

such as aragonite and calcite, that differ in their solubilities. Aragonite is the most soluble form and important for many calcifying organisms such as pteropods (Bednaršek et al., 2012). We define an event as a $[H^+]$ extreme event when the daily $[H^+]$ exceeds the 99th percentile, i.e. a one-in-a-hundred days event. Similarly, we define a Ω_A extreme event when the daily Ω_A falls below the 1st percentile. The percentiles are calculated for each grid cell from daily data of the 500-year long preindustrial
130   control simulation. In contrast to absolute thresholds, relative thresholds, such as those used here, allow the characterization of extreme events over regions with different statistical properties. In addition, biases in the simulated variables already alter the definition of relative thresholds and should thus have a smaller effect on projections of changes in extreme events based on these thresholds compared to projections based on absolute thresholds (see also Frölicher et al. (2018)).

We calculate four extreme event metrics: (a) the number of extreme event days per year (in days; number of days above the
135   99th percentile for $[H^+]$ and below the 1st percentile for Ω_A), (b) the annual mean duration (in days; the average number of days above the 99th percentile for $[H^+]$ and below the 1st percentile for Ω_A of single events within a year), (c) the annual mean maximal intensity (in $nmol\,kg^{-1}$ or Ω_A unit; maximum $[H^+]$ or Ω_A anomalies with respect to the percentile threshold over the duration of a single ocean acidification extreme event and then averaged over all events within a year), and (d) the mean volume covered by individual extreme events in the upper $200\,meters$ (in km^3; mean of clusters of connected grid cells that are
140   above the 99th percentile for $[H^+]$ or below the 1st percentile for Ω_A, calculated using the *scikit-image* library for *Python* for each day, these daily means are then averaged annually). The number of days, duration, and maximal intensity are calculated for individual grid cells at the surface and at $200\,m$ depth. While the truncation of extremes between years alters the results for duration and intensity, it allows for the calculation of annual extreme event characteristics. We focus our analysis not only on the surface, but also on $200\,m$ depth to study changes in extreme events within the thermocline. In Section 3.1, these grid cell
145   based characteristics are then aggregated globally.

The aim of this study is to assess how changes in $[H^+]$ and Ω_A variability lead to changes in different extreme event characteristics. Therefore, we isolate the effect of changes in variability by subtracting the secular trends at each grid cell and in each individual ensemble member prior to the calculation of the different extreme event characteristics (Figure 1). The secular trend is calculated as the five-member ensemble mean, which has been additionally smoothed with a 365-day running mean to
150   keep the seasonal signal in the data (further information in Appendix A). This calculation process ensures that day-to-day to interannual variability can change over the simulation period while the mean state stays approximately constant, as depicted for one grid cell in Figure 1. Thus, changes in the different extreme event characteristics are only caused by changes in variability.

2.2.2   Decomposition of $[H^+]$ variability into different components

In order to assess whether changes in low or high frequency variability cause changes in extreme events and their characteristics, we decompose the total variability in $[H^+]$ into interannual, seasonal, and residual daily variability. First, we calculate the climatological seasonal cycle from the daily data by averaging each calendar day over all years in the time period of interest. Seasonal variability is then identified with the time-series variance and standard deviation of this 365-day long seasonal cycle.
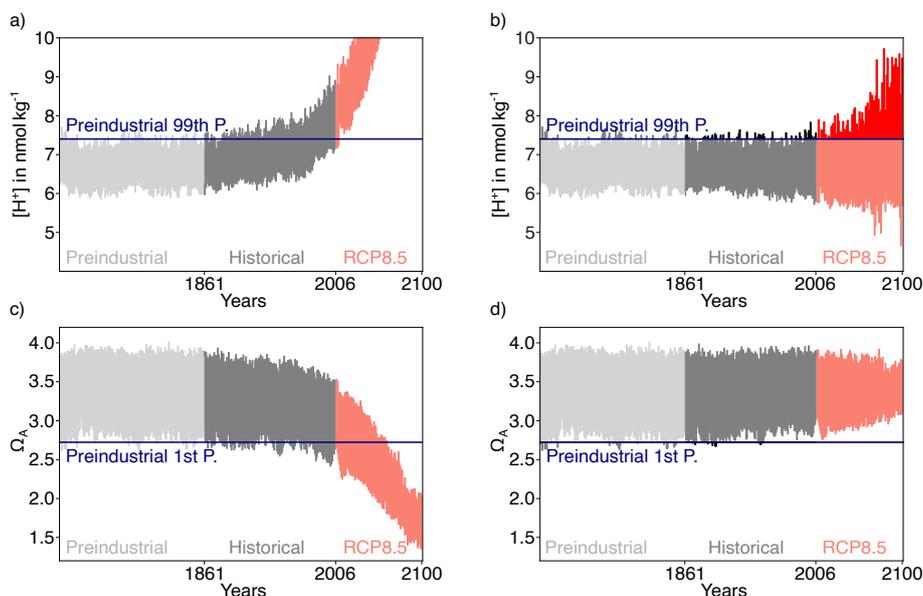
Figure 1. Simulated daily surface [H$^+$] (a) and Ω_A (c) at 40°N and 30°W in the North Atlantic for one ensemble member over the prein-dustrial, the 1861-2005 historical period, and the 2006-2100 period under RCP8.5. (b,d) Same as (a,c), but the ensemble mean has been subtracted. For [H$^+$], the preindustrial 99$^\mathrm{th}$ percentile threshold (horizontal blue line in panels a and b) is increasingly exceeded even when subtracting the ensemble mean change, because [H$^+$] variability increases. In contrast, a reduction in Ω_A variability leads to a reduced undershooting of the preindustrial 1$^\mathrm{st}$ percentile (panel d).

160 As described above, the secular trend in the daily data has been removed with the five-member ensemble mean before doing the analysis. In a next step, we subtract the seasonal cycle from the data and estimate the spectral density (Chatfield, 1996) of this residual time series using the *periodogram* function from the *scipy* python library. From the spectral density we then calculate the variance and standard deviation arising from variations on interannual and sub-annual timescales to obtain interannual and residual daily variability (further information is given in Appendix B). In section 3.3, we use the variance to report the

165 contributions from interannual, seasonal, and residual daily variability to global-mean variability change, but for the spatial changes in Figure 7, we use the standard deviation.

2.2.3   Taylor deconvolution method to identify mechanistic controls of [H$^+$] variability changes

To understand the processes behind the simulated changes in [H$^+$] variability and extremes, we decompose the changes into contributions from changes in temperature (T), salinity (S), total alkalinity (A$_\mathrm{T}$), and total dissolved inorganic carbon (C$_\mathrm{T}$).

170 Assuming linearity, the difference of [H$^+$] from its mean at time step i can be decomposed into contributions from the drivers by employing a first order Taylor expansion:

$$H^+(i) - \overline{H}^+ \simeq \frac{\partial H^+}{\partial C_T}\bigg|_{\overline{C_T},\overline{A_T},\overline{T},\overline{S}} \left(C_T(i) - \overline{C_T}\right) + \frac{\partial H^+}{\partial A_T}\bigg|_{\overline{C_T},\overline{A_T},\overline{T},\overline{S}} \left(A_T(i) - \overline{A_T}\right)$$
$$+ \frac{\partial H^+}{\partial T}\bigg|_{\overline{C_T},\overline{A_T},\overline{T},\overline{S}} \left(T(i) - \overline{T}\right) + \frac{\partial H^+}{\partial S}\bigg|_{\overline{C_T},\overline{A_T},\overline{T},\overline{S}} \left(S(i) - \overline{S}\right). \tag{1}$$

The partial derivatives are evaluated at $\overline{T}$, $\overline{S}$, $\overline{C_T}$, and $\overline{A_T}$, the temporal mean values of the drivers in the period of interest. While it is important to take into account the climatological total phosphate and total silicate concentrations for calculating the partial derivatives (Orr and Epitalon, 2015), one introduces only small errors by neglecting variations in phosphate and silicate.

175 The partial derivatives in Equation 1 are evaluated using the *Mocsy 2.0* package (Orr and Epitalon, 2015).

Using the Taylor decomposition (Equation 1), one can for example express the seasonal variation in [H$^+$] as a function of the drivers' seasonal variations (Kwiatkowski and Orr, 2018). In this study however, we analyze the time-series standard deviation of [H$^+$] that also includes variability on other time scales (see Section 2.2.2) and the drivers of its changes. By making the Taylor approximation (Equation 1) and from the definition of variance (the squared standard deviation, e.g. Coles (2001)), it

180 follows that the standard deviation of [H$^+$] can be written as a function of the partial derivatives with respect to the drivers (sensitivities), the standard deviations of the drivers, and their pairwise correlation coefficients:

$$\sigma_{H^+}^2 = \frac{\partial H^+}{\partial C_T}\sigma_{C_T}^2 + \frac{\partial H^+}{\partial A_T}\sigma_{A_T}^2 + \frac{\partial H^+}{\partial T}\sigma_T^2 + \frac{\partial H^+}{\partial S}\sigma_S^2$$
$$+ 2\frac{\partial H^+}{\partial C_T}\frac{\partial H^+}{\partial A_T}\text{cov}\left(C_T, A_T\right) + 2\frac{\partial H^+}{\partial C_T}\frac{\partial H^+}{\partial T}\text{cov}\left(C_T, T\right)$$
$$+ 2\frac{\partial H^+}{\partial C_T}\frac{\partial H^+}{\partial S}\text{cov}\left(C_T, S\right) + 2\frac{\partial H^+}{\partial A_T}\frac{\partial H^+}{\partial T}\text{cov}\left(A_T, T\right)$$
$$+ 2\frac{\partial H^+}{\partial A_T}\frac{\partial H^+}{\partial S}\text{cov}\left(A_T, S\right) + 2\frac{\partial H^+}{\partial T}\frac{\partial H^+}{\partial S}\text{cov}\left(T, S\right), \tag{2}$$

where the pairwise covariances are functions of the standard deviations and correlation coefficients according to $\text{cov}(x,y) = \sigma_x\sigma_y\rho_{x,y}$ and the partial derivatives are again evaluated at the temporal mean values $\overline{T}$, $\overline{S}$, $\overline{C_T}$, and $\overline{A_T}$. We aim at quantifying the contribution from mean changes in the drivers, which change the sensitivities, compared to the contributions from changes

185 in the drivers' standard deviations and correlations to [H$^+$] standard deviation changes. Unfortunately, these contributions can not be separated into summable terms because [H$^+$] standard deviation is a nonlinear function of those. However, we can analyse how much standard deviation change can be explained when considering only parts of the changes in the drivers. We hence first analyze the change in [H$^+$] standard deviation arising only from mean changes in the drivers. As a next step we also change the drivers' standard deviations and identify the additional change in [H$^+$] standard deviation compared to

190 the case of only changes in the means. Finally, we identify the contribution from changes in the phasing of the drivers as the additional change in [H$^+$] standard deviation by also taking into account changes in the drivers' correlation coefficients (further information is given in Appendix C).
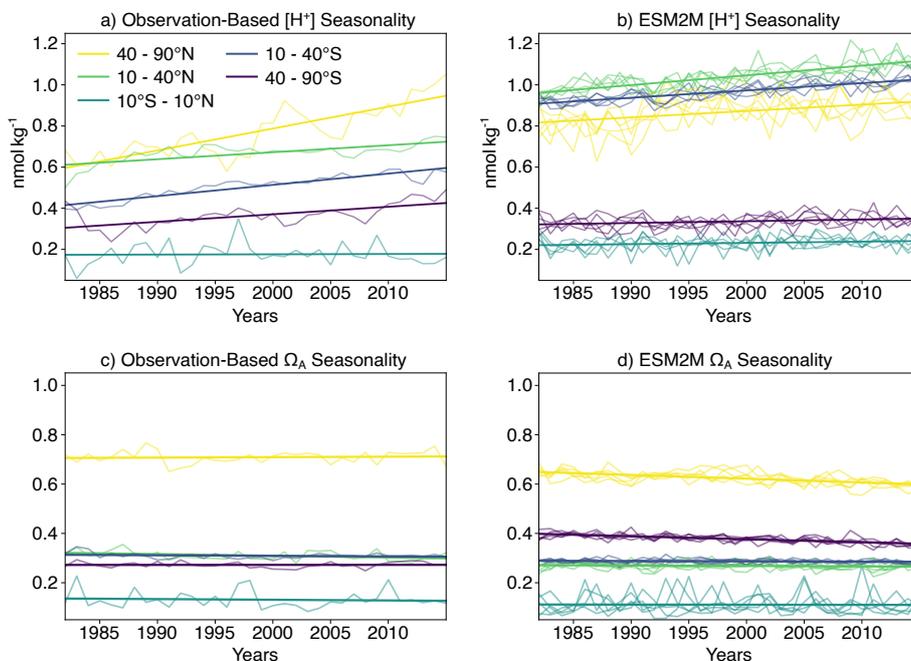
Figure 2. (a,b) Seasonal amplitude of [H$^+$] over the period 1982-2015 averaged over five different latitude bands for a) the observation-based estimate and b) the GFDL ESM2M historical (1982-2005) and RCP8.5 (2006-2015) ensemble simulations. (c,d) The same as (a,b), but for Ω_A. Linear trends in all panels are overlaid as thick lines. The linear trend of the simulated changes is calculated as the mean of the five individual ensemble trends.

2.3 Model evaluation

The focus of our analysis is on changes in variability in [H$^+$] and Ω_A. As observational-based daily data of the inorganic
195 carbon chemistry at the global scale is not available, we limit the evaluation of the Earth system model simulation to the representation of the seasonal cycles of [H$^+$] and Ω_A, and especially on its changes over the 1982-2015 period. To do so, we created an observation-based dataset for surface monthly [H$^+$] and Ω_A covering the 1982-2015 period using monthly surface salinity, temperature, pCO$_2$, and A$_T$ fields. Monthly salinity and temperature data are taken from the Hadley Centre EN.4.2.1 analysis product (Good et al., 2013). A$_T$ is then calculated using the *LIARv2* total alkalinity regression from salinity and tem-
200 perature (Carter et al., 2018). For pCO$_2$, we use the neural-network-interpolated monthly data from Landschützer et al. (2016), which is based on SOCATv4 (Bakker et al., 2016). Although not fully capturing pCO$_2$ variability in regions with only few observations (Landschützer et al., 2016), the pCO$_2$ dataset is generally well suited for analyzing pCO$_2$ seasonality and changes therein (Landschützer et al., 2018). [H$^+$] and Ω_A are then calculated from salinity, temperature, A$_T$, and pCO$_2$ using the *co2sys* carbonate chemistry package (van Heuven et al., 2011).

205

The GFDL ESM2M captures the observed trends in the seasonal [H$^+$] and Ω_A amplitudes for different latitudinal bands over the 1982-2015 period relatively well (Figure 2). The simulated ensemble mean trends in the seasonal [H$^+$] amplitude (calculated as the average of the individual trends in the ensemble) are positive for the northern low and high latitudes as well as the southern low latitudes (Figure 2, Supplementary Table A1), consistent with the observation-based estimates. In addition,

210 both the model and the observational-based estimates show no significant change in [H$^+$] seasonality in the equatorial region between 10 °S and 10 °N (purple thick lines in Figure 2a,b). However, the simulated [H$^+$] seasonality trends tend to be smaller than estimated from observations in the northern (yellow thick lines in Figure 2a,b) and southern high latitudes (red thick lines in Figure 2a,b), where the trends from the model ensemble are 0.031 ± 0.030 nmol kg^{-1} decade^{-1} ($\pm$ 90 % confidence interval for each ensemble member individually and then averaged) and 0.009 ± 0.010 nmol kg^{-1} per decade, compared to

215 the observational-based trends of 0.106 ± 0.040 nmol kg^{-1} decade^{-1} and 0.037 ± 0.028 nmol kg^{-1} decade^{-1}, respectively. In contrast, the simulated trend in [H$^+$] seasonality in the northern low latitudes (0.047 ± 0.012 nmol kg^{-1} decade^{-1}; blue thick line in Figure 2b) is slightly larger than the observation-based trend (0.034 ± 0.032 nmol kg^{-1} decade^{-1}; blue thick line in Figure 2a). For Ω_A, the model simulates negative trends in seasonal amplitude that are not seen in the observation-based estimate in the northern high latitudes (-0.015 ± 0.010 vs 0.002 ± 0.009 Ω_A units) and in the southern high latitudes (-0.012 ± 0.004

220 vs 0.000 ± 0.005 Ω_A units per decade). Conversely, a slight negative trend is seen in northern low latitudes in the observation-based estimate (-0.007 ± 0.006 Ω_A units per decade) that is not present in the simulations. In addition, GFDL ESM2M also captures the observed mean seasonal cycle [H$^+$] and Ω_A well in most regions and in particular for Ω_A (the mean values of the seasonal amplitudes in Figure 2). However, potential biases in the mean seasonal amplitudes do not directly have an effect on projected changes in extreme events, as we base the extreme events definition on relative thresholds.

225

In summary, taking into account additional evaluations not shown here of the mean states of [H$^+$] and Ω_A and the underlying drivers (Bopp et al., 2013; Kwiatkowski and Orr, 2018), the model performs well against a number of key seasonal performance metrics. However, the model slightly underestimates past changes in seasonal amplitude of [H$^+$], especially in the northern and southern high latitudes. Furthermore, it overestimates negative trends in Ω_A seasonal amplitude in those regions. Even though

230 we lack the daily observational-based data to undertake a full assessment, we consider our model to be well suited to assess changes in open ocean ocean acidification extreme events.

3 Results

We start by discussing the simulated changes in different ocean acidity extreme characteristics at the global scale, before we analyze changes at the regional scale and identify the drivers of changes. We recall that the large secular increase in [H$^+$] and

235 the large secular decrease in Ω_A was removed for the analysis as we focus on changes in variability and their impact on extreme event characteristics.

3.1 Global changes in ocean acidity extremes

In preindustrial times GFDL ESM2M suggests that a typical surface $[H^+]$ extreme event had an intensity of 0.08 nmol kg^{-1} (Figure 3c, Supplementary Table A2) and lasted 11 days (Figure 3e). Ocean acidity extremes occur coherently with a typical
240   volume in the upper 200 m of $2.7 \cdot 10^3$km^3 (Figure 3g). Over the historical period (from preindustrial to 1986-2005), the model projects that the number of surface $[H^+]$ extreme days per year increases from 3.65 days per year to 10.0 days per year (9.5-10.4 days per year; Figure 3a). The intensity and duration are projected to increase to 0.12 nmol kg^{-1} and 15 days. Compared to preindustrial conditions, this corresponds to a 173% (160-184%) increase in number of days per year, a 44% (39-47%) increase in the intensity and a 45% (41-48%) increase in the duration of $[H^+]$ extreme events. The volume of single events is
245   projected to increase by 20% (14-27%) over the historical period.

Over the 21st century, ocean acidity extreme events are projected to further increase in frequency, intensity, duration, and volume (Figure 3). By 2081-2100 under the RCP8.5 scenario, the number of $[H^+]$ extreme days per year at surface is projected to increase to 50.1 days (50.0-50.3 days; corresponding to a 1273% increase). The intensity is projected to increase to 0.38
250   nmol kg^{-1} (0.37-0.39 nmol kg^{-1}; 371%), the duration to 32 days (31-32; 199%) and the volume to $13.9 \cdot 10^3$km^3 (13.8-14.1; 414%).

At 200 m depth, $[H^+]$ extreme events are in general more intense (0.17 nmol kg^{-1}; Figure 3d) and longer-lasting (38 days; Figure 3f) than at surface during preindustrial conditions. The stronger extreme events are caused by the overall larger variability at 200 m than at surface. The longer duration is connected to the more pronounced contribution from interannual variability
255   compared to the surface (see Section 3.3). However, projected changes over the historical period and the 21st century are smaller at 200 m depth than at surface and with larger year-to-year variations across the ensembles. Under present-day (1986-2005), the number of extreme days per year at 200 m depth is 4.3 days per year (3.7-5.1 day per year; corresponding to a 18% increase since preindustrial), the intensity 0.20 nmol kg^{-1} (0.19-0.21 nmol kg^{-1}; 18%), and the duration 46 days (43-50 days; 21%). By the end of the 21st century under the RCP8.5 scenario, the number of $[H^+]$ extreme days per year is projected to increase
260   to 32.1 days per year (30.9-34.8 days per year), the intensity to 0.34 nmol kg^{-1} (0.33-0.34 nmol kg^{-1}) and the duration to 99 days (95-102) days. Notably, $[H^+]$ extreme events are projected to become less intense at 200 m depth than at surface (0.34 nmol kg^{-1} vs 0.38 nmol kg^{-1}) by the end of the century under RCP8.5, even though they were more intense in preindustrial times at depth. In contrast, surface $[H^+]$ extreme events remain shorter in duration at the end of the century than at 200 m depth.

265   Under the RCP2.6 scenario and by the end of the century, the magnitude of changes in the different $[H^+]$ extreme event characteristics would be substantially reduced. This reduction is especially pronounced at the surface (blue lines in Figure 3). There, the number of extreme days per year, intensity, and duration under the RCP2.6 are projected to be only 46% (44-47), 43% (43-44) and 75% (73-77) of that under the RCP8.5 scenario. At depth, the differences between the RCP2.6 and RCP8.5 scenario are less pronounced and only emerge in the second half of the 21st century. In contrast to the surface, the number
270   of $[H^+]$ extreme days per year, the intensity, and the volume of events at depth are projected to considerably increase even
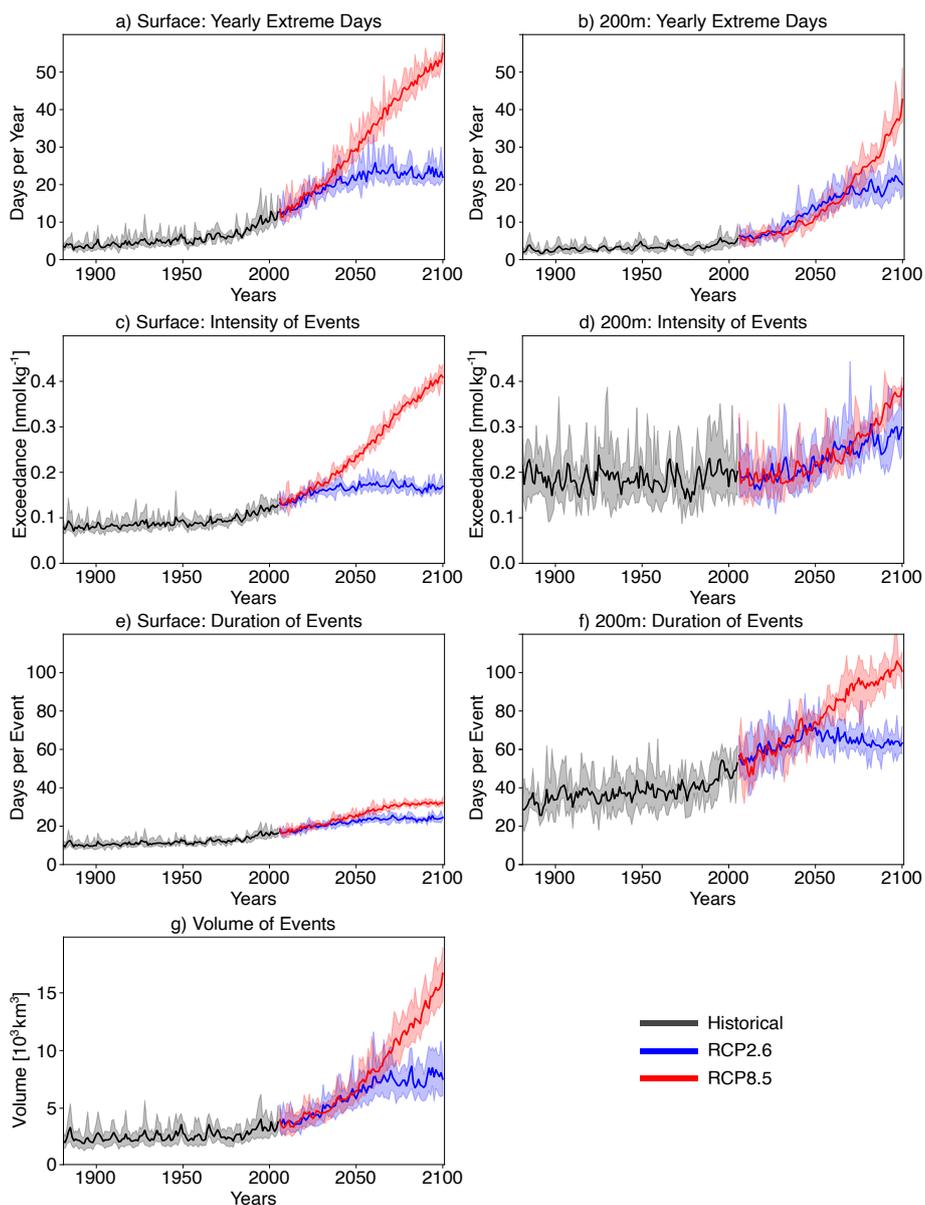
Figure 3. Simulated changes in globally averaged [H$^+$] extreme event characteristics over the 1861-2100 period following historical (black lines) and future RCP8.5 (red) and RCP2.6 scenario (blue). Frequency, intensity, and duration are shown for the surface (a,c,e) and for 200 m depth (b,d,f). Volume is shown in (g). The thick lines display the five-member ensemble means and the shaded areas represent the maximum and minimum ranges of the individual ensemble members.
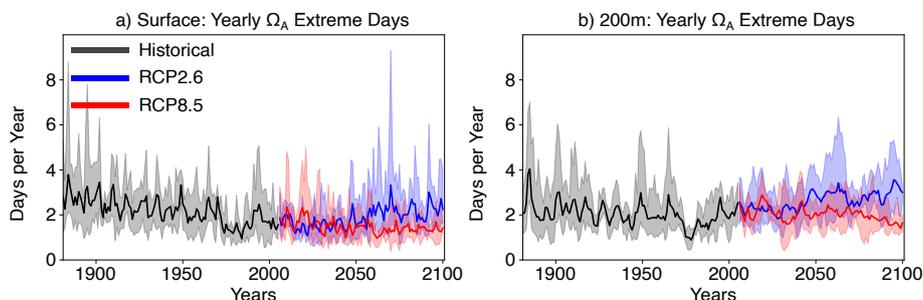
Figure 4. Globally averaged simulated number of extreme Ω_A days per year from 1861 to 2100 following historical (black lines) and future RCP2.6 (blue) and RCP8.5 (red) scenarios at (a) the surface and (b) 200 m depth. The thick lines display the five-member ensemble means and the shaded areas represent the maximum and minimum range of the individual ensemble members.

after the atmospheric CO_2 concentration stabilize in RCP2.6 around year 2050. This delayed response at subsurface is due to the relatively slow surface-to-deep ocean transport of carbon. However, this is not the case for the duration, which slightly decreases in the second half of the 21st century at depth.

275   In contrast to [H$^+$] extreme events, the yearly number of Ω_A extreme days is projected to decrease over the historical and the 21st century under both the RCP8.5 and RCP2.6 scenario (Figure 4, Supplementary Table A3). The simulated decrease is slightly stronger under RCP8.5 than under RCP2.6. The number of surface Ω_A extreme days per year by the end of the century is projected to be 63% (-70,-54) smaller under RCP8.5 and 39% (-49,-20) smaller under RCP2.6 than at preindustrial. Projected changes at depth are less pronounced than at surface, again with larger decreases under RCP8.5 than under RCP2.6.

280   **3.2   Regional changes in ocean acidity extremes**

Surface [H$^+$] extremes are projected to become more frequent in 87% of the surface ocean area by the end of the 21st century under the RCP8.5 scenario. However, the projected changes in ocean acidity extremes are not uniform over the globe but show distinct spatial patterns (Figure 5; Supplementary Figure A2). The largest increases in the number of [H$^+$] extreme days per year are projected in the Arctic Ocean (up to +120 days per year), in the subtropical gyres (up to +60 days per year), in parts of the Southern Ocean and near Antarctica. There are also some regions including the eastern equatorial Pacific and parts of the Southern Ocean, where the number of yearly extreme days in surface [H$^+$] is projected to decrease. These are in general also the regions where the seasonality in [H$^+$] is projected to decrease (see section 3.3 below). The largest changes in intensity of surface [H$^+$] extremes (Figure 5c) are projected for the subtropics, especially in the Northern Hemisphere. For example, events become up to 0.8 nmol kg^{-1} more intense in the subtropical North Pacific and Atlantic. Large changes are also projected for the Arctic Ocean and around Antarctica. Regions with large increases in the number of yearly extreme days tend to show also large increases in the duration of extreme events (Figure 5e). The Arctic Ocean is an exception. Although the number of yearly extreme days increases sharply, the increase in duration is not as pronounced. This is because extremes are already long-lasting,

but rare at preindustrial times (Supplementary Figure A2). So even though extreme events are projected to reoccur annually by the end of the century under RCP8.5, the absolute increase in duration is relatively small.

295

At 200 m depth, the projected pattern of changes in yearly extreme event days generally resembles the changes at the surface (Figure 5b). The largest increases in yearly extreme event days are projected for parts of the subtropics, the Southern Ocean, and the Arctic Ocean. In contrast to the surface, [H$^+$] extremes are projected to become less frequent in the equatorial Atlantic, the northern Indian Ocean, the North Pacific and in large parts of the Southern Ocean. The regions showing a decline in

300 [H$^+$] extremes at depth include also some of the eastern boundary current systems, such as the the Humboldt, California, and Benguela Current systems. In most of these regions, extreme [H$^+$] days are projected to disappear in the RCP8.5 ensemble by the end of this century (grey regions in Figure 5b). The largest increases in subsurface event intensity are projected in the subtropics (Figure 5d), whereas the duration of [H$^+$] extremes is projected to strongly increase in many regions of the mid-to-high latitudes of both hemispheres (Figure 5f). The projected increases in duration at 200m depth are much larger than at surface.

305

The increase in the number of extreme days per year, the intensity, and the duration is smaller under RCP2.6 compared to RCP8.5 for most of the ocean (Supplementary Figure A1). The largest increases in occurrence of extremes under RCP2.6 are simulated for the Arctic Ocean, similar as under RCP8.5, and for parts of the Southern Ocean. The regions in the Southern Ocean where the occurrence of extremes is projected to decrease largely overlap with those for RCP8.5, at surface and at depth.

310 On the other hand, unlike under RCP8.5, a decrease in extreme event occurrence is only projected for a small fraction of the tropical oceans under RCP2.6.

Extreme events in Ω_A are projected to become less frequent throughout most of the ocean (89% of surface area under RCP8.5 at the end of the 21st century; Fig 6a). In many regions, Ω_A extreme events are projected to disappear by 2081-2100 under the

315 RCP8.5 scenario (grey regions in Fig 6a). However, the frequency of surface Ω_A extremes is projected to increase by 10 or more days per year in the subtropical gyres, especially in the western parts of the subtropical gyres. At depth, no extreme events are projected for most of the ocean during 2081-2100 under RCP8.5 (Fig 6b).

3.3 Decomposing [H$^+$] variability changes into interannual, seasonal, and residual daily variability changes

The underlying changes in [H$^+$] variability and extreme events in [H$^+$] may arise from changes in interannual variability, sea-

320 sonal variability, and residual daily variability. We therefore decompose the total variability into these three components (see Section 2.2.2). Figure 7 shows their contribution to the overall change in variability from preindustrial to 2081-2100 following the RCP8.5 scenario at surface and at depth. Changes in seasonality clearly dominate the overall change in variability at surface with 87% contribution to the overall variance change in the global mean (Figure 7b). Changes in interannual variability (3% contribution to overall variance change; Figure 7a) and residual daily variability (10%; Figure 7c) play a minor role. The largest

325 increases in variability for all three variability types are projected for the northern high latitudes, where also the number of extreme event days increases most strongly. The increases in extreme events around Antarctica and the southern end of South
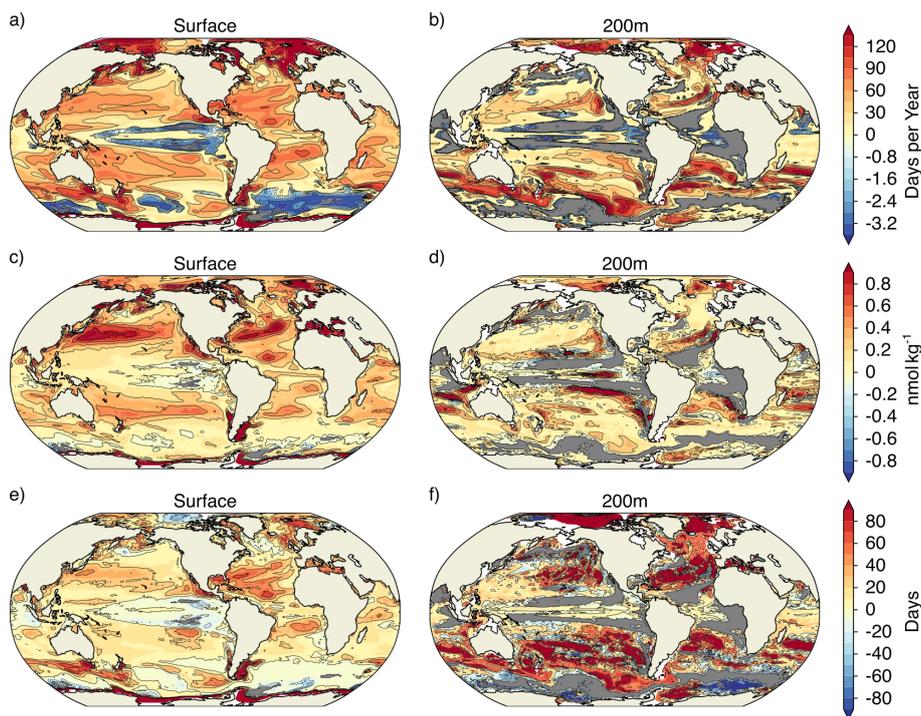
Figure 5. Simulated regional changes in [H$^+$] extreme event characteristics from preindustrial to 2081-2100 under the RCP8.5 scenario at surface and at depth for (a,b) the number of extreme event days in days per year, (c,d) the maximal intensity of events in nmol kg^{-1}, and (e,f) the duration of events in days. Shown are changes averaged over all five ensemble members. Grey colors represent areas, where no extremes occur during 2081-2100 and the black lines highlight pattern structures.
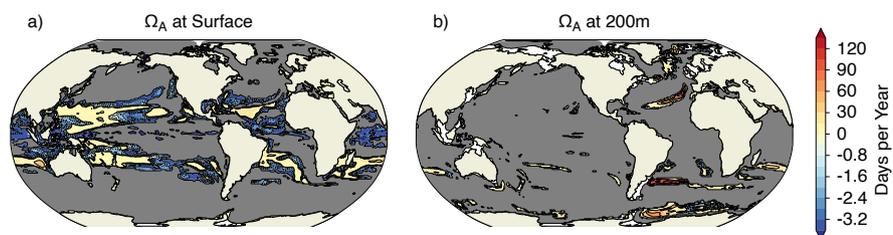
Figure 6. Simulated regional changes in the number of Ω_A extreme days per year from preindustrial to 2081-2100 under the RCP8.5 scenario (a) at the surface and (b) at 200 m depth. Shown are changes averaged over all five ensemble members. The black lines highlight the pattern structure and grey colors represent regions where no ensemble member simulates extremes during 2081-2100.
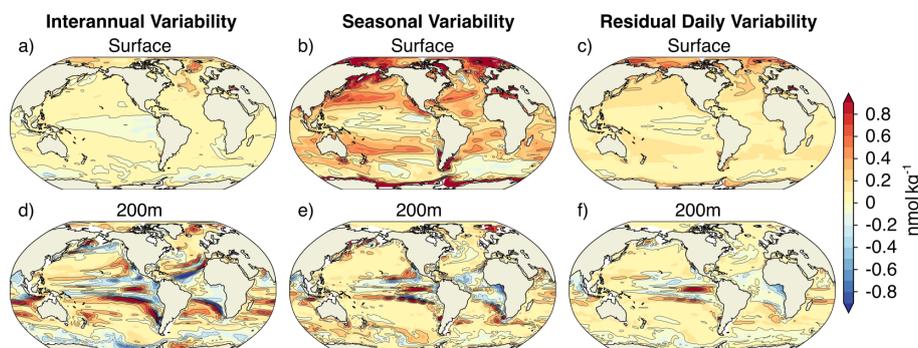
Figure 7. Contribution to projected changes in [H$^+$] standard deviation from (a,d) interannual variability, (b,e) seasonal variability, and (c,f) residual daily variability between the preindustrial and the 2081-2100 period following the RCP8.5 scenario. Shown are the ensemble mean changes. The black lines highlight the pattern structure.

America (Figure 5a) are mainly caused by increases in seasonal variability (Figure 7b). The regions that are projected to experience a decline in extremes (Figure 5a) coincide with those of decreasing interannual and seasonal variability (Figure 7a,b).

330   In contrast to the surface, changes in interannual and to a lesser extent residual daily variability at 200 m depth are also important for explaining the changes in [H$^+$] variability and extremes (Figure 7d,f). Even though the changes in seasonal variability still dominate the overall changes in variance at the global scale (with 42% contribution), the interannual variability is almost equally important (37%). Changes in residual daily variability also contribute substantially to changes in total variability (20%). The patterns of variability changes are very similar across the three types of variability. The largest increases in
335   [H$^+$] variability are simulated in the subtropics. In those regions the model also projects an increase in [H$^+$] extreme events (Figure 5b). Furthermore, these regions tend to be already more variable at preindustrial (see Supplementary Figure A2a). However, the model also projects an increase in variability for less variable regions at preindustrial, such as the Arctic Ocean, leading to increases in extremes there. All three variability types are projected to decrease in the tropics and the Southern Ocean, where the occurrence of extreme events is projected to largely decrease (c.f. Figure 5b). The variability decrease in
340   those regions is most pronounced for interannual variability (Figure 7d).

3.4   Drivers of changes in [H$^+$] variability

In this section, we investigate the drivers of variability changes in [H$^+$]. Changes in [H$^+$] variability can be attributed to three factors: (i) changes in the mean states of the drivers that control the sensitivities (changing the partial derivatives in Equation 2 that affect the amplitude of the variation in [H$^+$] for a given deviation of a driving variable x from its mean $(x(i) - \overline{x})$ in
345   Equation 1), (ii) changes in the variabilities of the drivers (the drivers' standard deviations in Equation 2), and (iii) changes in the phasing between the drivers (the correlation coefficients in Equation 2 that describe the degree to which the drivers co-vary with each other) (see Section 2.2.3).
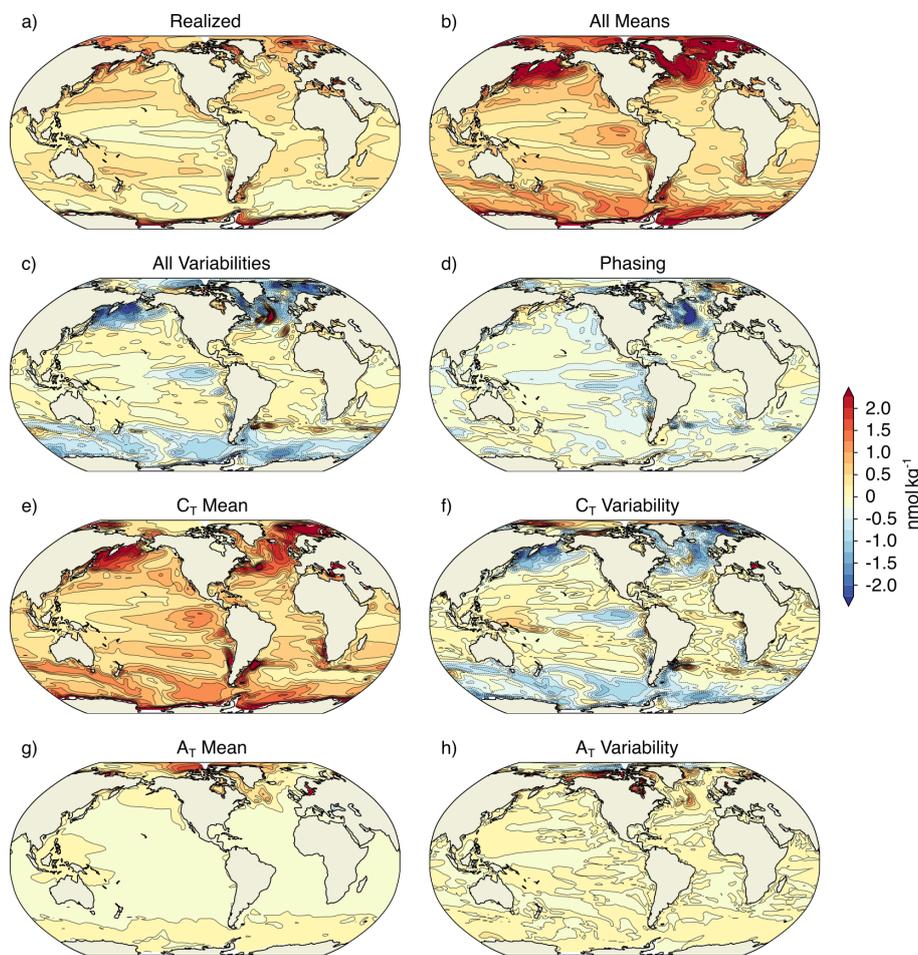
Figure 8. (a) The realized change in $[\mathrm{H}^+]$ standard deviation at surface from preindustrial to 2081-2100 following the RCP8.5 scenario. It is decomposed into (b) the change in $[\mathrm{H}^+]$ standard deviation from changes in the drivers' mean values, (c) the $[\mathrm{H}^+]$ standard deviation change from additionally changing the drivers variabilities, and (d) the contribution from additional changes in the phasing of the drivers (pairwise correlations between the drivers). (e) The change in $[\mathrm{H}^+]$ standard deviation obtained when only changing the mean state of C_T. (f) The change in $[\mathrm{H}^+]$ standard deviation from additionally changing the variability of C_T. (g,h) Same as (e,f), but for A_T. Temperature and salinity contributions are small and not shown here. The black contours highlight the pattern structures.

We investigate the first contribution (impact of mean state changes on the sensitivities) by changing only the mean states of the drivers (in Equation 2) from the preindustrial values to those of the 2081-2100 period following the RCP8.5 scenario. The

350   drivers' standard deviations and correlation coefficients stay at the preindustrial levels. Figure 8b shows that the mean changes in the drivers cause an overall increase in surface $[\mathrm{H}^+]$ variability. Global $[\mathrm{H}^+]$ variability changes due to mean changes (0.67 nmol kg^{-1}; Figure 8b) are more than twice as large as the total realized changes (0.30 nmol kg^{-1}; Figure 8a). Especially

in the higher latitudes, the changes due to mean changes are much larger than the total realized changes in [H^+] variability.

355   What causes this large increase [H^+] variability? In general, an increase in mean C_T, temperature, and salinity would lead to an increase in [H^+] variability, whereas an increase in mean A_T would lead to a decrease. GFDL ESM2M projects an increase in mean C_T over the entire surface ocean (Supplementary Figure A4) and therefore an increase [H^+] variability (Figure 8e). Surface A_T is projected to increase globally, especially in the low-to-mid latitudes of the Atlantic (Supplementary Figure A4), and therefore dampens slightly the overall increase in [H^+] variability (Figure 8g). This is not the case in the Arctic Ocean, 360   where A_T is projected to decrease and amplifies the increase in [H^+] variability caused by changes in C_T (Figure 8g). Changes in temperature and salinity play a minor role (not shown in Figure 8).

    Next, we investigate the second contribution, i.e. the impact of changes in variability of the drivers on [H^+] variability (Figure 8c). At the global scale, the projected change in the variability of the drivers (Supplementary Figure A5) causes a decrease 365   in [H^+] variability and counteracts to some degree the increase in [H^+] variability due to increase in the mean drivers. However, this decrease due to the drivers' variability is spatially not uniform. In the low latitudes (except in the eastern equatorial Pacific), changes in the variability of the drivers lead to a slight increase in [H^+] variability, whereas in the high latitudes they lead to a decrease in [H^+] variability. The increase in the low latitudes and the decrease in the high latitudes and the equatorial Pacific is mainly caused by increases and decreases in C_T variability (Figure 8f; Supplementary Figure A5). Again, the contri- 370   bution from changes in A_T variability (Figure 8h) is important in the Arctic Ocean, while temperature and salinity play a minor role. The additional variability change from the third contribution (changes in the phasing of the drivers; Figure 8d) is overall smaller, except in the North Atlantic.

    At 200 m depth, the picture looks similar as at the surface. Again, the projected increase in [H^+] variability (Figure 9a) 375   caused by the mean changes in the drivers (Figure 9b; Supplementary Figure A4) is damped by the decrease in the variability of the drivers (Figure 9c; Supplementary Figure A5). The changes in mean and variability of C_T (Figure 9e,f) are the most important drivers of changes. Increases in mean A_T somewhat dampen the overall increase in the low latitudes (Figure 9g). Again, changes in temperature and salinity are negligible. In contrast to the surface, however, the individual changes in [H^+] variability due to the mean and variability changes of the drivers, in particular of C_T, are much larger at 200 m depth. The 380   largest individual changes are projected for the southern edges of the subtropical gyres in the north and for the northern edges of the subtropical gyres in the south. The changes in the Arctic Ocean at 200 m depth are overall smaller than at the surface. Changes in the phasing of the drivers are overall less important than changes in the variabilities and mean states.

4   Discussion and conclusions

We provide a first quantification of the historical and future changes in short-term ocean acidity extreme events at global 385   and regional scale by analyzing daily 3D output from a ensemble simulation of a comprehensive Earth system model. In our
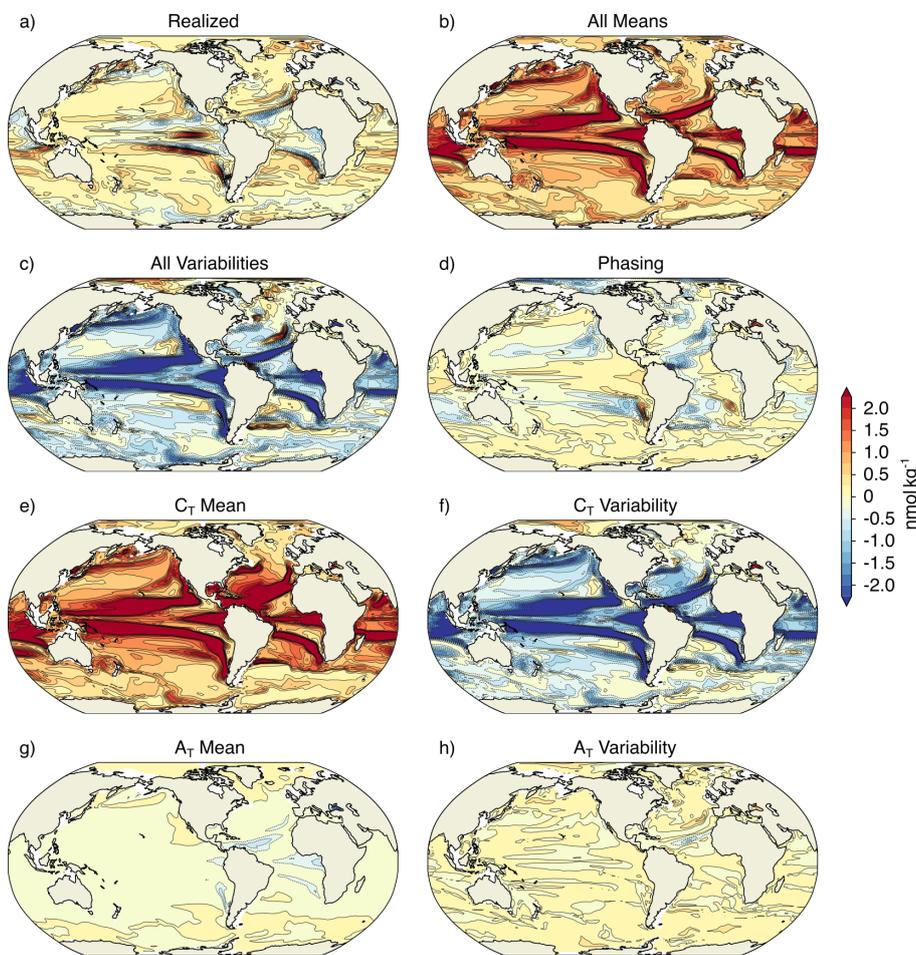
Figure 9. Same as Figure 8 but at 200 m depth.

analysis, we focus on changes in extreme events that arise only from changes in daily to interannual variability. Secular CO_2 emission-induced trends in the mean state were removed from the model output before the extreme events analysis. We show that extreme $[H^+]$ events are projected to become more frequent, longer lasting, more intense and spatially more extensive under increasing atmospheric CO_2 concentration, both at surface and also within the thermocline. These changes in $[H^+]$ ex-
390   treme event characteristics are substantially reduced under the RCP2.6 scenario compared to RCP8.5. The increase in $[H^+]$ variability and extreme events is a consequence of its nonlinear dependence on the drivers. It is mainly driven by the projected increase in mean C_T and additionally altered by changes in C_T variability and A_T mean and variability. Extreme events in Ω_A are projected to become less frequent in the future. This is because Ω_A becomes less sensitive to variations in the physical and biogeochemical state of seawater under elevated atmospheric CO_2.

395
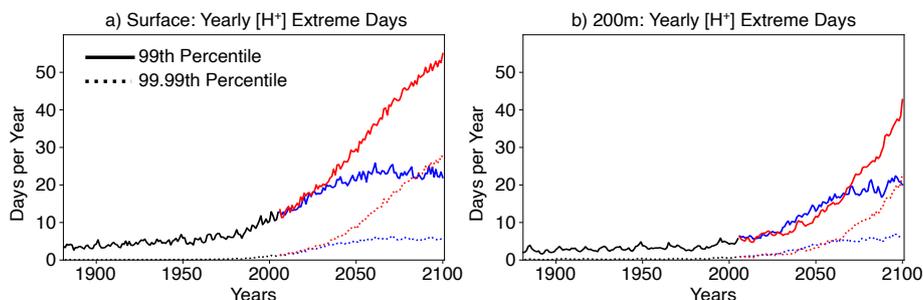
Figure 10. Globally averaged number of extreme event days for $[H^+]$ over the historical (black lines), RCP2.6 (blue), and RCP8.5 (red) simulations for the preindustrial 99th (solid lines) and 99.99th percentile (dotted lines) at (a) the surface and (b) 200m depth

In this study, we use the 99th percentile of the distribution from a preindustrial simulation for the definition of an extreme $[H^+]$ event, but the results may depend on the choice of this threshold. We tested the sensitivity of our results by using also the 99.99th percentile threshold. The relative increase in the numbers of extreme $[H^+]$ days per year is larger for these very rare extremes (Figure 10). For example, nearly every second day with $[H^+]$ exceeding the 99th percentile (red solid lines in Figure 10) is also a day with $[H^+]$ exceeding the 99.99th percentile (red dotted lines in Figure 10) by the end of the 21st century under RCP8.5, both at surface and at depth. In other words, an event that occurs every 27 years at preindustrial becomes almost as frequent in the future as an event that occurs every hundred days at preindustrial. As a result of this large relative increase in rare extremes, the model projects as many days with $[H^+]$ exceeding the 99.99th percentile by the end of the century under RCP8.5 (red dotted lines in Figure 10) as it projects days exceeding the 99th percentile under RCP2.6 (blue solid lines in Figure 10).

The projected increase in $[H^+]$ variability and decrease in Ω_A variability also alters the occurrence of extreme events based on absolute thresholds. An often used threshold is $\Omega_A = 1$ below which seawater is corrosive with respect to the calcium carbonate mineral aragonite (Bednaršek et al., 2012). We assess the influence of the general decline in Ω_A variability at the point in time where a grid cell falls below $\Omega_A = 1$ for the first time. To do so, we compare these points in time within the historical and RCP8.5 ensemble to the ones for the hypothetical case where Ω_A variability stays at the preindustrial level but mean Ω_A undergoes the ensemble mean evolution. We find that the decline in Ω_A variability, which is observed in the historical and RCP8.5 ensemble, leads to an average delay of the first occurrence of undersaturation by about 11 years at the surface and about 16 years at 200 m depth. At surface, these delays of undersaturation occur throughout the high latitudes (Figure 11a). At depth, the delays are most pronounced in the tropics (Figure 11b), but delays also occur in the high latitudes. Absolute thresholds for $[H^+]$ are generally not well established. Tests with arbitrarily set absolute thresholds suggest that these thresholds are generally exceeded a few years earlier because of increasing variability, but the temporal difference is less pronounced than that for Ω_A.
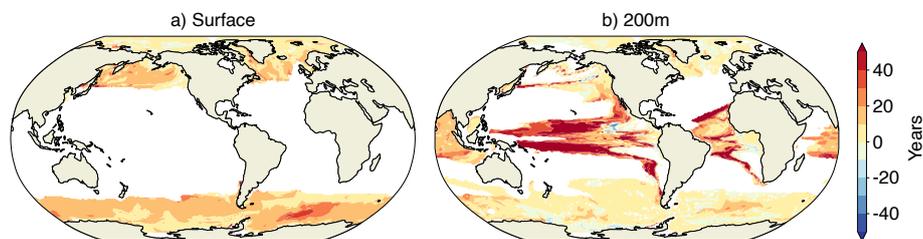
Figure 11. The temporal difference in years between the first occurrence of aragonite undersaturation in the historical and RCP8.5 ensemble and a hypothetical simulation with the mean changes observed in the historical and RCP8.5 ensemble but with preindustrial variability. Positive values (yellow and red) indicate a delayed onset of undersaturation resulting from declines in Ω_A variability.

420   Previous studies have shown that the seasonal cycle of surface ocean $p\mathrm{CO_2}$ and surface $[\mathrm{H^+}]$ will be strongly amplified under increasing atmospheric $\mathrm{CO_2}$ (Gallego et al., 2018; Landschützer et al., 2018; McNeil and Sasse, 2016; Kwiatkowski and Orr, 2018). Here we show that the changes in the seasonal cycle of $[\mathrm{H^+}]$ translate into large increases in short-term extreme acidity events, at surface as well as at $200\,\mathrm{m}$ depth. In addition to earlier studies, we also show that changes in residual daily variability contribute to changes in extreme $[\mathrm{H^+}]$ events under increasing atmospheric $\mathrm{CO_2}$. It is therefore critical to use daily
425   temporal output to assess extreme events in ocean biogeochemistry. Currently, ocean biogeochemical variables from models that participate in the sixth phase of the Coupled Model Intercomparison Project Phase 6 are routinely stored with a monthly resolution on the Earth system grid (Jones et al., 2016). We therefore recommend to save out and use higher than monthly resolution to analyze variability in the surface and sub-surface ocean carbonate systems, in particular for studying extreme events.

430

Even though we consider our results as robust, a number of potential caveats remain. First, the horizontal resolution of the ocean model in ESM2M is rather coarse and cannot represent critical scales of small-scale circulation structures (e.g. Turi et al. (2018)). In addition, the biogeochemical processes included in ESM2M are designed for the open ocean, but do not capture the highly variable coastal processes (Hofmann et al., 2011). High resolution ocean models with improved process represen-
435   tations are therefore needed to explore extreme events in ocean carbonate chemistry, especially in coastal regions. However, observation-based carbonate system data on daily time scale would also be necessary to thoroughly evaluate the models' capability to represent daily variations in carbonate chemistry. Secondly, our results at the local scale might depend on the model formulation. The GFDL ESM2M is largely able to project $[\mathrm{H^+}]$ variability changes (Figure 2) arising from the nonlinear dependence on the drivers' mean states as the ocean carbonate chemistry is known and the mean changes in the drivers match
440   observational records relatively well over the historical period (Bopp et al., 2013). However, it is currently rather uncertain how $[\mathrm{H^+}]$ variability changes as a result of changes in the drivers' variabilities. It is well known that current Earth system models have imperfect or uncertain representations of ocean variability over a range of timescales (Frölicher et al., 2016; Resplandy et al., 2015; Keller et al., 2014). A possible way forward would be to assess changes in ocean acidity extreme events within a multi-model ensemble, which would likely provide upper and lower bounds of future changes in these events. Finally, it is

445 assumed that physical and biogeochemical changes in the ocean will also increase diurnal variability. In particular in coastal areas, such diurnal variations can have amplitudes that are much larger than the projected changes over the 21st century (Hofmann et al., 2011). However, GFDL ESM2M does not fully resolve the diurnal variability. In the ocean biogeochemistry model (TOPAZv2) that is embedded in GFDL ESM2M, for example, phytoplankton growth is not able to consume internal nutrient stores in the absence of light during night and diel vertical migration is also not simulated. Future studies with Earth system
450 models that resolve diurnal chemistry extremes are needed to quantify the impacts of changes in the diurnal cycles on extreme acidity events.

Our analysis also has important consequences for the impact of ocean acidification on marine ecosystems. It implies that marine organisms will have less time to recover from very high [H$^+$] events in the future, as the frequency and the duration of
455 the ocean acidity extremes are projected to increase substantially over the 21st century. While coastal species may be adapted to large variability in ocean acidity, the large projected increase in [H$^+$] extreme events in the open ocean may push organisms and ecosystems that are commonly accustomed to a more steady environment to the limits of their resilience. The risks for substantial ecosystem impacts are aggravated by the fact that the frequency and intensity of marine heatwaves is also projected to substantially increase (Frölicher et al., 2018), which also negatively impact marine ecosystems (Wernberg et al., 2016; Smale
460 et al., 2019). The interactions of intensified multiple stressors has the potential to influence marine ecosystems and the ocean's biogeochemical cycles in an unprecedented manner (Gruber, 2011). However, further research is needed to understand the combined impacts of short term ocean acidity extremes and marine heatwaves on marine ecosystems.

In conclusion, our analysis reveals that marine organisms and ecosystems are projected to be exposed to less stable [H$^+$]
465 conditions in the future with more frequent occurrences of extreme [H$^+$] conditions. Such extremes events are projected to last longer, to be more intense and to cover larger volumes of seawater and therefore potentially add to the stress on organisms and ecosystems from mean ocean acidification.

Appendix A: Identifying and removing the secular trend in the model data

In this study, we analyze the changes in short-term extreme events in [H$^+$] and Ω_A that arise from day-to-day to interannual
470 variability changes in these variables. We therefore need to remove the secular trends from the data prior to analysis. We estimate the secular trends from the five-member ensemble means, assuming that sub-annual and interannual to decadal variations in the individual ensemble members are phased randomly and do not imprint on the ensemble means because they average out. A larger ensemble size would be necessary for this assumption to perfectly hold. However, this potential source of error does not qualitatively alter our results.

475

Unlike sub-annual and interannual variability, the seasonal cycle is a deterministic component within a climate time series. Its phase is not independent between ensemble members and it doesn't average out when calculating ensemble means. We

remove the seasonal cycles from the ensemble means by smoothing the ensemble means with a 365-day running mean filter, i.e. by calculating the convolution of the time series with a rectangular window of length 365 and height 1/365. This filter

480   also removes variability on sub-annual and interannual timescales and thereby also reduces the error we make due to the small ensemble size that is discussed above. We then subtract the running-mean-filtered ensemble means from the five ensemble members to remove the secular trends in the individual ensemble members.

Appendix B: Identifying interannual and residual daily variability

The spectral density describes how the variance in a time series is distributed over different frequencies ν_j. It is proportional to

485   the absolute value squared of the discrete Fourier transformation (DFT) of the time series. Defining the spectral density only for positive frequencies, it is given by

$$f(\nu_j) = 2\frac{\Delta t^2}{T}\left|\sum_{k=1}^{N} x_k \cdot \exp\left(-\mathrm{i}2\pi\nu_j \cdot \Delta t\, k\right)\right|^2,$$

with N the number of time steps, x_k the values of the time series at each time step, Δt the time interval between two time steps, $T = N \cdot \Delta t$, and the frequencies $\nu_j = j/T$. The autocovariance is the inverse Fourier transform of the spectral density (Wiener-

490   Khintchine theorem, Chatfield (1996))[1]. As a consequence, the variance within the time series, given by the autocovariance at lag zero, is obtained by integrating the spectral density over all positive frequencies, $\sigma^2 = \int_0^\infty f(\nu)\mathrm{d}\nu$. For a discrete time series, where the maximal resolved frequency is given by $\nu_{\mathrm{max}} = 1/2\Delta t$, the identity reads

$$\sigma^2 = \sum_{j=0}^{N/2} f(\nu_j)\frac{1}{N\Delta t}.$$

Based on this equation, one can separate the contributions to variance from low-frequency and high-frequency variations. In this

495   study, we determine interannual variability and the residual (sub-annual) daily variability. Interannual variability is calculated by summing over the contributions to variance from all frequencies up to a cycle of once per year, i.e. by evaluating the sum up to i_{cut} for which $\nu_{\mathrm{cut}} = 1/365\,\mathrm{day}^{-1}$. Accordingly, residual daily variability is obtained by evaluating the sum from $i_{\mathrm{cut}}+1$ to $N/2$. Prior to this separation, the seasonal variability is removed from the data by subtracting the 365-day climatology. We use the *periodogram* function from *scipy.signal* to estimate the spectral density.

---

[1]In the continuous case, the theorem states

$$\gamma(\tau) = \int_{-\infty}^{\infty} \tilde{f}(\nu)\exp(\mathrm{i}2\pi\nu\tau)\mathrm{d}\nu,$$

with the autocovariance function $\gamma(\tau)$ and the spectral density $\tilde{f}$ defined for positive and negative frequencies. Since the two-sided spectral density, $\tilde{f}$, is a real and even function, one can also use

$$\gamma(\tau) = \int_{0}^{\infty} f(\nu)\cos(2\pi\nu\tau)\mathrm{d}\nu$$

with the one-sided spectral density $f = 2 \cdot \tilde{f}$ that is used in this text.

Appendix C: Decomposition of [H$^+$] standard deviation change

According to Equation 2 in the main text, the variance in [H$^+$] can be approximated as a function of the four means of the drivers $\mu_{1,..4}$ (we assume that changes in total phosphate and total silicate are negligible), the four standard deviations of the drivers $\sigma_{1,..4}$, and the six pairwise correlation coefficients $\rho_{i\neq j=1,..4}$. The total change in [H$^+$] standard deviation (or variance) between two periods can thus be obtained by calculating

$$\sigma_{\mathrm{H}}(\mu_i^1, \sigma_i^1, \rho_{ij}^1) - \sigma_{\mathrm{H}}(\mu_i^0, \sigma_i^0, \rho_{ij}^0),$$

where the superscript 0 indicates the first period, here preindustrial, and the superscript 1 indicates the second period, here 2081-2100 under RCP8.5.

Based on this relation, we then assess how changes in the drivers' means, standard deviations, and correlation coefficients contribute to the change in [H$^+$] standard deviation. Firstly, we investigate how much [H$^+$] standard deviation change can be explained by solely adjusting the drivers' mean values to the values simulated for 2081-2100 under RCP8.5,

$$\sigma_{\mathrm{H}}(\mu_i^1, \sigma_i^0, \rho_{ij}^0) - \sigma_{\mathrm{H}}(\mu_i^0, \sigma_i^0, \rho_{ij}^0).$$

Changes in the mean values contribute largest to [H$^+$] variability changes. Secondly, we investigate how much standard deviation change can be additionally explained when also taking into account changes in the drivers' standard deviations,

$$\sigma_{\mathrm{H}}(\mu_i^1, \sigma_i^1, \rho_{ij}^0) - \sigma_{\mathrm{H}}(\mu_i^1, \sigma_i^0, \rho_{ij}^0).$$

Overall, standard deviation changes have the second largest imprint on [H$^+$] variability changes. Thirdly, the remaining contribution from changes in the pairwise correlation coefficient is calculated as

$$\sigma_{\mathrm{H}}(\mu_i^1, \sigma_i^1, \rho_{ij}^1) - \sigma_{\mathrm{H}}(\mu_i^1, \sigma_i^1, \rho_{ij}^0).$$

In analogy to the first two steps, the contributions from mean and standard deviation changes in C_{T} and A_{T} alone are also assessed.

Latitude	Obs. [H$^+$]	ESM2M [H$^+$]	Obs. Ω_{A}	ESM2M Ω_{A}
40°-90° N	0.106 ± 0.040	0.031 ± 0.030	1.9 ± 8.7	-15.1 ± 9.5
10°-40° N	0.034 ± 0.034	0.047 ± 0.012	-6.7 ± 5.6	-1.8 ± 4.7
10° S-10° N	0.001 ± 0.016	0.006 ± 0.011	-2.8 ± 10.7	-0.5 ± 13.1
40°-10° S	0.055 ± 0.014	0.035 ± 0.008	-2.4 ± 5.1	-1.2 ± 2.8
90°-40° S	0.037 ± 0.028	0.009 ± 0.010	0.1 ± 4.8	-12.2 ± 3.9

Table A1. Linear trends in seasonal amplitude of [H$^+$] (in nmol kg^{-1} decade^{-1}) and Ω_{A} (in 10^{-3} decade^{-1}) for five latitude bands over the period 1982-2015. Results are shown for the observational-based data (Obs.) and the five-member ensemble mean of the ESM2M simulations (ESM2M). The range ($\pm$) denotes the 90 % confidence interval.

	PI	1986-2005	EOC RCP2.6	EOC RCP8.5
Number Surf.	3.65	9.97 (9.49-10.38)	22.87 (21.93-23.45)	50.12 (49.98-50.30)
200m	3.65	4.32 (3.72-5.09)	19.88 (16.96-22.53)	32.10 (30.91-34.75)
Duration Surf.	10.64	15.38 (15.04-15.72)	23.79 (23.40-24.11)	31.78 (31.23-32.13)
200m	38.00	45.95 (42.84-49.96)	62.94 (60.49-66.11)	98.66 (95.06-102.01)
Intensity Surf.	0.08	0.12 (0.11-0.12)	0.17 (0.16-0.17)	0.38 (0.37-0.39)
200m	0.17	0.20 (0.19-0.21)	0.28 (0.25-0.30)	0.34 (0.33-0.34)
Volume	2709	3247 (3082-3451)	7654 (6873-8464)	13927 (13836-14109)

Table A2. Simulated global mean $[H^+]$ extreme event characteristics for the preindustrial, present day (1986-2005), RCP2.6 end of century (EOC; 2081-2100), and RCP8.5 end of century (2081-2100) periods. Numbers of yearly extreme days are given in days per year, durations in days, intensities in $\mathrm{nmol\,kg^{-1}}$, and the volumes in $\mathrm{km^3}$. The first values represent ensemble means, values in brackets denote ensemble minima and maxima.

	PI	1986-2005	EOC RCP2.6	EOC RCP8.5
Number Surf.	3.65	1.75 (1.50-2.20)	2.24 (1.86-2.93)	1.36 (1.09-1.69)
200m	3.65	1.98 (1.51-2.77)	3.01 (2.28-3.71)	1.72 (1.38-2.02)
Duration Surf.	19.70	17.84 (16.84-18.92)	19.37 (18.07-21.13)	29.28 (27.37-32.57)
200m	38.61	66.06 (59.74-18.92)	98.71 (89.01-109.01)	111.56 (106.62-122.70)
Intensity Surf.	2.92	3.42 (3.26-3.64)	3.21 (3.07-3.48)	1.51 (1.42-1.63)
200m	3.26	4.96 (3.87-6.67)	7.90 (6.05-11.06)	6.02 (2.85-9.13)
Volume	3640	3158 (2888-3460)	3662 (3021-4215)	3378 (3086-3714)

Table A3. The same as Table A2, but for Ω_A . Intensity is given in $10^{-3}\,\Omega_A$ units. The remaining units are identical to those in Table A2.

520 *Data availability.* The GFDL ESM2M simulations are available upon request.

Author contributions. FAB and TLF designed the study. FAB performed the simulations, assisted by TLF and JGJ. FAB performed the analysis and wrote the initial manuscript. All authors contributed significantly to the writing of the paper.

Competing interests. All authors declare no competing interests.

Disclaimer. The work reflects only the authors' view; the European Commission and their executive agency are not responsible for any use
525 that may be made of the information the work contains.
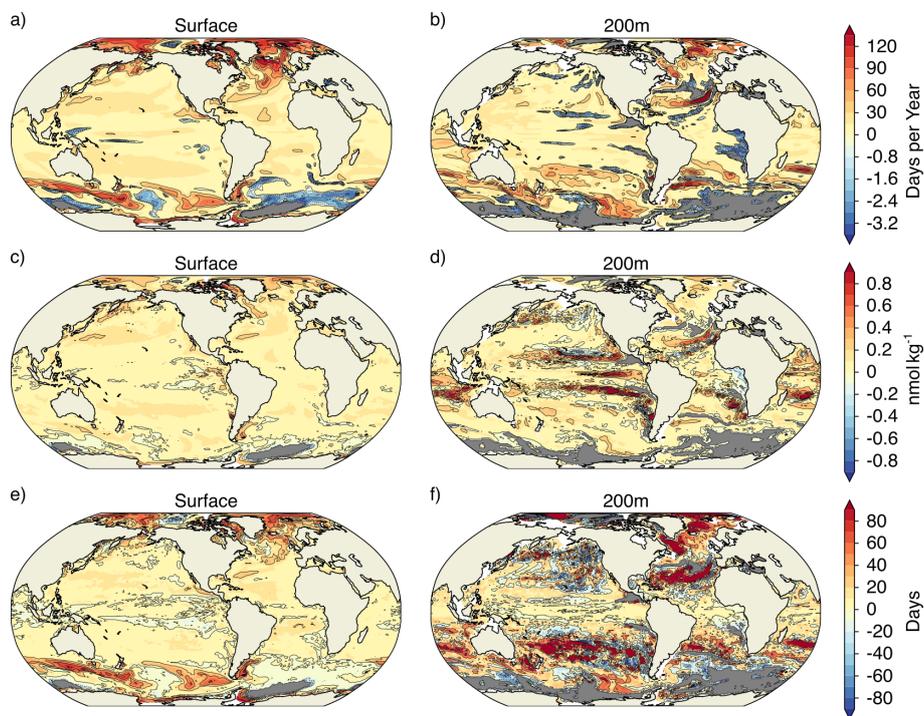
Figure A1. Simulated regional changes in (a,b) the number of extreme [H$^+$] days per year, (c,d) the maximal intensity extreme [H$^+$] events, and (e,f) the duration of extreme [H$^+$] events between preindustrial and 2081-2100 following the RCP2.6 scenario. Left panels show changes for the surface, whereas right panels show changes for 200 m depth. Shown are changes averaged over all five ensemble members. The black contours highlight the pattern structures. Grey areas represent areas with no extremes.

Acknowledgements. FAB and TLF have received funding the Swiss National Science Foundation (PP00P2_170687) and from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820989 (project COMFORT, Our common future ocean in the Earth system — quantifying coupled cycles of carbon, oxygen, and nutrients for determining and achieving safe operating spaces with respect to tipping points). FAB and TLF also thank the CSCS Swiss National Supercomputing Centre for computing resources. The authors thank Elizabeth Drenkard, Fortunat Joos, and Jens Terhaar for discussion and comments, and Rick Slater for the help in porting the ESM2M model code to CSCS.

530
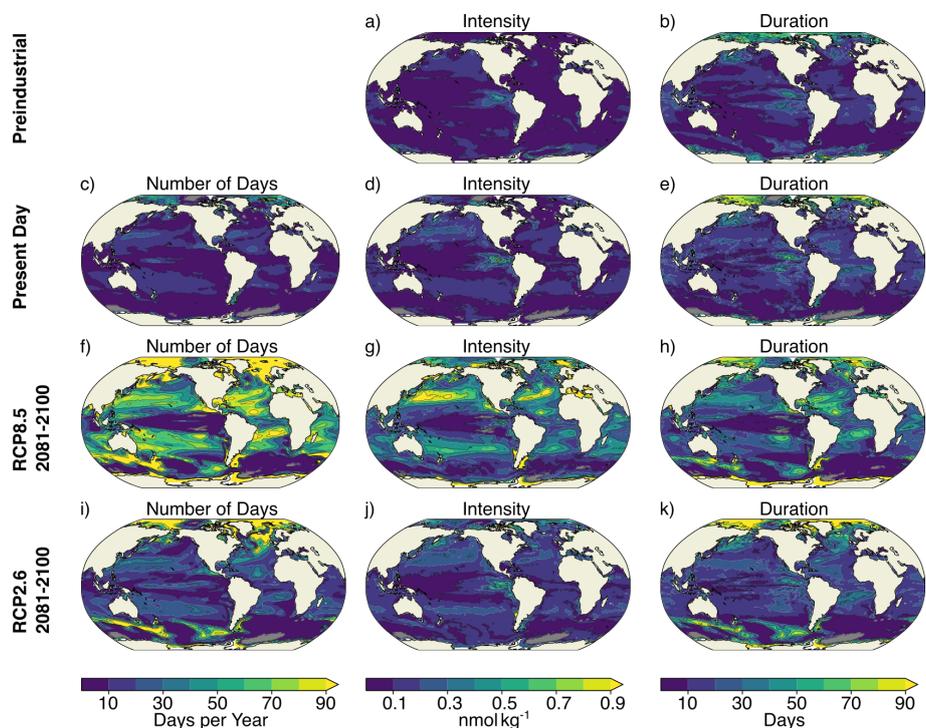
Figure A2. Simulated characteristics of surface [H$^+$] extreme events for preindustrial (a,b), present-day ensemble mean (1986-2005, c-e), RCP8.5 end of century ensemble mean (2081-2100, f-h), and RCP2.6 end of century ensemble mean (2081-2100, i-k). Grey colors represent regions where no ensemble member simulates extremes. The black contours highlight the pattern structures.

References

Anderson, J. L., Balaji., V., Broccoli, A. J., Cooke, W. F., Delworth, T. L., Dixon, K. W., Donner, L. J., Dunne, K. A., Freidenreich, S. M.,
Garner, S. T., Gudgel, R. G., Gordon, C. T., Held, I. M., Hemler, R. S., Horowitz, L. W., Klein, S. A., Knutson, T. R., Kushner, P. J.,
535 Langenhost, A. R., Cheung, L. N., Liang, Z., Malyshev, S. L., Milly, P. C. D., Nath, M. J., Ploshay, J. J., Ramaswamy, V., Schwarzkopf,
M. D., Shevliakova, E., Sirutis, J. J., Soden, B. J., Stern, W. F., Thompson, L. A., Wilson, R. J., Wittenberg, A. T., and Wyman, B. L.:
The New GFDL Global Atmosphere and Land Model AM2–LM2: Evaluation with Prescribed SST Simulations, J. Clim., 17, 4641–4673,
https://doi.org/10.1175/JCLI-3223.1, 2004.

Bakker, D. C. E., Pfeil, B., O’Brien, K. M., Currie, K. I., Jones, S. D., Landa, C. S., Lauvset, S. K., Metzl, N., Munro, D. R., Nakaoka,
540 S.-I., Olsen, A., Pierrot, D., Saito, S., Smith, K., Sweeney, C., Takahashi, T., Wada, C., Wanninkhof, R., Alin, S. R., Becker, M., Bellerby,
R. G. J., Borges, A. V., Boutin, J., Bozec, Y., Burger, E., Cai, W.-J., Castle, R. D., Cosca, C. E., DeGrandpre, M. D., Donnelly, M.,
Eischeid, G., Feely, R. A., Gkritzalis, T., González-Dávila, M., Goyet, C., Guillot, A., Hardman-Mountford, N. J., Hauck, J., Hoppema,
M., Humphreys, M. P., Hunt, C. W., Ibánhez, J. S. P., Ichikawa, T., Ishii, M., Juranek, L. W., Kitidis, V., Körtzinger, A., Koffi, U. K., Kozyr,
A., Kuwata, A., Lefèvre, N., Lo Monaco, C., Manke, A., Marrec, P., Mathis, J. T., Millero, F. J., Monacci, N., Monteiro, P. M. S., Murata,
545 A., Newberger, T., Nojiri, Y., Nonaka, I., Omar, A. M., Ono, T., Padín, X. A., Rehder, G., Rutgersson, A., Sabine, C. L., Salisbury,
J., Santana-Casiano, J. M., Sasano, D., Schuster, U., Sieger, R., Skjelvan, I., Steinhoff, T., Sullivan, K., Sutherland, S. C., Sutton, A.,
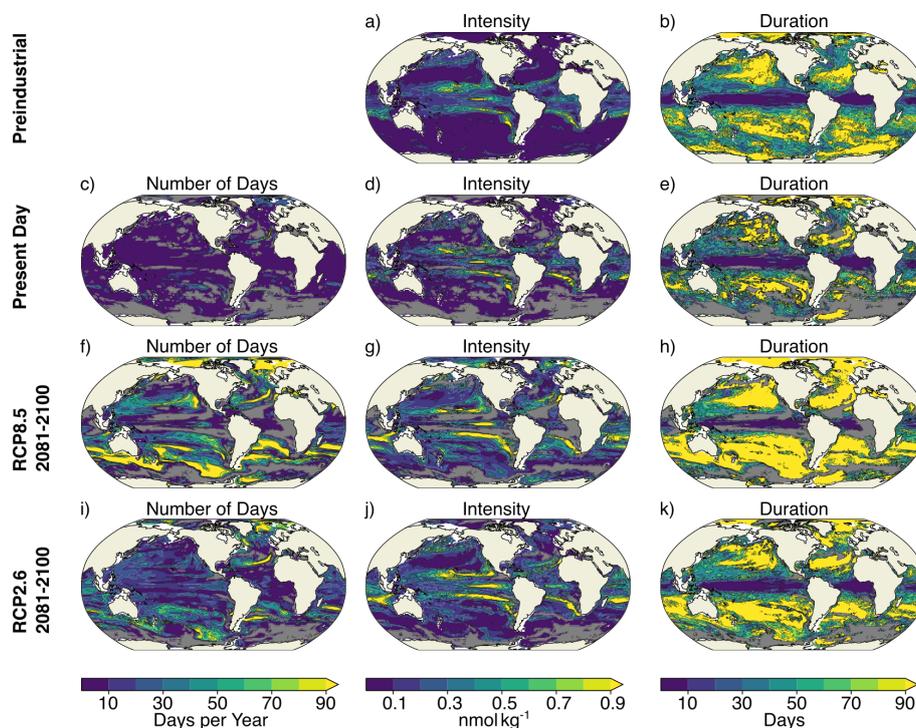
Figure A3. The same as Figure A2, but for 200 m depth.
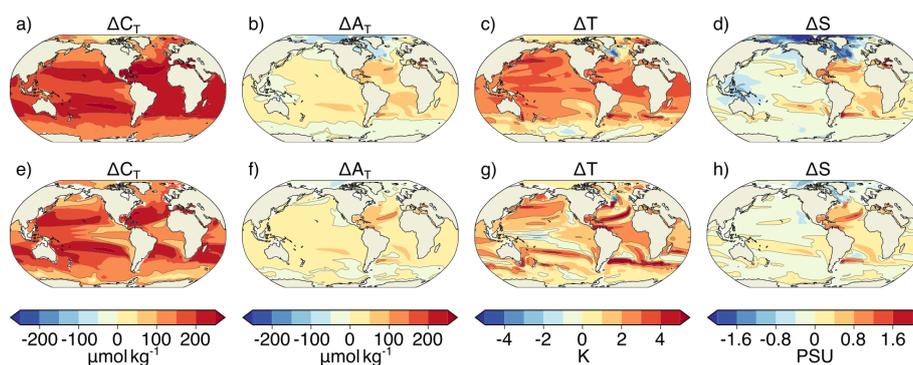
Figure A4. Simulated ensemble mean changes C_T (a,e), A_T (b,f), T (c,g), and S (d,h) from preindustrial to 2081-2100 following the RCP8.5 scenario. Shown are changes for (a-d) the surface and (e-h) at 200 m depth. The black contours highlight the pattern structures.

Tadokoro, K., Telszewski, M., Thomas, H., Tilbrook, B., van Heuven, S., Vandemark, D., Wallace, D. W., and Woosley, R.: Surface Ocean CO$_2$ Atlas (SOCAT) V4, https://doi.org/10.1594/PANGAEA.866856, 2016.

Bednaršek, N., Tarling, G. A., Bakker, D. C. E., Fielding, S., Jones, E. M., Venables, H. J., Ward, P., Kuzirian, A., Lézé, B., Feely, R. A., and
550   Murphy, E. J.: Extensive dissolution of live pteropods in the Southern Ocean, Nat. Geosci., 5, 881–885, https://doi.org/10.1038/ngeo1635, 2012.
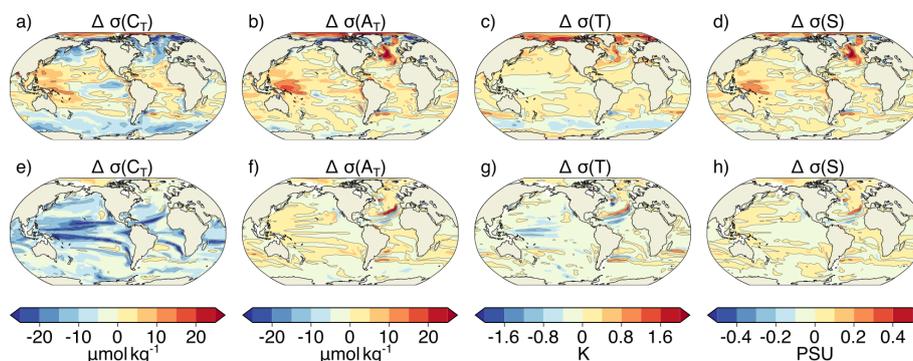
Figure A5. Simulated ensemble mean changes in the standard deviations of C_T (a,e), A_T (b,f), T (c,g), and S (d,h) from preindustrial to 2081-2100 under the RCP8.5 scenario. Shown are changes for (a-d) the surface and (e-h) at 200 m depth. The black contours highlight the pattern structures.

Bednaršek, N., Feely, R. A., Reum, J. C. P., Peterson, B., Menkel, J., Alin, S. R., and Hales, B.: *Limacina helicina* shell dissolution as an indicator of declining habitat suitability owing to ocean acidification in the California Current Ecosystem, Proc. R. Soc. B, 281, 20140 123, https://doi.org/10.1098/rspb.2014.0123, 2014.

555   Bindoff, N., Cheung, W., Kairo, J., Aristegui, J., Guinder, V., Hallberg, R., Hilmi, N., Jiao, N., Karim, M., Levin, L., O'Donoghue, S., Purca Cuicapusa, S., Rinkevich, B., Suga, T., Tagliabue, A., and Williamson, P.: Chapter 5: Changing Ocean, Marine Ecosystems, and Dependent Communities, IPCC Special Report on the Ocean and Cryosphere (SROCC), 2019.

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., and Vichi, M.: Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models, Biogeosciences, 10, 560   https://doi.org/10.5194/bg-10-6225-2013, 2013.

Britton, D., Cornwall, C. E., Revill, A. T., Hurd, C. L., and Johnson, C. R.: Ocean acidification reverses the positive effects of seawater pH fluctuations on growth and photosynthesis of the habitat-forming kelp, Ecklonia radiata, Sci. Rep., 6, 26 036, https://doi.org/10.1038/srep26036, 2016.

Caldeira, K. and Wickett, M. E.: Anthropogenic carbon and ocean pH, Nature, 425, 365–365, https://doi.org/10.1038/425365a, 2003.

565   Carter, B. R., Feely, R. A., Williams, N. L., Dickson, A. G., Fong, M. B., and Takeshita, Y.: Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate, Limnol. Oceanogr. Methods, 16, 119–131, https://doi.org/10.1002/lom3.10232, 2018.

Chatfield, C.: The analysis of time series - an introduction, Chapman & Hall/CRC, 5th edn., 1996.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer-Verlag London, 2001.

Collins, M., Sutherland, M., Bouwer, L., Cheong, S.-M., Frölicher, T., Des Combes, H., Koll, M., Losada, I., McInnes, K., Ratter, B., Rivera- 570   Arriga, E., Susanto, R., Swingedouw, D., and Tibig, L.: Chapter 6: Extremes, Abrupt Changes and Managing Risks, IPCC Special Report on the Ocean and Cryosphere (SROCC), 2019.

Doney, S. C., Fabry, V. J., Feely, R. A., and Kleypas, J. A.: Ocean Acidification: The Other CO_2 Problem, Annual Review of Marine Science, 1, 169–192, https://doi.org/10.1146/annurev.marine.010908.163834, 2009.

Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., Stouffer, R. J., Cooke, W., Dunne, K. A., Harrison, 575   M. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Phillipps, P. J., Sentman, L. T., Samuels, B. L., Spelman, M. J., Winton, M.,

Wittenberg, A. T., and Zadeh, N.: GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics, J. Clim., 25, 6646–6665, https://doi.org/10.1175/JCLI-D-11-00560.1, 2012.

Dunne, J. P., John, J. G., Shevliakova, E., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Sentman, L. T., Adcroft, A. J., Cooke, W., Dunne, K. A., Griffies, S. M., Hallberg, R. W., Harrison, M. J., Levy, H., Wittenberg, A. T., Phillips, P. J., and Zadeh, N.:
580   GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics, J. Clim., 26, 2247–2267, https://doi.org/10.1175/JCLI-D-12-00150.1, 2013.

Fassbender, A. J., Rodgers, K. B., Palevsky, H. I., and Sabine, C. L.: Seasonal Asymmetry in the Evolution of Surface Ocean pCO2 and pH Thermodynamic Drivers and the Influence on Sea-Air CO2 Flux, Global Biogeochem. Cycles, 32, 1476–1497, https://doi.org/10.1029/2017GB005855, 2018.

585 Feely, R., Chris, S., Hernandez-Ayon, J., Ianson, D., and Hales, B.: Evidence for Upwelling of Corrosive " Acidified" Water onto the Continental Shelf, Science, 320, 1490–1492, https://doi.org/10.1126/science.1155676, 2008.

Form, A. U. and Riebesell, U.: Acclimation to ocean acidification during long-term CO_2 exposure in the cold-water coral Lophelia pertusa, Glob. Change Biol., 18, 843–853, https://doi.org/10.1111/j.1365-2486.2011.02583.x, 2012.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C.,
590   Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon,
595   A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2019, Earth Syst. Sci. Data, 11, 1783–1838, https://doi.org/10.5194/essd-11-1783-2019, 2019.

Frölicher, T. L., Rodgers, K. B., Stock, C. A., and Cheung, W. W. L.: Sources of uncertainties in 21st century projections of potential ocean ecosystem stressors, Global Biogeoch Cycles, 30, 1224–1243, https://doi.org/10.1002/2015GB005338, 2016.

600 Frölicher, T. L., Fischer, E. M., and Gruber, N.: Marine heatwaves under global warming, Nature, 560, 360–364, https://doi.org/10.1038/s41586-018-0383-9, 2018.

Frölicher, T. L., Ramseyer, L., Raible, C. C., Rodgers, K. B., and Dunne, J.: Potential predictability of marine ecosystem drivers, Biogeosci. Discuss., 2020, 1–38, https://doi.org/10.5194/bg-2019-506, 2020.

Gallego, M. A., Timmermann, A., Friedrich, T., and Zeebe, R. E.: Drivers of future seasonal cycle changes in oceanic pCO_2, Biogeosciences,
605   15, 5315–5327, https://doi.org/10.5194/bg-15-5315-2018, 2018.

Gattuso, J.-P. and Buddemeier, R. W.: Calcification and CO_2, Nature, 407, 311–313, https://doi.org/10.1038/35030280, 2000.

Gehlen, M., Gruber, N., Gangstro, R., Bopp, L., and Oschlies, A.: Chapter 12: Biogeochemical Consequences of Ocean Acidification and Feedback to the Earth System, in: Ocean Acidification, edited by: Gattuso, J.-P. and Hansson, L., pp. 230–248, Oxford University Press, 2012.

610 Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, J. Geophys. Res. Oceans, 118, 6704–6716, https://doi.org/10.1002/2013JC009067, 2013.

Griffies, S.: ELEMENTS OF MOM4p1, GFDL ocean group technical report no.6, NOAA/Geophysical Fluid Dynamics Laboratory, Princeton University Forrestal Campus, 201 Forrestal Road, Princeton, NJ 08540-6649, 2009.

Gruber, N.: Warming up, turning sour, losing breath: ocean biogeochemistry under global change, Philos. Trans. R. Soc. A, 369, 1980–1996, https://doi.org/10.1098/rsta.2011.0003, 2011.

Hall-Spencer, J. M., Rodolfo-Metalpa, R., Martin, S., Ransome, E., Fine, M., Turner, S. M., Rowley, S. J., Tedesco, D., and Buia, M.-C.: Volcanic carbon dioxide vents show ecosystem effects of ocean acidification, Nature, 454, 96–99, https://doi.org/10.1038/nature07051, 2008.

Hauck, J. and Völker, C.: Rising atmospheric CO_2 leads to large impact of biology on Southern Ocean CO_2 uptake via changes of the Revelle factor, Geophys. Res. Lett., 42, 1459–1464, https://doi.org/10.1002/2015GL063070, 2015.

Hauri, C., Gruber, N., McDonnell, A. M. P., and Vogt, M.: The intensity, duration, and severity of low aragonite saturation state events on the California continental shelf, Geophys. Res. Lett., 40, 3424–3428, https://doi.org/10.1002/grl.50618, 2013.

Hofmann, G. E., Smith, J. E., Johnson, K. S., Send, U., Levin, L. A., Micheli, F., Paytan, A., Price, N. N., Peterson, B., Takeshita, Y., Matson, P. G., Crook, E. D., Kroeker, K. J., Gambi, M. C., Rivest, E. B., Frieder, C. A., Yu, P. C., and Martz, T. R.: High-Frequency Dynamics of Ocean pH: A Multi-Ecosystem Comparison, PloS One, 6, 1–11, https://doi.org/10.1371/journal.pone.0028983, 2011.

Joint, I., Doney, S. C., and Karl, D. M.: Will ocean acidification affect marine microbes?, ISME J., 5, 1–7, https://doi.org/10.1038/ismej.2010.79, 2011.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J. G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J. T., and Zaehle, S.: C4MIP – The Coupled Climate–Carbon Cycle Model Intercomparison Project: experimental protocol for CMIP6, Geosci. Model Dev., 9, 2853–2880, https://doi.org/10.5194/gmd-9-2853-2016, 2016.

Keller, K. M., Joos, F., and Raible, C. C.: Time of emergence of trends in ocean biogeochemistry, Biogeosciences, 11, 3647–3659, https://doi.org/10.5194/bg-11-3647-2014, 2014.

Kroeker, K. J., Micheli, F., Gambi, M. C., and Martz, T. R.: Divergent ecosystem responses within a benthic marine community to ocean acidification, Proc. Natl. Acad. Sci. U.S.A., 108, 14 515–14 520, https://doi.org/10.1073/pnas.1107789108, 2011.

Kroeker, K. J., Kordas, R. L., Crim, R., Hendriks, I. E., Ramajo, L., Singh, G. S., Duarte, C. M., and Gattuso, J.-P.: Impacts of ocean acidification on marine organisms: quantifying sensitivities and interaction with warming, Global Change Biology, 19, 1884–1896, https://doi.org/10.1111/gcb.12179, 2013.

Kroeker, K. J., Bell, L. E., Donham, E. M., Hoshijima, U., Lummis, S., Toy, J. A., and Willis-Norton, E.: Ecological change in dynamic environments: Accounting for temporal environmental variability in studies of ocean change biology, Glob. Change Biol., 26, 54–67, https://doi.org/10.1111/gcb.14868, 2020.

Kwiatkowski, L. and Orr, J. C.: Diverging seasonal extremes for ocean acidification during the twenty-first century, Nat. Clim. Change, 8, 141–145, https://doi.org/10.1038/s41558-017-0054-0, 2018.

Kwiatkowski, L., Gaylord, B., Hill, T., Hosfelt, J., Kroeker, K. J., Nebuchina, Y., Ninokawa, A., Russell, A. D., Rivest, E. B., Sesboüé, M., and Caldeira, K.: Nighttime dissolution in a temperate coastal ocean ecosystem increases under acidification, Sci. Rep., 6, 22 984, https://doi.org/10.1038/srep22984, 2016.

Landschützer, P., Gruber, N., and Bakker, D. C. E.: Decadal variations and trends of the global ocean carbon sink, Global Biogeochem. Cycles, 30, 1396–1417, https://doi.org/10.1002/2015GB005359, 2016.

Landschützer, P., Gruber, N., Bakker, D. C. E., Stemmler, I., and Six, K. D.: Strengthening seasonal marine CO_2 variations due to increasing atmospheric CO_2, Nat. Clim. Change, 8, 146–150, https://doi.org/10.1038/s41558-017-0057-x, 2018.

Lauvset, S. K., Carter, B. R., Perez, F. F., Jiang, L.-Q., Feely, R. A., Velo, A., and Olsen, A.: Processes Driving Global Interior Ocean pH Distribution, Global Biogeochem. Cycles, 34, e2019GB006 229, https://doi.org/10.1029/2019GB006229, 2020.

Leinweber, A. and Gruber, N.: Variability and trends of ocean acidification in the Southern California Current System: A time series from Santa Monica Bay, J. Geophys. Res. Oceans, 118, 3622–3633, https://doi.org/10.1002/jgrc.20259, 2013.

655 McNeil, B. I. and Sasse, T. P.: Future ocean hypercapnia driven by anthropogenic amplification of the natural CO_2 cycle, Nature, 529, 383–386, https://doi.org/10.1038/nature16156, 2016.

Mehrbach, C., Culberson, C. H., Hawley, J. E., and Pytkowicx, R. M.: MEASUREMENT OF THE APPARENT DISSOCIATION CONSTANTS OF CARBONIC ACID IN SEAWATER AT ATMOSPHERIC PRESSURE 1, Limnol. Oceanogr., 18, 897–907, https://doi.org/10.4319/lo.1973.18.6.0897, 1973.

660 Najjar, R. and Orr, J.: Design of OCMIP-2 simulations of chlorofluorocarbons, the solubility pump and common biogeochemistry, internal ocmip report, LSCE/CEA Saclay, Gif-sur-Yvette, France, ocmip5.ipsl.jussieu.fr/OCMIP/phase2/simulations/design.ps, 1998.

Oppenheimer, M., Glavovic, B., Hinkel, J., Van de Wal, R., Magnan, A., Abd-Elgawad, A., Cai, R., Cifuentes-Jara, M., Deconto, R., Ghosh, T., Hay, J., Isla, F., Marzeion, B., Meyssignac, B., and Sebesvari, Z.: Chapter 4: Sea Level Rise and Implications for Low Lying Islands, Coasts and Communities, IPCC Special Report on the Ocean and Cryosphere (SROCC), 2019.

665 Orr, J. C. and Epitalon, J.-M.: Improved routines to model the ocean carbonate system: mocsy 2.0, Geosci. Model Dev., 8, 485–499, https://doi.org/10.5194/gmd-8-485-2015, 2015.

Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Key, R. M., Lindsay, K., Maier-Reimer, E., Matear, R., Monfray, P., Mouchet, A., Najjar, R. G., Plattner, G.-K., Rodgers, K. B., Sabine, C. L., Sarmiento, J. L., Schlitzer, R., Slater, R. D., Totterdell, I. J., Weirig, M.-F., Yamanaka, Y., and Yool, A.: Anthropogenic ocean acidification
670 over the twenty-first century and its impact on calcifying organisms, Nature, 437, 681–686, https://doi.org/10.1038/nature04095, 2005.

Palter, J. B., Frölicher, T. L., Paynter, D., and John, J. G.: Climate, ocean circulation, and sea level changes under stabilization and overshoot pathways to 1.5 K warming, Earth Syst. Dyn., 9, 817–828, https://doi.org/10.5194/esd-9-817-2018, 2018.

Resplandy, L., Séférian, R., and Bopp, L.: Natural variability of CO_2 and O_2 fluxes: What can we learn from centuries-long climate models simulations?, J. Geophys. Res. Oceans, 120, 384–404, https://doi.org/10.1002/2014JC010463, 2015.

675 Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P.: RCP8.5 — A scenario of comparatively high greenhouse gas emissions, Clim. Change, 109, 33, https://doi.org/10.1007/s10584-011-0149-y, 2011.

Riebesell, U., Zondervan, I., Rost, B., Tortell, P. D., Zeebe, R. E., and Morel, F. M. M.: Reduced calcification of marine plankton in response to increased atmospheric CO_2, Nature, 407, 364–367, https://doi.org/10.1038/35030078, 2000.

Rodgers, K. B., Sarmiento, J. L., Aumont, O., Crevoisier, C., de Boyer Montégut, C., and Metzl, N.: A wintertime uptake window for
680 anthropogenic CO_2 in the North Pacific, Global Biogeochem. Cycles, 22, 1–16, https://doi.org/10.1029/2006GB002920, 2008.

Sarmiento, J. and Gruber, N.: Ocean Biogeochemical Dynamics, Princeton University Press, 2006.

Sentman, L. T., Shevliakova, E., Stouffer, R. J., and Malyshev, S.: Time Scales of Terrestrial Carbon Response Related to Land-Use Application: Implications for Initializing an Earth System Model, Earth Interact., 15, 1–16, https://doi.org/10.1175/2011EI401.1, 2011.

Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P. C. D., Caspersen, J. P., Sentman, L. T., Fisk, J. P., Wirth, C., and Crevoisier,
685 C.: Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink, Global Biogeochem. Cycles, 23, 1–16, https://doi.org/10.1029/2007GB003176, 2009.

Smale, D. A., Wernberg, T., Oliver, E. C. J., Thomsen, M., Harvey, B. P., Straub, S. C., Burrows, M. T., Alexander, L. V., Benthuysen, J. A., Donat, M. G., Feng, M., Hobday, A. J., Holbrook, N. J., Perkins-Kirkpatrick, S. E., Scannell, H. A., Sen Gupta, A., Payne, B. L., and

Moore, P. J.: Marine heatwaves threaten global biodiversity and the provision of ecosystem services, Nat. Clim. Change, 9, 306–312, https://doi.org/10.1038/s41558-019-0412-1, 2019.

Turi, G., Alexander, M., Lovenduski, N. S., Capotondi, A., Scott, J., Stock, C., Dunne, J., John, J., and Jacox, M.: Response of O_2 and pH to ENSO in the California Current System in a high-resolution global climate model, Ocean Sci., 14, 69–86, https://doi.org/10.5194/os-14-69-2018, 2018.

van Heuven, S., Pierrot, D., Rae, J., Lewis, E., and Wallace, D.: MATLAB Program Developed for CO_2 System Calculations, http://gts.sourceforge.net/, 2011.

van Vuuren, D. P., Stehfest, E., den Elzen, M. G. J., Kram, T., van Vliet, J., Deetman, S., Isaac, M., Klein Goldewijk, K., Hof, A., Mendoza Beltran, A., Oostenrijk, R., and van Ruijven, B.: RCP2.6: exploring the possibility to keep global mean temperature increase below 2°C, Clim. Change, 109, 95, https://doi.org/10.1007/s10584-011-0152-3, 2011.

Weiss, R.: Carbon dioxide in water and seawater: the solubility of a non-ideal gas, Mar. Chem., 2, 203 – 215, https://doi.org/https://doi.org/10.1016/0304-4203(74)90015-2, 1974.

Wernberg, T., Bennett, S., Babcock, R. C., de Bettignies, T., Cure, K., Depczynski, M., Dufois, F., Fromont, J., Fulton, C. J., Hovey, R. K., Harvey, E. S., Holmes, T. H., Kendrick, G. A., Radford, B., Santana-Garcon, J., Saunders, B. J., Smale, D. A., Thomsen, M. S., Tuckett, C. A., Tuya, F., Vanderklift, M. A., and Wilson, S.: Climate-driven regime shift of a temperate marine ecosystem, Science, 353, 169–172, https://doi.org/10.1126/science.aad8745, 2016.

Winton, M.: A Reformulated Three-Layer Sea Ice Model, J. Atmos. Ocean. Technol., 17, 525–531, https://doi.org/10.1175/1520-0426(2000)017<0525:ARTLSI>2.0.CO;2, 2000.

Wittenberg, A. T., Rosati, A., Delworth, T. L., Vecchi, G. A., and Zeng, F.: ENSO Modulation: Is It Decadally Predictable?, J. Clim., 27, 2667–2681, https://doi.org/10.1175/JCLI-D-13-00577.1, 2014.