



Potential sources of variability in ocean acidification mesocosm experiments

Maria Moreno de Castro¹, Markus Schartau², and Kai Wirtz¹

¹Helmholtz-Zentrum Geesthacht, Centre for Materials and Coastal Research

²GEOMAR Helmholtz Centre for Ocean Research Kiel

Correspondence to: M. M. de Castro (maria.moreno@hzg.de)

Abstract. Mesocosm experiments on phytoplankton dynamics under high CO₂ concentrations mimic the response of marine primary producers to future ocean acidification. However, potential acidification effects can be hindered by the high standard deviation typically found in the distribution of the replicates exposed to the same treatment. In experiments with multiple unresolved factors and a suboptimal number of replicates, post-processing statistical inference tools may fail to detect an effect. In such cases, model-based data analyses are suitable tools to unearth potential responses to the treatment and to identify which uncertainties may give rise to the observed divergences. As test cases, we use data showing high variability from two independent mesocosm experiments, where, according to statistical inference tools, biomass appeared insensitive to changing CO₂ conditions. Our simulations, in stead, show earlier and more intense phytoplankton blooms in modeled replicates at high CO₂ concentrations and suggest that uncertainties in average cell size, phytoplankton biomass losses and initial nutrient concentration potentially outweigh acidification effects by triggering strong variability during the bloom phase. We also estimate the thresholds below which uncertainties do not escalate into high variability. This information may help to interpret controversial results about acidification and to design future mesocosm experiments.

Keywords: Variability, Uncertainty Quantification, Mesocosms, Ocean Acidification.

1 Introduction

Oceans are a sink for about 30% of the excess atmospheric CO₂ generated by human activities (Sabine et al., 2004). Increasing carbon dioxide concentration in aquatic environments alters the balance of chemistry reactions and thereby adds acidity, which is known as ocean acidification (OA) (Caldeira and Wickett, 2003). Interestingly, the sensitivity of the photoautotrophic production of particulate organic mater (POC) to OA is less pronounced than previously thought. A general compilation of studies on CO₂ enrichment reveals an overall increase in POC (e.g. Schluter et al., 2014; Eggers et al., 2014; Zondervan et al., 2001; Riebesell et al., 2000) but also no CO₂ effects in POC abundance (e.g. Jones et al., 2014; Engel et al., 2014). In particular, high variances are often present in mesocosm experiments, even in replicates of similar CO₂ conditions (Paul et al., 2015; Schulz et al., 2008; Engel et al., 2008; Kim et al., 2006; Engel et al., 2005). This variability discloses a sever reduction in the ratio between acidification response signal and the variability in observations, which results in a low signal-to-noise ratio.



Mesocosms enclose plankton communities, therefore they comprise a more realistic experimental set-up compared to batch or chemostat experiments (Riebesell et al., 2008). However, mesocosm also contain a higher number of possible interactions, thus opportunities for uncontrolled heterogeneity to spread. Moreover, physiological states vary for different phytoplankton cells and environmental conditions, making independent experimental studies prone to divergent results. In addition, it is impracticable to account for every possible factor that occurs in the original environment or to fine-control initial community structure and ecophysiological states among replicates. Such unresolved ecological details may propagate over the course of the experiment and produce high standard deviations of same treatment replicates.

Our main working hypotheses on the origins of variability in mesocosm experiments are:

- differences among replicates can be interpreted as random variations that compromise similarity of the initial biological state among same treatment replicates
- such uncertainties can amplify during the experiment and generate considerable divergence in the treatment response
- which differences are more significant can be revealed by the decomposition of the variability of the outcomes in terms of the nature, intensity and timing of the amplified response they triggered.

The confirmation of these hypotheses entails two important aspects. First, to heuristically explore variability would require experiments designed to quantify the sensitivity of the mesocosm to variations in potentially relevant factors describing cell physiology and community structure. However, such design would require high-dimensional multi-factorial experimental setups, which are hardly, if at all, manageable, even for low number of replicates. Second, post-processing comparison of variances between treatments and variances within treatments may point to misleading conclusions. This is because variability in the outcomes from replicates of the same treatment may be enlarged by unresolved concomitant effects that are not symmetrically distributed. Such factors can give rise to aberrant outcomes, thus high standard deviations (Cottingham et al., 2005; Miller, 1988), especially for experiments where a suboptimal number of replicates lowers the likelihood for non treatment related dissimilarities to be canceled out by averaging (Ruxton and Colegrave, 2006; Peterman, 1990).

To investigate how a given system reacts to unresolved factor variations, a model-supported sensitivity analysis offers an alternative to statistical inference. Classical approaches involve testing the robustness of model results to variations of model control factors (as parameters and initial conditions) accounting for the system dynamics under environmental forcing (Klepper, 1997). Such analyses are often done point-wise around some reference values of the control factors. Typically, these reference values should be retrieved prior to the sensitivity analysis and should provide a model fit to the experimental or observational data. When replicates are available, model results are fitted to represent the mean of the data. A large sensitivity is revealed when small variations of a factor value induce pronounced changes in model results. In contrast, a low sensitivity is indicated by unaltered or slightly changed modeled outcomes in spite of large variations of a factor value. From a modeling perspective, these sensitivity analyses help to resolve uncertainties in model results. In this study, we quantify uncertainties by a similar approach, however, with different objective and interpretation. Its major rationale is to associate the variability in experimental observations to a variational range bounding the uncertainty of each control factor. The margins of the variational range of each



factor are thus confined by the ability of the dynamical model to reproduce the magnitude of the variability observed in two independent OA mesocosm experiments, namely Pelagic Enrichment CO₂ Experiment (PeECE II and III). These confidence intervals describe the tolerance thresholds below which uncertainties do not escalate into high variability in mesocosm experiments. This information can be important to ensure reproducibility, thus a comparison between results of different independent experiments, and to increase confidence about effects of OA on phytoplankton (Broadgate et al., 2013).

2 Method

2.1 Data integration and model description

We describe here the initial biological state that we use as reference dynamics for our hypotheses. Model equations are shown in Table 1. Reference values of the parameters are shown in Table 2. Initial conditions and parameters are in the following referred as 'model factors'. Although both experiments differ in their species composition, environmental conditions and nutrient supply (Engel et al., 2008; Schulz et al., 2008), same parameter set is employed for PeECE II and III, a feature showing the model validation skills. An exhaustive model documentation is given in Appendix B. The model simulates experimental data from the Pelagic Enrichment CO₂ Experiment (PeECE), a set of 9 outdoor mesocosms placed in coastal waters close to Bergen (Norway) during the spring 2003 (PeECE II) and the spring 2005 (PeECE III). In both experiments, blooms of the natural phytoplankton community were induced and treated in three replicates for future, present and past CO₂ conditions (Engel et al., 2008; Schulz et al., 2008; Riebesell et al., 2007, 2008). Experimental data are available through the data portal Pangaea (doi: 10.1594/PANGAEA.723045 for PeECE II and doi: 10.1594/PANGAEA.726955 for PeECE III).

Field data of aquatic CO₂ concentration, temperature and light were used as direct model inputs (see Appendix C). Measurements of POC, PON and DIN were used for model calibration. POC and PON data were adjusted for a direct comparison with model results (see Appendix D), since some contributions remain unresolved by our dynamical equations. State variables of our model comprise carbon and nitrogen content of phytoplankton, Phy_C and Phy_N and dissolved inorganic nitrogen, DIN, as representative for all nutrients. The dynamics of non phytoplanktonic components DH, i.e. detritus and heterotrophs, are distinguished by DH_C and DH_N. Then POC = Phy_C + DH_C and PON = Phy_N + DH_N.

Mean cell size in the community, here represented as the logarithm of the mean equivalent spherical diameter ESD, is used as model parameter. It determines specific ecophysiological features using allometric relations, relevant for the computation of subsistence quota, as well as nutrient and carbon uptake rates. About the latter, to resolve sensitivities to different DIC conditions, we seek for a relatively accurate description of carbon acquisition as function of DIC and size. It has been suggested by previous observations and models that ambient DIC concentration increases primary production (e.g. Schluter et al., 2014; Rost et al., 2003; Zondervan et al., 2001; Riebesell et al., 2000; Chen, 1994; Riebesell et al., 1993) and mean cell size in the community (Sommer et al., 2015; Eggers et al., 2014; Tortell et al., 2008). We adopted and simplified a biophysically explicit description for carbon uptake from Wirtz (2011), where the efficiency of intracellular DIC transport has been derived as a function of mean cell size $\ell = Ln(ESD/1\mu m)$ and CO₂ concentration. For very large cells, the formulation converges to the surface to volume ratio, that in our notation reads $e^{-\ell}$. By contrast, the allometric dependence of primary production on CO₂



does not apply to picophytoplankton so that together we have a non uniform allometric scaling $f_{\text{CO}_2}(\ell)$ in the carboxylation rate

$$f_{\text{CO}_2} = \left(\frac{1 - e^{-a_{\text{CO}_2} \cdot \text{CO}_2}}{1 + a^* \cdot e^{(\ell - a_{\text{CO}_2} \cdot \text{CO}_2)}} \right). \quad (1)$$

The specific carbon absorption coefficient a_{CO_2} reflects size-independent features of the DIC acquisition machinery, as carbon concentration mechanisms (Raven and Beardall, 2003). The coefficient a^* represents carboxylation depletion.

2.2 Uncertainty analysis

In this study we perform non intrusive forward propagation of uncertainty. Non intrusive means there is not coupling between the uncertainties and the temporal dynamics (since we consider uncertainties only present in the system initial set-up, in contrast to intrusive methods (Chantrasmi and Iaccarino, 2012) involving stochastic dynamical equations with time-varying uncertainties (Toral and Colet, 2014)). Forward refers to the fact that unresolved differences among replicates simulated as variations of the model control factors are propagated through the model to project the overall variability in the system response, (in contrast to backward methods of parameter estimation where the likelihood of inputs values is conditioned by the prior knowledge of the output distribution (Chantrasmi and Iaccarino, 2012; Larssen et al., 2006)). Then, the overall biomass variability induced by the factors uncertainties is compared with the variability in POC experimental data. The comparison between simulated and experimental variability in POC helps us to identify which changes in physiological state and in community structure are main potential contributors to the variability.

We consider model factors, ϕ_i , with $i = 1, \dots, N = 19$, made of 14 process parameters and 5 initial conditions for the state variables. Their reference values, $\langle \phi_i \rangle$, were adjusted to yield model solutions reproducing the mean of each treatment. Reference values are listed in Tables 1 and 2. Factor variations are introduced as random values within the variational range $[\langle \phi_i \rangle - \Delta \phi_i, \langle \phi_i \rangle + \Delta \phi_i]$, where $\Delta \phi_i$ is the standard deviation of the normal distribution of possible factor values. To calculate $\Delta \phi_i$, we first generate 10^4 simulations, each one with a different factor value, ϕ_i . The ensemble of model solutions simulates the potential experimental outcomes (see Appendix A). The factor value for each POC trajectory is randomly drawn from a normal distribution around the factor reference value $\langle \phi_i \rangle$. Such residuals distribution is assumed by popular parametric statistical inference tools as ANOVA (Field et al., 2008). We take the ensemble average over modeled replicates and calculate the standard deviation, $\Delta \text{POC}_i^{\text{mod}}$. Then $\Delta \phi_i$ is the standard deviation of the distribution of factor values such as $\Delta \text{POC}_i^{\text{mod}}$ does not exceed the standard deviation of the experimental POC data, $\Delta \text{POC}^{\text{exp}}$, at any mesocosm, at any time. The effect size of variations of ϕ_i on the variability is then

$$\varepsilon_i = \frac{(\Delta \text{POC}_i^{\text{mod}})}{\Delta \phi_i}. \quad (2)$$

This effect size expresses the maximum variability a factor can generate, $\Delta \text{POC}_i^{\text{mod}}$, relative to the associated range of factor variations, $\Delta \phi_i$, to ensure $\Delta \text{POC}_i^{\text{mod}}$ is the closest to $\Delta \text{POC}^{\text{exp}}$ at any time. More generally, ε_i relates the uncertainty of a dependent variable X (here $X = \text{POC}$) and the uncertainty of the input factors ϕ_i , a proxy of what is known as sensitivity coefficients $c_i = \frac{\partial X}{\partial \phi_i}$ in the indirect method to find the variance of a dependent variable from the uncertainties of direct



measurements (Ellison and Williams, 2012)

$$(\Delta X)^2 = \sum_{i=1}^N c_i^2 \cdot (\Delta \phi_i)^2.$$

This expression is based on the assumption that changes in X in response to variations in one factor ϕ_i are independent from those due to changes in another factor ϕ_j and that all changes are small, thus cross terms and higher order derivatives are negligible. Where no reliable mathematical description of the relationship $X(\phi_i)$ exists, c_i can be evaluated directly by experiment (Ellison and Williams, 2012). As in our case only the rate equation for POC changes is known, but not its analytical solution, and, as mentioned in the Introduction, such measurements are costly in mesocosm experiments, we obtain equivalent information through the numerical calculation of the corresponding effect sizes ε_i .

In the following, the standard deviation of the factors, i.e. factors uncertainty will be given as percentage of the reference values and will be called $\Delta \Phi_i$. The actual factor range is given by: $\Delta \phi_i = \frac{\Delta \Phi_i \cdot \phi_i}{100}$. Strong heteroscedasticity, i.e. irregularities in the standard deviations of experimental POC data (see, for instance, small $\Delta \text{POC}^{\text{exp}}$ at day 8 in Fig.2p), translates into drastically enhanced effect sizes if the model-data comparison would be done at a daily basis. For this reason we consider the temporal mean of the standard deviation per phase, i.e. prebloom, bloom, and postbloom. We inferred phases for PeECE II from Engel et al. (2008) and for PeECE III from Schulz et al. (2008) and Tanaka et al. (2008)).

To numerically calculate the ensemble of 10^4 POC trajectories that simulate potential outcomes from experimental replicates (Fig. 7), we apply the Heun integration method with a time step of $4 \cdot 10^{-4}$, (about 35 seconds of experimental time). The number of modeled POC time series is chosen ensuring convergence, such as a higher number of model realizations, i.e. a higher number of modeled replicates, will produce the same results.

Perturbations of the similarity among replicates produced by strong changes in environmental conditions (storms, dysfunctional devices,...) or by errors in sample procedures are not the scope of this work. After few decades, current state-of-the-art of experimental techniques for running plankton mesocosms is advanced. We believe such differences are of low impact or well-understood in terms of their consequences for final outcomes (Riebesell et al., 2010). Variations in model characterization including structural variability (Adamson and Morozov, 2014) and uncertainties in model parametrization (Kennedy and O'Hagan, 2001) requires extensive further analyses that are left as outlook. It is worth to stress that our analysis assumes a particular mechanistic description of phytoplankton dynamics that suggests sufficient (but not necessary) causes of controversial results in mesocosm experiments. Other kind of uncertainties that are not considered in this analysis may also contribute to the observed variability.

3 Results

3.1 The CO₂ effect on POC dynamics

Our model reproduces the mean of PON, POC and DIN experimental data per treatment, i.e. for future, present and past CO₂ conditions, in the two independent PeECE experiments (Fig. 2). For PeECE II, PON is moderately overestimated and



postbloom POC is slightly underestimated. Even so, the model represents the experimental data with similar precision than the mean of the same treatment experimental replicates (see Appendix E). The mean of the same treatment replicates and its associated standard deviation is typically used to describe experimental data (see Fig. 1b in Engel et al. (2008) for PeECE II or Fig. 8a in Schulz et al. (2008) for PeECE III). Means are in the foundations of the statistical inference tools that did not detect acidification responses for PeECE II and III. However, with our mechanistic model-based analysis, phytoplankton growth in future CO₂ POC time series shows an earlier and elevated bloom with respect to past CO₂ conditions. The future and past reference trajectories bound the range of the CO₂ enrichment effect, as shown by the dark gray area in Fig. 3. POC variability due to variations in model factors simulating experimental uncertainties is plotted in light gray area. Our results suggest that to avoid high POC standard deviations in experimental data potentially masking OA effects, factors variations triggering variability during the bloom need to be reduced.

3.2 Variability decomposition

Our method allows for a factor specific variability decomposition ($\Delta\text{POC}_i^{\text{mod}}$) of the total variability shown in in Fig. 3. The effect of factors variations simulating experimental differences among replicates is classified depending on its timing (Fig.4). POC variability during the prebloom phase can be explained mainly by differences of factors related to subsistence quota, i.e. Q_{subs}^* and α_Q , in both PeECE II and III experiments (left column in Fig. 4). This means that differences in subsistence quota first intensify the divergence of POC trajectories, to be damped few days later by the system dynamics. These subsistence parameters only need to vary about 6% and 8% among replicates (see Table 3), to maximize their contribution to the $\Delta\text{POC}^{\text{exp}}$, thus their effect size is the highest (see Fig.6). Differences in initial nutrient concentration, $\text{DIN}(0)$, mean cell size, ℓ , and phytoplankton biomass loss coefficient, L^* , generate the modeled variability mainly during the bloom (with just about 20% differences among replicates, see Table 3 and second column in Fig. 4) showing high values of effect size (gray highlight in Fig. 6). Amplified variability in the postbloom phase (third column in Fig. 4) emerges from uncertainties in the reference temperature T_{ref} for the Arrhenius equation, Eq. (B2), in sinking loss or export flux, s , and in remineralization and excretion, r^* . Effect size of T_{ref} is high, with just about 12% variation. To generate the high divergence during the postbloom, a strong perturbation of parameters relevant for the non phytoplanktonic biomass is needed (about 81% of the reference value for sinking and 96% for remineralization and excretion, see Table 3), which translates to a relatively low effect size. POC variability throughout all bloom phases (right column in Fig. 4) follows from varying carbon and nitrogen initial conditions, Phy_C and Phy_N , nutrient uptake related factors, V_{max}^* , α_V and Aff , and protein allocation for photosynthetic machinery, f_p . About the latter, high standard deviations of the tolerance (see Table 3) suggests non conclusive results.

Interestingly, effect size ε_i is low for carbon acquisition a_{CO_2} and light absorption a_{PAR} . Perturbations of the initial detritus concentration, $\text{DH}_C(0)$ and $\text{DH}_N(0)$ also have no impact on the dynamics as long as they were within reasonable ranges ($\Delta\Phi_i < 100$). In fact, more than tenfold differences among replicates in such non relevant factors were necessary to achieve a perceptible variability ($\Delta\text{POC}_i^{\text{mod}}$).



4 Discussion

In this study we adapted a sensitivity analysis to assess factor variations that can affect experimental data, a perspective different to the classical description of the mean system behavior. We make use of the methodology with a low-complexity model that describes major features of phytoplankton growth dynamics and fits the mean of mesocosm experimental PeECE II and III data with high accuracy for all CO₂ treatment levels. Then we applied our approach to decompose POC variability and we confirmed the working hypotheses (Figs. 4,6).

We showed that small differences in initial nutrient concentration, mean cell size and phytoplankton biomass losses are sufficient to generate the experimentally observed bloom variability (ΔPOC)^{exp}.

4.1 Nutrient concentration

Differences among replicates in initial nutrient concentration substantially contribute to POC variability, a sensitivity that is, interestingly, not well expressed when varying the initial cellular carbon or nitrogen content of the algae, $\text{Phy}_C(0)$ and $\text{Phy}_N(0)$. The relevance of accuracy in the initial nutrient concentration in replicated mesocosms was already pointed in Riebesell et al. (2008). Under constant growth rate, $\text{DIN}(0)$ determines the timing of nutrient depletion, therefore differences in initial nutrient concentrations may also translate into temporal variations in the succession of species. We not only showed that such dependence also holds in more general dynamics, but our method can also bound the variational range for differences in initial DIN concentration for experiments with low number of replicates. The standard deviation of $\text{DIN}(0)$ in the experimental set-up for PeECE III was 50% of the mean, significantly above our tolerance threshold (see Table 3 for initial DIN concentration). Following Riebesell et al. (2007), we took day 2 as initial condition, when the measured DIN was $14 \pm 2 \mu\text{mol-C L}^{-1}$, as showed in Table 1. As $2 \mu\text{mol-C L}^{-1}$ is approximately the 14% of $14 \mu\text{mol-C L}^{-1}$, replicates variability at day 2 was about a 14%. Therefore, experimental differences in initial nutrient concentration were similar to the tolerance threshold for initial DIN established to avoid high variability ($(20 \pm 6)\%$ in Table 3), which represents an explanation to the high divergence observed in POC measurements.

For PeECE II, experimentally measured DIN concentration at day 0 was $10.7 \pm 0.8 \mu\text{mol-C L}^{-1}$, meaning a 7.5% difference among replicates, below our projected tolerance level (7.5 is out of the range [14,26]). Same applies to day 2, with DIN concentration equal to $8 \pm 0.5 \mu\text{mol-C L}^{-1}$ (Table 1). Our approach shows that differences in initial nutrient concentration in PeECE II were not high enough to trigger the experimentally observed POC variability. Incidentally, there was a phosphate re-addition on day 8 of the experiment, establishing new initial nutrient concentration for the subsequent days. When the dynamics in one replicate significantly diverges from the mean dynamics of the treatment, even if the re-addition occurs at the same time and at the same concentration in all the replicates, the mesocosm with that outlier trajectory will not response as the others, and with the new nutrient condition, the divergence may be further amplified. In that case, nutrient re-addition has the same impact on the systems as variations in initial conditions of nutrient concentration. Also for PeECE II, variability in POC is about 30% higher than variability in PON, as shown in Fig. 2. We attribute the temporal decoupling between C and N dynamics to the break of symmetry among replicates by the nutrient re-addition due to the strong sensitivity of the



system to initial nutrient concentrations and a concomitant change in subsistence N:C quota, which is a sensitive parameter, especially during the prebloom phase (Fig. 4 and Fig. 6). Increase of POC : PON ratios under nitrogen deficiency has been observed frequently during experimental studies (e.g. Antia et al., 1963; Biddanda and Benner, 1997) and has been attributed to preferential PON degradation and to intracellular decrease of the N:C ratio (Schartau et al., 2007). We hence confirm that nutrient re-addition during the course of the experiments results in a significant disturbance, as also mentioned in Riebesell et al. (2008), although a complete analysis would require a model explicitly accounting for other nutrients, as phosphate and silicate.

4.2 Mean cell size as proxy for community structure

We found a limited tolerance to variations in mean cell size of the community, ℓ , which has a threshold of about 22% variation (see Table 3). If we take the averaged mean cell size of PeECE II, $\langle \ell \rangle = 1.6$, and III, $\langle \ell \rangle = 1.8$, from Table 2, we obtain $\langle \ell \rangle = 1.7$. Then the absolute standard deviation is $\Delta \ell = 22 \cdot \frac{1.7}{100} \sim 0.4$. Therefore, our methodology shows that variations within the range limited by $\langle \ell \rangle \pm \Delta \ell$, that is $[1.3, 2.1]$, are sufficient to reproduce the observed experimental POC variability during the bloom. As ℓ is in log-scale, the corresponding ESD increment is within the variational range $\langle \text{ESD} \rangle \pm \Delta \text{ESD}$, that is $[3.7, 8.1] \mu\text{m}$ (or $[25, 285] \mu\text{m}^3$ in volume). These values are easily reached in the course of species succession, and supports studies showing that community composition outweighs ocean acidification (Eggers et al., 2014; Kroeker et al., 2013; Kim et al., 2006).

4.3 Phytoplankton loss

Another major contributor to POC variability during the bloom phase is phytoplankton biomass loss, L^* . With a standard deviation of about 20% (Table 3), uncertainties in L^* generate variability larger than the model response to OA, in particular at the end of the growth phase and the beginning of the decay phase. Unresolved details in phytoplankton loss rate include, among others, replicate differences in cell aggregation or damage by collisions, mortality by virus, parasites, morphologic malformations, or grazing by non filtered mixotrophs or micro-zooplankton.

4.4 Consequences for the design of mesocosm experiments

Our model projections show that a suitable target variable to detect OA effects is the slope of the growth phase. We also provide thresholds to uncertainties that can be used for improving future sampling strategies with low number of replicates. Tolerances given in Table 3 quantify how much replicates similarity can be compromised before the variability of the outcomes outweighs potential acidification effects. Some tolerances indicate to maximal variations in observable quantities, as nutrient concentration and community composition. We show that a better control of such dissimilarities among replicates can help to keep the variability below the range of the acidification effect, specially during the bloom. Strategies to reduce $(\Delta \text{POC})^{\text{mod}}$ should similarly apply to lower $(\Delta \text{POC})^{\text{exp}}$. For example, a complete characterization of phytoplankton biomass losses, which includes aggregation and grazing, limits the amplification of simulated POC variability. Uncertainties in physiological states,



like differences in affinity to nutrients and subsistence quota are more difficult to measure. However, with our analysis we provide plausible explanations for negative results in the detection potential acidification effects (Paul et al., 2015; Schulz et al., 2008; Engel et al., 2008; Kim et al., 2006; Engel et al., 2005). In this manner, our study also proves the limitation of hypothesis testing tools commonly used to assess the statistical significance of effect detectability.

5 5 Conclusions

Our model projections indicate that phytoplankton responses to OA mainly expected to occur during the bloom phase, presenting a higher and earlier bloom under acidification conditions. Moreover, we found that amplified POC variability during the bloom can be explained by variations in initial DIN concentration, mean cell size and the phytoplankton loss rate.

The model-based analysis provides information that can be considered for refinements of experimental design and sampling strategies. We have found specific ecophysiological factors that should be confined, otherwise acidification responses are likely to be masked by variability in POC.

With our approach we reverse the question of how experimental data can constrain model parameter estimates and instead determine the range of variability in experimental data that can be explained by means of modeling with variational ranges bounding uncertainties in specific control factors. We tested the hypothesis whether small differences among replicates have the potential to generate higher variability in biomass time series than the response that can be attributed to the effect of CO₂. We therefore conclude that modeling studies that integrate data from acidification experiments should resolve physiological regulation capacities on cellular and community level. In fact, modeling the propagation of uncertainties revealed cell size to be a major contributor to the phytoplankton biomass variability. This motivates to use of adaptive size-trait based dynamics since such approaches allows for the resolution of ecophysiological trait shifts in non stationary scenarios (Wirtz and Eckhardt, 1996; Wirtz, 2013). The role of intracellular protein allocation can be also clarified by a trait-based approach, since our results about the impact of its variations were non conclusive.

With this study we established a foundation for further model-based analysis for uncertainties propagation that can be generalized to any kind of experiments in biogeoscience. Extensions comprising a representation of the stochastic nature of variations by introducing a new random value for parameters at every time step are straight forward to implement (Toral and Colet, 2014). Another possible extension can include covariance matrices, which show the interaction of variations in two factors simultaneously. Finally, we argue that a more explicit description of uncertainty quantification is pivotal in our interpretation and generalization of experimental results.

Appendix A: Model representation of replicates

The possibility of simulating a high number of replicates is one of the unique strengths of modeling. For each factor, we simulate possible realizations of the same treatment with slight variations of the factor reference value (simulating differences



among replicates in physiological states and community structure). We generate a total of 10^4 modeled dynamical solutions for the state variables. Examples for two factors are shown in Fig. 7.

Appendix B: Definition of relative growth rate

Relative growth rate μ is calculated from primary production rate subtracting respiration and mortality losses, $\mu = P - R - L$.

5 Primary production

Primary production rate reflects the limiting effects of light, dissolved inorganic carbon (DIC), temperature and nutrient internal quota,

$$P = P_{\max} \cdot f_{\text{PAR}} \cdot f_{\text{CO}_2} \cdot f_T \cdot f_Q \cdot f_p. \quad (\text{B1})$$

P_{\max} is the maximum primary production rate, (Table 2). Specific light limitation f_{PAR} depends on light and CO_2 . For the attenuation coefficient a_z , we consider that in coastal regions light intensity is reduced to 1% of its surface value typically in 5 m (Denman and Gargett, 1983) and we obtain $a_z = 0.75\text{m}^{-1}$. Then, PAR experienced by cells at mixed layer depth (MLD = 4.5 m, Engel et al. (2008)), is calculated from radiation at water surface, PAR_0 (see Appendix C), following an exponential decay described by the Lambert-Beer law

$$\text{PAR} = \text{PAR}_0 \int_0^{\text{MLD}} e^{-a_z \cdot z} dz.$$

15 The relationship between photosynthesis and irradiance can be formulated referring to a cumulative one-hit Poisson distribution (Ley and Mauzerall, 1982; Dubinsky et al., 1986). With the temperature and carbon acquisition dependence, it yields

$$f_{\text{PAR}} = \left(1 - e^{-\frac{a_{\text{PAR}} \cdot \text{PAR}}{P_{\max} \cdot f_{\text{CO}_2} \cdot f_T}}\right),$$

where a_{PAR} is the effective absorption related to the chloroplast cross-section and saturation response time for receptors (Geider et al. (1998a), Wirtz and Pahlow (2010)), the carbon acquisition term f_{CO_2} is described in Section 2.1, Eq. (1).

20 f_T is the temperature dependence. We consider that all metabolic rates depend on protein folding that increases with rising temperature following the Arrhenius equation (Scalley and Baker, 1997) as described in Geider et al. (1998b) or Schartau et al. (2007)

$$f_T = e^{-E_a \cdot \left(\frac{1}{T} - \frac{1}{T_{\text{ref}}}\right)}, \quad (\text{B2})$$

with activation energy $E_a = \frac{T_{\text{ref}}^2}{10} \cdot \log(Q_{10})$ as in Wirtz (2013), where we use $Q_{10} = 1.88$ for phytoplankton (Eppley, 1972; 25 Brush et al., 2002) and T_{ref} is the mean measured temperature (see Appendix C).



The allometric factor α_Q quantifies the scaling relation of subsistence quota and cell size. We use a Droop dependency on nutrient N:C ratio (Droop, 1973) which has been recently mechanistically derived (Wirtz and Pahlow, 2010; Pahlow and Oschlies, 2013)

$$f_Q = \left(1 - \frac{Q_{\text{subs}}}{Q}\right)$$

- 5 where $Q = \frac{\text{Phy}_N}{\text{Phy}_C}$. Its lower reference, the subsistence quota $Q_{\text{subs}} = Q_{\text{subs}}^* \cdot e^{-\alpha_Q \cdot \ell}$, is considered size dependent to reflect a lower protein demand for uptake mechanisms in large cells (Litchman et al., 2007).

The last term in Eq. (B1) accounts for an energy allocation trade-off in phytoplankton cells: protein allocation for photosynthetic compounds as RuBisCo and pigments, f_p , versus allocation for nutrient uptake, f_v , expressed by $f_p + f_v = 1$ (Wirtz and Pahlow, 2010; Pahlow and Oschlies, 2013). We simplify detailed partition models by setting the trait fractions constant.

10 Respiratory cost and nutrient uptake rates

Efforts related to nutrient uptake V are represented by a respiration term. Other expenses as biosynthetic costs are neglected (Pahlow, 2005). Respiration rate is then calculated as

$$R = \zeta \cdot V$$

where ζ expresses the specific respiratory cost of nitrogen assimilation (Raven, 1980; Aksnes and Egge, 1991; Pahlow, 2005).

- 15 For simplicity, our model merges the set of potentially limiting nutrients (e.g. P, Si and N) to a single resource only, that is DIN. We follow Aksnes and Egge (1991) as described in Pahlow (2005) for the maximum uptake rate

$$V_{\text{max}} = \frac{1}{\frac{1}{V_{\text{max}}^* \cdot f_T} + \frac{1}{\text{Aff} \cdot \text{DIN}}},$$

- comprising the maximum uptake coefficient, V_{max}^* , and the nutrient affinity, Aff. Besides adding a temperature dependence of nutrient uptake as given in Schartau et al. (2007), we assume that respiratory costs decrease with increasing cell size (Edwards et al., 2012) which leads to an allometric scaling in nutrient uptake (Wirtz, 2013) with exponent α_V . We also account for the static proteins allocation trade-off between photosynthetic machinery, f_p , and nutrients uptake, $f_v = 1 - f_p$. Then the nutrient uptake term yields

$$V = (1 - f_p) \cdot V_{\text{max}} \cdot e^{-\alpha_V \cdot \ell}.$$

Loss rates

- 25 To describe the loss rate of phytoplankton biomass we use a density dependent term

$$L = L^* \cdot (\text{Phy}_C + \text{DH}_C).$$



The resulting matter flux increases the biomass of detritus and heterotrophs (DH) and a fraction of it becomes part of the remineralizable pool. A temperature dependent remineralization term (Schartau et al., 2007)

$$r = r^* \cdot f_T$$

describes any kind of DIN production, as hydrolysis and remineralization of organic matter, excretion of ammonia directly by zooplankton and rapid remineralization of fecal pellets produced also by zooplankton. The other fraction of the non phytoplanktonic biomass is removed by settling with a rate related to the sinking coefficient, s , given in Tables 1 and 2. Our model is calibrated with experimental data from enclosed mesocosms where aquarium pumps ensured mixing. Therefore we assume that wealthy enough organisms were able to achieve neutral buoyancy (Boyd and Gradmann, 2002) thus sinking does not directly affect phytoplankton biomass.

10 Appendix C: Forcings

As mentioned in the introduction, we used measured aquatic CO_2 , temperature and photosynthetic active radiation, PAR, as model inputs (see Fig. 8). For the two PeECE experiments the photon flux density was measured by the Geophysical Institute of the University of Bergen. To calculate the surface radiation inside the mesocosms, PAR_0 , we follow (Schulz et al., 2008) and consider that 80% of incident PAR passed through the gas tight tents of which up to 15% penetrated into c.a. 2.5 m depth, the center of the mixed surface layer in PeECE III. Daily carbon dioxide data were interpolated and PAR signal was filtered by singular spectrum analysis to avoid sudden changes that could be detrimental for the performance of the numerical calculation since Heun method requires differentiable functions.

Appendix D: POC adjustment

The applied model equations attribute phytoplankton, detritus, and herbivorous heterotrophs to particulate organic matter. Measurements of particulate organic carbon also include some fractions of large bacterioplankton, carnivorous zooplankton, as well as extracellular gel particles like transparent exopolymer particles. These additional organic contributions to POC measurements are not explicitly resolved in our model. For comparisons between simulation results and observations we therefore have to adjusted POC data. We used data of transparent exopolymer particles (TEP) of Egge et al. (2009) for adjusting PeECE III POC measurements. For PeECE II, $\text{POC} = \text{POC}' - \text{POC}''$, where POC are represented by dots in Figs. 2 and 4, POC' are raw data from PANGAEA and POC'' are the difference between particle abundance, PA, of the Coulter Counter measurements and the Flow Cytometry data in Engel et al. (2008):

$$\text{POC}'' = \beta \cdot (\text{PA}_{\text{Coulter Counter}} - \text{PA}_{\text{Flow Cytometry}}).$$

The scaling parameter $\beta=0.000065 \mu\text{mol-C}^{-1} \text{L}$ was tuned to provide reductions between 40 and 50% from total POC, in agreement with adjustments of PeECE III.



Appendix E: Residuals of the model-data fit

We calculate the cumulative residuals (Table 4) with respect to the mean of experimental replicates per treatment, time and mesocosm. For experimental data, residuals are calculated as

$$E = \sum_{\text{treat,rep,day}} |Y_{\text{treat,rep,day}}^{\text{exp}} - \langle Y_{\text{treat,day}}^{\text{exp}} \rangle| / \eta$$

5 and for model results

$$M = \sum_{\text{treat,rep,day}} |Y_{\text{treat,rep,day}}^{\text{mod}} - \langle Y_{\text{treat,day}}^{\text{exp}} \rangle| / \eta$$

with $\eta = 9$ the total number of mesocosms.

Author contributions. K.W., M.S and M.M.C. developed the model code, M.M.C performed the simulations and prepared the manuscript, which was revised by K.W. and M.S.

10 *Acknowledgements.* Acknowledgements will be added after the review process.



References

- Adamson, M. and Morozov, A.: Defining and detecting structural sensitivity in biological models: developing a new framework, *Journal of Mathematical Biology*, 69, 1815–1848, doi:10.1007/s00285-014-0753-3, <http://dx.doi.org/10.1007/s00285-014-0753-3>, 2014.
- Aksnes, D. L. and Egge, J. K.: A theoretical model for nutrient uptake in phytoplankton, *Marine Ecology Progress Series*, 70, 65–72, 1991.
- 5 Antia, N. J., MacAllistel, C. D., Parsons, T. R., Stephens, K., and Strickland, J. D. H.: Further measurements of primary production using a large-volume plastic sphere, *Limnology and Oceanography*, 8, 166 – 173, doi:<http://dx.doi.org/10.4319/lo.1963.8.2.0166>, 1963.
- Biddanda, B. and Benner, R.: Carbon, nitrogen, and carbohydrate fluxes during the production of particulate and dissolved organic matter by marine phytoplankton, *Limnology and Oceanography*, 42, 506–518, doi:10.4319/lo.1997.42.3.0506, <http://dx.doi.org/10.4319/lo.1997.42.3.0506>, 1997.
- 10 Boyd, C. M. and Gradmann, D.: Impact of osmolytes on buoyancy of marine phytoplankton, *Marine Biology*, 141, 605–618, 2002.
- Broadgate, W., Riebesell, U., Armstrong, C., Brewer, P., Denman, K., Feely, R., Gao, K., Gatusso, J. P., Isensee, K., Kleypas, J., Laffoley, D., Orr, J., Pöetner, H. O., de Rezende, C. E., Schimdt, D., urban, E., Waite, A., and Valdés, L.: Ocean acidification summary for policymakers - Third Symposium on the ocean in a high-CO₂ world., *International Geosphere-Biosphere Programme, Sweden*, p.26, 2013.
- Brush, M., Brawley, J., Nixon, S., and Kremer, J.: Modeling phytoplankton production: problems with the Eppley curve and an empirical alternative, *Marine Ecology Progress Series*, 238, 31–45, doi:doi:10.3354/meps238031, 2002.
- 15 Caldeira, K. and Wickett, M. E.: Oceanography: Anthropogenic carbon and ocean pH, *Nature*, 425, 365–365, doi:10.1038/425365a, <http://dx.doi.org/10.1038/425365a>, 2003.
- Chantramsi, T. and Iaccarino, G.: Forward and backward uncertainty propagation for discontinuous system response using the Pade-Legendre method, *International Journal for Uncertainty Quantification*, 2, 125–143, 2012.
- 20 Chen, C. Y.: Effect of pH on the growth and carbon uptake of marine phytoplankton, *Mar. Ecol. Prog. Ser.*, 109, 83–94, 1994.
- Cottingham, K. L., Lennon, J. T., and Brown, B. L.: Knowing when to draw the line: designing more informative ecological experiments, *Frontiers in Ecology and the Environment*, doi:10.1890/1540-9295(2005)003[0145:KWTDTL]2.0.CO;2, [http://dx.doi.org/10.1890/1540-9295\(2005\)003\[0145:KWTDTL\]2.0.CO;2](http://dx.doi.org/10.1890/1540-9295(2005)003[0145:KWTDTL]2.0.CO;2), 2005.
- Denman, K. L. and Gargett, A. E.: Time and space scales of vertical mixing and advection of phytoplankton in the upper ocean, *Limnology and Oceanography*, 28, 801–815, 1983.
- 25 Droop, M. R.: Some thoughts on nutrient limitation in algae, *Journal of Phycology*, 9, 264–272, doi:10.1111/j.1529-8817.1973.tb04092.x, <http://dx.doi.org/10.1111/j.1529-8817.1973.tb04092.x>, 1973.
- Dubinsky, Z., Falkowski, P. G., and Wyman, K.: Light harvesting and utilization by phytoplankton, *Plant Cell Physiology*, 21, 1335–134, 1986.
- 30 Edwards, K., Klausmeier, C. A., and Litchman, E.: Allometric scaling and taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton, *Limnology and Oceanography*, 57, 554–556, 2012.
- Egge, J. K., Thingstad, T. F. Larsen, A., Engel, A., Wohlers, J., Bellerby, R. G. J., and Riebesell, U.: Primary production during nutrient-induced blooms at elevated CO₂ concentrations, *Biogeosciences*, 6, 877–885, <http://www.biogeosciences.net/6/877/2009/>, 2009.
- Eggers, S. L., Lewandowska, A. M., Barcelos e Ramos, J., Blanco-Ameijeiras, S., Gallo, F., and Matthiessen, B.: Community composition has greater impact on the functioning of marine phytoplankton communities than ocean acidification, *Global Change Biology*, 20, 713–723, doi:10.1111/gcb.12421, 2014.
- 35 Ellison, S. L. R. and Williams, A.: *Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement*, third edn., p.26, 2012.



- Engel, A., Schulz, K. G., Riebesell, U., Bellerby, R., Delille, B., and Schartau, S.: Effects of CO₂ on particle size distribution and phytoplankton abundance during a mesocosm bloom experiment (PeECE II), *Biogeosciences*, 5, 509–521, 2008.
- Engel, A., Cisternas Nova, C., Wurst, M., Endres, S., Tang, T., Schartau, M., and Lee, C.: No detectable effect of CO₂ on elemental stoichiometry of *Emiliana huxleyi* in nutrient-limited, acclimated continuous cultures, *Marine Ecology Progress Series*, 507, 15–30, <http://www.int-res.com/abstracts/meps/v507/p15-30/>, 2014.
- Engel, A. et al.: Testing the direct effect of CO₂ concentration on a bloom of the coccolithophorid *Emiliana huxleyi* in mesocosm experiments, *Limnology and Oceanography*, 50, 493–507, 2005.
- Eppley, R. W.: Temperature and phytoplankton growth in the sea, *Fishery Bulletin*, 1972.
- Field, A., Miles, J., and Field, Z.: *Discovering statistics using R*, SAGE, 2008.
- Geider, R., Macintyre, Graziano, L., and McKay, R. M.: Responses of the photosynthetic apparatus of *Dunaliella tertiolecta* (Chlorophyceae) to nitrogen and phosphorus limitation, *European Journal of Phycology*, 33, 315–332, doi:10.1080/09670269810001736813, <http://dx.doi.org/10.1080/09670269810001736813>, 1998a.
- Geider, R. J., MacIntyre, H. L., and Kana, T. M.: A dynamic regulatory model of phytoplankton acclimation to light, nutrients, and temperature, *Limnology and Oceanography*, 43, 679–694, doi:10.4319/lo.1998.43.4.0679, <http://dx.doi.org/10.4319/lo.1998.43.4.0679>, 1998b.
- Jones, B. M., Iglesias-Rodriguez, M. D., Skipp, P. J., Edwards, R. J., Greaves, M. J., et al.: Responses of the *Emiliana huxleyi* Proteome to Ocean Acidification, *PLoS ONE*, 8, 2857–2869, doi:10.1371/journal.pone.0061868, 2014.
- Kennedy, M. C. and O’Hagan, A.: Bayesian Calibration of Computer Models, *Journal of the Royal Statistical Society, Series B*, 63, 425–464, 2001.
- Kim, J.-M., Lee, K., Shin, K., Kang, J.-H., Lee, H.-W., Kim, M., Jang, P.-G., and Jang, M.-C.: The effect of seawater CO₂ concentration on growth of a natural phytoplankton assemblage in a controlled mesocosm experiment, *Limnology and Oceanography*, 51, 1629–1636, 2006.
- Klepper, O.: Multivariate aspects of model uncertainty analysis: tools for sensitivity analysis and calibration, *Ecological Modelling*, 101, 1–13, doi:[http://dx.doi.org/10.1016/S0304-3800\(96\)01922-9](http://dx.doi.org/10.1016/S0304-3800(96)01922-9), <http://www.sciencedirect.com/science/article/pii/S0304380096019229>, 1997.
- Kroeker, K. J., Kordas, R. L., Crim, R., Hendriks, I. E., Ramajo, L., Singh, G. S., Duarte, C. M., and Gattuso, J.-P.: Impacts of ocean acidification on marine organisms: quantifying sensitivities and interaction with warming, *Global Change Biology*, 19, 1884–1896, doi:10.1111/gcb.12179, <http://dx.doi.org/10.1111/gcb.12179>, 2013.
- Larssen, T., Huseby, R. B., Cosby, B. J., Høst, G., Høgåsen, T., and Aldrin, M.: Forecasting acidification effects using a Bayesian calibration and uncertainty propagation approach., *Environmental Science Technology*, 40, 2006.
- Ley, A. C. and Mauzerall, D. C.: Absolute absorption cross-sections for photosystem II and the minimum quantum requirement for photosynthesis in *Chlorella vulgaris*, *Biochimica et Biophysica Acta*, 680, 95–106, 1982.
- Litchman, E., Klausmeier, C. A., Schofield, O., and Falkowski, P.: The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level, *Ecology Letters*, 10, 1170–1181, 2007.
- Miller, R. G. J.: *Beyond ANOVA, Basics of Applied Statistics*. Wiley, New York – Chichester – Brisbane – Toronto – Singapore 1986, 317 S., £ 30.65; ISBN 0-471-81922-0, *Biometrical Journal*, 30, 874–874, doi:10.1002/bimj.4710300741, <http://dx.doi.org/10.1002/bimj.4710300741>, 1988.
- Pahlow, M.: Linking chlorophyll–nutrient dynamics to the Redfield N:C ratio with a model of optimal phytoplankton growth, *Marine Ecology Progress Series*, 287, 33–43, 2005.



- Pahlow, M. and Oschlies, A.: Optimal allocation backs Droop's cell-quota model, *Marine Ecology Progress Series*, 473, 1–5, 2013.
- Paul, C., Matthiessen, B., and Sommer, U.: Warming, but not enhanced CO₂ concentration, quantitatively and qualitatively affects phytoplankton biomass, *Marine Ecology Progress Series*, 528, 39–51, doi:10.3354/meps11264, <http://www.int-res.com/abstracts/meps/v528/p39-51/>, 2015.
- 5 Peterman, R. M.: The importance of reporting statistical power: the forest decline and acidic deposition example, *Ecology*, 71, 2024–2027, 1990.
- Raven, J. and Beardall, J.: Carbon Acquisition Mechanisms of Algae: Carbon Dioxide Diffusion and Carbon Dioxide Concentrating Mechanisms, in: *Photosynthesis in Algae*, edited by Larkum, A., Douglas, S., and Raven, J., vol. 14 of *Advances in Photosynthesis and Respiration*, pp. 225–244, Springer Netherlands, doi:10.1007/978-94-007-1038-2_11, http://dx.doi.org/10.1007/978-94-007-1038-2_11, 2003.
- 10 Raven, J. A.: Nutrient transport in microalgae, *Adv. Microb. Physiol.*, 21, 47–226, 1980.
- Riebesell, U., Wolf-Gladrow, D. A., and Smetacek, V.: Carbon dioxide limitation of marine phytoplankton growth rates, *Nature*, 361, 249–251, doi:10.1038/361249a0, <http://dx.doi.org/10.1038/361249a0>, 1993.
- 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 CO₂, *Nature*, 407, 364–367, doi:10.1038/35030078, <http://dx.doi.org/10.1038/35030078>, 2000.
- 15 Riebesell, U., Schulz, K. G., Bellerby, R. G. J., Botros, M., Fritsche, P., Meyerhofer, M., Neill, C., Nondal, G., Oschlies, A., Wohlers, J., and Zollner, E.: Enhanced biological carbon consumption in a high CO₂ ocean, *Nature*, 450, 545–548, doi:10.1038/nature06267, <http://dx.doi.org/10.1038/nature06267>, 2007.
- Riebesell, U., Bellerby, R. G. J., Grossart, H. P., and Thingstad, F.: Mesocosm CO₂ perturbation studies: from organism to community level, *Biogeosciences*, 5, 1157–1164, doi:XXXXXX/bg-5-1157-2008, <http://www.biogeosciences.net/5/1157/2008/>, 2008.
- 20 Riebesell, U., Fabry, V. J., Hansson, L., and Gattuso, J. P.: Guide to best practices for ocean acidification research and data reporting, Publications Office of the European Union., 2010.
- Rost, B., Riebesell, U., Burkhardt, S., and Suetemeyer, D.: Carbon acquisition of bloom-forming marine phytoplankton, *Limnology and Oceanography*, 48, 55–67, 2003.
- Ruxton, G. D. and Colegrave, N.: *Experimental design for the life sciences*, Oxford: Oxford University Press, 2006.
- 25 Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T.-H., Kozyr, A., Ono, T., and Rios, A. F.: The Oceanic Sink for Anthropogenic CO₂, *Science*, 305, 367–371, doi:10.1126/science.1097403, <http://www.sciencemag.org/content/305/5682/367.abstract>, 2004.
- Scalley, M. L. and Baker, D.: Protein folding kinetics exhibit an Arrhenius temperature dependence when corrected for the temperature dependence of protein stability, *Proceedings of the National Academy of Sciences*, 94, 10 636–10 640, doi:10.1073/pnas.94.20.10636, <http://www.pnas.org/content/94/20/10636.abstract>, 1997.
- 30 Schartau, M., Engel, A., Schröter, J., Thoms, S., Völker, C., and Wolf-Gladrow, D.: Modelling carbon overconsumption and the formation of extracellular particulate organic carbon, *Biogeosciences*, 4, 433–454, doi:10.5194/bg-4-433-2007, <http://www.biogeosciences.net/4/433/2007/>, 2007.
- Schluter, L., Lohbeck, K. T., Gutowska, M. A., Groger, J. A., Riebesell, U., and Reusch, T. B. H.: Adaptation of a globally important coccolithophore to ocean warming and acidification, *Nature Climate Change*, 4, 1024–1030, doi:10.1038/nclimate2379, <http://dx.doi.org/10.1038/nclimate2379>, 2014.

**Table 1.** States variables and their dynamics.

State variable	dynamical equation	ini. cond.	units
phytoplankton carbon	$\frac{d\text{Phy}_C}{dt} = (P-R-L) \cdot \text{Phy}_C$	2.5	$\mu\text{mol-C L}^{-1}$
phytoplankton nitrogen	$\frac{d\text{Phy}_N}{dt} = V \cdot \text{Phy}_C - L \cdot \text{Phy}_N$	0.4	$\mu\text{mol-N L}^{-1}$
nutrients concentration	$\frac{d\text{DIN}}{dt} = r \cdot \text{DH}_N - V \cdot \text{Phy}_C$	8 ± 0.5 (*) 14 ± 2 (**)	$\mu\text{mol-N L}^{-1}$ $\mu\text{mol-N L}^{-1}$
detritus and heterotrophs C	$\frac{d\text{DH}_C}{dt} = L \cdot \text{Phy}_C - (s \cdot \text{DH}_C + r) \cdot \text{DH}_C$	0.1	$\mu\text{mol-C L}^{-1}$
detritus and heterotrophs N	$\frac{d\text{DH}_N}{dt} = L \cdot \text{Phy}_N - (s \cdot \text{DH}_N + r) \cdot \text{DH}_N$	0.01	$\mu\text{mol-N L}^{-1}$

* PeECE II, ** PeECE III

Schulz, K. G., Riebesell, U., Bellerby, R. G. J., Biswas, H., Meyerhöfer, M., Müller, M. N., Egge, J. K., Nejtgaard, J. C., Neill, C., Wohlers, J., and Zöllner, E.: Build-up and decline of organic matter during PeECE III, *Biogeosciences*, 5, 707–718, doi:XXXXXX/bg-5-707-2008, <http://www.biogeosciences.net/5/707/2008/>, 2008.

5 Sommer, U., Paul, C., and Moustaka-Gouni, M.: Warming and Ocean Acidification Effects on Phytoplankton - From Species Shifts to Size Shifts within Species in a Mesocosm Experiment, *PLOS ONE*, 10, 39–51, doi:10.1371/journal.pone.0125239, 2015.

Tanaka, T., Thingstad, T. F., Løvdal, T., Grossart, H.-P., Larsen, A., Allgaier, M., Meyerhöfer, M., Schulz, K. G., Wohlers, J., Zöllner, E., and Riebesell, U.: Availability of phosphate for phytoplankton and bacteria and of glucose for bacteria at different pCO₂ levels in a mesocosm study, *Biogeosciences*, 5, 669–678, doi:10.5194/bg-5-669-2008, <http://www.biogeosciences.net/5/669/2008/>, 2008.

Toral, R. and Colet, P.: *Stochastic Numerical Methods*, Wiley-VCH, 2014.

10 Tortell, P. D., Payne, C. D., Li, Y., Trimborn, S., Rost, B., Smith, W. O., Riesselman, C., Dunbar, R. B., Sedwick, P., and DiTullio, G. R.: CO₂ sensitivity of Southern Ocean phytoplankton, *Geophysical Research Letters*, 35, n/a–n/a, doi:10.1029/2007GL032583, <http://dx.doi.org/10.1029/2007GL032583>, 104605, 2008.

Wirtz, K. W.: Non-uniform scaling in phytoplankton growth rate due to intracellular light and CO₂ decline, *Journal of Plankton Research*, 33, 1325–1341, 2011.

15 Wirtz, K. W.: Mechanistic origins of variability in phytoplankton dynamics: Part I: Niche formation revealed by a Size-Based Model, *Marine Biology*, 2013.

Wirtz, K. W. and Eckhardt, B.: Effective variables in ecosystem models with an application to phytoplankton succession, *Ecological Modelling*, 92, 33–53, 1996.

20 Wirtz, K. W. and Pahlow, M.: Dynamic chlorophyll and nitrogen:carbon regulation in algae optimizes instantaneous growth rate, *Marine Ecology Progress Series*, 402, 81–9, 2010.

Zondervan, I., Zeebe, R. E., Rost, B., and Riebesell, U.: Decreasing marine biogenic calcification: A negative feedback on rising atmospheric pCO₂, *Global Biogeochemical Cycles*, 15, 507–516, doi:10.1029/2000GB001321, <http://dx.doi.org/10.1029/2000GB001321>, 2001.

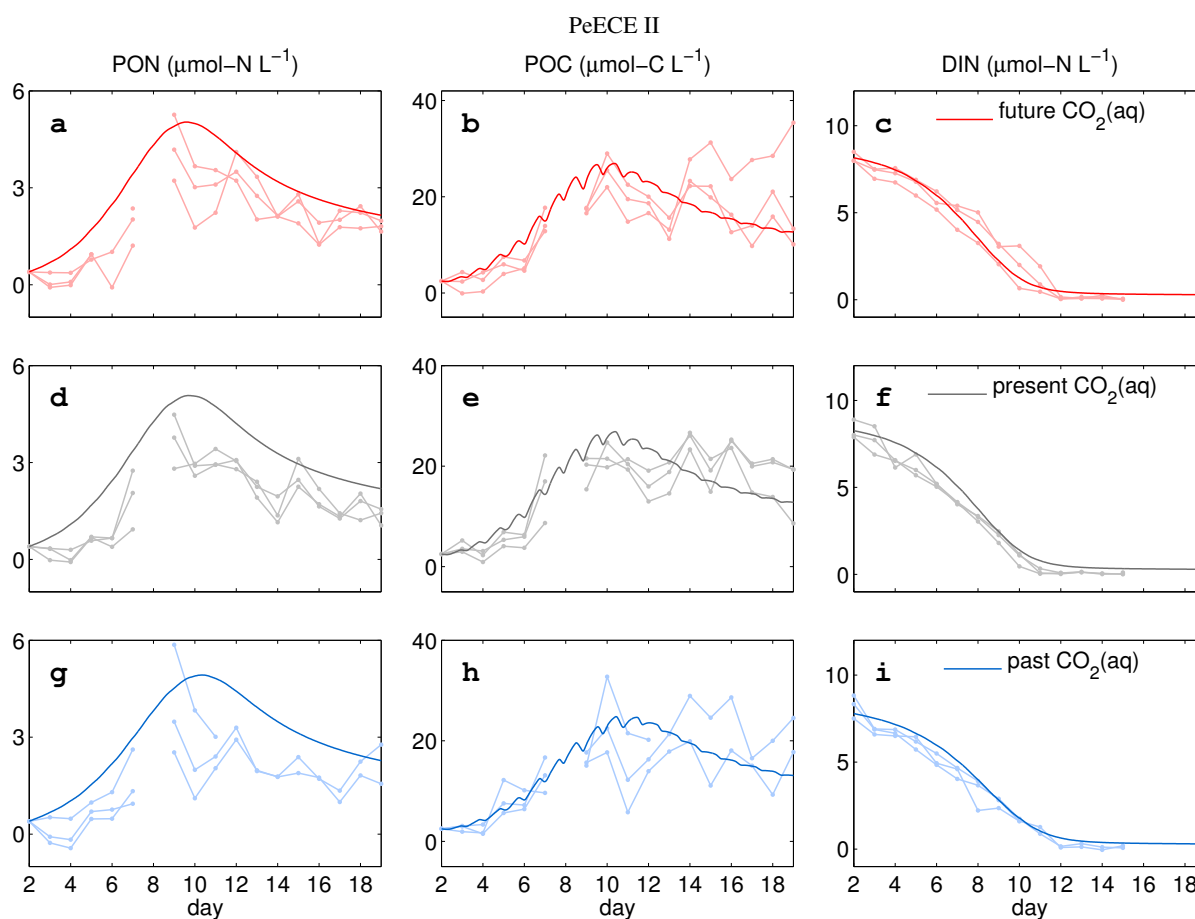


Figure 1. Solid lines show reference runs for POC, PON and DIN simulating the mean of the replicates per treatment, with different colors for the three experimental CO₂ set-ups. Dots are replicated data from the Pelagic Enrichment CO₂ Experiment (PeECE II) for newly produced POC and PON, i.e. starting values at day 2 were subtracted from subsequent measurements as in Riebesell et al. (2007).



Table 2. Parameter values used for the reference run, $\langle \phi_i \rangle$. All values are common to both PeECE II and III experiments, only the mean temperature (determined by environmental forcing) and the averaged cell size in the community are different since different species composition succeeded in the experiments (*Emiliania huxleyi* was the major contributor to POC in PeECE II (Engel et al., 2008) but also diatoms significantly bloomed during PeECE III (Schulz et al., 2008)).

Parameter	Value	Units	Variable	Reference
a_{CO_2} carbon acquisition	0.15	$(\mu\text{mol-C})^{-1}\text{L}$	Phy_C	this study
a_{PAR} light absorption	0.7	$\mu\text{mol phot}^{-1}\text{m}^2$	"	this study
a^* carboxylation depletion	0.15	μm^{-1}	"	this study
P_{\max} max. photosyn. rate	12	d^{-1}	"	this study
Q_{subs}^* subsist. quota offset	0.33	$\text{mol-N}(\text{mol-C})^{-1}$	"	this study
α_Q Q_{subs} allometry	0.4	-	"	this study
ζ costs of N assimil.	2	$\text{mol-C}(\text{mol-N})^{-1}$	"	Raven (1980)
l size $\text{Ln}(\text{ESD}/1\mu\text{m})$	1.6	-	$\text{Phy}_C, \text{Phy}_N, \text{DIN}$	PeECE II data
	1.8	-		PeECE III data
f_p fraction of protein in photosyn. machinery	0.4	-	"	this study
V_{\max}^* max. nutrients uptake	0.5	$\text{mol-N}(\text{mol-C d})^{-1}$	"	this study
A_{ff} nutrient affinity	0.2	$(\mu\text{mol-C d})^{-1}\text{L}$	"	this study
α_V V_{\max} allometry	0.45	-	"	Edwards et al. (2012)
L^* phyto. losses coeff.	$11 \cdot 10^{-3}$	$(\mu\text{mol-C d})^{-1}$	$\text{Phy}_C, \text{Phy}_N$ and DH_C, DH_N	this study
r^* DIN remin. & excret.	1.5	d^{-1}	DH_C, DH_N	this study
s DH sinking	10	$\text{L}(\mu\text{mol-C d})^{-1}$	"	this study
T_{ref} referen. temperature	8.3	Celsius	$\text{Phy}_C, \text{Phy}_N$ and	PeECE II data
	10.1	Celsius	$\text{DIN}, \text{DH}_C, \text{DH}_N$	PeECE III data



Table 3. Tolerance of the mesocosms to differences among replicates, given as a percentage of the reference factor value listed in Tables 1 and 2. According to our model projections, above these thresholds ($\Delta\text{POC}_i^{\text{mod}}$ exceeds $(\Delta\text{POC})^{\text{exper}}$. In bold, main contributors to the modeled variability ($\Delta\text{POC}^{\text{mod}}$ during the bloom (see Sec. 3).

factor ϕ_i		$\Delta\Phi_i$ (%)						averaged tolerance (%)
		PeECE II			PeECE III			
		Future	Present	Past	Future	Present	Past	
$\text{Phy}_C(0)$	initial phyto C biomass	68	49	46	78	60	100	67 ± 6
$\text{Phy}_N(0)$	initial phyto N biomass	26	19	22	21	16	29	22 ± 4
DIN(0)	initial DIN	20	28	29	17	11	18	20 ± 6
a_{CO_2}	carbon absorption	89	46	23	86	63	46	59 ± 23
a_{PAR}	light absorption	>100	>100	98	>100	>100	92	> 100
P_{max}	maximum photosyn. rate	27	18	16	22	16	28	21 ± 5
Q_{subs}^*	subsistence quota offset	6	5	6	5	4	9	6 ± 1
α_Q	Q_{subs} allometry	9	7	8	7	5	10	8 ± 2
ℓ	size Ln(ESD/$1\mu\text{m}$)	25	20	29	19	14	22	22 ± 5
f_p	fraction of protein in photosyn. machinery	92	75	44	36	17	38	50 ± 25
V_{max}^*	maximum nutrient uptake	13	11	14	10	8	14	12 ± 2
Aff	nutrients affinity	39	31	42	38	36	55	40 ± 7
α_V	V_{max} allometry	14	11	15	10	8	14	12 ± 2
L^*	phytoplankton losses	22	30	28	12	10	15	20 ± 8
r^*	DIN remineralization	73	99	98	128	37	52	81 ± 31
s	DH sinking	> 100	> 100	96	> 100	61	79	>100
T_{ref}	reference temperature	17	12	14	9	7	14	12 ± 3

Table 4. Cumulative residuals for PeECE III.

	Y	E	M	units
POC	35.1	37.4		$\mu\text{mol-C L}^{-1}$
PON	6.0	9.1		$\mu\text{mol-N L}^{-1}$
DIN	6.7	9.2		$\mu\text{mol-N L}^{-1}$

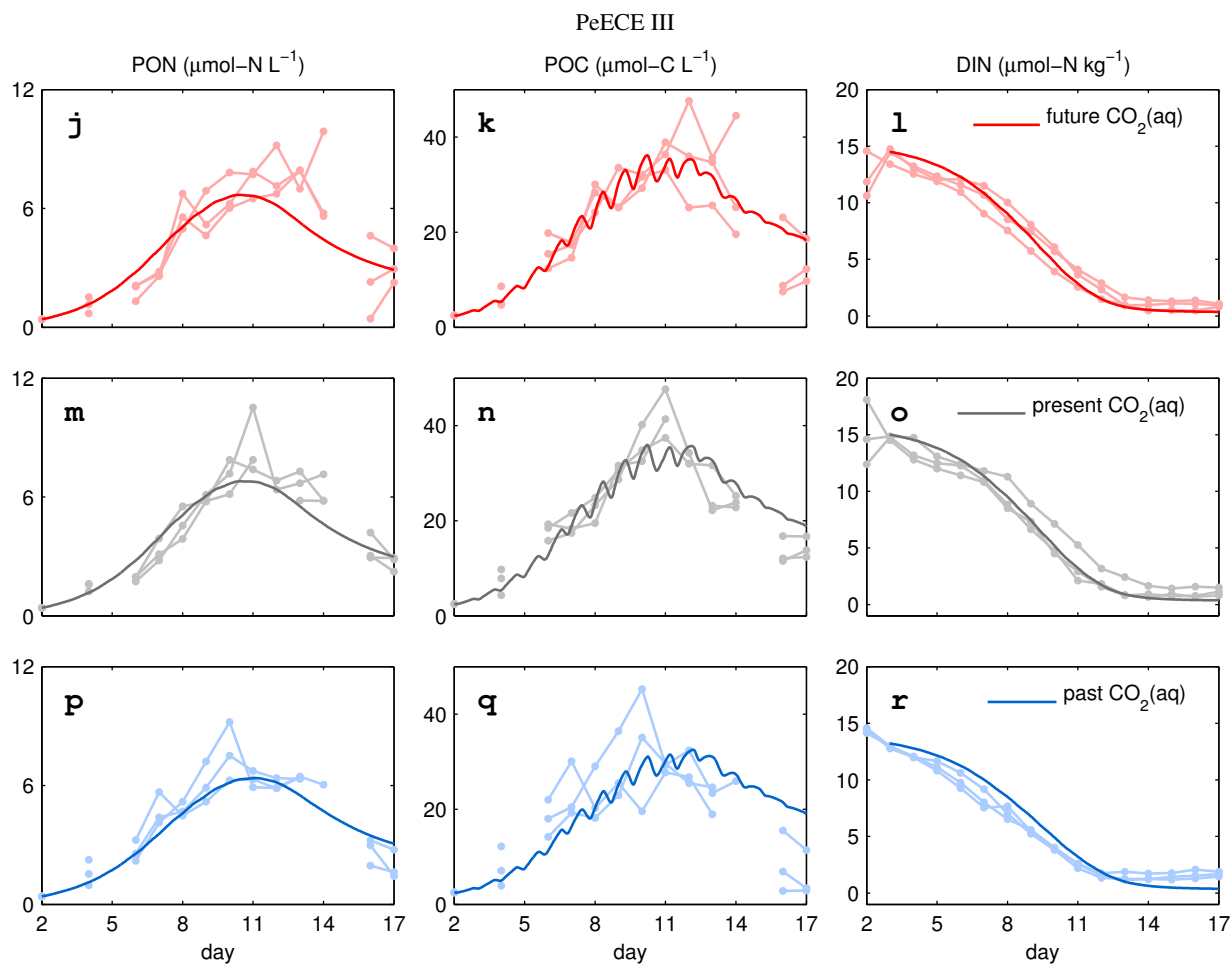


Figure 2. As in Fig. 2 for PeECE III.

List of changes

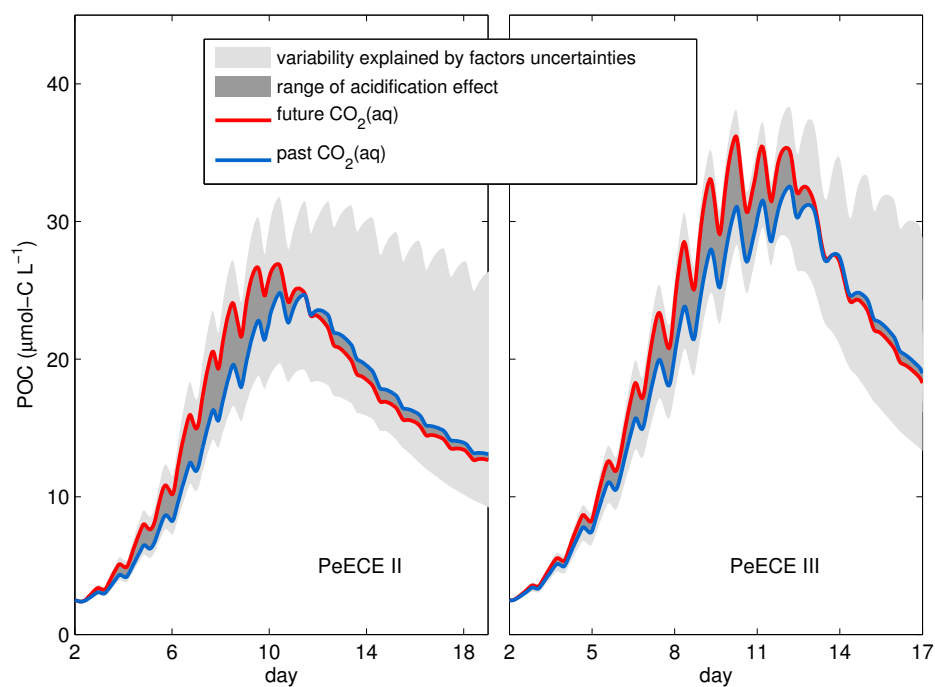


Figure 3. Reference simulation of POC for high CO₂ (red) and low CO₂ (blue) experimental conditions bound the range of acidification effect (dark gray) according to our model projections. Light gray is limited by the modeled POC variability, $(\Delta\text{POC})^{\text{mod}}$, quantified as the standard deviation of numerically simulated replicates calculated with differences in model factors simulating experimental uncertainties.

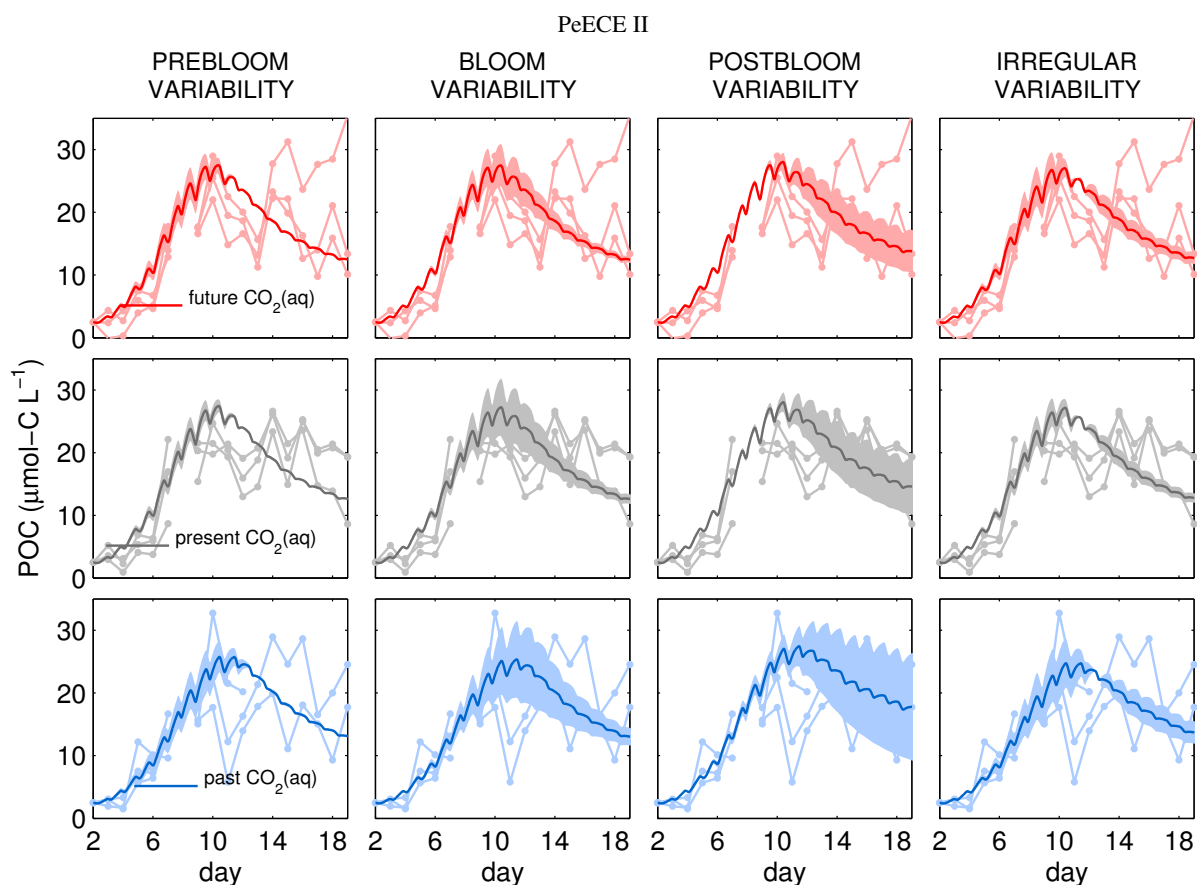


Figure 4. POC variability decomposition per factor, $(\Delta\text{POC})_i^{\text{mod}}$ for PeECE II. Shaded areas are bounded by the standard deviation of 10^4 modeled POC time series (see Sec. 2), around the mean trajectory of the ensemble (solid line). The timing of the amplification of the variability determines four separated kinds of behavior: factor uncertainties generating variability during the prebloom, bloom, postbloom or at irregular phase (see Sec. 3).

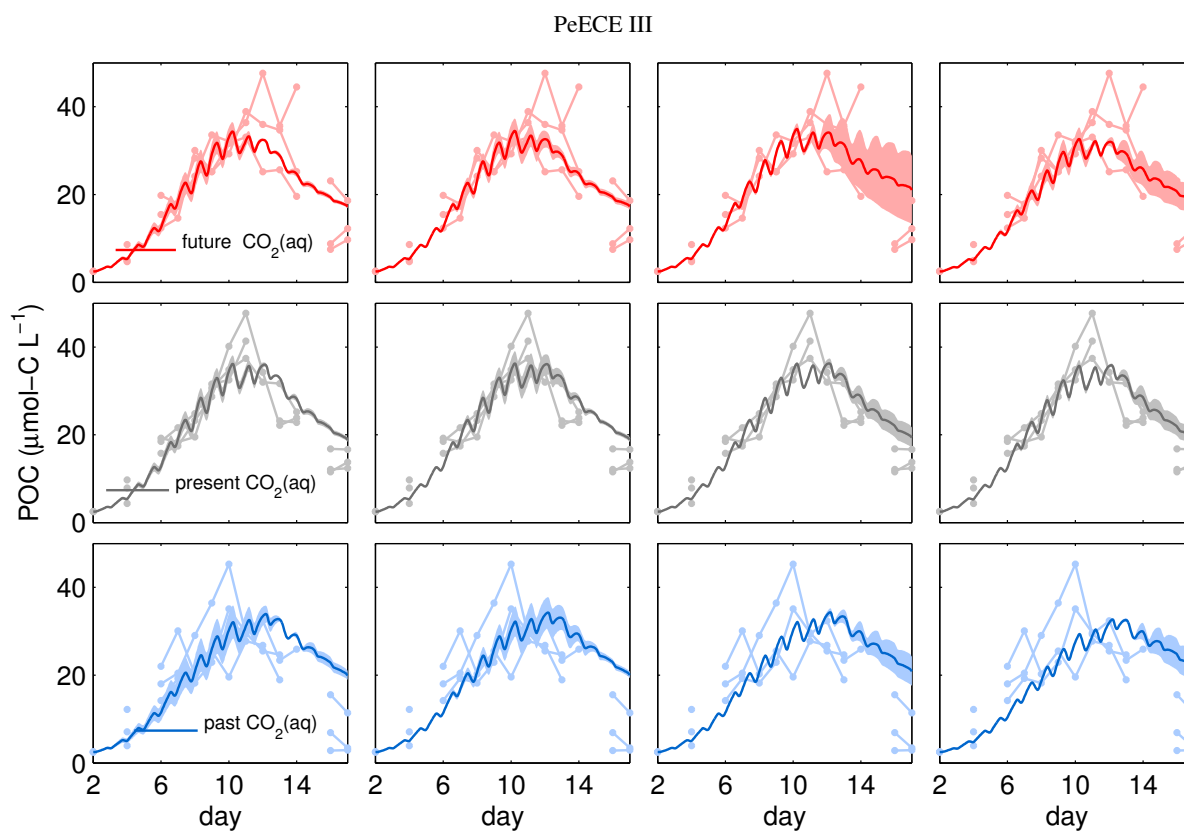


Figure 5. As Fig. 4, for PeECE III.

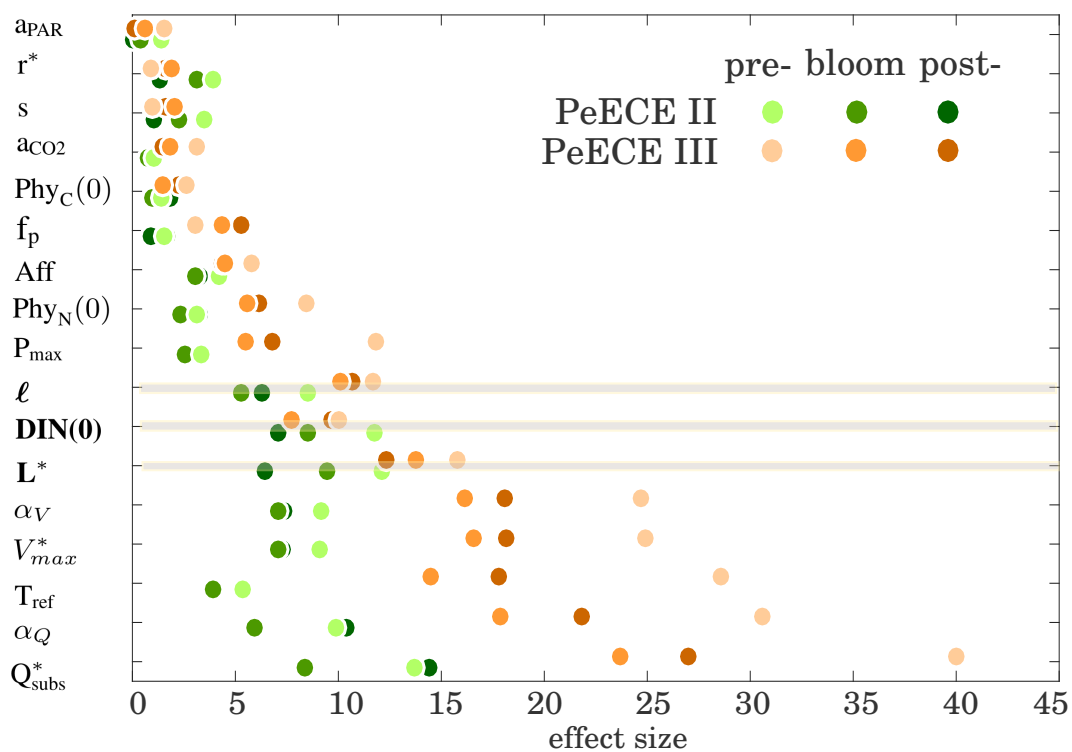


Figure 6. Effect size ε_i , Eq. (2) of variations in factors ϕ_i listed in Tables 1 and 2 for different bloom phases in two OA independent mesocosm experiments. Factors whose variations trigger variability mainly during the bloom (Fig. 4) and potentially mask acidification effects (Fig. 3) are highlighted.

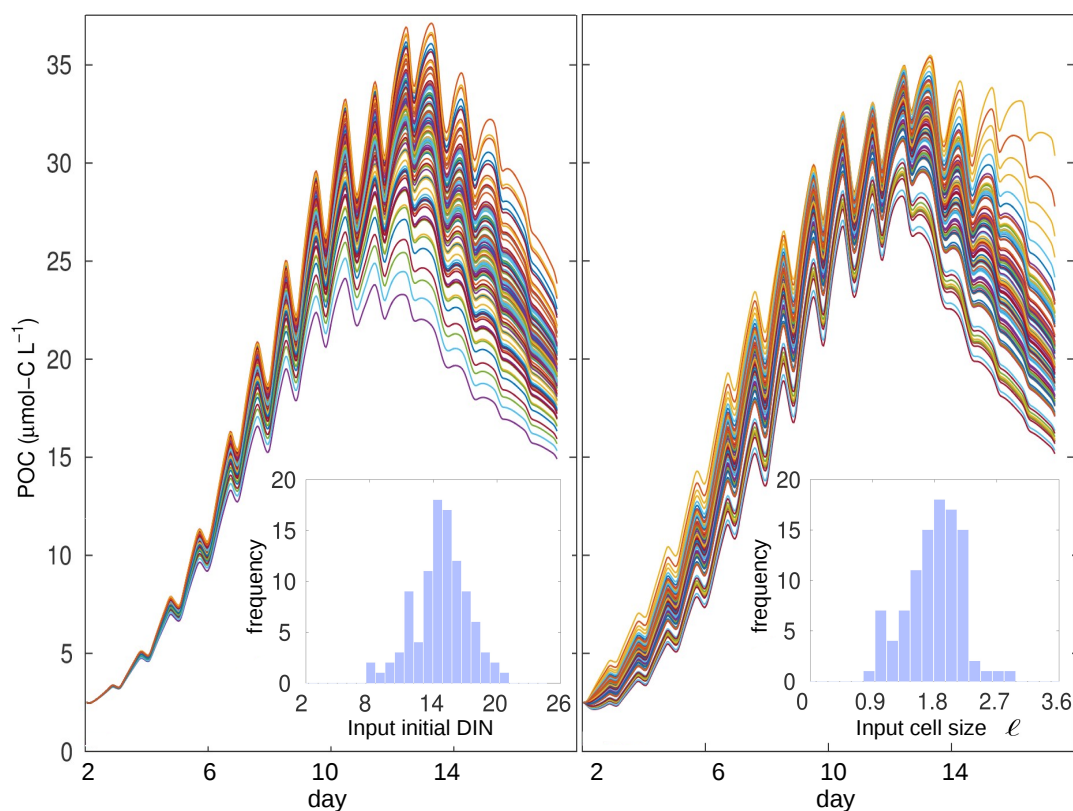


Figure 7. Examples of 10^2 model projections of POC for the past CO_2 treatment in PeECE III. In the left panel, model outputs with slight variations in initial nutrient concentration, $\langle \text{DIN}(0) \rangle \pm \Delta \text{DIN}(0)$, and in the right panel, with slight input variations in mean cell size, $\langle \ell \rangle \pm \Delta \ell$, where the means and standard deviations are given in Tables 2 and 3 respectively. To calculate system tolerances and uncertainties effect sizes, we numerically calculate 10^4 model realizations, thus the distribution of input factors is closer to normal.

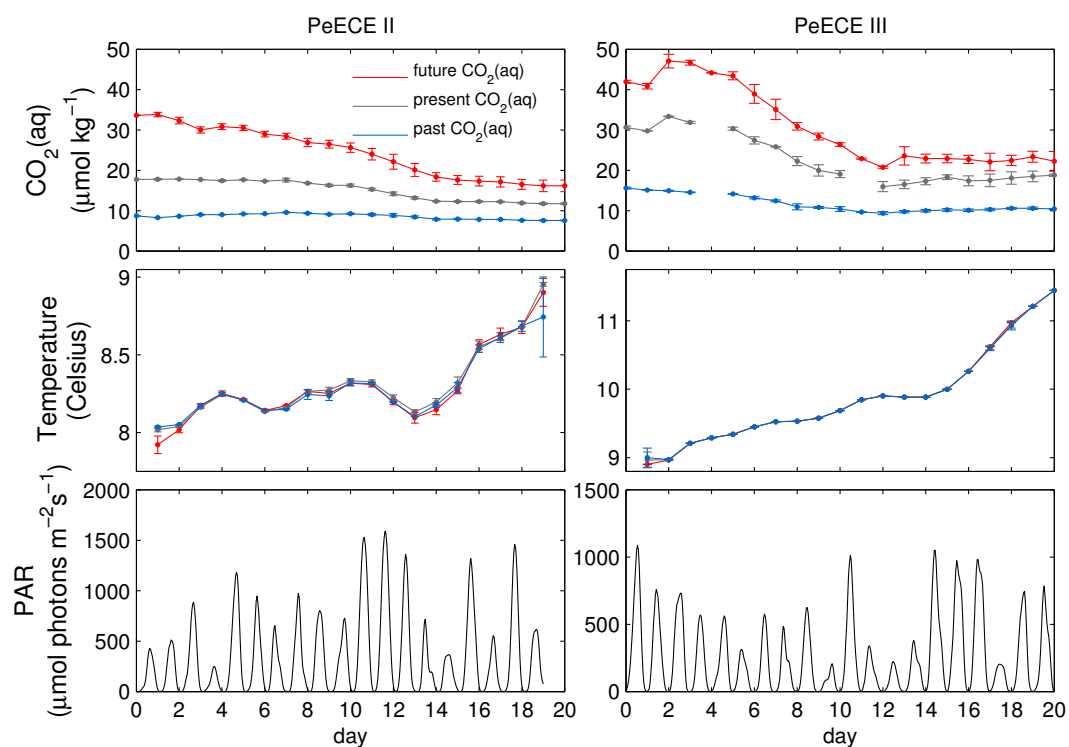


Figure 8. Environmental data from PeECE II and III are taken as model inputs. Error bars denote the standard deviation of the same treatment replicates.