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 long-term ocean acidification, but also due to increases in ocean acidity variability. Here, we use daily mean output from ensemble simulations of a comprehensive Earth system model under a low and high CO_2 emission scenario to isolate and quantify the impact of changes in variability on changes in variability-driven ocean acidity extreme events. Globally, the number of days with variability-driven extremely high $[H^+]$ 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 variability-driven extreme 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. Under a low emission scenario, the large increases in ocean acidity extreme event characteristics are substantially reduced. At surface, the changes are mainly driven by increases in $[H^+]$ seasonality, whereas changes in interannual variability are also important in the thermocline. Increases in $[H^+]$ variability and variability extremes arise predominantly from increases in the sensitivity of $[H^+]$ to variations in its drivers due to the increase in oceanic anthropogenic carbon. In contrast to $[H^+]$, the occurrence of variability-driven extremes in low aragonite saturation state is projected to decrease. The increase in $[H^+]$ variability and associated increase in extreme variability 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.

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 slows 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 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 $[HCO_3^-]$ from $[CO_3^{2-}]$. The associated decline in $[CO_3^{2-}]$ reduces the calcium carbonate saturation state $\Omega = [Ca^{2+}]\,[CO_3^{2-}] / \left([Ca^{2+}]\,[CO_3^{2-}]\right)_{sat}$, i.e.

the product of calcium ($[Ca^{2+}]$) and carbonate ion concentrations relative to the product at saturation. Undersaturated waters with $\Omega < 1$ are corrosive for calcium carbonate minerals. The calcium carbonate saturation state Ω differs between different mineral forms of calcium carbonate, such as aragonite and calcite, which differ in their solubilities. 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 (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 decadal- to centennial-scale ocean acidification trend are short-term extreme variability events on daily to monthly timescales, during which ocean pH and/or Ω are extremely low (Hofmann et al., 2011; Joint et al., 2011; Hauri et al., 2013). These events can be driven by a range of different processes, such as ocean mixing, biological production and remineralization, mineral dissolution, temperature and air-sea gas exchange variations, 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 (i.e. the depth between the supersaturated upper ocean and the undersaturated deep ocean (Feely et al., 2008; Leinweber and Gruber, 2013)). Ocean pH can also rapidly change as a consequence of microbial activity (Joint et al., 2011). Phytoplankton blooms and accompanying respiration drastically increase the partial pressure of CO_2 (pCO_2) and reduce pH in the thermocline (Sarmiento and Gruber, 2006). Such extreme variability events may have pH levels that are much lower than the mean pH conditions projected for the near future (Hofmann et al., 2011).

Most of the scientific literature on ocean acidification has focused on gradual changes in the mean state in ocean chemistry (Orr et al., 2005; Bopp et al., 2013; Frölicher et al., 2016). However, to understand the full consequences of ocean acidification for marine organisms and ecosystem services, it is also necessary to understand how variability extremes in ocean acidity 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 the species have evolved in a chemically stable or a highly variable environment (Rivest et al., 2017; Cornwall et al., 2020). Furthermore, if the frequency and intensity of short-term extreme variability events in ocean acidity strongly increase, some organisms may have difficulties to adapt, especially if key CO_2 system variables cross some critical thresholds, e.g. from calcium carbonate saturation to undersaturation. Key plankton species such as coccolithophores (Riebesell et al., 2000), foraminifera and pteropods (Bednaršek et al., 2012) were found to be adversely affected by low carbonate concentrations. After only several days of aragonite 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 in which deep-water corals are exposed to low-pH waters for a week, some corals exhibit reduced calcification, 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 growing evidence that the organism response to variability in ocean acidity could change with ocean acidification (Britton et al., 2016).

Therefore, understanding the temporal variability of ocean carbonate chemistry and the changes therein is of critical importance for understanding the impacts of ocean acidification on marine organisms and ecosystems (Hofmann et al., 2011).

Changes in extremes arise from changes in the mean, variability, or shape of the probability distribution (Coles, 2001). Under continued long-term ocean acidification (i.e. changes in the mean), one can expect that extreme events in $[H^+]$ and Ω will become more frequent and intense (Hauri et al., 2013). In addition to the changes in the mean, recent studies suggest that the seasonal cycles in $[H^+]$ and Ω are also strongly modulated under elevated atmospheric CO_2. Higher background concentrations of dissolved inorganic carbon and warmer temperatures produce stronger departures from mean state values for a given change in pertinent physical or chemical drivers for $[H^+]$ and weaker departures for Ω (Kwiatkowski and Orr, 2018; Fassbender et al., 2018). Other studies have also addressed the changes in the seasonal cycle of pCO_2 (Landschützer et al., 2018; Gallego 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 project that the winter-summer difference in surface $[H^+]$ will increase by 81%, whereas the seasonal amplitude for aragonite saturation state (Ω_A) is projected to decrease by 9 % on global 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 acidity events.

Unlike for marine heatwaves (Frölicher et al., 2018; Collins et al., 2019) and extreme sea level events (Oppenheimer et al., 2019), little is known about the characteristics and changes of extreme ocean acidity events and if so, only on seasonal timescales (Kwiatkowski and Orr, 2018). A global view of how extreme events in ocean chemistry due to changes in variability 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 extreme variability events in ocean acidity 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 extreme variability events in ocean acidity is a pressing issue.

In this study, we use daily mean output of a five-member ensemble simulation under a low and high CO_2 emissions scenario with a comprehensive Earth system model to investigate how changes in interannual, seasonal, and subannual variability under rising atmospheric CO_2 levels affect the occurrence, intensity, duration and volume of $[H^+]$ and Ω extreme variability events.

2 Methods

2.1 Model & experimental design

The simulations used in this study were made with the fully coupled carbon-climate Earth system model 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 modules, 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. The MOM4p1 model has a free surface and the surface level is centered around about 5 m depth and the spacing between consecutive levels is about 10 m down to a depth of about 230 m (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 of $2^\circ \times 2.5^\circ$ 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 horizontal 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, alkalinity, 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). The dissociation constants for carbonic acid and bicarbonate ions are from Dickson and Millero (1987), which are based on Mehrbach et al. (1973), and the carbon dioxide solubility is calculated according to Weiss (1974). Total alkalinity in ESM2M includes contributions from phosphoric and silicic acids and their conjugate bases. TOPAZv2 also simulates diurnal variability in ocean physics as well as in phytoplankton growth. While diurnal variations in open ocean pH 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 ran a 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, global 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 preindustrial and 2081-2100. The RCP2.6 scenario represents a low emission, high mitigation future (van Vuuren et al., 2011) with a simulated warming in GFDL ESM2M 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 a multi-century long preindustrial control simulation, that was extended with 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 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 (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). In addition, an accompanying 500-year preindustrial control simulation was performed.

2.2 Analysis methods

2.2.1 Extreme event definition and characterization

We analyze daily mean data of $[H^+]$ and the aragonite saturation state Ω_A in the upper 200 m of the water column. $[H^+]$ is on the total scale and hence the sum of the concentrations of free protons and sulfate ions. 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 mean data of the 500-year preindustrial 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 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 m (in km^3; mean volume of 3D clusters of connected grid cells that are above the 99th percentile for $[H^+]$ or below the 1st percentile for Ω_A, calculated using the *measure.label* function from 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. While the truncation of extremes between years alters the results for duration and maximal 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 to study changes in extreme events within the thermocline, where most organisms susceptible to ocean acidification are found, such as reef-forming corals and calcifying phytoplankton.

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 keep the seasonal signal in the data (further information in Appendix A). The removal of the secular trend ensures that the mean state in the processed data stays approximately constant while day-to-day to interannual variability can change over the simulation period (depicted for one grid cell in Figure 1). Thus, in our study, changes in the different extreme event characteristics are only

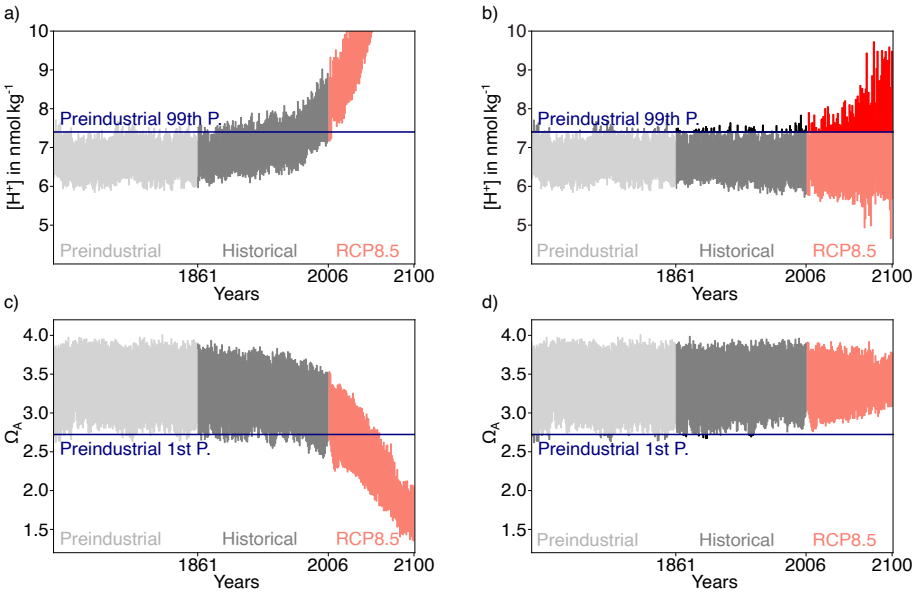

Figure 1. Simulated daily surface [H^+] (a) and Ω_A (c) at 40°N and 30°W in the North Atlantic for one ensemble member over the preindustrial, the 1861-2005 historical period, and the 2006-2100 period under RCP8.5. (b,d) Same as (a,c), but the ensemble-mean change with respect to the average of the 500-year long preindustrial control simulation has been subtracted. For [H^+], the preindustrial 99th 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 1st percentile (panel d).

caused by changes in variability and we call these events extreme variability events. In the Discussion section, we compare the simulated changes in extreme variability events to the total changes in extremes, which include changes due to secular trend in ocean acidity.

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

In order to assess whether changes in low or high frequency variability cause changes in extreme variability events and their characteristics, we use three steps to decompose the total variability in [H^+] into interannual, seasonal, and subannual variability (Figure 2). In a first step, we calculate the climatological seasonal cycle from the daily mean 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 of this 365-day long seasonal cycle. As described above, the secular trend in the daily mean data has been removed with the five-member ensemble mean before doing the analysis. In a second 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.signal* python library. In a third step, we calculate the variance arising from variations on interannual and subannual timescales from the spectral density to obtain interannual and subannual variability (further information is given in Appendix B). Following

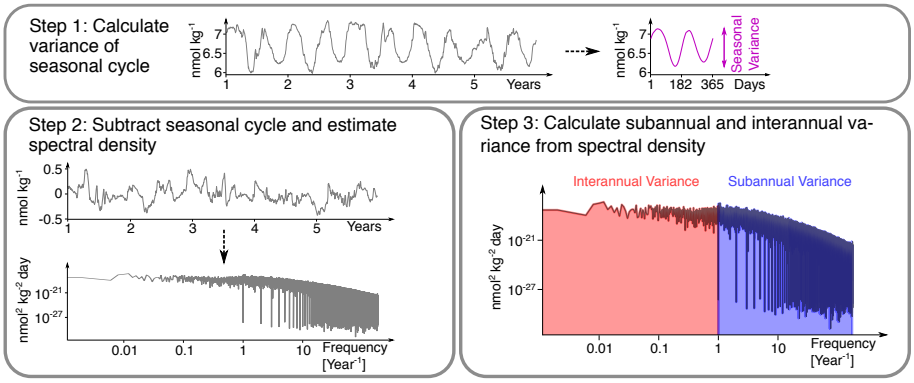

Figure 2. The three-step decomposition of $[H^+]$ variance into interannual, seasonal, and subannual variance, exemplified for a surface grid cell at 40° N and 30° W in the North Atlantic at preindustrial. In a first step, the climatological seasonal cycle is determined (over the whole period, only five years are depicted here) and its variance is calculated. Note that the seasonal cycle in this grid cell has two minima and maxima. In a second step, the spectral density of the anomalies with respect to the seasonal cycle is calculated. In a third step, interannual and subannual variance is estimated from the spectral density.

this methodology, subannual variability comprises all variations in daily mean data with periodicities of less than a year that are not part of the seasonal cycle.

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

To understand the processes behind the simulated changes in $[H^+]$ variability and variability extremes, we decompose these changes into contributions from changes in temperature (T), salinity (S), total alkalinity (A_T), and total dissolved inorganic carbon (C_T). 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

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

and analogously for Ω_A. 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. The partial derivatives in Equation 1 are evaluated using *Mocsy 2.0* (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 variance of $[H^+]$ and Ω_A 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 (e.g. Coles (2001)), it follows that the variance 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:

$$\begin{aligned}\sigma^2_{H^+} = &\left(\frac{\partial H^+}{\partial C_T}\right)^2 \sigma^2_{C_T} + \left(\frac{\partial H^+}{\partial A_T}\right)^2 \sigma^2_{A_T} + \left(\frac{\partial H^+}{\partial T}\right)^2 \sigma^2_{T} + \left(\frac{\partial H^+}{\partial S}\right)^2 \sigma^2_{S} \\ &+2\frac{\partial H^+}{\partial C_T}\frac{\partial H^+}{\partial A_T}\mathrm{cov}\left(C_T, A_T\right) + 2\frac{\partial H^+}{\partial C_T}\frac{\partial H^+}{\partial T}\mathrm{cov}\left(C_T, T\right) \\ &+2\frac{\partial H^+}{\partial C_T}\frac{\partial H^+}{\partial S}\mathrm{cov}\left(C_T, S\right) + 2\frac{\partial H^+}{\partial A_T}\frac{\partial H^+}{\partial T}\mathrm{cov}\left(A_T, T\right) \\ &+2\frac{\partial H^+}{\partial A_T}\frac{\partial H^+}{\partial S}\mathrm{cov}\left(A_T, S\right) + 2\frac{\partial H^+}{\partial T}\frac{\partial H^+}{\partial S}\mathrm{cov}\left(T, S\right), \end{aligned} \tag{2}$$

where the pairwise covariances are functions of the variances and correlation coefficients according to $\mathrm{cov}(x, y) = \sigma_x \sigma_y \rho_{x,y}$ and the partial derivatives are again evaluated at the temporal mean values $\overline{T}$, $\overline{S}$, $\overline{C_T}$, and $\overline{A_T}$. This methodology has also been used to propagate uncertainties in carbonate system calculations (Dickson and Riley, 1978; Orr et al., 2018) and to identify drivers of potential predictability in carbonate system variables (Frölicher et al., 2020). Based on Equation 2 and the analogous result for Ω_A, a change in variance in $[H^+]$ and Ω_A can be attributed to changes in the sensitivities that arise from changes in the drivers mean states, to changes in the drivers standard deviations, and to changes in the pairwise correlations between the drivers. We do so by calculating the full Taylor series of Equation 2 that has contributions up to the fifth order. We then identify the variance change from mean changes in the drivers as the sum of all terms in the expansion that describe the contributions of sensitivity changes to the overall change in variance ($\Delta_s \sigma^2_{H^+}$). Likewise, we identify the contribution from standard deviation changes in the drivers ($\Delta_\sigma \sigma^2_{H^+}$). We further group terms in the expansion that stem from simultaneous changes in the sensitivities and standard deviations ($\Delta_{s\sigma} \sigma^2_{H^+}$) and the remaining terms that arise either from correlation changes alone or mixed contributions from correlation changes and changes in sensitivities and standard deviations ($\Delta_{\rho+} \sigma^2_{H^+}$). Since these four components contain all terms in the Taylor series, they exactly reproduce a change in variance represented by Equation 2,

$$\Delta\sigma^2_{H^+} = \Delta_s \sigma^2_{H^+} + \Delta_\sigma \sigma^2_{H^+} + \Delta_{s\sigma} \sigma^2_{H^+} + \Delta_{\rho+} \sigma^2_{H^+}. \tag{3}$$

However, Equation 2 itself is an approximation to the simulated $[H^+]$ and Ω_A variance, leading to a small mismatch between the sum of the components introduced above and simulated variance change (black lines vs. grey dashed lines in the zonal mean plots in Figures 8, 9, and 10).

We also assess the contributions from sensitivity changes arising only from mean changes in C_T ($\Delta_s \sigma^2_{H^+}\big|_{C_T}$), the contribution from standard deviation changes in C_T alone ($\Delta_\sigma \sigma^2_{H^+}\big|_{C_T}$), and the contribution from simultaneous mean and standard

deviation changes in C_T ($\Delta_{s\sigma}\sigma^2_{H^+}\big|_{C_T}$). Further information on the decomposition is given in Appendix C.

215

2.3 Model evaluation

The focus of our analysis is on changes in variability in $[H^+]$ and Ω_A. As observation-based daily data of the inorganic 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. We developed an observation-
220 based dataset for surface monthly $[H^+]$ and Ω_A using monthly surface salinity, temperature, pCO_2, and A_T fields. 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 temperature (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 appears to
225 be generally well suited for analyzing pCO_2 seasonality and changes therein (Landschützer et al., 2018). An exception is the Southern Ocean where data-based pCO_2 products are uncertain due to sparse data in winter (Gray 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). Uncertainties in the derived seasonal cycles for $[H^+]$ and Ω_A that arise from uncertainties in the observation-based input variables are not quantified in this study.

230

In most regions, the GFDL ESM2M captures the observation-based mean seasonal cycle in $[H^+]$ and Ω_A well, in particular for Ω_A (the mean values of the seasonal amplitudes in Figure 3). 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.

235

We then compare the simulated ensemble-mean trends in seasonal amplitude with the observation-based estimates (further information on the methodology is given in Appendix D). Similar as for the mean seasonal cycle, 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 3). The ensemble-mean trends in the simulated seasonal $[H^+]$ amplitudes are positive for all latitude
240 bands (Figure 3, Table 1), consistent with the observation-based estimates. While the estimates for the simulated trends are significantly larger than zero for all latitude bands, this is not the case for the observation-based trends in the equatorial region (10°S - 10°N) and the northern low latitudes (10°N - 40°N) (Table 1). The simulated $[H^+]$ seasonality trends are significantly smaller (with 90% confidence level) than estimated from observations in the northern high (40°N - 90°N; orange thick lines in Figure 3a,b) and southern low latitudes (10°S - 40°S; blue thick lines in Figure 3a,b), where the trends from the model en-
245 semble are 0.031 ± 0.012 nmol kg^{-1} decade^{-1} and 0.035 ± 0.003 nmol kg^{-1} per decade, compared to the observational-based trends of 0.106 ± 0.040 nmol kg^{-1} decade^{-1} and 0.055 ± 0.014 nmol kg^{-1} decade^{-1}, respectively. The simulated ensemble

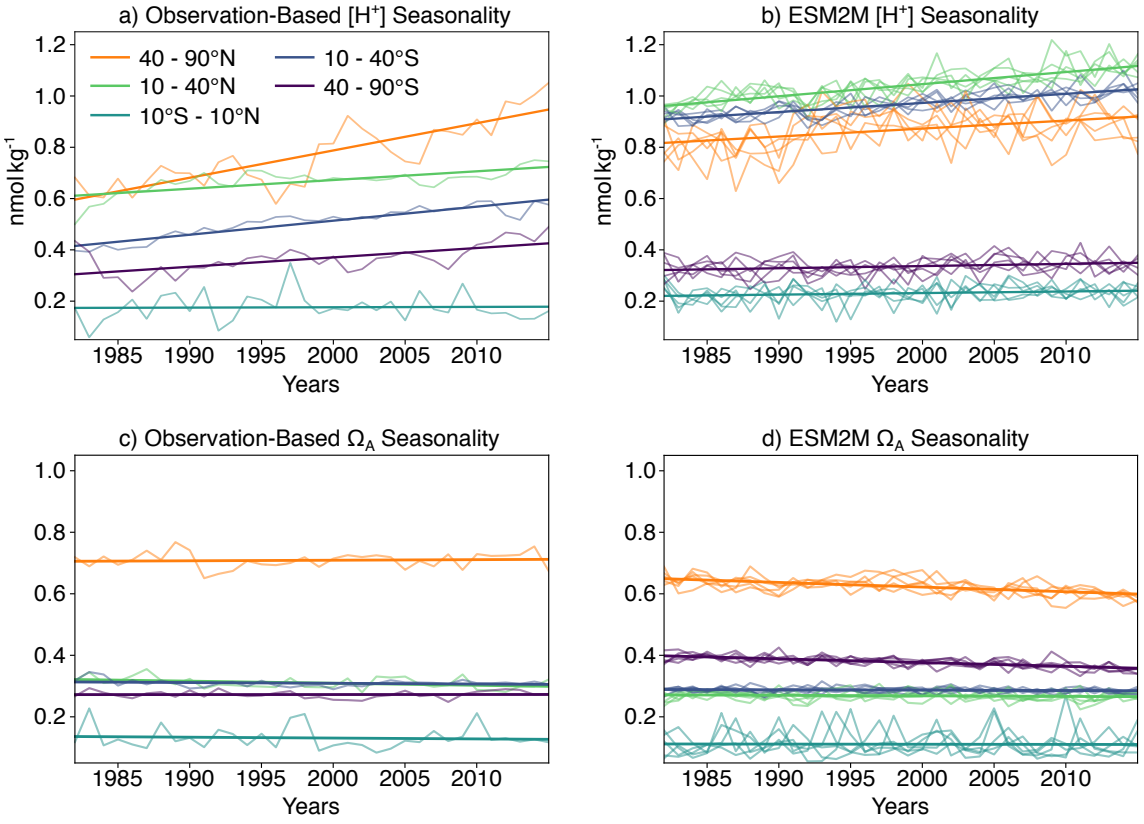

Figure 3. (a,b) Seasonal amplitude of $[H^+]$ over the period 1982-2015 averaged over five different latitude bands for a) the observation-based estimate and b) the GFDL ESM2M historical (1982-2005) and RCP8.5 (2006-2015) ensemble simulations. (c,d) The same as (a,b), but for Ω_A. Linear trends in all panels are overlaid as thick lines. The linear trend of the simulated changes is calculated as the mean of the five individual ensemble trends.

mean trends for the remaining latitude bands are not significantly different from the observation-based trend estimates.

For Ω_A, we find a significant negative trend in the observation-based data in the northern low latitudes and significant negative trends in the simulations in the northern and southern high latitudes (Table 1). The negative trends in seasonal amplitude in the simulations are significantly different from the observation-based trends in the northern high latitudes (-0.015 ± 0.004 vs. 0.002 ± 0.009 Ω_A units) and in the southern high latitudes (-0.012 ± 0.002 vs. 0.000 ± 0.005 Ω_A units per decade).

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 increases in seasonal amplitude of $[H^+]$, especially in the northern and southern high latitudes. In contrast to the observation-based data, the model also projects negative trends in the Ω_A seasonal amplitude there. Nevertheless, the observation-based trends in the northern and especially southern high latitudes are rather

Latitude	Obs. $[H^+]$	ESM2M $[H^+]$	Obs. Ω_A	ESM2M Ω_A
40°N - 90°N	0.106 ± 0.040	0.031 ± 0.012	1.9 ± 8.7	-15.1 ± 3.8
10°N - 40°N	0.034 ± 0.034	0.047 ± 0.005	-6.7 ± 5.6	-1.8 ± 2.0
10°S-10°N	0.001 ± 0.016	0.006 ± 0.005	-2.8 ± 10.7	-0.5 ± 5.3
40°S - 10°S	0.055 ± 0.014	0.035 ± 0.003	-2.4 ± 5.1	-1.2 ± 1.2
90°S - 40°S	0.037 ± 0.028	0.009 ± 0.004	0.1 ± 4.8	-12.2 ± 1.7

Table 1. 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) following the RCP8.5 scenario over 2006-2015. The range ($\pm$) denotes the 90 % confidence interval.

uncertain because winter time data is sparse there. Even though we lack the daily observational-based data to undertake a full assessment, it appears that the GFDL ESM2M model is adequate to assess changes in open ocean ocean acidification extreme events.

3 Results

We start by discussing the simulated changes in different ocean acidity extreme variability event characteristics at the global scale (i.e. grid cell based characteristics are aggregated globally), before we analyze changes at the local to regional scale and identify the drivers of changes. We recall that the large secular increase in $[H^+]$ and the large secular decrease in Ω_A was removed for the analysis as we focus on changes in variability and their impact on extreme variability event characteristics.

3.1 Global changes in ocean acidity variability extremes

In preindustrial times, the GFDL ESM2M suggests that an average surface $[H^+]$ extreme variability event had a maximal intensity of 0.08 nmol kg^{-1} (Figure 4c, Table 2) and lasted 11 days (Figure 4e). Ocean acidity extremes in the upper 200 m occur with a typical volume of $2.7 \cdot 10^3$km^3, which is about 0.004 % of the total ocean volume in the upper 200 m (Figure 4g). Over the historical period (from preindustrial to 1986-2005), the model projects that the number of surface $[H^+]$ extreme days increases from 3.65 days per year to 10.0 days per year (Figure 4a, ensemble ranges are given in Table 2). The maximal 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% increase in number of days per year, a 44% increase in the maximal intensity and a 45% increase in the duration of $[H^+]$ extreme variability events. The volume of individual events is projected to increase by 20% over the historical period. Over the 21st century, extreme variability events in ocean acidity are projected to further increase in frequency, intensity, duration, and volume (Figure 4). By 2081-2100 under the RCP8.5 scenario, the number of $[H^+]$ extreme days per year at surface is projected to increase to 50 days (corresponding to a 1273% increase relative to the preindustrial). The maximal intensity is projected to increase to 0.38 nmol kg^{-1} (371% increase), the duration to 32 days (199% increase) and the volume to $13.9 \cdot 10^3$km^3 (414%

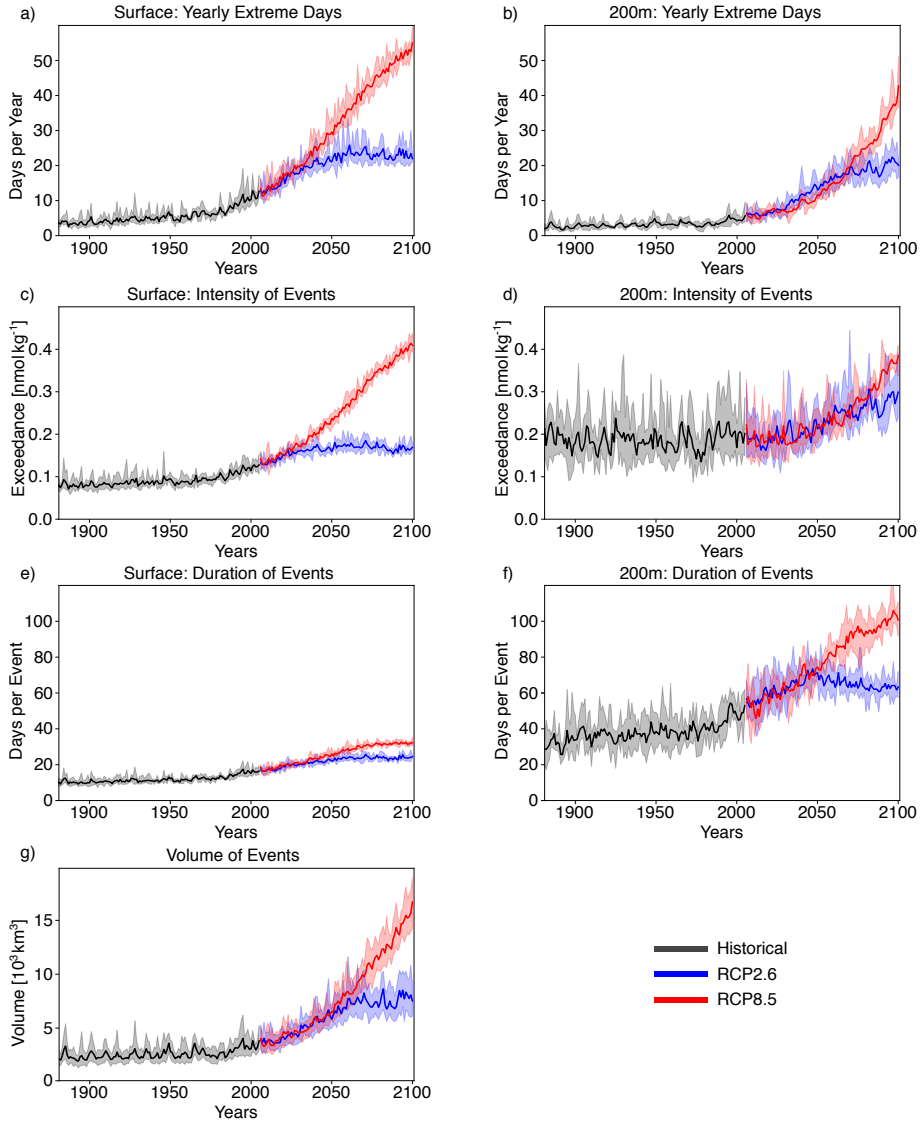

Figure 4. Simulated changes in globally averaged $[H^+]$ extreme variability event characteristics over the 1861-2100 period following historical (black lines) and future RCP8.5 (red) and RCP2.6 scenario (blue). Frequency, maximal intensity, and duration are shown for the surface (a,c,e) and for 200 m (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.

increase).

280

At 200 m, $[H^+]$ extreme variability events are in general more intense (0.17 nmol kg^{-1}; Figure 4d) and longer-lasting (38 days; Figure 4f) than at surface during preindustrial conditions. The stronger extreme events are caused by the overall larger

variability at 200 m than at surface in the preindustrial. The longer duration is connected to the more pronounced contribution from interannual variability (see Section 3.3). However, projected relative changes over the historical period and the 21st century are smaller at 200 m than at surface and with larger year-to-year variations across the ensembles. Under present-day conditions (1986-2005), the number of extreme days per year at 200 m is 4.3 days per year (corresponding to a 18% increase since preindustrial), the maximal intensity 0.20 nmol kg^{-1} (18% increase), and the duration 46 days (21% increase). By the end of the 21st century under the RCP8.5 scenario, the number of $[H^+]$ extreme days per year is projected to increase to 32.1 days per year, the maximal intensity to 0.34 nmol kg^{-1} and the duration to 99 days. Notably, extreme variability events in $[H^+]$ are projected to become less intense at 200 m 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 variability events remain shorter in duration at the end of the century than at 200 m.

	PI	1986-2005	2081-2100 RCP2.6	2081-2100 RCP8.5
Number Surf.	3.65	9.97 (9.49-10.38)	22.87 (21.93-23.45)	50.12 (49.98-50.30)
200 m	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)
200 m	38.00	45.95 (42.84-49.96)	62.94 (60.49-66.11)	98.66 (95.06-102.01)
Maximal Intensity Surf.	0.08	0.12 (0.11-0.12)	0.17 (0.16-0.17)	0.38 (0.37-0.39)
200 m	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 2. Simulated global ensemble-mean $[H^+]$ extreme variability event characteristics for the preindustrial (PI), present day (1986-2005), and the end of this century (2081-2100) for both RCP2.6 and RCP8.5. Numbers of yearly extreme days are given in days per year, durations in days, intensities in nmol kg^{-1}, and volumes in km^3. Values in brackets denote ensemble minima and maxima.

Under the RCP2.6 scenario and by the end of the century, the magnitude of changes in the different $[H^+]$ extreme variability event characteristics are substantially reduced compared to the RCP8.5 scenario. This reduction is especially pronounced at the surface (blue lines in Figure 4). There, the number of extreme days per year, maximal 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 of $[H^+]$ extreme days per year and the maximal intensity at depth as well as the volume of events are projected to increase significantly even after the atmospheric CO_2 concentration stabilizes in RCP2.6 around year 2050. This delayed response at subsurface is due to the relatively slow surface-to-subsurface 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 (Figure 4f). This decrease in duration mainly occurs in the subtropics, where events generally last long (Figure A2b). It is connected to an increase in the contribution from high-frequency variability to total variability in those regions over that period.

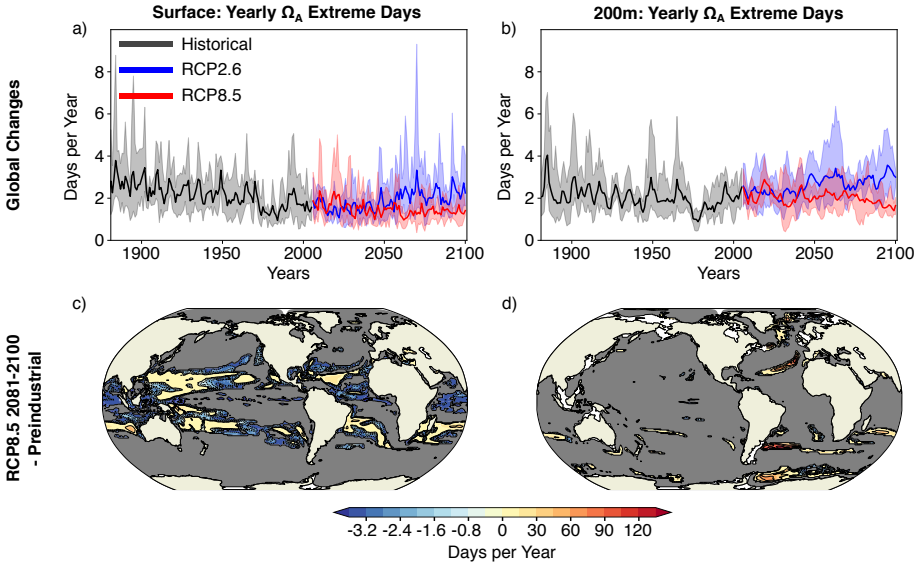

Figure 5. Simulated changes in the yearly number of Ω_A extreme variability days. Panels a-b) show the globally averaged simulated number of extreme variability days per year in Ω_A from 1861 to 2100 following historical (black lines) and future RCP2.6 (blue) and RCP8.5 (red) scenarios at (a) the surface and (b) 200 m. The thick lines display the five-member ensemble means and the shaded areas represent the maximum and minimum range of the individual ensemble members. Panels c-d) show the simulated regional changes in the number of extreme variability days per year in Ω_A from preindustrial to 2081-2100 under the RCP8.5 scenario (c) at the surface and (d) at 200 m. 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 variability extremes during 2081-2100.

In contrast to $[H^+]$ extreme variability events, the yearly number of Ω_A extreme variability days is projected to decrease over the historical and the 21st century under both the RCP8.5 and RCP2.6 scenario (Figure 5a-b, Supplementary Table A1). The number of surface Ω_A extreme variability days per year by the end of this century is projected to be 63% smaller under RCP8.5 and 39% smaller under RCP2.6 than at preindustrial (ensemble ranges are given in Supplementary Table A1). Projected changes at depth are less pronounced than at surface, again with larger decreases under RCP8.5 than under RCP2.6. It should be noted that, despite this decline in extreme variability events, the long-term decline in the mean state of Ω_A still leads to more frequent occurrence of low values in Ω_A (see Discussion section).

3.2 Regional changes in ocean acidity variability extremes

Surface $[H^+]$ variability 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 these ocean acidity extremes are not uniform over the globe (Figure 6; 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 $[H^+]$ variability extremes (Figure 6c) 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 variability events (Figure 6e). 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 variability events are projected to occur each year by the end of the century under RCP8.5, the increase in duration is relatively small.

At 200 m, the projected pattern of changes in yearly extreme event days generally resembles that at the surface (Figure 6b). 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, $[H^+]$ variability extremes at 200 m 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 indicating a decline in $[H^+]$ variability 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 variability events are projected to disappear in the RCP8.5 scenario by the end of this century (grey regions in Figure 6b). The largest increases in subsurface event intensity are projected in the subtropics (Figure 6d), whereas the duration of $[H^+]$ variability extremes is projected to increase strongly in many regions of the mid-to-high latitudes of both hemispheres (Figure 6f). The projected increases in duration at 200 m are much larger than at surface.

The increase in the number of extreme days per year, the maximal 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 variability 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 extreme variability events is projected to decrease largely overlap with those for RCP8.5, at surface and at depth. On the other hand, unlike under RCP8.5, a decrease in extreme variability event occurrence is only projected for a small fraction of the tropical oceans under RCP2.6.

While the decline in mean Ω_A generally leads to lower values in Ω_A, extreme variability 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; Figure 5c). In many regions, extreme variability events in Ω_A are projected to disappear by 2081-2100 under the RCP8.5 scenario (grey regions in Figure 5c). However, the frequency of surface Ω_A variability 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 variability events are projected for most of the ocean during 2081-2100 under RCP8.5 (Figure 5d).

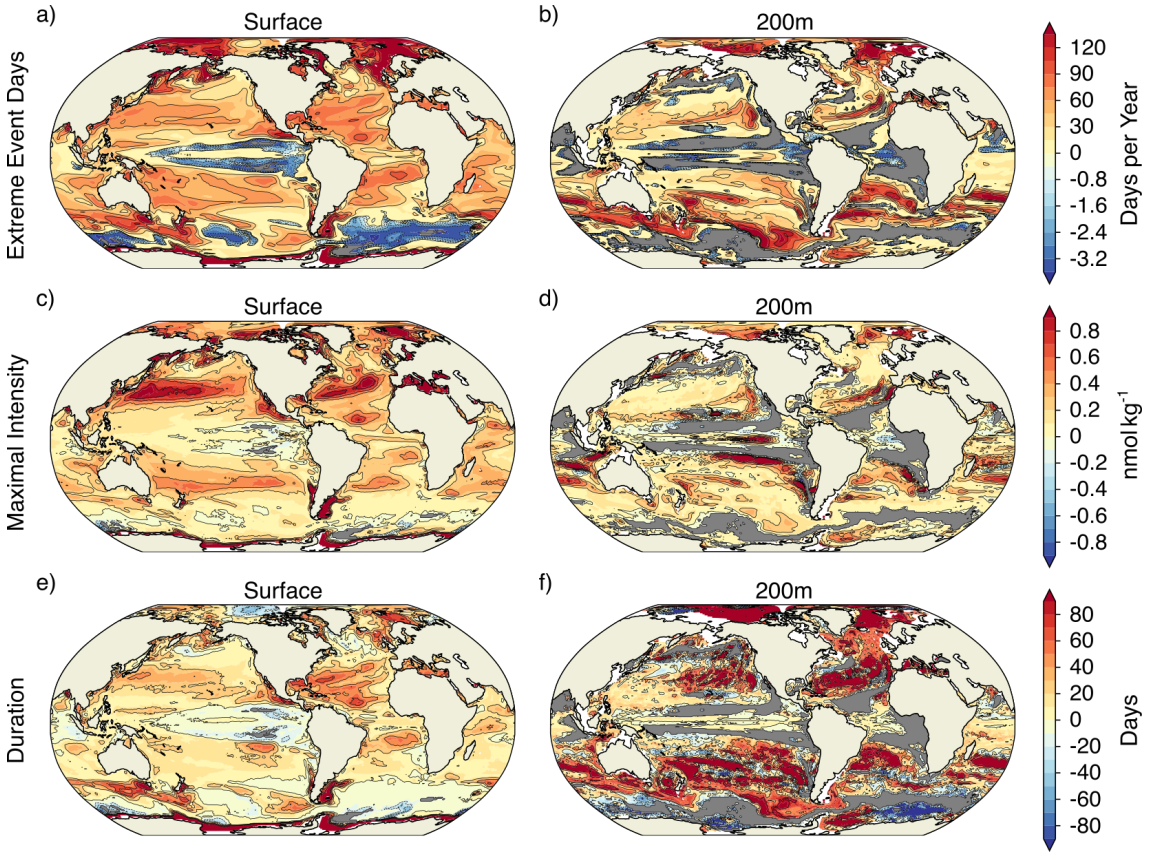

Figure 6. Simulated regional changes in $[H^+]$ extreme variability event characteristics from preindustrial to 2081-2100 under the RCP8.5 scenario at surface and at depth for (a,b) the number of extreme event days in days per year, (c,d) the maximal intensity of events in nmol kg^{-1}, and (e,f) the duration of events in days. Shown are changes averaged over all five ensemble members. Grey colors represent areas, where no variability extremes occur during 2081-2100 and the black lines highlight pattern structures.

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

The underlying changes in $[H^+]$ variability may arise from changes in interannual variability, seasonal variability, and subannual variability. We therefore decompose the total variability into these three components (see Section 2.2.2). For the preindustrial, the model simulates overall larger $[H^+]$ variance at depth than at the surface (0.42 $nmol^2 kg^{-2}$ vs. 0.15 $nmol^2 kg^{-2}$, not shown). Seasonality has the largest contribution at the surface (81 % of total variance). At 200 m, interannual variability has the largest contribution (63 %), and also subannual variability is more important compared to the surface (15% vs. 8%).

From preindustrial to the end of this century under the RCP8.5 scenario, 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,d). Changes in interannual variability (3% contribution to overall variance change; Figure 7a,d) and subannual variability (10%;

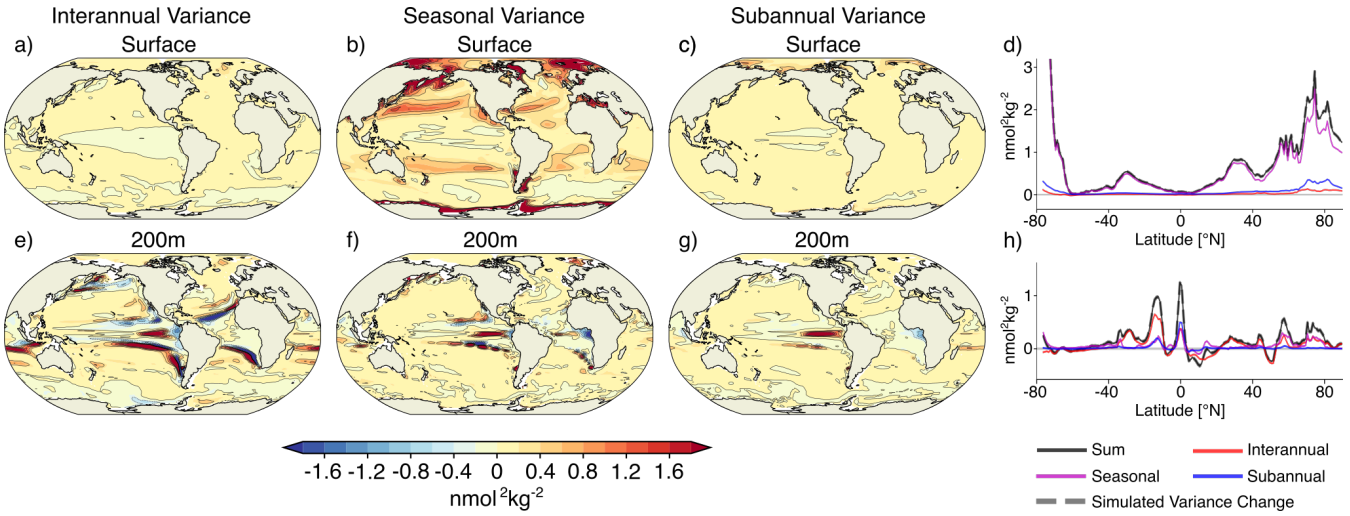

Figure 7. Contribution to projected changes in [H^+] variance from (a,e) interannual variability, (b,f) seasonal variability, and (c,g) subannual variability between the preindustrial and the 2081-2100 period following the RCP8.5 scenario at surface and at 200 m. Shown are the ensemble mean changes. The black lines highlight the pattern structure. Zonal mean contributions are shown for the surface (d) and for 200 m (h). The sum of the three components (black lines) accurately reproduces the simulated variance change (grey dashed lines).

Figure 7c,d) play a minor role. The largest increases in variability for all three variability types are projected for the northern high latitudes, where also the number of extreme variability event days increases most strongly. The increases in extreme events around Antarctica and the southern end of South America (Figure 6a) are mainly caused by increases in seasonal variability (Figure 7b). The regions that are projected to experience a decline in variability extremes (Figure 6a) coincide with those of decreasing interannual and seasonal variability (Figure 7a,b).

In contrast to the surface, changes in interannual and to a lesser extent subannual variability at 200 m are also important for explaining the changes in [H^+] variability extremes (Figure 7e,g,h). Changes in interannual variability contribute most to overall variance change at the global scale (with 42% contribution). Seasonal variability changes are almost equally important (37%), and changes in subannual 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 [H^+] variability are simulated north and south of the equator. In those regions the model also projects an increase in [H^+] extreme variability events (Figure 6b). Furthermore, these regions tend to be already more variable during the preindustrial (see Supplementary Figure A2a). However, the model also projects an increase in variability for less variable regions at preindustrial, such as northern high latitudes, leading to increases in variability extremes there. All three variability types are projected to decrease in the tropics and parts of the Southern Ocean, where the occurrence of extreme variability events is projected to largely decrease (c.f. Figure 6b). The variability decrease in those regions is most pronounced for interannual variability (Figure 7e).

3.4 Drivers of $[H^+]$ and Ω_A variability changes

In this section, we investigate the drivers of variability changes in $[H^+]$ and Ω_A. We attribute changes in variability to four factors (see Section 2.2.3 for further details): (i) changes in the mean states of the drivers that control the sensitivities ($\Delta_s\sigma^2_{H^+}$), (ii) changes in the variabilities of the drivers ($\Delta_\sigma\sigma^2_{H^+}$), (iii) simultaneous changes in the mean states and variabilities of the drivers ($\Delta_{s\sigma}\sigma^2_{H^+}$; this contribution arises because both mean states and variabilities change, and can neither be attributed to (i) nor (ii) alone), and (iv) changes in the correlations between the drivers, also including mixed contributions from correlation changes together with mean state and variability changes ($\Delta_{\rho+}\sigma^2_{H^+}$). In other words, (iv) describes the change in variability that arises because the correlations between the drivers also change, and not only their mean states and variabilities.

The drivers' mean changes between the preindustrial and 2081-2100 under RCP8.5 cause a strong increase in surface $[H^+]$ variability between the periods that is most pronounced in the northern and southern high latitudes ($\Delta_s\sigma^2_{H^+}$; Figure 8b, pink line in Figure 8h). On global average, these variance changes due to the mean changes in the drivers ($\Delta_s\sigma^2_{H^+}$=1.3 nmol^{2}kg^{-2}) are much larger than the total simulated variance change in $[H^+]$ ($\Delta\sigma^2_{H^+}$=0.5 nmol^{2}kg^{-2}, Figure 8a, dashed grey line in Figure 8h). In general, an increase in mean C_T, temperature, and salinity would lead to an increase in $\Delta_s\sigma^2_{H^+}$, 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 $\Delta_s\sigma^2_{H^+}\big|_{C_T}$ (Figure 8f, green line in Figure 8h). In the high latitudes, a relatively small increase in mean C_T leads a large increase in $\Delta_s\sigma^2_{H^+}\big|_{C_T}$, because $[H^+]$ is very sensitive to changes in C_T due to the low buffer capacity there. Decreases in mean A_T further contributes to the increase in $\Delta_s\sigma^2_{H^+}$ (not shown). This is not the case in the low-to-mid latitudes, where mean surface A_T is projected to increase, in particular in the Atlantic Ocean (Figure A4), and therefore dampens slightly the overall increase in $\Delta_s\sigma^2_{H^+}$. The changes in A_T are largely due to changes in freshwater cycling that also manifest in salinity changes (Supplementary Figure A4, Carter et al. (2016)). Mean changes in temperature and salinity play a minor role for explaining the large increase in $\Delta_s\sigma^2_{H^+}$ (not shown).

Why is the increase in $\Delta\sigma^2_{H^+}$ smaller than that following from the mean changes in the drivers (i.e. $\Delta_s\sigma^2_{H^+}$)? In the high latitudes, the projected change in the variability of the drivers (Supplementary Figure A5) contributes negatively to the $[H^+]$ variability change and counteracts to some degree the increase in $\Delta_s\sigma^2_{H^+}$. These variability changes alone would have a small impact on $\Delta_\sigma\sigma^2_{H^+}$ (Figure 8c), but the variability changes dampen the increases from the mean changes ($\Delta_{s\sigma}\sigma^2_{H^+}$, Figure 8d). The latter contribution is large in regions where mean changes would else lead to a strong increase (see anticorrelated patterns in Figures 8b and d). In the high latitudes, decreases in C_T variability (Supplementary Figure A5a) together with increases in mean C_T (Supplementary Figure A4a) can explain a large part of the negative contribution from $\Delta_{s\sigma}\sigma^2_{H^+}$ (Figure 8g and golden line in Figure 8h). In the northern high latitudes, also mean and variability changes in A_T are important for $\Delta_{s\sigma}\sigma^2_{H^+}$ (not shown). The additional contribution from changes in the correlations between the drivers ($\Delta_{\rho+}\sigma^2_{H^+}$; Figure 8e) also tends to contribute negatively to $[H^+]$ variability changes, especially in the North Atlantic.

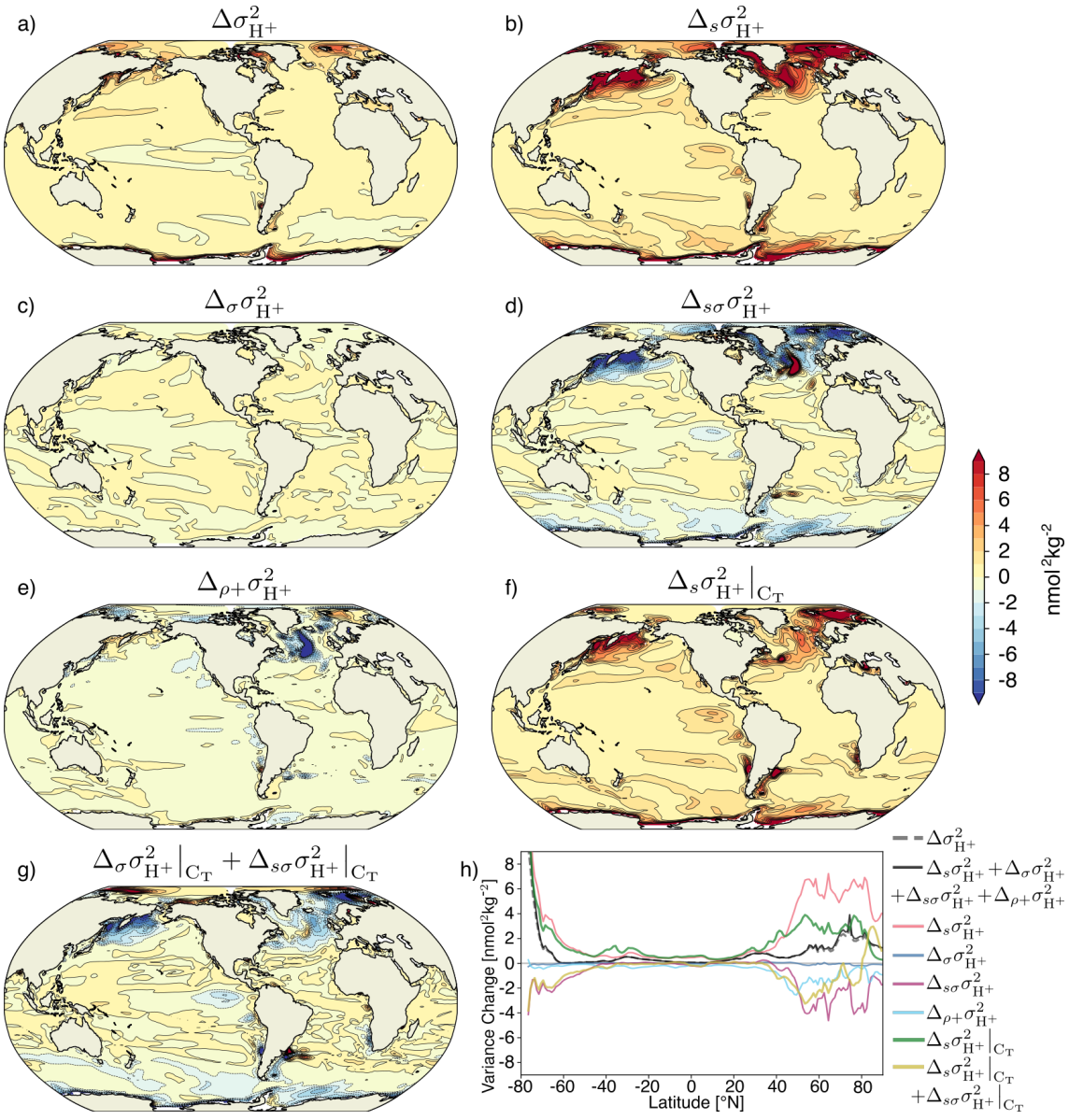

Figure 8. Decomposition of surface $[H^+]$ variability changes into different drivers. Shown are changes from preindustrial to 2081-2100 following the RCP8.5 scenario. The simulated change in $[H^+]$ variance ($\Delta\sigma^2_{H+}$) (a) is decomposed into the contribution from changes in the sensitivities that arise from changes in the drivers' mean values ($\Delta_s\sigma^2_{H+}$) (b), the contribution from changes in the drivers' standard deviations ($\Delta_\sigma\sigma^2_{H+}$) (c), the contribution from simultaneous changes in the sensitivities and the drivers' standard deviations ($\Delta_{s\sigma}\sigma^2_{H+}$) (d), and the contribution from correlation changes alone and simultaneous changes in correlations and sensitivities and standard deviations ($\Delta_{\rho+}\sigma^2_{H+}$) (e). Furthermore, the contribution to $[H^+]$ variance change from mean changes in C_T alone ($\Delta_s\sigma^2_{H+}|_{C_T}$) (f) and that from standard deviation changes in C_T together with simultaneous changes in mean state and standard deviation of C_T ($\Delta_\sigma\sigma^2_{H+}|_{C_T} + \Delta_{s\sigma}\sigma^2_{H+}|_{C_T}$) (g) is shown. The black contours in a-g) highlight the pattern structures. The zonal mean contribution of panels a-g) is shown in panel h).

At 200 m, the projected increase in $\Delta\sigma^2_{H^+}$ (Figure 9a) is also a result of the large increase due to the mean changes in the drivers ($\Delta_s\sigma^2_{H^+}$;Figure 9b; Supplementary Figure A4) and the decrease due to the interplay between mean changes and decreases in the variability ($\Delta_{s\sigma}\sigma^2_{H^+}$; Figure 9d). Similar to the surface, the changes in mean and variability of C_T are the most important drivers of changes (Figure 9f,g; green and golden lines in Figure 9h). Again, changes in temperature and salinity are of minor importance in most areas (not shown). In contrast to the surface, however, the individual compensating contributions to $[H^+]$ variability change from mean and variability changes in the drivers, in particular those in C_T, are much larger at 200 m. The global average variance change due to the mean changes in the drivers ($\Delta_s\sigma^2_{H^+} = 3.7$ nmol^{2}kg^{-2}) is much larger than the overall simulated variance change ($\Delta\sigma^2_{H^+} = 0.1$ nmol^{2}kg^{-2}). The 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. There, the preindustrial background $[H^+]$ variability is also the largest (Figure A3a). As a result, an increase in the sensitivities due to an increase in mean C_T has the largest effect there. The contribution from changes in the correlations between the drivers is overall small (Figure 9e and cyan line in Figure 9h).

Unlike for $[H^+]$, both mean changes ($\Delta_s\sigma^2_\Omega$; red lines in Figure 10) and variability changes in the drivers ($\Delta_\sigma\sigma^2_\Omega$: blue lines in Figure 10) lead to a decrease in Ω_A variability ($\Delta\sigma^2_\Omega$; black dashed lines in Figure 10). At 200 m, variability changes are even the dominant driver for reductions in Ω_A variability. Simultaneous changes in means and variabilities ($\Delta_{s\sigma}\sigma^2_\Omega$; purple lines in Figure 10) contribute positively and dampen the reduction in Ω_A variability from mean and variability changes alone. Mean and variability changes in C_T are the main drivers for changes in Ω_A variability as indicated by the tight relation between the dashed and solid red, blue, and purple lines in Figure 10, in particular at 200 m. An exception is the northern high latitudes, where A_T changes also play a substantial role at the surface (not shown). Correlation changes in the drivers ($\Delta_{\rho+}\sigma^2_\Omega$; cyan lines in Figure 10) are of similar relative importance as for $[H^+]$ and again have the largest imprint in the northern mid-to-high latitudes at the surface.

4 Discussion and conclusions

We provide a first quantification of the historical and future changes in extreme variability events in ocean acidity by analyzing daily mean 3D output from a ensemble simulation of a comprehensive Earth system model. In our 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 analyzing extremes. We show that extreme variability events in $[H^+]$ 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 $[H^+]$ extreme event characteristics are substantially reduced under the RCP2.6 scenario compared to RCP8.5. The increase in $[H^+]$ extreme variability events is a consequence of increased sensitivity of $[H^+]$ to variations in its 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 as well as changes in the correlations between the drivers. Extreme variability events in Ω_A are projected to become less frequent in the future. It is because

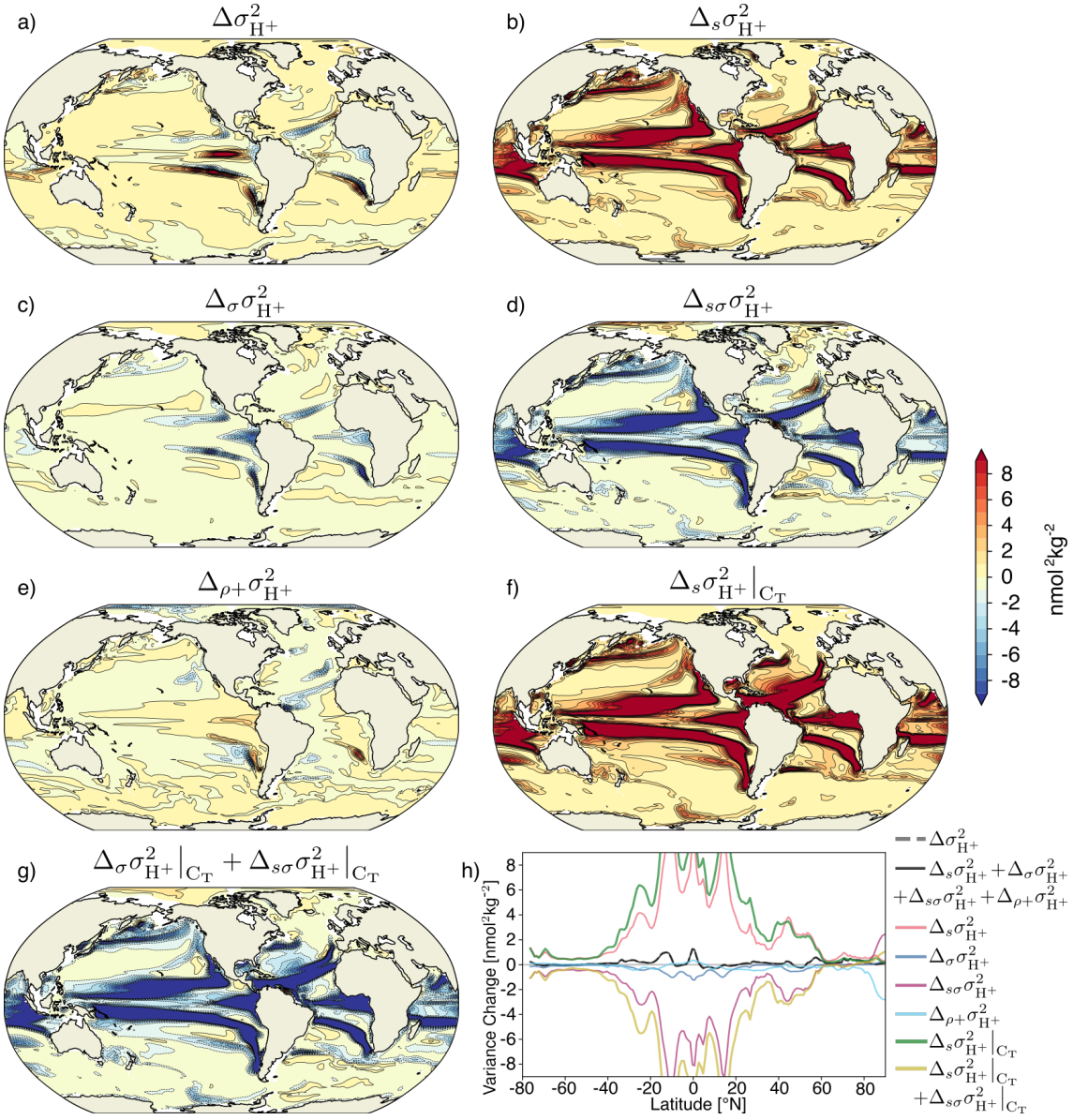

Figure 9. Same as Figure 8 but at 200 m.

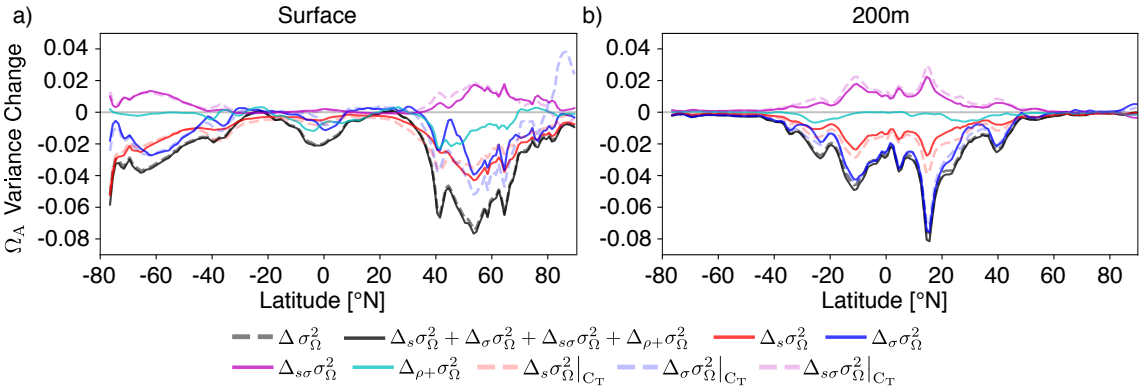

Figure 10. Decomposition of Ω_A variability changes into different drivers. The simulated zonal mean contribution to variance changes in Ω_A (black dashed lines, $\Delta\sigma^2_\Omega$) from preindustrial to 2081-2100 (RCP8.5) at the surface (a) and at 200 m (b). Shown is the contribution from sensitivity changes (due to mean changes in the drivers) (red lines, $\Delta_s\sigma^2_\Omega$), standard deviation changes in the drivers (blue lines, $\Delta_\sigma\sigma^2_\Omega$), simultaneous changes in sensitivities and standard deviations (purple lines, $\Delta_{s\sigma}\sigma^2_\Omega$), and all contributions that involve changes in the drivers' correlations (cyan lines, $\Delta_{\rho+}\sigma^2_\Omega$). Furthermore, contributions from mean changes, standard deviation changes and simultaneous mean and standard deviation changes in C_T alone are shown (dashed red, blue, and purple lines, respectively). In contrast to Figures 8 and 9 and due to their large contribution, we also show the zonal mean contribution from variability changes in C_T alone here.

Ω_A, unlike $[H^+]$, becomes less sensitive to variations in the drivers with the mean increase in C_T. Furthermore, the projected reductions in the drivers' variabilities, mainly in C_T, significantly add to the reduced occurrence of Ω_A variability extremes.

In this study, we analyze changes in extreme variability events that are defined relative to a shifting baseline. If the long-term increase in ocean acidity and decrease in Ω_A is taken into account, i.e. defining the extremes with respect to a fixed preindustrial baseline (here the preindustrial 99th percentile for $[H^+]$ and the preindustrial 1st percentile for Ω_A), the changes in $[H^+]$ and Ω_A extremes are much larger (cyan lines in Figure 11). Under the RCP8.5 scenario, every day becomes an extreme event day in year 2051 at surface and in year 2067 at 200 m depth (Figure 11a). The model also projects year-round extreme conditions for Ω_A at the surface and at 200 m by the end of the 21st century under RCP8.5 (Figure 11b). Comparing the two frameworks for surface $[H^+]$ extremes under present-day conditions, the annual number of extreme event days as defined in this study (i.e. with shifting baseline; black line in Figure 11) is on global average only 3.8 % of that also including the mean changes (i.e. with fixed preindustrial baseline; cyan line in Figure 11). This fraction differs regionally and reaches more than 10 % in the North Pacific, the North Atlantic, and the Arctic Ocean. Interestingly, the GFDL ESM2M projects that surface mean $[H^+]$ overshoots the preindustrial 99th percentile in year 1975 on global average. Thereafter, higher variability actually reduces the number of extreme event days that are above the preindustrial percentile. Surface mean Ω_A falls below the preindustrial 1st percentile in year 1990. After that, lower variability further increases the number of extreme event days below the preindustrial percentile.

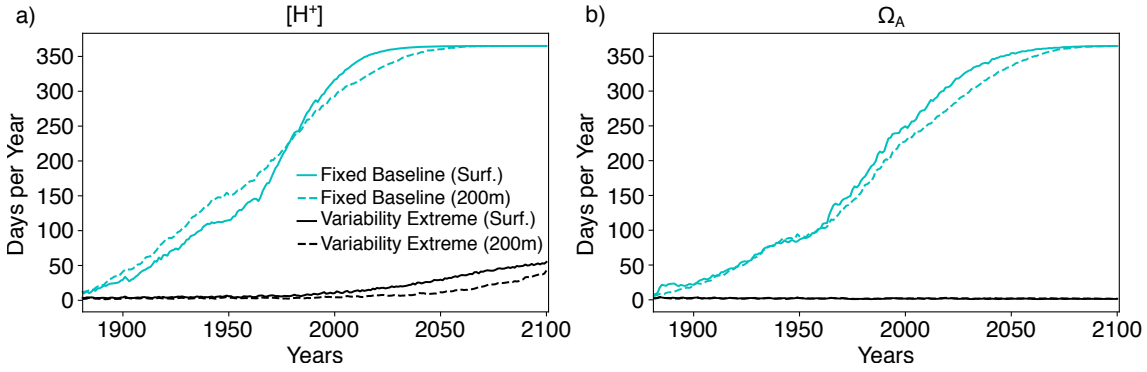

Figure 11. Simulated globally averaged number of extreme event days per year defined with a shifting baseline (black lines) and with a fixed preindustrial baseline (cyan lines) for $[H^+]$ using the 99th percentile (a) and for Ω_A using the 1st percentile (b) over the 1861-2100 period following the RCP8.5 scenario. Solid lines show results at the surface and dashed lines at 200 m.

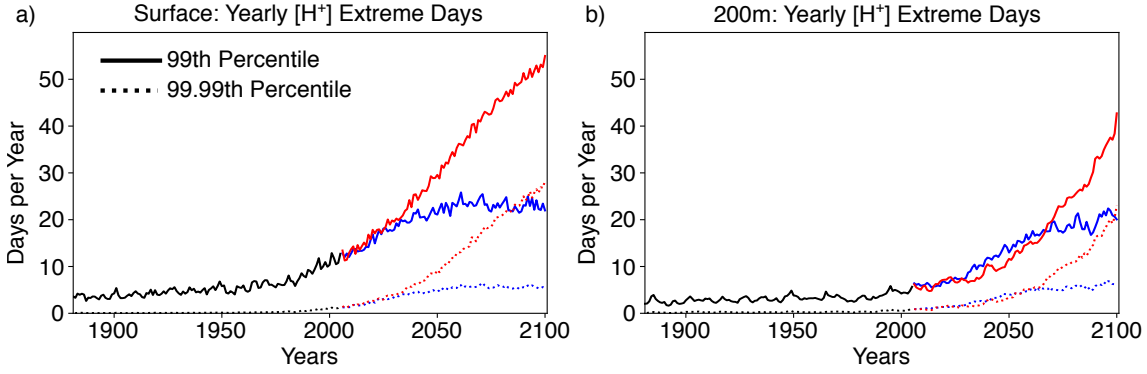

Figure 12. Globally averaged number of extreme variability event days for $[H^+]$ over the historical (black lines), RCP2.6 (blue), and RCP8.5 (red) simulations for the preindustrial 99th (solid lines) and 99.99th percentile (dotted lines) at (a) the surface and (b) 200 m.

We use the 99th percentile of the distribution from a preindustrial simulation for the definition of an extreme $[H^+]$ variability 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 variability extremes (Figure 12). For example, nearly every second day with $[H^+]$ exceeding the 99th percentile (red solid lines in Figure 12) is also a day with $[H^+]$ exceeding the 99.99th percentile (red dotted lines in Figure 12) 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 variability 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 12) as it projects days exceeding the 99th percentile under RCP2.6 (blue solid lines in Figure 12).

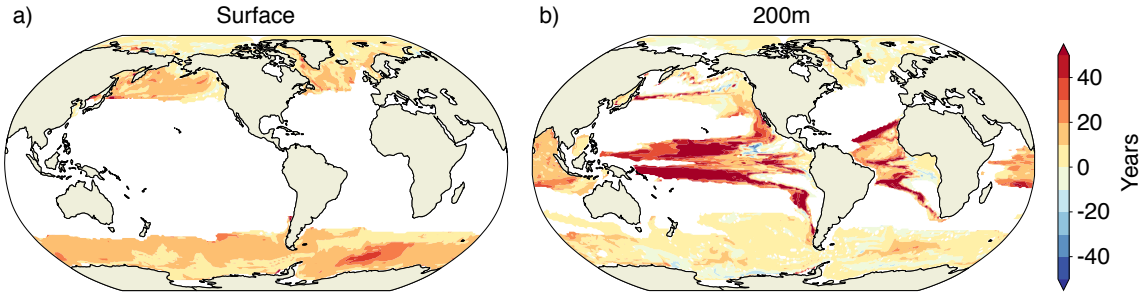

Figure 13. The temporal difference in years between the first occurrence of aragonite undersaturation in the historical and RCP8.5 ensemble and a hypothetical simulation where variability does not change over the 1861-2100 period, but only the mean changes. Positive values (yellow and red) indicate a delayed onset of undersaturation resulting from declines in Ω_A variability.

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 time where a grid cell falls below $\Omega_A = 1$ for the first time. To do so, we compare these times 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. At surface, these delays of undersaturation occur throughout the high latitudes (Figure 13a). At depth, the delays are most pronounced in the tropics (Figure 13b), but delays also occur in the high latitudes. Assuming unchanged seasonality, McNeil and Matear (2008) found that seasonal aragonite undersaturation of surface waters in the Southern Ocean may occur 30 years earlier than annual mean aragonite undersaturation. However, our simulation shows that the reduction in Ω_A variability delays the onset of undersaturation by about 10 to 15 years in the Southern Ocean relative to a hypothetical simulation where variability does not change. Therefore, changes in variability need to be taken into account when projecting the onset of seasonal undersaturation, especially in the high latitudes and in the thermocline of the tropics.

Previous studies have shown that the seasonal cycle of surface ocean pCO_2 will be strongly amplified under increasing atmospheric CO_2 (Gallego et al., 2018; Landschützer et al., 2018; McNeil and Sasse, 2016) and that a similar amplification is expected for surface $[H^+]$ (Kwiatkowski and Orr, 2018). Here we show that the changes in the seasonal cycle of $[H^+]$ translate into large increases in short-term extreme acidity events, at surface as well as at 200 m. In addition to earlier studies, we also show that changes in subannual variability contribute to changes in extreme $[H^+]$ variability events under increasing atmospheric CO_2 and that the average duration of extreme variability events at the surface and at present-day is about 15 days. It is therefore critical to use daily temporal output to assess extreme events in ocean acidity. Currently, ocean carbonate system variables from models that participate in the sixth phase of the Coupled Model Intercomparison Project are routinely stored with a monthly frequency on the Earth system grid (Jones et al., 2016). We therefore recommend to store and use high-frequency

output to study extreme events in the ocean carbonate systems.

Even though we consider our results as robust, a number of potential caveats remain. First, the horizontal resolution of the ocean model in GFDL 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 GFDL 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 representations are therefore needed to explore extreme events in ocean carbonate chemistry, especially in coastal regions. 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, in particular at the local scale, might depend on the model formulation. As the mean increases in C_T mainly drive the increases in extreme $[H^+]$ variability events (see Figure 8f), we expect that models with larger oceanic uptake of anthropogenic carbon show larger changes in extreme variability events than models with lower anthropogenic carbon uptake. The GFDL ESM2M matches observation-based estimates of historical global anthropogenic CO_2 uptake relatively well, but still has difficulties in representing the regional patterns in storage (Frölicher et al., 2015). Therefore, the exact regional patterns of C_T changes may differ from model to model and further studies focusing on the physical processes that lead to the regional C_T changes may help to better constrain the regional patterns in changes of acidity extremes. In addition, it is currently rather uncertain how $[H^+]$ and Ω_A variability changes as a result of changes in the drivers' variabilities. We have demonstrated that this factor is particularly important at depth for $[H^+]$ and for Ω_A. 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 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. Future studies with Earth system models that resolve diurnal processes are needed to quantify changes in diurnal variability and the impacts of these changes on extreme acidity events.

Our results also have important consequences for our understanding of the impact of ocean acidification on marine ecosystems. The projected increase in the frequency and the duration of ocean acidity variability extremes implies that marine organisms will have less time to recover from high $[H^+]$ events in the future. The large projected increase in $[H^+]$ extreme variability 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 are also projected to substantially increase (Frölicher et al., 2018), which also negatively impact marine ecosystems (Wernberg et al., 2016; Smale 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 shows that marine organisms and ecosystems are projected to be exposed to less stable [H^+] conditions in the future with more frequent occurrences of variability-driven short-term 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 the long-term increase in ocean acidity.

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

In this study, we analyze the changes in extreme variability events in [H^+] and Ω_A that arise from day-to-day to interannual variability changes in these variables. We therefore need to remove the secular trends from the data prior to analysis. We estimate the secular trend in a simulation from the five-member ensemble mean, assuming that subannual and interannual to decadal variations in the individual ensemble members are phased randomly and do not imprint on the ensemble mean 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. We remove the seasonal cycle, here defined as the 365-day long mean evolution over the course of a year, 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 also removes variability on subannual 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 subannual variability

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

$$f(\nu_j) = 2\frac{\Delta t^2}{T}\left|\sum_{k=1}^{N} x_k \cdot \exp\left(-\mathrm{i}2\pi\nu_j \cdot \Delta t\, k\right)\right|^2, \tag{B1}$$

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-Khintchine theorem, Chatfield (1996))[1]. As a consequence, the variance within the time series, given by the autocovariance

[1] In the continuous case, the theorem states

$$\gamma(\tau) = \int_{-\infty}^{\infty} \tilde{f}(\nu)\exp(\mathrm{i}2\pi\nu\tau)\mathrm{d}\nu, \tag{B2}$$

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 $\nu_{\max} = 1/2\Delta t$, the identity reads

$$\sigma^2 = \sum_{j=0}^{N/2} f(\nu_j)\frac{1}{N\Delta t}. \tag{B4}$$

Based on this equation, one can separate the contributions to variance from low-frequency and high-frequency variations. In this study, we determine interannual variability and subannual 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 $\nu_{\text{cut}} = 1/365\,\text{day}^{-1}$. Accordingly, subannual variability is obtained by evaluating the sum from $i_{\text{cut}}+1$ to $N/2$. Prior to this separation, the seasonal variability is removed from the data by subtracting the 365-day climatology.

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

Following Equation 2 in the main text, the variance in $[H^+]$ (or Ω_A) can be approximated as a function of the four sensitivities

$$\boldsymbol{s} = \left(\frac{\partial \mathrm{H}^+}{\partial \mathrm{A_T}}, \frac{\partial \mathrm{H}^+}{\partial \mathrm{C_T}}, \frac{\partial \mathrm{H}^+}{\partial \mathrm{S}}, \frac{\partial \mathrm{H}^+}{\partial \mathrm{T}}\right)^\mathsf{T} \tag{C1}$$

that in turn depend on the mean values of the drivers, the four standard deviations of the drivers

$$\boldsymbol{\sigma} = (\sigma_{\mathrm{A_T}}, \sigma_{\mathrm{C_T}}, \sigma_{\mathrm{S}}, \sigma_{\mathrm{T}})^\mathsf{T}, \tag{C2}$$

and the six pairwise correlation coefficients, in matrix notation given by

$$\boldsymbol{\rho} = \begin{pmatrix} 1 & \rho_{AC} & \rho_{AS} & \rho_{AT} \\ \rho_{AC} & 1 & \rho_{CS} & \rho_{CT} \\ \rho_{AS} & \rho_{CS} & 1 & \rho_{ST} \\ \rho_{AT} & \rho_{CT} & \rho_{ST} & 1 \end{pmatrix}. \tag{C3}$$

Based on this notation, we can rewrite Equation 2 of the main text as

$$\sigma^2_{\mathrm{H}^+} = \sum_{i=1}^{4}\sum_{j=1}^{4} s_i s_j \sigma_i \sigma_j \rho_{ij}. \tag{C4}$$

We use Equation C4 to decompose the variability change between the preindustrial and 2081-2100 into the contributions from changes in $\boldsymbol{s}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\rho}$. Since it is a polynomial of fifth order, its Taylor series has five orders, too[2]. In the following, all terms

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

$$\gamma(\tau) = \int_0^\infty f(\nu)\cos(2\pi\nu\tau)\mathrm{d}\nu \tag{B3}$$

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

[2]We use the drivers' standard deviations instead of their variances for the decomposition. With the latter, the Taylor expansion would have infinite terms and could not be decomposed exactly as it is done in the following. However, it would asymptotically lead to the same decomposition of $[H^+]$ variance change into $\Delta_s\sigma^2_{\mathrm{H}^+}$, $\Delta_\sigma\sigma^2_{\mathrm{H}^+}$, $\Delta_{s\sigma}\sigma^2_{\mathrm{H}^+}$, and $\Delta_{\rho+}\sigma^2_{\mathrm{H}^+}$ that is presented below.

of the Taylor series are given. We denote the sum of first order terms that contain changes in the four sensitivities $\Delta s_{1,\dots 4}$ by $\Delta_s^{(1)}\sigma_{\mathrm{H}^+}^2$, the sum of second order terms that contain changes in the sensitivities and standard deviations by $\Delta_{s\sigma}^{(2)}\sigma_{\mathrm{H}^+}^2$, and so on.

The first order is given by $\Delta^{(1)}\sigma_{\mathrm{H}^+}^2 = \Delta_s^{(1)}\sigma_{\mathrm{H}^+}^2 + \Delta_\sigma^{(1)}\sigma_{\mathrm{H}^+}^2 + \Delta_\rho^{(1)}\sigma_{\mathrm{H}^+}^2$ with

$$\begin{aligned}
\Delta_s^{(1)}\sigma_{\mathrm{H}^+}^2 =& 2\sum_{k=1}^{4}\sum_{j=1}^{4} s_j\sigma_k\sigma_j\rho_{kj}\Delta s_k \\
\Delta_\sigma^{(1)}\sigma_{\mathrm{H}^+}^2 =& 2\sum_{k=1}^{4}\sum_{j=1}^{4} s_k s_j\sigma_j\rho_{kj}\Delta\sigma_k \\
\Delta_\rho^{(1)}\sigma_{\mathrm{H}^+}^2 =& \sum_{k=1}^{4}\sum_{l=1}^{4} s_k s_l\sigma_k\sigma_l\Delta\rho_{kl}.
\end{aligned} \tag{C5}$$

The second order contains

$$\begin{aligned}
\Delta_{ss}^{(2)}\sigma_{\mathrm{H}^+}^2 =& \sum_{k=1}^{4}\sum_{l=1}^{4} \sigma_k\sigma_l\rho_{kl}\Delta s_k\Delta s_l \\
\Delta_{\sigma\sigma}^{(2)}\sigma_{\mathrm{H}^+}^2 =& \sum_{k=1}^{4}\sum_{l=1}^{4} s_k s_l\rho_{kl}\Delta\sigma_k\Delta\sigma_l \\
\Delta_{s\sigma}^{(2)}\sigma_{\mathrm{H}^+}^2 =& 2\sum_{k=1}^{4}\sum_{l=1}^{4} (s_l\sigma_l\rho_{kl}\Delta s_k\Delta\sigma_k + s_l\sigma_k\rho_{kl}\Delta s_k\Delta\sigma_l) \\
\Delta_{s\rho}^{(2)}\sigma_{\mathrm{H}^+}^2 =& 2\sum_{k=1}^{4}\sum_{l=1}^{4} s_l\sigma_k\sigma_l\Delta s_k\Delta\rho_{kl} \\
\Delta_{\sigma\rho}^{(2)}\sigma_{\mathrm{H}^+}^2 =& 2\sum_{k=1}^{4}\sum_{l=1}^{4} s_k s_l\sigma_l\Delta\sigma_k\Delta\rho_{kl}.
\end{aligned} \tag{C6}$$

The third order terms read

$$\begin{aligned}
\Delta_{ss\sigma}^{(3)}\sigma_{\mathrm{H}^+}^2 =& 2\sum_{k=1}^{4}\sum_{l=1}^{4} \sigma_l\rho_{kl}\Delta s_k\Delta s_l\Delta\sigma_k \\
\Delta_{s\sigma\sigma}^{(3)}\sigma_{\mathrm{H}^+}^2 =& 2\sum_{k=1}^{4}\sum_{l=1}^{4} s_l\rho_{kl}\Delta s_k\Delta\sigma_k\Delta\sigma_l \\
\Delta_{ss\rho}^{(3)}\sigma_{\mathrm{H}^+}^2 =& \sum_{k=1}^{4}\sum_{l=1}^{4} \sigma_k\sigma_l\Delta s_k\Delta s_l\Delta\rho_{kl} \\
\Delta_{\sigma\sigma\rho}^{(3)}\sigma_{\mathrm{H}^+}^2 =& \sum_{k=1}^{4}\sum_{l=1}^{4} s_k s_l\Delta\sigma_k\Delta\sigma_l\Delta\rho_{kl} \\
\Delta_{s\sigma\rho}^{(3)}\sigma_{\mathrm{H}^+}^2 =& 2\sum_{k=1}^{4}\sum_{l=1}^{4} (s_l\sigma_k\Delta s_k\Delta\sigma_l\Delta\rho_{kl} + s_l\sigma_l\Delta s_k\Delta\sigma_k\Delta\rho_{kl}).
\end{aligned} \tag{C7}$$

The fourth order reads

$$\Delta^{(4)}_{ss\sigma\sigma}\sigma^2_{\mathrm{H^+}} = \sum_{k=1}^{4}\sum_{l=1}^{4}\rho_{kl}\Delta s_k \Delta s_l \Delta\sigma_k \Delta\sigma_l$$

$$\Delta^{(4)}_{ss\sigma\rho}\sigma^2_{\mathrm{H^+}} = 2\sum_{k=1}^{4}\sum_{l=1}^{4}\sigma_l \Delta s_k \Delta s_l \Delta\sigma_k \Delta\rho_{kl}$$

$$\Delta^{(4)}_{s\sigma\sigma\rho}\sigma^2_{\mathrm{H^+}} = 2\sum_{k=1}^{4}\sum_{l=1}^{4}s_l \Delta s_k \Delta\sigma_k \Delta\sigma_l \Delta\rho_{kl} \tag{C8}$$

and the fifth order is given by

$$\Delta^{(5)}_{ss\sigma\sigma\rho}\sigma^2_{\mathrm{H^+}} = \sum_{k=1}^{4}\sum_{l=1}^{4}\Delta s_k \Delta s_l \Delta\sigma_k \Delta\sigma_l \Delta\rho_{kl}. \tag{C9}$$

We identify the variance change from changes in the sensitivities as

$$\Delta_s\sigma^2_{\mathrm{H^+}} = \Delta^{(1)}_s\sigma^2_{\mathrm{H^+}} + \Delta^{(2)}_{ss}\sigma^2_{\mathrm{H^+}}, \tag{C10}$$

the change from standard deviation changes as

$$\Delta_\sigma\sigma^2_{\mathrm{H^+}} = \Delta^{(1)}_\sigma\sigma^2_{\mathrm{H^+}} + \Delta^{(2)}_{\sigma\sigma}\sigma^2_{\mathrm{H^+}}, \tag{C11}$$

the change from simultaneous changes in sensitivities and standard deviations as

$$\Delta_{s\sigma}\sigma^2_{\mathrm{H^+}} = \Delta^{(2)}_{s\sigma}\sigma^2_{\mathrm{H^+}} + \Delta^{(3)}_{ss\sigma}\sigma^2_{\mathrm{H^+}} + \Delta^{(3)}_{s\sigma\sigma}\sigma^2_{\mathrm{H^+}} + \Delta^{(4)}_{ss\sigma\sigma}\sigma^2_{\mathrm{H^+}}, \tag{C12}$$

and that from correlation changes and mixed contributions that include correlation changes as

$$\Delta_{\rho+}\sigma^2_{\mathrm{H^+}} = \Delta^{(1)}_\rho\sigma^2_{\mathrm{H^+}} + \Delta^{(2)}_{s\rho}\sigma^2_{\mathrm{H^+}} + \Delta^{(2)}_{\sigma\rho}\sigma^2_{\mathrm{H^+}} + \Delta^{(3)}_{ss\rho}\sigma^2_{\mathrm{H^+}} + \Delta^{(3)}_{\sigma\sigma\rho}\sigma^2_{\mathrm{H^+}} + \Delta^{(3)}_{s\sigma\rho}\sigma^2_{\mathrm{H^+}} + \Delta^{(4)}_{ss\sigma\rho}\sigma^2_{\mathrm{H^+}} + \Delta^{(4)}_{s\sigma\sigma\rho}\sigma^2_{\mathrm{H^+}} + \Delta^{(5)}_{ss\sigma\sigma\rho}\sigma^2_{\mathrm{H^+}}. \tag{C13}$$

Finally, we calculate $\Delta_s\sigma^2_{\mathrm{H^+}}\big|_{\mathrm{C_T}}$, $\Delta_\sigma\sigma^2_{\mathrm{H^+}}\big|_{\mathrm{C_T}}$, and $\Delta_{s\sigma}\sigma^2_{\mathrm{H^+}}\big|_{\mathrm{C_T}}$, the analogues for Equations C10-C12 that only take into account changes in $\mathrm{C_T}$. This is done by calculating $\Delta s_{1,\ldots4}$ only based on mean changes in $\mathrm{C_T}$ and by setting the standard deviation changes for $\mathrm{A_T}$, S, and T to zero.

Appendix D: Comparison of simulated ensemble-mean trends in seasonal amplitude to observation-based trends

We construct confidence intervals for the observation-based slope estimates following Hartmann et al. (2013). For the simulations, we use the arithmetic average of the five ensemble-member slope estimates as the estimator,

$$\hat{\bar{b}} = \frac{1}{5}\sum_{k=1}^{5}\hat{b}_k \tag{D1}$$

with estimated variance

$$\hat{\sigma}^2_{\bar{b}} = \frac{1}{5^2}\sum_{k=1}^{5}\hat{\sigma}^2_{b_k}. \tag{D2}$$

We then construct the confidence interval for $\hat{\bar{b}}$ as

$$(\hat{\bar{b}} - q \cdot \hat{\sigma}_{\bar{b}}, \hat{\bar{b}} + q \cdot \hat{\sigma}_{\bar{b}}), \tag{D3}$$

with q the $(1+p)/2$-quantile (we use $p = 0.9$) of the t-distribution with $5 \cdot (N-2)$ degrees of freedom. We correct the sample size N (34, the number of years we use for the fits) to a reduced sample size N_r when we find positive lag-one autocorrelation in the residuals of the fits (data - linear regression model). Lag-one autocorrelation is estimated as the average of the five ensemble-member lag-one autocorrelation estimates

$$\hat{\bar{\rho}} = \frac{1}{5} \sum_{k=1}^{5} \hat{\rho}_k. \tag{D4}$$

and we obtain $N_r = N \cdot (\hat{\bar{\rho}} - 1)/(\hat{\bar{\rho}} + 1)$. Positive $\hat{\bar{\rho}}$ is only found in the northern high latitudes. This is in contrast to the observation-based case, where we find large positive $\hat{\rho}_o$ (up to 0.7) in the residuals of all latitude bands besides the tropical region.

For testing the significance of a difference between the simulation slope estimate $\hat{\bar{b}}$ and the observation-based estimate $\hat{b}_o$, we use Welch's test that assumes different variances for the two estimates (Andrade and Estévez-Pérez, 2014). The variance of the simulation slope estimate is calculated by dividing the ensemble-averaged slope variance by the ensemble size (Equation D2) and is hence smaller than the observation-based slope variance. If the absolute value of the test statistic

$$\frac{\hat{\bar{b}} - \hat{b}_o}{\sqrt{\hat{\sigma}_{\bar{b}}^2 + \hat{\sigma}_o}} \tag{D5}$$

is larger than the $(1+p)/2$-quantile of the t distribution with (Andrade and Estévez-Pérez, 2014)

$$\frac{\left(\hat{\sigma}_{\bar{b}}^2 + \hat{\sigma}_{b_o}^2\right)^2}{\hat{\sigma}_{\bar{b}}^4/(5 \cdot (N_r - 2)) + \hat{\sigma}_{b_o}^4/(N_{r,o} - 2)} \tag{D6}$$

degrees of freedom, we consider the observation-based and simulation slope to be different from each other with confidence level $p = 0.9$.

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.

	PI	1986-2005	2081-2100 RCP2.6	2081-2100 RCP8.5
Number Surf.	3.65	1.75 (1.50-2.20)	2.24 (1.86-2.93)	1.36 (1.09-1.69)
200 m	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)
200 m	38.61	66.06 (59.74-18.92)	98.71 (89.01-109.01)	111.56 (106.62-122.70)
Maximal Intensity Surf.	2.92	3.42 (3.26-3.64)	3.21 (3.07-3.48)	1.51 (1.42-1.63)
200 m	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 A1. Simulated global ensemble-mean Ω_A extreme variability event characteristics for the preindustrial (PI), present day (1986-2005), and the end of this century (2081-2100) for both RCP2.6 and RCP8.5. Numbers of yearly extreme days are given in days per year, durations in days, intensities in $10^{-3}\,\Omega_A$ units and volumes in km^3. Values in brackets denote ensemble minima and maxima.

Disclaimer. The work reflects only the authors' view; the European Commission and their executive agency are not responsible for any use that may be made of the information the work contains.

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.

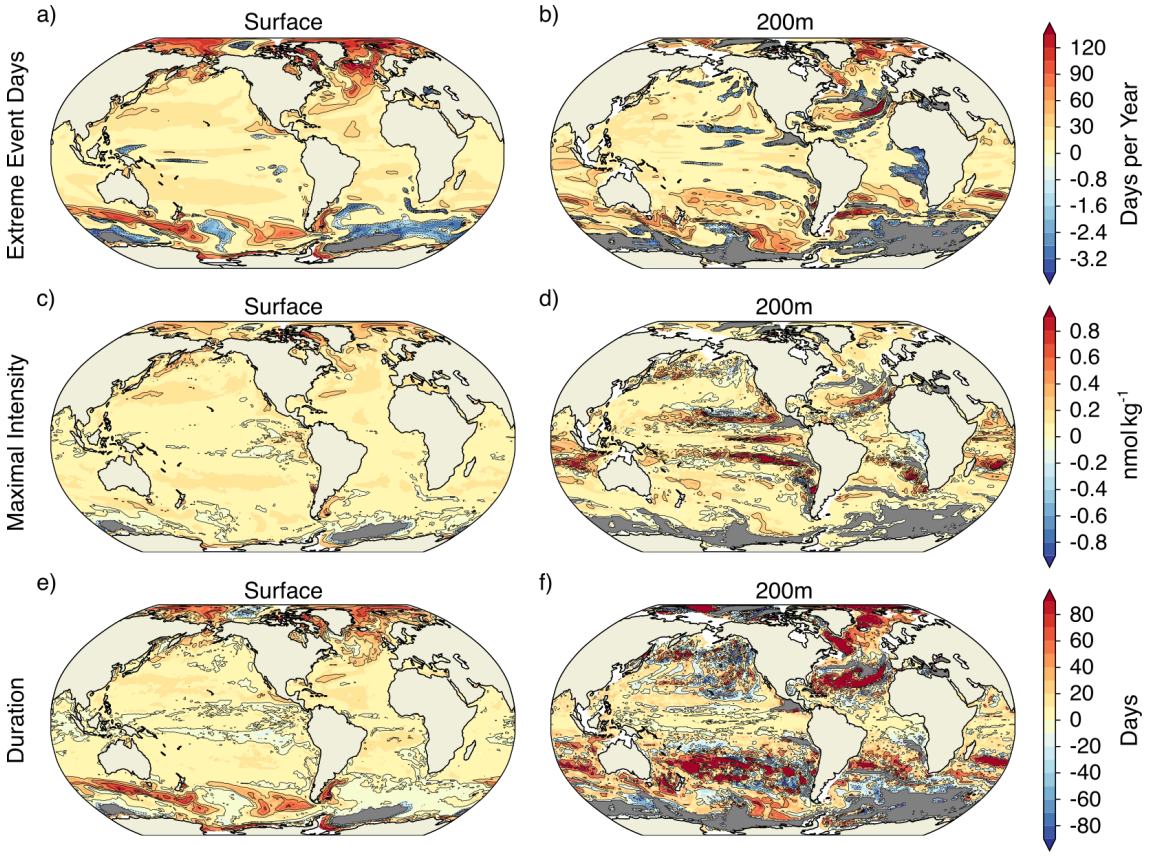

Figure A1. Simulated regional changes in (a,b) the number of extreme $[H^+]$ days per year, (c,d) the maximal intensity of extreme $[H^+]$ variability events, and (e,f) the duration of extreme $[H^+]$ variability events between preindustrial and 2081-2100 following the RCP2.6 scenario. Left panels show changes for the surface, whereas right panels show changes for 200 m. Shown are changes averaged over all five ensemble members. The black contours highlight the pattern structures. Grey areas represent areas with no variability extremes during 2081-2100.

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., 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.

Andrade, J. and Estévez-Pérez, M.: Statistical comparison of the slopes of two regression lines: A tutorial, Anal. Chim. Acta, 838, 1 – 12, https://doi.org/https://doi.org/10.1016/j.aca.2014.04.057, 2014.

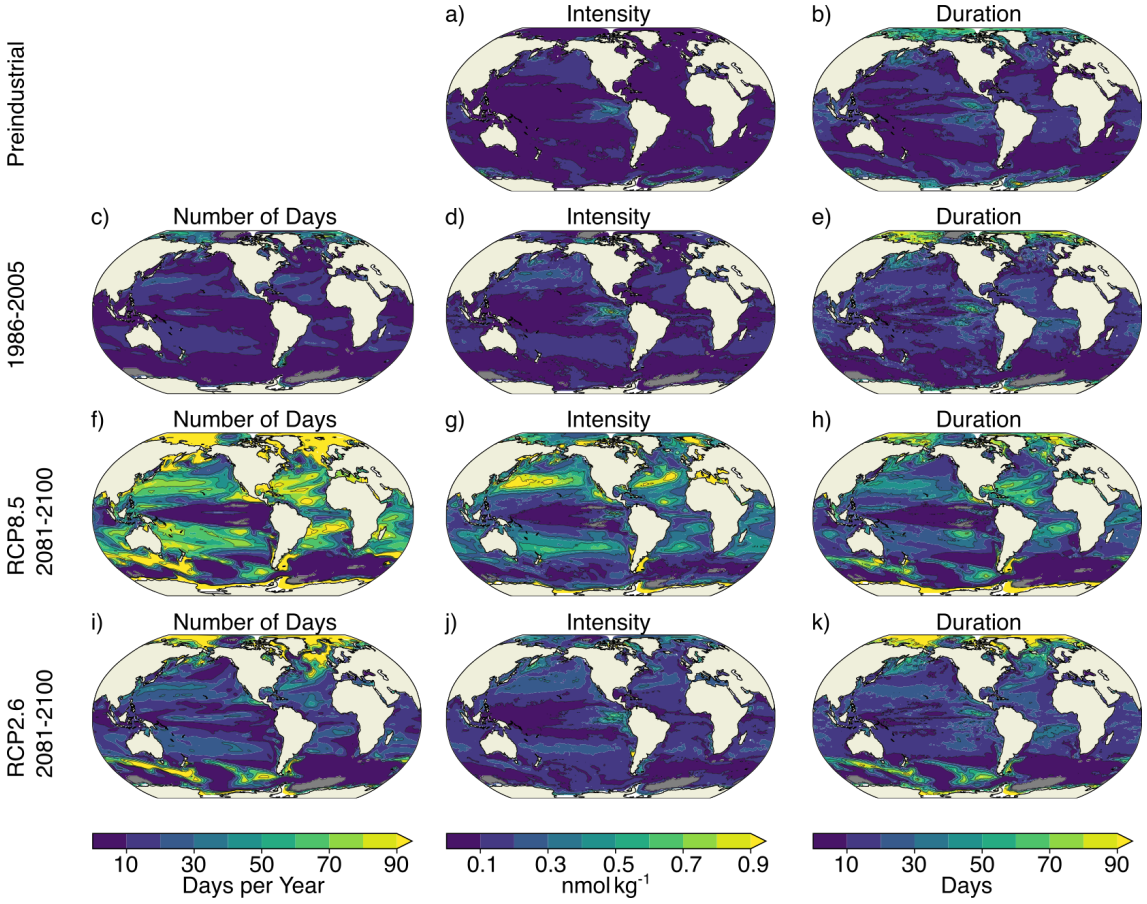

Figure A2. Simulated characteristics of surface $[H^+]$ extreme variability events for preindustrial (a,b), 1986-2005 ensemble mean (c-e), RCP8.5 2081-2100 ensemble mean (f-h), and RCP2.6 2081-2100 ensemble mean (i-k). Grey colors represent regions where no ensemble member simulates variability extremes.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

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, 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, 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., 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.

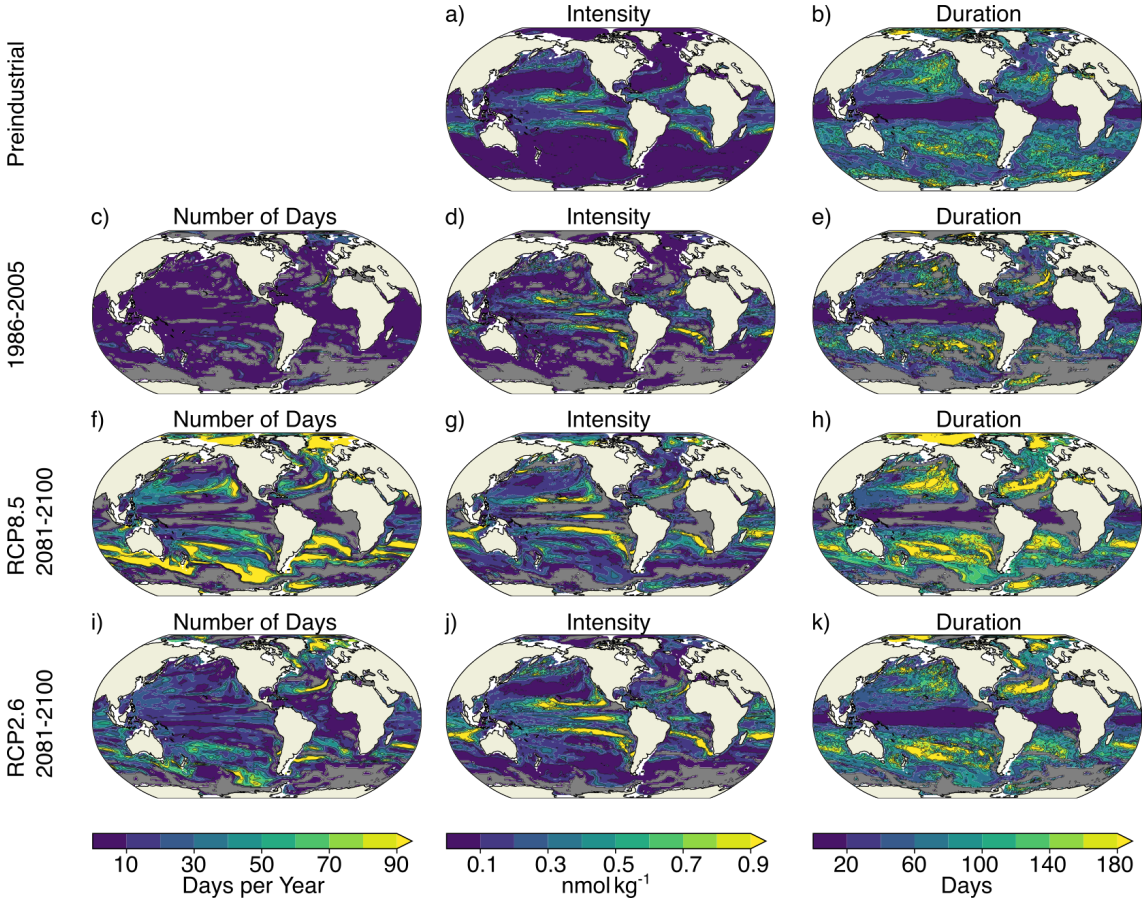

Figure A3. The same as Figure A2, but for 200 m. The color scale for the duration plots changed with respect to that in Figure A2.

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 Murphy, E. J.: Extensive dissolution of live pteropods in the Southern Ocean, Nat. Geosci., 5, 881–885, https://doi.org/10.1038/ngeo1635, 2012.

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.

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, https://doi.org/10.5194/bg-10-6225-2013, 2013.

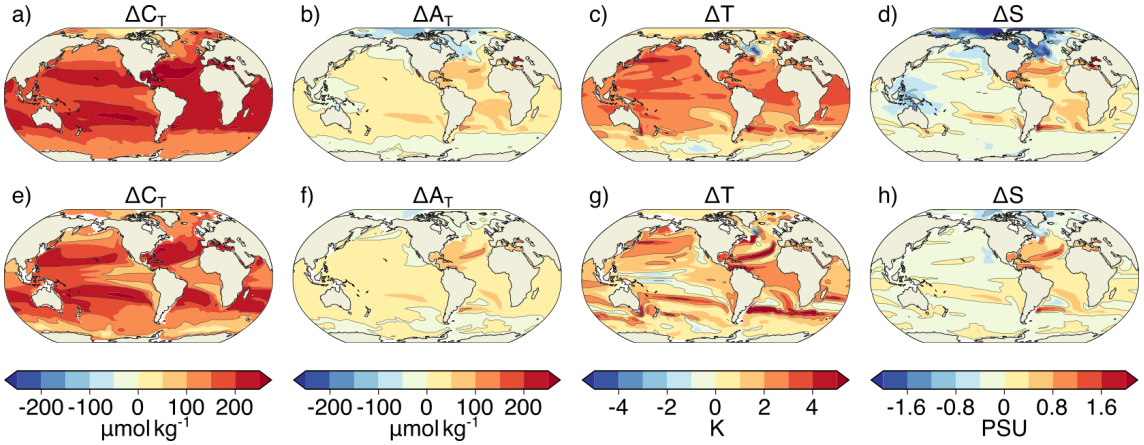

Figure A4. Simulated ensemble mean changes in C_T (a,e), A_T (b,f), T (c,g), and S (d,h) from preindustrial to 2081-2100 following the RCP8.5 scenario. Shown are changes for (a-d) the surface and (e-h) at 200 m. The black contours highlight the pattern structures.

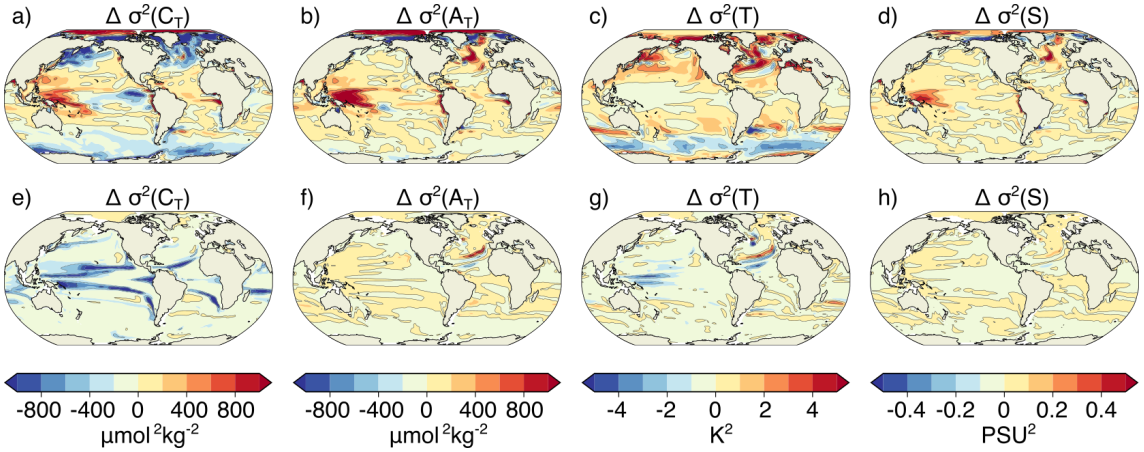

Figure A5. Simulated ensemble mean changes in the variances of C_T (a,e), A_T (b,f), T (c,g), and S (d,h) from preindustrial to 2081-2100 under the RCP8.5 scenario. Shown are changes for (a-d) the surface and (e-h) at 200 m. The black contours highlight the pattern structures.

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.

Carter, B. R., Frölicher, T. L., Dunne, J. P., Rodgers, K. B., Slater, R. D., and Sarmiento, J. L.: When can ocean acidification impacts be detected from decadal alkalinity measurements?, Global Biogeochem. Cycles, 30, 595–612, https://doi.org/10.1002/2015GB005308, 2016.

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-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.

Cornwall, C. E., Comeau, S., DeCarlo, T. M., Larcombe, E., Moore, B., Giltrow, K., Puerzer, F., D'Alexis, Q., and McCulloch, M. T.: A coralline alga gains tolerance to ocean acidification over multiple generations of exposure, Nat. Clim. Change, 10, 143–146, https://doi.org/10.1038/s41558-019-0681-8, 2020.

Dickson, A. and Millero, F.: A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media, DEEP-SEA RES PT A, 34, 1733 – 1743, https://doi.org/https://doi.org/10.1016/0198-0149(87)90021-5, 1987.

Dickson, A. and Riley, J.: The effect of analytical error on the evaluation of the components of the aquatic carbon-dioxide system, Mar. Chem., 6, 77 – 85, https://doi.org/https://doi.org/10.1016/0304-4203(78)90008-7, http://www.sciencedirect.com/science/article/pii/0304420378900087, 1978.

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, 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.: 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.

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., 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, 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.

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, Biogeosciences, 17, 2061–2083, https://doi.org/10.5194/bg-17-2061-2020, 2020.

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., and Winton, M.: Dominance of the Southern Ocean in Anthropogenic Carbon and Heat Uptake in CMIP5 Models, J. Clim., 28, 862–886, https://doi.org/10.1175/JCLI-D-14-00117.1, 2015.

Gallego, M. A., Timmermann, A., Friedrich, T., and Zeebe, R. E.: Drivers of future seasonal cycle changes in oceanic pCO_2, Biogeosciences, 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.

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.

Gray, A. R., Johnson, K. S., Bushinsky, S. M., Riser, S. C., Russell, J. L., Talley, L. D., Wanninkhof, R., Williams, N. L., and Sarmiento, J. L.: Autonomous Biogeochemical Floats Detect Significant Carbon Dioxide Outgassing in the High-Latitude Southern Ocean, Geophys. Res. Lett., 45, 9049–9057, https://doi.org/10.1029/2018GL078013, 2018.

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.

Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B., Thorne, P., Wild, M., and Zhai, P.: Observations: Atmosphere and Surface Supplementary Material, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 2013.

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, Glob. Change Biol., 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.

McNeil, B. I. and Matear, R. J.: Southern Ocean acidification: A tipping point at 450-ppm atmospheric CO_2, Proc. Natl. Acad. Sci. U.S.A., 105, 18 860–18 864, https://doi.org/10.1073/pnas.0806318105, 2008.

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.

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.

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 over the twenty-first century and its impact on calcifying organisms, Nature, 437, 681–686, https://doi.org/10.1038/nature04095, 2005.

Orr, J. C., Epitalon, J.-M., Dickson, A. G., and Gattuso, J.-P.: Routine uncertainty propagation for the marine carbon dioxide system, Mar. Chem., 207, 84 – 107, https://doi.org/https://doi.org/10.1016/j.marchem.2018.10.006, http://www.sciencedirect.com/science/article/pii/S030442031830149X, 2018.

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.

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.

Rivest, E. B., Comeau, S., and Cornwall, C. E.: The Role of Natural Variability in Shaping the Response of Coral Reef Organisms to Climate Change, Curr. Clim. Change Rep., 3, 271–281, https://doi.org/10.1007/s40641-017-0082-x, 2017.

Rodgers, K. B., Sarmiento, J. L., Aumont, O., Crevoisier, C., de Boyer Montégut, C., and Metzl, N.: A wintertime uptake window for 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, 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.