Manuscript prepared for Biogeosciences
with version 2.3 of the LaTeX class copernicus.cls.
Date: 21 July 2009

Projected 21st century decrease in marine productivity: a multi-model analysis

M. Steinacher[1,2], F. Joos[1,2], T. L. Frölicher[1,2], L. Bopp[3], P. Cadule[3], S. C. Doney[4], M. Gehlen[3], B. Schneider[5], and J. Segschneider[6]

[1]Climate and Environmental Physics, Physics Institute, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland
[2]Oeschger Centre for Climate Change Research, University of Bern, Erlachstrasse 9a, CH-3012 Bern, Switzerland
[3]Laboratoire du Climat et de l'Environnement (LSCE), L'Orme des Merisiers Bât. 712, F-91191 Gif sur Yvette, France
[4]Dept. of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543-1543, USA
[5]Institute of Geosciences, University of Kiel, Ludewig-Meyn-Str. 10, D-24098 Kiel, Germany
[6]Max-Planck-Institut für Meteorologie, Bundesstrasse 53, D-20146 Hamburg, Germany

Correspondence to: M. Steinacher (steinacher@climate.unibe.ch)

Abstract. Changes in marine net primary productivity and export of particulate organic carbon are projected over the 21st century with three global coupled carbon cycle-climate models. These include representations of marine ecosystems and the carbon cycle of different structure and complexity. All three models show a decrease in global mean marine productivity and export production between 7 and 20% by 2100 relative to preindustrial conditions, for the SRES A2 emission scenario. Two different regimes for productivity changes are consistently identified in all three models. The first chain of mechanisms is dominant in the low- and mid-latitude ocean and in the North Atlantic: reduced input of macro-nutrients into the euphotic zone related to enhanced stratification, reduced mixed layer depth, and slowed circulation causes a decrease in macro-nutrient concentrations and in productivity and export of particulate organic carbon. The second regime is projected for parts of the Southern Ocean: an alleviation of light and/or temperature limitation leads to an increase in primary and export production as productivity is fueled by a sustained nutrient input. A region of disagreement among the models is the Arctic, where two models project an increase in productivity while one model projects a decrease. Projected changes in seasonal and interannual variability are modest in most regions. Regional model skill metrics are proposed to generate multi-model mean fields that show an improved skill in representing observations compared to a simple multi-model average. Model results are compared to recent productivity projections with three different algorithms, usually applied to infer primary production from satellite observations.

1 Introduction

Marine productivity and the marine biological cycle are important elements of the climate system. Biological processes influence, among other Earth system properties, the atmospheric abundance of radiative agents such as CO_2 (e.g. Volk and Hoffert, 1985; Siegenthaler and Wenk, 1984), N_2O (Suntharalingam and Sarmiento, 2000; Goldstein et al., 2003; Schmittner and Galbraith, 2008), dimethylsulphate (Bopp et al., 2003) and aerosols as well as the bio-optical properties of seawater and upper ocean physics (Timmermann and Jin, 2002; Manizza et al., 2008). However, the representation of ocean ecosystems (Six and Maier-Reimer, 1996; Moore et al., 2004; Le Quéré et al., 2005; Maier-Reimer et al., 2005; Aumont and Bopp, 2006; Vichi et al., 2007) and biogeochemical cycles in comprehensive atmosphere-ocean general circulation models (AOGCMs; Bopp et al., 2001; Fung et al., 2005; Wetzel et al., 2006; Crueger et al., 2008) is a relatively new field that requires further development to provide matured and robust results.

The goal of this study is to provide a multi-model estimate of long-term trends in net primary productivity (PP) and export of organic material (EP) using global warming simulations from three fully coupled atmosphere-ocean general circulation models and to identify the mechanisms behind these changes. These are the IPSL-CM4-LOOP model (IPSL), the COSMOS Earth System Model of the Max-Planck Institute for Meteorology (MPIM), and the Climate System Model CSM1.4-carbon of the National Center for Atmospheric Research (NCAR). The focus of the analysis is on how decadal-to-century scale changes in physical factors and nutrient availability affect global and regional productivity and export. The motivation is to provide an account on the performance of current climate-ecosystem models under global warming and to derive a best estimate of changes in productivity using regional model skill metrics. Our interest is further fueled by the contradicting projections for global productivity from mechanistic models, as used here, and a recent statistical model approach (Sarmiento et al., 2004).

A general finding across the hierarchy of mechanistic models is that marine global productivity and organic matter export decreases in 21st century global warming simulations (Klepper and De Haan, 1995; Maier-Reimer et al., 1996; Joos et al., 1999; Matear and Hirst, 1999; Plattner et al., 2001; Bopp et al., 2001; Fung et al., 2005; Frölicher et al., 2009). Increased stratification and a slowed thermohaline circulation in response to surface warming and freshening cause a decrease in the delivery of nutrients to the surface. As a consequence, global productivity and export is reduced. In these models, the marine biological cycle is fully closed in the sense that nutrient uptake by phytoplankton, export of organic material into the thermocline, remineralization of organic material and transport of inorganic nutrients by the circulation is represented. In the simpler models, productivity is tied to the availability of nutrients (such as phosphate or iron), light and temperature without considering food web dynamics, whereas in the more complex models the growth of phyto- and zooplankton, nitrogen fixation, and food web interactions and floristic shifts are explicitly taken into account, albeit in a simplified way. Globally, the change in nutrient supply is the dominant mech-

anism for productivity changes in 21st century global warming simulations, whereas other factors such as changes in light availability and the growing season length due to sea ice retreat, altered oceanic mixing conditions, and cloud characteristics, or the direct impact of elevated temperature on physiology considerably affect regional responses in productivity (Bopp et al., 2001). A decrease in global primary and new production by 5 to 8% is also projected in an off-line simulation with an ecosystem model (Moore et al., 2002) driven by the climate induced changes in ocean physics from an AOGCM simulation of the SRES A1 mid-range emission scenario (Boyd and Doney, 2002); the decrease is primarily attributed to the prescribed reduction in subsurface nutrients. In contrast, Sarmiento et al. (2004) projects an increase in global primary productivity by 0.7 to 8.1% using an empirical model approach. We also note that a model that incorporates a pCO_2-sensitive biotic carbon-to-nitrogen relation yields an increase in productivity end export production in 21st century CO_2 scenarios (Schmittner et al., 2008; Oschlies et al., 2008).

Schneider et al. (2008) present results for the same suite of three Earth System models as used in this study. They provide detailed information on the performance of the three models under current climate conditions and compare modeled physical (temperature, salinity, mixed layer depth, meridional overturning, ENSO variability) and biological (primary and export production, chlorophyll concentration) results with observation-based data. Of particular interest is the model performance with respect to seasonal and interannual variability as changes on these time scales may be linked to the century scale changes examined here. The models capture the general distribution of higher absolute primary productivity and higher seasonal variability in the intermediate to high latitudes, though all models overestimate seasonal variability in intermediate southern latitudes. Interannual variability is largely controlled by the permanently stratified low-latitude ocean in all three models consistent with satellite data (Behrenfeld et al., 2006). However, the MPIM model strongly overestimates the amplitude and frequency of interannual productivity variations, while the variability amplitude is slightly too low in the NCAR model. Only the IPSL model is able to capture the correlation between observation-based productivity, sea surface temperature and stratification in the low-latitude, stratified ocean. The MPIM model, and to a lesser degree, the NCAR model, suffer from a too strong iron limitation compared to the real ocean.

A challenge for any multi-model analysis is how to extract and distill the information contained in the individual models in a quantitative way. Ideally, the strengths of each individual model would be combined while weaknesses and failures would be removed to obtain an optimal multi-model mean. Here, we use regional weights to compute multi-model mean fields in in productivity and productivity changes.

In this paper we analyze centennial-scale changes in marine productivity under anthropogenic climate warming. Unlike earlier studies, we make use of three interactively coupled global carbon cycle-climate models that include iron cycling and representations of the marine biogeochemistry of different complexities. The use of a multi-model ensemble increases the robustness of the results.

The models are forced with prescribed CO_2 emissions from reconstructions (1860-2000 AD) and a high emission scenario, SRES A2 (2000-2100 AD). In the next section, models and experimental setup are described. In the result section, we first present projections for marine productivity. Then, we investigate underlying physical and biogeochemical mechanisms, quantify model sensitivities, and also address changes in the seasonal cycle. Regional model skill metrics are used to compute multi-model mean changes. In the discussion section, results of the mechanistic models are compared with those of Sarmiento et al. (2004) and discussed in the light of earlier studies. In the following, the variables PP and EP are used to represent net primary productivity and export of particulate organic carbon (POC), respectively.

2 Methods

2.1 Models

All models used in this study are fully coupled 3-D atmosphere-ocean climate models that contributed to the IPCC Fourth Assessment Report (Solomon et al., 2007; Meehl et al., 2007). The models include carbon cycle modules for the terrestrial and oceanic components (Friedlingstein et al., 2006).

2.1.1 IPSL

The IPSL-CM4-LOOP (IPSL) model consists of the Laboratoire de Météorologie Dynamique atmospheric model (LMDZ-4) with a horizontal resolution of about $3^\circ \times 3^\circ$ and 19 vertical levels (Hourdin et al., 2006), coupled to the OPA- 8 ocean model with a horizontal resolution of $2^\circ \times 2^\circ \cdot \cos\phi$ and 31 vertical levels and the LIM sea ice model (Madec et al., 1998). The terrestrial biosphere is represented by the global vegetation model ORCHIDEE (Krinner et al., 2005) and the marine carbon cycle is simulated by the PISCES model (Aumont et al., 2003). PISCES simulates the cycling of carbon, oxygen, and the major nutrients determining phytoplankton growth (PO_4^{3-}, NO_3^-, NH_4^+, Si, Fe). Phytoplankton growth is limited by the availability of nutrients, temperature, and light. The model has two phytoplankton size classes (small and large), representing nanophytoplankton and diatoms, as well as two zooplankton size classes (small and large), representing microzooplankton and mesozooplankton. For all species the C:N:P ratios are assumed constant (122:16:1; Takahashi et al., 1985), while the internal ratios of Fe:C, Chl:C, and Si:C of phytoplankton are predicted by the model. Iron is supplied to the ocean by aeolian dust deposition and from a sediment iron source. Iron is also added at the surface if the iron concentration falls below a lower limit of 0.01 nM. Iron is taken up by the plankton cells and released during remineralization of organic matter. Scavenging of iron onto particles is the sink for iron to balance external input. There are three non-living components of organic carbon in the model: semi-labile dissolved organic carbon (DOC), with a lifetime of several weeks to years, as well as large and small detrital particles, which are fueled by

mortality, aggregation, fecal pellet production and grazing. Small detrital particles sink through the water column with a constant sinking speed of 3 m day^{-1}, while for large particles the sinking speed increases with depth from a value of 50 m day^{-1} at the depth of the mixed layer, increasing to a maximum sinking speed of 425 m day^{-1} at 5000 m depth. For a more detailed description of the PISCES model see Aumont and Bopp (2006) and Gehlen et al. (2006). Further details and results from the fully coupled model simulation of the IPSL-CM4-LOOP model are given in Friedlingstein et al. (2006).

2.1.2 MPIM

The Earth System Model employed at the Max-Planck- Institut für Meteorologie (MPIM) consists of the ECHAM5 (Roeckner et al., 2006) atmospheric model of 31 vertical levels with the embedded JSBACH terrestrial biosphere model and the MPIOM physical ocean model, which includes a sea ice model (Marsland et al., 2003) and the HAMOCC5.1 marine biogeochemistry model (Maier-Reimer, 1993; Six and Maier-Reimer, 1996; Maier-Reimer et al., 2005). The coupling of the marine and atmospheric model components, and in particular the carbon cycles, is achieved by using the OASIS coupler.

HAMOCC5.1 is implemented into the MPIOM physical ocean model configuration using a curvilinear coordinate system with a 1.5° nominal resolution where the North Pole is placed over Greenland, thus providing relatively high horizontal resolution in the Nordic Seas. The vertical resolution is 40 layers, with higher resolution in the upper part of the water column (10 m at the surface to 13 m at 90 m). The marine biogeochemical model HAMOCC5.1 is designed to address large-scale, long-term features of the marine carbon cycle, rather than to give a complete description of the marine ecosystem. Consequently, HAMOCC5.1 is a NPZD model with one phytoplankton group (implicitly divided into coccolithophorids and diatoms) and one zooplankton species. Detritus is formed from dead phytoplankton and zooplankton, and zooplankton fecal pellets. Furthermore, dissolved organic matter is produced by phytoplankton exudation and zooplankton excretion. The carbonate chemistry is identical to the one described in Maier-Reimer (1993). A more detailed description of HAMOCC5.1 can be found in Maier-Reimer et al. (2005), while here only the main features relevant for the described experiments will be outlined.

Phytoplankton growth depends on the availability of light (I) and nutrients. The local light supply is calculated from the temporally and spatially varying solar radiation at the sea surface, $I(0,t)$, as provided by the OGCM. Below the surface, light intensity is reduced due to attenuation by sea water (k_w) and chlorophyll (k_c) using a constant conversion factor for C:Chl, $R_{\mathrm{C:Chl}}$:

$$I(z,t) = I(0,t)\, e^{-(k_w + k_c\, \mathrm{PHY}\, 12\, R_{\mathrm{C:P}}/R_{\mathrm{C:Chl}})z} \tag{1}$$

Phytoplankton growth depends linearly on the availability of light, without saturation of growth rates for stronger irradiance (I). The growth rate $J(I(z,t))$, is calculated as $J(I) = \alpha_{\mathrm{PHY}}\, I(z,t)$,

where α_{PHY} is the slope of the P-vs-I-curve (production vs. light intensity). The light limited phytoplankton growth rate is then multiplied by the nutrient limitation factor, which is calculated from a simple Monod function, limited by the least available nutrient (either phosphate, nitrate, or iron).

Silicate concentrations are used to distinguish the growth of diatoms and coccolithophorides: if silicate is abundant, diatoms grow first, thereby reducing the amount of nutrients available for coccolithophoride growth. Only the shells (opal and calcium carbonate) that are part of detritus (DET) are considered for the partition into the fractions P_{sil} and P_{car} of production:

$$P_{\mathrm{sil}} = \min\left(\frac{\Delta \mathrm{DET}}{\Delta t}\, R_{\mathrm{Si:P}} \frac{\mathrm{SI(OH)_4}}{K_{\mathrm{PHY}}^{\mathrm{SI(OH)_4}} + \mathrm{SI(OH)_4}}, 0.5\,\mathrm{SI(OH)_4}\right), \tag{2}$$

where $\frac{\Delta \mathrm{DET}}{\Delta t}$ is the export production, $R_{\mathrm{Si:P}} = 25$ denotes the Si:P ratio required by diatoms, $K_{\mathrm{PHY}}^{\mathrm{SI(OH)_4}} = 1\,\mathrm{mmol\,m^{-3}}$ the half-saturation constant for silicate uptake. The remaining fraction of photosynthesis is by coccolithophorids. Again, as for opal we only account for the sinking part of calcite production:

$$P_{\mathrm{car}} = \frac{\Delta \mathrm{DET}}{\Delta t}\, R_{\mathrm{Ca:P}}\, \frac{K_{\mathrm{PHY}}^{\mathrm{SI(OH)_4}}}{K_{\mathrm{PHY}}^{\mathrm{SI(OH)_4}} + \mathrm{SI(OH)_4}}, \tag{3}$$

with $R_{\mathrm{Ca:P}} = 35$ being the $CaCO_3$ to PO_4 ratio.

The model also includes cyanobacteria that take up nitrogen from the atmosphere and transform it directly into nitrate. In the model version used here, biological production is temperature-independent, assuming that phytoplankton acclimate to local conditions. Global dust deposition fields are used to define the source function of bioavailable iron. Removal of dissolved iron occurs through biological uptake and export, and by scavenging which is described as a relaxation to the deep-ocean iron concentration of 0.6 nM if the local concentration exceeds this value. In the experiments used here, export of particulate matter is simulated using prescribed settling velocities for opal (30 m day^{-1}), calcite shells (30 m day^{-1}) and organic carbon (10 m day^{-1}). Remineralization of organic matter depends on the availability of oxygen. In anoxic regions, remineralization occurs via denitrification. HAMOCC5.1 also includes an interactive module to describe the sediment-water flux at the sea floor. This component further simulates pore water chemistry, the solid sediment fraction and interactions between the sediment and the oceanic bottom layer as well as between solid sediment and pore water constituents.

2.1.3 NCAR

The physical core of the NCAR CSM1.4 carbon climate model (Doney et al., 2006; Fung et al., 2005) is a modified version of the NCAR CSM1.4 coupled physical model, consisting of ocean, atmosphere, land and sea ice components integrated via a flux coupler without flux adjustments (Boville et al., 2001; Boville and Gent, 1998). The atmospheric model CCM3 is run with a horizontal

resolution of 3.75° and 18 levels in the vertical (Kiehl et al., 1998). The ocean model is the NCAR CSM Ocean Model (NCOM) with 25 levels in the vertical and a resolution of 3.6° in longitude and 0.8° to 1.8° in latitude (Gent et al., 1998). The sea ice component model runs at the same resolution as the ocean model, and the land surface model runs at the same resolution as the atmospheric model.

The CSM1.4-carbon model includes a modified version of the terrestrial biogeochemistry model CASA (Carnegie- Ames-Stanford Approach; Randerson et al., 1997), and a derivate of the OCMIP-2 (Ocean Carbon-Cycle Model Intercomparison Project Phase 2) ocean biogeochemistry model (Najjar et al., 2007). In the ocean model, the biological source-sink term has been changed from a nutrient restoring formulation to a prognostic formulation, and thus biological productivity is modulated by temperature (T), surface solar irradiance (I), mixed layer depth (MLD), and macro- and micro-nutrients (PO_4^{3-}, and iron):

$$\mathrm{PP} = \frac{T + 2°\mathrm{C}}{T + 10°\mathrm{C}} \; \min\left(\frac{[\mathrm{PO_4}]}{[\mathrm{PO_4}] + \kappa_{\mathrm{PO_4}}}, \frac{[\mathrm{Fe}]}{[\mathrm{Fe}] + \kappa_{\mathrm{Fe}}}\right) \; \frac{I}{I + \kappa_I} \cdot \min\left([\mathrm{PO_4}], \frac{[\mathrm{Fe}]}{r_{\mathrm{Fe:P}}}\right) \; \max\left(1, \frac{z_{\mathrm{MLD}}}{z_c}\right) \; \frac{1}{\tau}, \tag{4}$$

where $\kappa_{\mathrm{PO_4}} = 0.05\,\mu\mathrm{mol/l}$, $\kappa_{\mathrm{Fe}} = 0.03\,\mathrm{nmol/l}$, $\kappa_I = 20\,\mathrm{W/m^2}$, $r_{\mathrm{Fe:P}} = 5.85 \cdot 10^{-4}$, $\tau = 15\,\mathrm{days}$, and $z_c = 75\,\mathrm{m}$.

Following the OCMIP-2 protocols (Najjar et al., 2007) total biological productivity is partitioned 1/3 into sinking POC flux, here taken to be equivalent to export productivity (EP), and 2/3 into the formation of dissolved or suspended organic matter, where much of the latter is remineralized within the model euphotic zone. Total productivity thus contains both new and regenerated production, though the regenerated contribution is probably lower than in the real ocean, as only the turnover of semi-labile dissolved organic matter (DOM) is considered. NCAR primary productivity (PP) thus represents, rather, the carbon flux associated with net nutrient uptake and is not strictly equivalent to primary production as measured by ^{14}C methods. It is a reasonable proxy for the time and space variability of PP if somewhat underestimating the absolute magnitude. For reasons of simplicity, net nutrient uptake times the C:P ratio of 117 (Anderson and Sarmiento, 1994) is considered here as PP, even though it is not exactly the same. The ocean biogeochemical model includes the main processes of the organic and inorganic carbon cycle within the ocean and air-sea CO_2 flux. A parametrization of the marine iron cycle (Doney et al., 2006) includes atmospheric dust deposition/iron dissolution, biological uptake, vertical particle transport and scavenging. The CSM1.4-carbon source code is available electronically (see http://www.ccsm.ucar.edu/workinggroups/Biogeo/csm1_bgc/) and is described in detail in Doney et al. (2006).

2.2 Experiments

The models are forced by anthropogenic CO_2 emissions due to fossil fuel burning and land-use changes as reconstructed for the industrial period and following the SRES A2 emission scenario after 2000 AD. The NCAR and MPIM models also include CH_4 and CFCs. N_2O, volcanic emissions, and

changes in solar radiation are additionally taken into account by the NCAR model as described by Frölicher et al. (2009). All models were integrated for more than one thousand years for spin up as described by Schneider et al. (2008). For analysis, all variables have been interpolated onto a $1^\circ \times 1^\circ$ grid using a Gaussian interpolation. Control simulations in which CO_2 emissions are set to zero and other forcings are set to constant preindustrial levels are used to detrend model results. Slight trends in temperature, salinity, and nutrient concentrations have been removed from the IPSL and NCAR results. For the IPSL model the variables PP and EP have been detrended additionally.

3 Results

3.1 Projected annual mean primary productivity and export production under SRES A2

We briefly discuss the magnitude and spatio-temporal patterns of net primary production (PP) in comparison with satellite-based estimates (see Schneider et al. (2008) for a more comprehensive analysis) before addressing long-term changes in PP. Global annual PP ranges between 24 GtC yr^{-1} (MPIM) and 31 GtC yr^{-1} (IPSL) for modern conditions. This is considerably lower than the satellite-based range of 35 to 70 GtC yr^{-1} (Behrenfeld et al., 2006; Carr et al., 2006). The very low PP in the MPIM model is likely linked to an overall too strong limitation of PP by iron (Schneider et al., 2008). NCAR PP represents carbon uptake associated with net nutrient uptake, rather than overall primary productivity, and is thus underestimating real primary production by design. There are also deficiencies in the regional representation of PP (Fig. 1). High productivity along coastal margins is not adequately represented in coarse resolution models. The MPIM model underestimates productivity outside the equatorial regions, and the NCAR model has too low productivity in the equatorial Pacific. These deficiencies are related to the iron cycle of the two models. IPSL appears to underestimate productivity in high northern latitudes. The skill of individual models to represent the satellite-based productivity field is rather low with correlations between modeled and satellite-based fields of less than 0.6 (Fig. 1b). The errors in the simulated PP fields reflect both deficiencies in the simulated physical fields and in the representation of ecosystem processes in the coupled AOGCM. Results from ocean only models with prescribed surface forcing compare typically better with observations. We note that the satellite derived estimates also have uncertainties. For example, Carr et al. (2006) report that global PP estimates from twenty-four ocean-color-based models range over a factor of two, but correlations among the resulting fields are typically high.

Despite the deficiencies of individual models, the models as a class represent the pertinent features of the satellite-based observations such as a low productivity in the oligotrophic gyres and the southern high latitudes (all models), high productivity features in the North Atlantic (NCAR, IPSL), in the North Pacific (IPSL), around 30°S to 50°S (NCAR, IPSL), and in the equatorial and eastern boundary upwelling systems, high seasonal variability in the North Atlantic and in southern intermediate latitudes (all), as well as low seasonal variability around the equator (NCAR) and in mid

latitudes (all), and the correlation of temperature and stratification with productivity on the interannual time scales for the low-latitude, permanently stratified ocean (IPSL) or the Nino3 region (IPSL, NCAR). This comparison with satellite data allows us to continue with some confidence as well as with caution to the discussion of 21st century projections.

All three models show a reduction in the globally integrated annual mean primary production and POC export in the simulations from 1860 AD to 2100 AD under SRES A2 (Fig. 2, Table 1). The IPSL model, which also yields the highest preindustrial and present PP, shows the biggest changes. In that model PP declines by 4.6 GtC/yr by the end of this century, which is a reduction of the simulated preindustrial production by 13%. The other two models show reductions of 10% (2.3 GtC/yr; MPIM) and 7% (1.9 GtC/yr; NCAR). In the MPIM and NCAR models, the relative reduction in POC export follows closely the reduction in PP, while in the IPSL model the decline in POC export is more pronounced and amounts to a reduction of 20% with respect to the preindustrial value by 2100. The main reason for this decoupling of productivity and export in the IPSL model is a shift from diatoms and macrozooplankton to the smaller nanophytoplankton and microzooplankon (Bopp et al., 2005).

The projected PP decrease by the end of the century depends on the magnitude of the projected climate change and thus on the climate sensitivity of the models. A linear regression between global PP and global mean surface air temperature is used to normalize PP changes with respect to climate change (Fig. 2c). This yields a slope, i.e., the global PP decrease per °C warming, of 1.4 GtC yr^{-1} $°C^{-1}$ for the IPSL model, but only 0.6 GtC yr^{-1} $°C^{-1}$ for the MPIM and NCAR models.

We identify a number of regions with large reductions (more than 50 mg-C m^{-2} day^{-1}) in PP (Fig. 1). These correspond to high productivity areas. A large reduction in PP is found in the North Atlantic in the IPSL and NCAR model, around 35°S in IPSL and less pronounced in the NCAR, in the upwelling regions off Africa in all models and in the equatorial Pacific in the MPIM and IPSL model. These reductions are qualitatively consistent across the models with the obvious caveat that no major reductions can be expected in regions where an individual model fails to simulate a significant preindustrial productivity (e.g. MPIM outside the equator, NCAR in the equatorial Pacific). Consistent moderate increases in productivity are simulated in the high latitude Southern Ocean (all models) and around Svalbard, indicating that the high productivity zone in the North Atlantic is moving northward with climate warming and sea ice retreat. An increase in productivity is simulated in the North Pacific in the IPSL and NCAR models; we note that sea ice extent is unrealistically high in this area in the NCAR model (Steinacher et al., 2009; Weatherly et al., 1998). In summary, the model results suggest that PP will be reduced in most equatorial and mid-latitude regions and in the North Atlantic, and moderately enhanced in polar regions.

3.2 Mechanisms of long term shifts in PP under climate change

3.2.1 Attribution of productivity changes to individual drivers in the NCAR model

In order to identify links between long term shifts in PP and climate change, we first focus on the NCAR CSM1.4-carbon results. This model features the simplest formulations for biological production among the three models. PP is determined by the product $\mathrm{PP} \propto F_N \cdot F_I \cdot F_T \cdot B$ (Eq. 4), where the first three factors represent nutrient, light, and temperature limitation and B is a biomass proxy derived from phosphate and iron concentrations. The relative changes in these factors (Fig. 3a-d) directly yield the relative changes in PP (Fig. 3e). Light availability is tied to the mixed layer depth and sea ice fraction in the NCAR model. It increases when the mixed layer depth (MLD) exceeds 75 m. This unrealistic feature affects light limitation in the South Pacific (increased MLD/light availability) around 45°S and in a number of grid cells in the North Atlantic. We recall that the biomass proxy corresponds to the phosphate or (scaled) iron concentration (which ever is smaller) and thus directly represents nutrient concentrations.

The biomass proxy decreases in most areas of the world ocean (Fig. 3d). This can be attributed to a more efficient utilization of nutrients under global warming as found in previous work (e.g. Plattner et al., 2001). Reduced nutrient concentrations in combination with reduced export are indicative of reduced nutrient input from the thermocline into the mixed layer. Such conditions prevail in the Atlantic between 20°S and 65°N, in the western part of the Indian Ocean, and around 30°N and 35°S in the Pacific between 160°E and 140°W. PP shows little or no response to climate change in the tropical and subtropical Pacific, where PP is low due to an unrealistically strong iron limitation. On the other hand, reduced nutrient concentrations in combination with increased export are indicative of a sustained nutrient input into the euphotic zone. Sea ice retreat and warming in the Arctic alleviate the strong limitations by light and temperature and enhance Arctic productivity. Similarly, a reduction in temperature limitation boosts productivity around Antarctica in the model.

In the North Atlantic, where the largest PP changes occur, the PP decrease is dominated by a decrease in the biomass proxy. The reduction in nutrient concentration is linked to a reduction in the North Atlantic thermohaline circulation (Frölicher et al., 2009). Nutrients are used up more efficiently, and productivity decreases likely in response to less surface-to-deep exchange. The model also simulates an increase in light limitation, caused by changes in cloudiness and changes in mixed layer depth, and a somewhat stronger limitation by iron in the east and by phosphate in the west. The slight increase in PP in some areas in the Indian Ocean, around Australia, and in the South Atlantic around 25°S can mainly be attributed to an increased nutrient supply due to stronger upwelling.

In conclusion, PP changes in the NCAR model are tightly linked to changes in nutrient input into the euphotic zone in combination with an alleviation of light and temperature limitations in high latitudes. A reduced nutrient input into the surface is expected in climate change scenarios as surface stratification tends to increase in response to warming and freshening. Next, we will investigate

changes in physical factors such as stratification and upwelling and in nutrient availability and their link to productivity for all three models.

3.2.2 Basin-scale changes in productivity, physical properties, and nutrient concentrations

There is a surprisingly good overall consistency in projected trends among the models on the basin-scale and for a range of variables. Figure 4 shows projected changes in selected large regions for PP, EP, related physical properties, and nutrient concentrations for all three models. This comparison between changes in PP and in potential drivers is indicative of underlying mechanisms, albeit it does not allow for a stringent attribution as done in the previous section for the NCAR model. Overall, the results are qualitatively consistent across models and regions. PP, EP, MLD, and surface nutrient concentrations are projected to decrease in all models and in almost all regions, while sea surface temperature (SST) and stratification increase. This suggests that the mechanisms identified for the NCAR model are also key for the productivity changes in the IPSL and MPIM model. Namely, a reduced nutrient input related to enhanced stratification, reduced MLD, and a slowed circulation tends to decrease productivity and export of organic material under transient global warming.

All models exhibit pronounced changes in MLD and stratification in the North Atlantic, which transform to strong reductions in surface macro-nutrient concentrations. Consequently, PP and EP decrease in the IPSL and NCAR models by about 40% and 30%, respectively. In the MPIM model, preindustrial PP in the North Atlantic is unrealistically small due to too strong iron limitation and the 21st century reduction in PP is thus small as well.

All models show an increase in stratification and a decrease in MLD and macro-nutrients in the stratified ocean (SST $> 15^{\circ}$C). We again link this tentatively to a reduced nutrient input into the euphotic zone under global warming. Productivity and export decrease accordingly in all models.

In the Southern Ocean ($<45^{\circ}$S), relative PP trends are smaller than in other regions and vary in sign between different regions within the Southern Ocean. Changes that favor production, such as increased SST and light, and changes that tend to reduce production, such as reduced nutrient input, balance to some extent on the regional average. In the IPSL and NCAR simulations, PP increases on average, while MPIM shows a decrease of about 5%.

There are also some qualitative inconsistencies in projected trends between models. Most notable are the following two. IPSL simulates a decrease in PP and EP in the Arctic Ocean, in contrast to MPIM and NCAR that project an increase (Fig. 4). Surface iron concentration is projected to increase in IPSL in all regions, but to decrease in MPIM and NCAR in most regions (Fig. 4h). In the Arctic Ocean, light availability in the surface ocean is strongly enhanced in all models due to sea ice retreat. The annual mean sea ice cover in the Arctic is reduced by 32% (IPSL), 25% (MPIM), and 23% (NCAR) with respect to preindustrial conditions. This leads, together with an increase in SST and MLD, to a strong increase in PP and EP in the MPIM (+130%) and NCAR (+215%) simulations, despite the strong (NCAR, +90%) and moderate (MPIM, +20%) increase in

stratification and reduced surface nutrient concentrations. Although insolation and SST increase also strongly in IPSL, this model shows an opposite response in PP and EP. This can be explained with a strong increase in stratification of about 90% and the reduction in MLD and surface macro-nutrients of 50-70%.

The increase in surface iron concentration simulated by the IPSL model (20% in the global mean) is a consequence of the parametrization of the elemental ratio in phytoplankton. The ratio between carbon and nitrogen or phosphor is kept constant. In contrast, the iron-to-carbon ratio of phytoplankton is assumed to decrease with increasing nutrient (and light) limitation. Consequently, lower macro-nutrient concentrations in the euphotic zone lead to a relatively lower uptake of iron compared to other nutrients by plankton and to a lower iron-to-carbon ratio in organic material. In turn, less iron is exported out of the euphotic zone and iron concentrations increase, while macro-nutrient concentrations decrease. In the IPSL model, surface iron concentrations are restored to a minimum value of 0.01 nM. This influences the interannual variability in PP (Schneider et al., 2008). However, this potential artificial iron source does not contribute significantly to the long-term trend in surface iron because, first, the number of grid cells and months where iron is restored is reduced during the simulation, and second, these regions do not correspond to the regions where large changes in surface iron are simulated. In the NCAR and MPIM model, the iron-to-carbon and other elemental ratios are constant and iron concentration tends to decrease in parallel with macro-nutrient concentrations in the surface ocean.

3.2.3 Local correlations between changes in productivity and potential drivers

In this section, we address to which extent the features identified on the basin-scale are also evident on the local scale. We correlate simulated changes in annual mean PP with annual mean changes in SST, stratification, MLD, and shortwave radiation, as well as with phosphate and iron for each single grid cell (Fig. 5) and compare projected changes along two transects through the Atlantic (and Arctic), and the Pacific (Figs. 6 and 7). The transects, indicated in Figure 1, are selected to cover major productivity features in the two basins. The results tend to confirm the findings from the two previous sections, although the links between stratification, mixed layer depth and macro-nutrient concentrations are often somewhat obscured on the grid cell scale as evidenced by the small regression coefficient (R^2) found for many cells.

In the IPSL simulation, the PP decrease in the Pacific, North Atlantic and Indian Ocean correlates with enhanced stratification and decreased surface phosphate concentrations (Fig. 5). Changes in MLD correlate only weakly with PP trends; only in the North Atlantic and south-eastern Pacific are some relevant correlations found. Surface iron concentrations correlate positively with PP because surface iron increases almost everywhere in the IPSL simulation. Correlations for EP are similar (not shown).

The MPIM model shows generally weak correlations, which can be explained with the strong

iron limitation in that model. Under present climate conditions, PP is iron-limited in all regions except the tropical Atlantic (Schneider et al., 2008). Because surface iron concentrations decrease only slightly in most regions, no significant correlations are found. Exceptions are the low and mid latitudes of the Pacific, where surface iron concentrations decrease by about 20% and correlations of PP changes are found with surface iron (mainly in the subtropical gyres). Also, the PP decrease in the western tropical Pacific correlates with increased stratification and reduced MLD.

In the NCAR simulation, increased stratification correlates to some extent with reduced PP and EP in the tropical and southern Pacific, as well as in the North Atlantic. This model shows a stronger correlation between PP and MLD than the other two. The latter may be an artifact of the model light limitation. Significant positive correlations are found in the North Atlantic, North Pacific, and in the Southern Ocean. Reduced surface nutrient concentrations mainly correlate where the respective nutrient is limiting; PO_4 in the low- and mid-latitude Atlantic and in the northern Indian Ocean, iron in the Pacific and southern Indian.

In conclusion the multi-model analysis confirms important conclusions obtained by attributing changes in PP and EP to individual drivers in the NCAR model. We identify two different regimes for productivity changes in all models. First, a decrease in the concentrations of the limiting nutrient in combination with a decrease in productivity is indicative of reduced nutrient input from the thermocline into the mixed layer. This first regime is dominant in the low- and mid-latitude ocean and in the North Atlantic in all three models and in the Arctic for the IPSL model. This regime is for example indicated by the positive slope between productivity (PP and EP) and limiting nutrient (yellow and red color in the panels for PO_4 and Fe in Fig. 5) and the negative slope between productivity and stratification (blue color in the STRAT panel of Fig. 5) in areas where productivity is decreasing. For the second regime, an alleviation of light and temperature limitation leads to an increase in productivity, while productivity is fueled by a sustained or even increased nutrient input into the euphotic zone. This second regime is found in the Arctic in the NCAR and MPIM model and in parts of the Southern Ocean in all three models. Globally, the first regime is most important and global productivity and export production decreases in our 21st century global warming simulations.

3.3 A weighted multi-model mean of projected productivity changes

In the previous sections, it is shown that the models as a class represent most of the pertinent features of the satellite-based productivity estimates and that the underlying mechanisms for changes in productivity are broadly consistent across the range of models. However, individual models clearly fail to represent certain regional features.

The challenge is to combine the information from several models into a quantitative projection. In the assessments of the Intergovernmental Panel on Climate Change this has been achieved by averaging the results from individual models (Meehl et al., 2007). In this way, each model, whether skillful or not, is given equal weight. Obviously, such an approach is less than ideal as unrealistic

features of a particular model influence the multi-model mean. For example, if one of the models simulates rainfall in a desert region, the multi-model mean will also show rainfall in the desert. An alternative would be to rely on the model with the best skill score with respect to suitable observations. However, this seems also less than ideal as each model has certain weaknesses and useful information from the other models is lost. Here, we suggest the use of regional skill scores as weights to compute a 'best' or 'optimal' estimate of projected changes. The goal is to take advantage of the skill of individual models in simulating regional features and to exclude or minimize the influence of regional results where a model is in conflict with observational evidence.

Technically, the multi-model mean is computed following the skill score metric developed by Taylor (2001). For each model m and grid cell at coordinates (i, j) a skill score

$$S_{m,i,j} = \frac{2(1 + R_{i,j})}{(\sigma_{i,j} + 1/\sigma_{i,j})^2}, \tag{5}$$

is calculated (Taylor, 2001), where $R_{i,j}$ is the distance-weighted correlation coefficient between the satellite-based estimates ($\mathrm{PP}_{\mathrm{obs}}$) and the simulated productivity (PP_m; average 1998-2005) and $\sigma_{i,j}$ is the corresponding standard deviation normalized by the standard deviation of the observations. This metric penalizes models that have normalized standard deviations either greater than or less than one by reducing the skill score. The weights are calculated using a two-dimensional Gaussian function

$$w(x, y)_{i,j} = \exp\left\{-\left(\frac{(x - x_{i,j})^2}{2\rho^2} + \frac{(y - y_{i,j})^2}{2\rho^2}\right)\right\} \frac{A(x, y)}{\sum_{x,y} A(x, y)}, \tag{6}$$

where $x_{i,j}$ and $y_{i,j}$ are the longitude and latitude of the grid cell (i, j), $A(x, y)$ is the area of the grid cell at coordinates (x, y), and $\rho = 10^\circ$ characterizes the width of the distribution (the distance at which the weight has decreased from one to $1/\sqrt{e}$). The multi-model mean then is calculated in proportion to these regional skill scores (Fig. 8 a-c):

$$\mathrm{PP}^S_{i,j} = \sum_m \frac{S_{m,i,j}}{\sum_m S_{m,i,j}} \mathrm{PP}_{m,i,j} \tag{7}$$

Where no observation-based data is available to calculate a skill score (e.g. in the Arctic) the model results are averaged using equal weights.

The above skill score metric emphasizes pattern similarities, but does not penalize offsets between the mean of the fields. Therefore, we also investigate an alternative metric, E, based on mean square errors:

$$E_{m,i,j} = \sum_{x,y} w(x, y)_{i,j} (\mathrm{PP}_{\mathrm{obs}}(x, y) - \mathrm{PP}_m(x, y))^2 \tag{8}$$

The weights $w(x, y)_{i,j}$ used here are the same as given above. The multi-model mean with this second metric is calculated as

$$\mathrm{PP}^E_{i,j} = \sum_m \frac{E^{-1}_{m,i,j}}{\sum_m E^{-1}_{m,i,j}} \mathrm{PP}_{m,i,j}. \tag{9}$$

In addition, we have computed the arithmetic mean from all models ($\mathrm{PP}^{\mathrm{ave}}$) as well as the mean obtained by weighting individual models with their global ($\rho = \infty$) skill score ($\mathrm{PP}^{\mathrm{S_{glob}}}$).

Next, global skill scores (S_{glob}) and global root mean square errors (RMSE) are computed for the individual model results and for the multi-model fields obtained by the four different averaging methods (Table 1). The global skill score for the first field (PP^{S}) is considerably higher than for the others. All averaging methods result in a lower global skill score than that of the best model (IPSL). However, the RMSE is lower for the PP^{S} field than for each individual model and for the other multi-model fields. In the following, we discuss results from this metric only. We note that differences in the results obtained by the first two metrics (PP^{S} and PP^{E}) are generally small.

This skill score method accounts for the different skills of the models at reproducing regional features of the satellite based estimates, while not degrading the overall skill in representing the satellite-based field compared to the best individual model. For example, the NCAR model reproduces the high productivity tongue around 40°N in the North Atlantic. The IPSL model captures most of the high productivity features along the coasts of South America and Africa. The MPIM model has a high skill in the central Pacific and the most realistic latitudinal extension of the equatorial productivity belt. Therefore these models dominate the mean in those regions (Fig. 8d), and all these features are present in the multi-model mean (Fig. 9a). There remain weaknesses. All models underestimate productivity in the Arabian Sea and off the west coast of Central and North America. Consequently, the multi-model mean also misses these features. Overall, this method improves the multi-model mean significantly compared to simpler averaging methods (Table 1).

Regional skill scores are applied to calculate the multi-model mean of preindustrial PP and of the projected changes by the end of the 21st century (Fig. 9) and as a function of the global mean surface air temperature ($\mathrm{SAT_{glob}}$, Fig. 10d). The globally integrated annual mean PP decreases from 32.2 GtC yr^{-1} (preindustrial) to 28.7 GtC yr^{-1} by 2100 AD (-3.5 GtC yr^{-1}; -11%) for the multi-model mean (Fig. 2, Table 1). Large decreases in PP are projected for the North Atlantic, off the coast of Africa in the South Atlantic, in the equatorial Pacific, and in the South Pacific around 40°S; a slight increase in PP is found in the Southern Ocean and in the Arctic (Fig. 9b). Calculating the mean by 2100 has the disadvantage that PP changes are merged that correspond to different temperature changes as the models have different climate sensitivities. One way to avoid this is to calculate the regression slope $\Delta\mathrm{PP}/\Delta\mathrm{SAT_{glob}}$ for each grid cell (Fig. 10a-c) as done for the global PP in Fig. 2c. The patterns of the resulting PP change per centigrade SAT increase are broadly consistent with the patterns of the projected PP change by 2100.

3.4 Changes in the seasonal cycle

One aspect of the simulations to explore is how the seasonal cycle and interannual variability are modified under global warming. Here, we compare the simulated maximum seasonal PP amplitudes (annual maximum minus annual minimum) and their interannual variations for the decades 1860-

1869 and 2090-2099 along the two sections in the Atlantic and the Pacific shown in Fig. 1 and for the global zonal mean (Fig. 11).

In the global zonal mean, the seasonal amplitude is projected to decrease everywhere in the IPSL simulation. Largest reductions of about 200 mgC m^{-2} day^{-1} can be found at 40°N-70°N, where the reduction is clearly distinguishable from preindustrial interannual variability. Further, a marked reduction is found at low latitudes around 20°-30° and in the Arctic Ocean north of 80°N. The reduction in the north is linked to a large reduction in productivity in the Atlantic between 30°N and 60°N from April to July and in the Pacific between 55°N and 70°N from April to September. Not only the seasonal amplitude, but also the interannual variability in PP is projected to decrease for most latitudes.

The zonally averaged seasonal PP amplitude in the MPIM simulation is also reduced between 70°N and 60°S. Largest reductions of about 200 mgC m^{-2} day^{-1} are located in the Southern Ocean and around the equator. South of 60°S and north of 70°N the seasonal amplitude increases, consistent with an increase in productivity in these areas. The MPIM model exhibits a larger interannual variability than the other two models, and at most latitudes the projected changes are within the range of preindustrial interannual variability. Maximum changes in productivity occur from December to February in the Southern Ocean and during July/August in the Arctic Ocean.

In the NCAR model the zonally averaged seasonal PP amplitude is reduced by up to 300 mgC m^{-2} day^{-1} between 40°N and 60°N. An increase is found north of 60°N, in the Southern Ocean (40°S-60°S), and in the Arctic Ocean. Changes are small in other regions. The changes in the north are dominated by the Atlantic where PP is strongly reduced between 40°N and 60°N (March-June) and enhanced between 60°N and 70°N (April-June).

In summary, changes in seasonal cycle amplitude are relatively small, though there are exceptions. The seasonal amplitude tends to become smaller when overall productivity decreases. Interannual variability in the seasonal amplitude is substantial and projected to decrease in many areas in all three models.

4 Discussion and conclusions

The trends in ocean productivity in response to anthropogenic climate change have been analyzed with three coupled carbon cycle-climate models that incorporate marine biogeochemical-ecosystem models of different complexity. The decreasing trend in global primary production and particulate organic carbon export is a robust result, but relative and absolute magnitudes differ among models and regions.

The underlying mechanisms of change are qualitatively consistent across the models, except in the Arctic. All three models show a consistent change in physical drivers, surface concentrations of macro-nutrients, and productivity when considering regional averages (Fig. 4). Namely, the

models project an increase in sea surface temperature and stratification in all regions and an increase in available light in the Arctic in response to sea ice retreat. Macro-nutrient concentrations in the euphotic zone are projected to decrease in all regions and for all models. Two different regimes for change in productivity are identified, that were already discussed previously in the literature (Bopp et al., 2001; Sarmiento et al., 1998). First, all models indicate a decrease in productivity in the low- and mid-latitude ocean and in the North Atlantic in response to reduced nutrient delivery to the surface ocean linked to enhanced stratification, reduced mixed-layer depth and slowed ocean circulation. This is broadly consistent with earlier projections using box models, Earth System Models of Intermediate Complexity or general circulation models (Klepper and De Haan, 1995; Maier-Reimer et al., 1996; Joos et al., 1999; Matear and Hirst, 1999; Plattner et al., 2001; Bopp et al., 2001; Fung et al., 2005; Frölicher et al., 2009). Second, light and temperature limitation is reduced in the high-latitude ocean, whereas nutrient supply remains sufficient to support an increase in productivity. This second regime is found in the Arctic in the NCAR and MPIM model and in parts of the Southern Ocean in all three models. A qualitative difference among models is found in the Arctic, where IPSL projects a decrease in productivity related to a reduced supply of macro-nutrients, whereas NCAR and MPIM project a productivity increase due to reduced light and temperature limitation. In any case, absolute changes in productivity in the Arctic and the Southern Ocean are small in all three models. The models project also a different evolution of iron. The MPIM and NCAR models use constant elemental ratios in their production algorithms and consequently surface iron concentration are decreasing in parallel with macro-nutrient concentrations. In the IPSL model, the iron-to-carbon ratio of assimilated material is reduced under nutrient stress. As a consequence, iron concentration increases in the euphotic zone as less iron is exported to depth in the form of organic matter.

Quantitatively, the three models show large differences in regional responses. These are often linked to differences in the simulation of the mean productivity fields. For example, iron limitation is too strong in the MPIM in the low and mid-latitude ocean and in the NCAR model in the equatorial Pacific. Consequently, productivity in these regions is very low for these models and the projected decrease is also small by necessity. Other differences are related to the climate sensitivity of the models. The NCAR model has the smallest climate sensitivity and shows a smaller surface warming and smaller changes in low-latitude stratification than the IPSL and MPIM model. The comparison between observation-based productivity estimates and simulated productivity (Fig. 1; Schneider et al., 2008) suggests that it is not advisable to simply average the results from the three models as obvious shortcomings of the models would unfavorably influence the multi-model mean projection.

We have applied regional model skill metrics as weights in the computation of multi-model means. Here, we have used the satellite-based productivity estimates (average of annual mean PP for the period 1998 to 2005) of Behrenfeld et al. (2006) as a target against which the performance of individual models is assessed. A scale length is introduced for the regional skill score calculation that can be

adjusted for the problem considered. Here, the scale length has been selected to be representative for the spatial scale of marine biogeographical provinces ($\approx 10^{\circ}$); the exact choice of the numerical value is not crucial for our application. The multi-model mean PP changes are expressed as PP change per a nominal increase in global mean surface air temperature of 1°C to account for the different climate sensitivities of the models. The use of regional metrics has advantages. It results in an improved skill in representing the satellite-based productivity field compared to a conventional, IPCC-type multi-model average where each model is given equal weight. Most weight is attached to the model that represents an individual regional feature best, whereas little weight is attached to the models that fail to reproduce the regional feature. The regional metrics quantify the regional performance of each model (Fig. 8). Features that all models fail to represent as evidenced by low skills can be flagged in the multi-model average. Disadvantages are that suitable target fields have to be defined and scale lengths to be determined. The choice of an annual mean climatological field as a target is debatable. Additional targets including seasonal or interannual variability (Santer et al., submitted) may be applied. Most preferable would be observation-based data that include decadal scale trends when evaluating projections of the 21st century. Further, our approach, as any weighting scheme, is based on the assumption that the relative skill of the model remains about the same over time. A more fundamental caveat is worth mentioning. Each individual model provides an internally consistent representation of heat and mass fluxes, nutrient cycling, and ecosystem dynamics taking fully into account first order principles such as mass and energy conservation. By using regional weights, regional features from different models are combined to a new global mean field which may lack internal consistency. We believe that our regional weight approach is preferable compared to the conventional 'one model, one vote' approach to generate a multi-model mean projection of PP. However, we caution that this might not be the case for other applications.

Our results are contradictory to the results of Sarmiento et al. (2004) on the global scale and in most regions (Fig. 12). Sarmiento et al. (2004) project an increase in global primary productivity by 0.7 to 8.1% and not a decrease. These authors rely on an empirical model approach in combination with output for physical variables from AOGCM global warming simulations. The cycling of nutrients and nutrient concentrations are not explicitly considered. Seven physics-based diagnostics (surface temperature, salinity and density, upwelling and vertical density gradient in the top layers, mixed layer depth, and ice cover) are used to define 33 biogeographical provinces. An empirical chlorophyll model, describing chlorophyll as an exponential function of temperature, salinity, mixed layer depth and growing season length, is fitted to the SeaWiFS chlorophyll data for each province and used to project 21st century changes in chlorophyll from the AOGCM output. Finally, primary productivity is estimated from the chlorophyll concentration for three different productivity algorithms (Behrenfeld and Falkowski, 1997a; Carr, 2002; Marra et al., 2003). This chain of models yield an increase in productivity almost in the entire ocean for the Marra et al. algorithm and, to a lesser extent for the Carr algorithm, whereas the Behrenfeld and Falkowski algorithm yields a de-

crease in productivity in low and mid latitudes and an increase in high-latitudes. Only the projected decrease in low and mid latitudes with the Behrenfeld and Falkowski algorithm is consistent with this and an earlier process-based model study (Bopp et al., 2001).

What are the reasons for the discrepancies between results from the empirical approach and those from process-based climate-biogeochemical-ecosystem models? A fundamental conceptual difference is that the cycling of nutrients and nutrient availability is explicitly considered in the process-based models, whereas nutrient limitation is only implicitly included in the empirical approach of Sarmiento et al. (2004) and the satellite productivity algorithms. As nutrients are a key factor for phytoplankton growth and productivity, it appears necessary to take the decadal-to-century scale evolution of nutrient cycling into account as done in the process-based models. As discussed by Sarmiento et al. (2004), projected changes in chlorophyll are small for their empirical approach, and their changes in productivity depend critically on the applied satellite algorithm. Sarmiento et al. (2004) highlight the importance of the assumed relationship between temperature and productivity for a given chlorophyll concentration. This temperature sensitivity of productivity is very different among the satellite algorithms. For example, productivity increases with temperature by a factor of about two between 18°C and 30°C for the Marra et al. algorithm, but decreases by a factor of two over the same temperature range for the Behrenfeld and Falkowski algorithm. Consequently, productivity is projected to decrease in low and mid latitudes with the Behrenfeld and Falkowski algorithm and to increase with the Marra et al. algorithm in transient warming scenarios. These discrepancies between algorithms may reflect the difficulties to separate light and nutrient effects on productivity (Behrenfeld et al., 2008). We note that observation-based changes in global chlorophyll and inferred global productivity by Behrenfeld et al. (2006) evolve in parallel. An implicit assumption in the empirical approach is that the spatial relationship between productivity and physical forcing found for the modern ocean can be applied to temporal changes into the future. However, Schneider et al. (2008) find that the relationship between productivity and temperature in the low-latitude ocean is different for interannual variations of the last decades and the century-scale trends in transient warming simulations.

Process-based models are far from perfect (Schneider et al., 2008) and their results must be interpreted with some caution. However, it appears evident from our analysis that the cycling of nutrients and changes in the supply to the surface and in the concentration of nutrients must be realistically represented to project changes in productivity with some realism. What is required for further progress is to combine satellite, field, and laboratory observations, empirical approaches and process-based models to further improve our quantitative understanding. Novel metrics such as (multivariate) regional skill scores may prove useful to synthesize results from models and observational studies in a quantitative and transparent way. As far as modeling is concerned, factorial experiments dedicated to quantify the link between PP and individual parameters will be helpful to improve the understanding of model behavior and to compare model results with experimental data. Improved parametrizations

of ecosystem processes that take into account emerging results from field and laboratory studies are required to close gaps in understanding.

Acknowledgements. We would like to thank J. Sarmiento and R. Slater for providing us with the data from their study. We also thank the SeaWiFS Project and M. Behrenfeld for providing data from satellite observations. Furthermore, we thank J. Sarmiento and M. Behrenfeld for their valuable comments that helped to improve the manuscript and K. Taylor for sharing his expertise on model skill metrics. This work was funded by the European Union projects CARBOOCEAN (511176-2) and EUROCEANS (511106-2) and is a contribution to the "European Project on Ocean Acidification" (EPOCA) which received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 211384. Additional support was received from the Swiss National Science Foundation. SCD acknowledges support from the NASA Ocean Biology and Biogeochemistry Program (NNX07AL80G).

References

Anderson, L. A. and Sarmiento, J. L.: Redfield ratios of remineralization determined by nutrient data analysis, Global Biogeochem. Cy., 8, 65–80, 1994.

Aumont, O. and Bopp, L.: Globalizing results from ocean in situ iron fertilization studies, Global Biogeochem. Cy., 20, 2006.

Aumont, O., Maier-Reimer, E., Blain, S., and Monfray, P.: An ecosystem model of the global ocean including Fe, Si, P colimitations, Global Biogeochem. Cy., 17, 1060, doi:10.1029/2001GB001745, 2003.

Behrenfeld, M. J. and Falkowski, P. G.: A consumer's guide to phytoplankton primary productivity models, Limnol. Oceanogr., 42, 1479–1491, 1997a.

Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, Limnol. Oceanogr., 42, 1–20, 1997b.

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., Milligan, A. J., Falkowski, P. G., Letelier, R. M., and Boss, E. S.: Climate-driven trends in contemporary ocean productivity, Nature, 444, 752–755, 2006.

Behrenfeld, M. J., Halsey, K. H., and Milligan, A. J.: Evolved physiological responses of phytoplankton to their integrated growth environment, Philos. T. R. Soc. B, 363, 2687–2703, doi:10.1098/rstb.2008.0019, 2008.

Bopp, L., Monfray, P., Aumont, O., Dufresne, J. L., Le Treut, H., Madec, G., Terray, L., and Orr, J. C.: Potential impact of climate change on marine export production, Global Biogeochem. Cy., 15, 81–99, 2001.

Bopp, L., Aumont, O., Belviso, S., and Monfray, P.: Potential impact of climate change on marine dimethyl sulfide emissions, Tellus B, 55, 11–22, 2003.

Bopp, L., Aumont, O., Cadule, P., Alvain, S., and Gehlen, M.: Response of diatoms distribution to global warming and potential implications: a global model study, Geophys. Res. Lett., 32, 4 pp., 2005.

Boville, B. A. and Gent, P. R.: The NCAR Climate System Model, version one, J. Climate, 11, 1115–1130, 1998.

Boville, B. A., Kiehl, J. T., Rasch, P. J., and Bryan, F. O.: Improvements to the NCAR CSM-1 for transient climate simulations, J. Climate, 14, 164–179, doi:10.1029/2002JD003026, 2001.

Boyd, P. W. and Doney, S. C.: Modelling regional responses by marine pelagic ecosystems to global climate change, Geophys. Res. Lett., 29, doi:10.1029/2001GL014130, 2002.

Carr, M.-E.: Estimation of potential productivity in Eastern Boundary Currents using remote sensing, Deep-Sea Res. Pt. II, 49, 59–80, 2002.

Carr, M.-E., Friedrichs, M. A. M., Schmeltz, M., Aita, M. N., Antoine, D., Arrigo, K. R., Asanuma, I., Aumont, O., Barber, R., Behrenfeld, M., Bidigare, R., Buitenhuis, E. T., Campbell, J., Ciotti, A., Dierssen, H., Dowell, M., Dunne, J., Esaias, W., Gentili, B., Gregg, W., Groom, S., Hoepffner, N., Ishizaka, J., Kameda, T., Le Quéré, C., Lohrenz, S., Marra, J., Melin, F., Moore, K., Morel, A., Reddy, T. E., Ryan, J., Scardi, M., Smyth, T., Turpie, K., Tilstone, G., Waters, K., and Yamanaka, Y.: A comparison of global estimates of marine primary production from ocean color, Deep-Sea Res. Pt. II, 53, 741–770, doi:10.1016/j.dsr2.2006.01.028, 2006.

Crueger, T., Roeckner, E., Raddatz, T., Schnur, R., and Wetzel, P.: Ocean dynamics determine the response of oceanic CO_2 uptake to climate change, Clim. Dynam., 31, 151–168, doi:10.1007/s00382-007-0342-x, 2008.

Doney, S., Lindsay, K., Fung, I., and John, J.: Natural variability in a stable, 1000-yr global coupled climate-

carbon cycle simulation, J. Climate, 19, 3033–3054, 2006.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate-Carbon Cycle Feedback Analysis: Results from the C^4MIP Model Intercomparison: Evolution of carbon sinks in a changing climate, J. Climate, 19, 3337–3353, 2006.

Frölicher, T. L., Joos, F., Plattner, G. K., Steinacher, M., and Doney, S. C.: Natural variability and anthropogenic trends in oceanic oxygen in a coupled carbon cycle-climate model ensemble, Global Biogeochem. Cy., 23, 1–15, doi:10.1.1029/2008GB003316, 2009.

Fung, I., Doney, S., Lindsay, K., and John, J.: Evolution of carbon sinks in a changing climate, P. Natl. Acad. Sci. USA, 102, 11 201–11 206, 2005.

Gehlen, M., Bopp, L., Ernprin, N., Aumont, O., Heinze, C., and Raguencau, O.: Reconciling surface ocean productivity, export fluxes and sediment composition in a global biogeochemical ocean model, Biogeosciences, 3, 521–537, 2006.

Gent, P. R., Bryan, F. O., Danabasoglu, G., Doney, S. C., Holland, W. R., Large, W. G., and McWilliams, J. C.: The NCAR Climate System Model global ocean component, J. Climate, 11, 1287–1306, 1998.

Goldstein, B., Joos, F., and Stocker, T. F.: A modeling study of oceanic nitrous oxide during the Younger Dryas cold period, Geophys. Res. Lett., 30, doi:10.1029/2002GL016418, 2003.

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., Levan, P., Li, Z.-X., and Lott, F.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, Clim. Dynam., 19, 3445–3482, 2006.

Joos, F., Plattner, G. K., Stocker, T. F., Marchal, O., and Schmittner, A.: Global warming and marine carbon cycle feedbacks and future atmospheric CO_2, Science, 284, 464–467, 1999.

Kiehl, J. T., Hack, J. J., Bonan, G. B., Boville, B. A., Williamson, D. L., and Rasch, P. J.: The National Center for Atmospheric Research Community Climate Model, J. Climate, 11, 1151–1178, 1998.

Klepper, O. and De Haan, B. J.: A sensitivity study of the effect of global change on ocean carbon uptake, Tellus B, 47, 490–500, 1995.

Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cy., 19, 2005.

Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Da Cunha, L. C., Geider, R., Giraud, X., Klaas, C., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J., and Wolf-Gladrow, D.: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, Glob. Change Biol., 11, 2016–2040, doi:10.1111/j.1365-2468.2005.01004.x, 2005.

Madec, G., Delecluse, P., Imbard, M., and Levy, C.: OPA 8.1 ocean general circulation model reference manual, Notes du Pôle de Modélisation 11, Institut Pierre Simon Laplace des Sciences de l'Environment Global, Paris, France, 1998.

Maier-Reimer, E.: Geochemical cycles in an ocean general circultion model. Preindustrial tracer distributions,

Global Biogeochem. Cy., 7, 645–677, 1993.

Maier-Reimer, E., Mikolajewicz, U., and Winguth, A.: Future ocean uptake of CO_2: interaction between ocean circulation and biology, Clim. Dynam., 12, 63–90, 1996.

Maier-Reimer, E., Kriest, I., Segschneider, J., and Wetzel, P.: The HAMburg Ocean Carbon Cycle model HAMOCC5.1, Berichte zur Erdsystemforschung 14/2005, Max Planck-Institut für Meteorologie, Hamburg, Germany, 2005.

Manizza, M., Le Quéré, C., Watson, A. J., and Buitenhuis, E. T.: Ocean biogeochemical response to phytoplankton-light feedback in a global model, J. Geophys. Res.-Oceans, 113, doi:10.1029/2007JC004478, 2008.

Marra, J., Ho, C., and Trees, C. C.: An alternative algorithm for the calculation of primary productivity from remote sensing data, Tech. rep., Lamont-Doherty Earth Obs., Palisades, N.Y., USA, 2003.

Marsland, S. J., Haak, H., Jungclaus, J. H., Latif, M., and Roske, F.: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates, Ocean Model., 5, 91–127, 2003.

Matear, R. J. and Hirst, A. C.: Climate change feedback on the future oceanic CO_2 uptake, Tellus B, 51, 722–733, 1999.

Meehl, G. A., Stocker, T. F., Collins, W. D., Friedlingstein, P., Gaye, A. T., Gregory, J. M., Kitoh, A., Knutti, R., Murphy, J. M., Noda, A., Raper, S. C. B., Watterson, I. G., Weaver, A. J., and Zhao, Z.-C.: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, chap. Global Climate Projections, pp. 747–846, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Moore, J. K., Doney, S. C., Kleypas, J. A., Glover, D. M., and Fung, I. Y.: An intermediate complexity marine ecosystem model for the global domain, Deep-Sea Res. Pt. II, 49, 403–462, doi:10.1016/S0967-0645(01)00108-4, 2002.

Moore, J. K., Doney, S. C., and Lindsay, K.: Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model, Global Biogeochem. Cy., 18, doi:10.1029/2004GB002220, 2004.

Najjar, R. G., Jin, X., Louanchi, F., Aumont, O., Caldeira, K., Doney, S. C., Dutay, J.-C., Follows, M., Gruber, N., Joos, F., Lindsay, K., Maier-Reimer, E., Matear, R. J., Matsumoto, K., Mouchet, A., Orr, J. C., Sarmiento, J. L., Schlitzer, R., Weirig, M. F., Yamanaka, Y., and Yool, A.: Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from Phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2), Global Biogeochem. Cy., 21, doi:10.1029/2006GB002857, 2007.

Oschlies, A., Schulz, K. G., Riebesell, U., and Schmittner, A.: Simulated 21st century's increase in oceanic suboxia by CO_2-enhanced biotic carbon export, Global Biogeochem. Cy., 22, doi:10.1029/2007GB003147, 2008.

Plattner, G.-K., Joos, F., Stocker, T. F., and Marchal, O.: Feedback mechanisms and sensitivities of ocean carbon uptake under global warming, Tellus B, 53, 564–592, 2001.

Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y., and Field, C. B.: The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, Global Biogeochem. Cy., 11, 535–560, 1997.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U.,

and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, J. Climate, 19, 3771–3791, 2006.

Santer, B. D., Taylor, K. E., Gleckler, P. J., Bonfils, C., Barnett, T. P., Pierce, D. W., Wigley, T. M. L., Mears, C., Wentz, F. J., Brüggemann, W., Gillett, N. P., Klein, S. A., Solomon, S., Stott, P. A., and Wehner, M. F.: Incorporating Model Quality Information in Climate Change Detection and Attribution Studies, P. Natl. Acad. Sci. USA, submitted.

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

Sarmiento, J. L., Slater, R., Barber, R., Bopp, L., Doney, S. C., Hirst, A. C., Kleypas, J., Matear, R., Mikolajewicz, U., Monfray, P., Soldatov, V., Spall, S. A., and Stouffer, R.: Response of ocean ecosystems to climate warming, Global Biogeochem. Cy., 18, doi:10.1029/2003GB002134, 2004.

Schmittner, A. and Galbraith, E. D.: Glacial greenhouse-gas fluctuations controlled by ocean circulation changes, Nature, 456, 373–376, doi:10.1038/nature07531, 2008.

Schmittner, A., Oschlies, A., Matthews, H. D., and Galbraith, E. D.: Future changes in climate, ocean circulation, ecosystems, and biogeochemical cycling simulated for a business-as-usual CO_2 emission scenario until year 4000 AD, Global Biogeochem. Cy., 22, doi:10.1029/2007GB002953, 2008.

Schneider, B., Bopp, L., Gehlen, M., Segschneider, J., Frölicher, T. L., Cadule, P., Friedlingstein, P., Doney, S. C., Behrenfeld, M. J., and Joos, F.: Climate-induced interannual variability of marine primary and export production in three global coupled climate carbon cycle models, Biogeosciences, 5, 597–614, 2008.

Siegenthaler, U. and Wenk, T.: Rapid atmospheric CO_2 variations and ocean circulation, Nature, 308, 624–626, 1984.

Six, K. D. and Maier-Reimer, E.: Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model, Global Biogeochem. Cy., 10, 559–583, 1996.

Solomon, S., Qin, D., Manning, M., and et al.: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, chap. Technical Summary, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Steinacher, M., Joos, F., Froelicher, T. L., Plattner, G.-K., and Doney, S. C.: Imminent ocean acidification in the Arctic projected with the NCAR global coupled carbon cycle-climate model, Biogeosciences, 6, 515–533, 2009.

Suntharalingam, P. and Sarmiento, J. L.: Factors governing the oceanic nitrous oxide distribution: Simulations with an ocean general circulation model, Global Biogeochem. Cy., 14, 429–454, 2000.

Takahashi, T., Broecker, W. S., and Langer, S.: Redfield ratio based on chemical-data from isopycnal surfaces, J. Geophys. Res.-Oceans, 90, 6907–6924, 1985.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram., J. Geophys. Res.-Atmos., 106, 7183–7192, 2001.

Timmermann, A. and Jin, F. F.: Phytoplankton influences on tropical climate, Geophys. Res. Lett., 29, doi:10.1029/2002GL015434, 2002.

Vichi, M., Pinardi, N., and Masina, S.: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: Theory, J. Marine Syst., 64, 89–109, doi:10.1016/j.jmarsys.2006.03.006, 2007.

Volk, T. and Hoffert, M. I.: The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present, chap. Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2 changes., pp. 99–110, Geophys. Monogr. Ser. 32, AGU, Washington, D.C., USA, 1985.

Weatherly, J. W., Briegleb, B. P., Large, W. G., and Maslanik, J. A.: Sea Ice and Polar Climate in the NCAR CSM, J. Climate, 11, 1472–1486, 1998.

Wetzel, P., Maier-Reimer, E., Botzet, M., Jungclaus, J., Keenlyside, N., and Latif, M.: Effects of ocean biology on the penetrative radiation in a coupled climate model, J. Climate, 19, 3973–3987, 2006.

Table 1. Simulated global annual primary production (PP) and POC export (EP) for the three models IPSL, MPIM and NCAR under SRES A2. PP values are also given for weighted means of the three models derived from regional skill scores (PP^S), mean square errors (PP^E), and global skill scores ($PP^{S_{glob}}$), as well as for the arithmetic average (PP^{ave}). Global skill scores (S_{glob}) and root mean square errors (RMSE) indicate the ability of the individual models and the multi-model means to reproduce the satellite-based estimates of PP (average 1998-2005, see main text for details). Values are averaged over the periods 1860-1869 (1865), 1985-2004 (2000), and 2090-2099 (2095). ΔPP and ΔEP indicate changes between corresponding periods.

Primary production	*IPSL*	*MPIM*	*NCAR*	PP^S	PP^E	$PP^{S_{glob}}$	PP^{ave}
PP_{glob} 1865 [GtC yr^{-1}]	34.9	23.9	27.5	32.2	31.6	30.2	28.5
PP_{glob} 2000 [GtC yr^{-1}]	33.8	23.7	26.6	31.2	30.8	29.3	27.8
PP_{glob} 2095 [GtC yr^{-1}]	30.3	21.6	25.6	28.7	28.1	26.9	25.6
ΔPP_{glob} 1865-2000 [GtC yr^{-1}]	-1.1 (-3%)	-0.2 (-1%)	-0.9 (-3%)	-1.0 (-3%)	-0.8 (-3%)	-0.9 (-3%)	-0.7 (-2%)
ΔPP_{glob} 2000-2095 [GtC yr^{-1}]	-3.5 (-10%)	-2.1 (-9%)	-1.0 (-4%)	-2.5 (-8%)	-2.7 (-9%)	-2.4 (-8%)	-2.2 (-8%)
ΔPP_{glob} 1865-2095 [GtC yr^{-1}]	-4.6 (-13%)	-2.3 (-10%)	-1.9 (-7%)	-3.5 (-11%)	-3.5 (-11%)	-3.3 (-11%)	-2.9 (-10%)
S_{glob}	0.49	0.16	0.37	0.39	0.31	0.29	0.23
RMSE [mgC m^{-2} day^{-1}]	284	353	334	275	282	286	298
POC export	*IPSL*	*MPIM*	*NCAR*				
EP_{glob} 1865 [GtC yr^{-1}]	9.1	5.0	9.1				
EP_{glob} 2000 [GtC yr^{-1}]	8.7	5.0	8.8				
EP_{glob} 2095 [GtC yr^{-1}]	7.3	4.5	8.4				
ΔEP_{glob} 1865-2000 [GtC yr^{-1}]	-0.4 (-4%)	0.0 (0%)	-0.3 (-3%)				
ΔEP_{glob} 2000-2095 [GtC yr^{-1}]	-1.4 (-16%)	-0.5 (-10%)	-0.4 (-5%)				
ΔEP_{glob} 1865-2095 [GtC yr^{-1}]	-1.8 (-20%)	-0.5 (-10%)	-0.7 (-8%)				

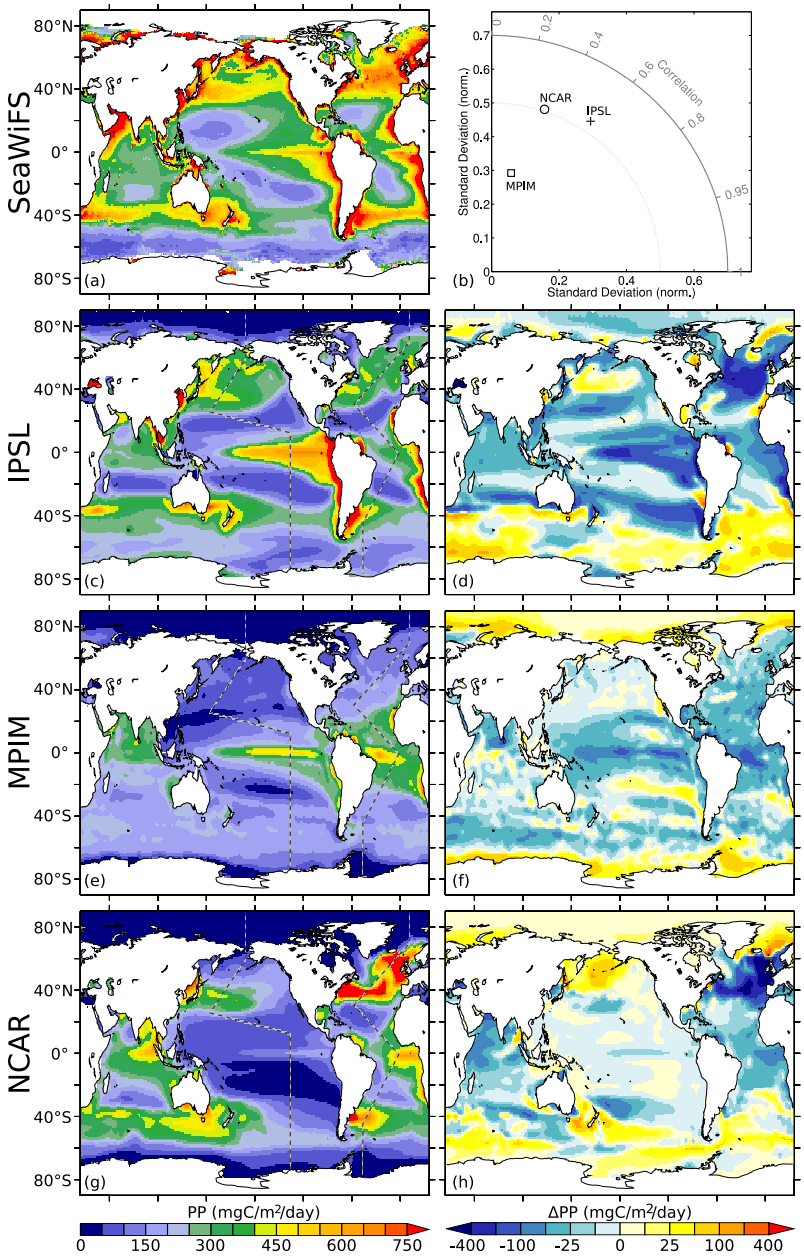

Fig. 1. Productivity (left) and projected changes by 2090-2099 (right). Vertically integrated annual mean primary production (PP, $\frac{\mathrm{mgC}}{\mathrm{m}^2\ \mathrm{day}}$) derived from ocean color **(a)** (SeaWiFS; Behrenfeld et al., 2006; Behrenfeld and Falkowski, 1997b) and simulated by IPSL **(c)**, MPIM **(e)**, and NCAR **(g)** under preindustrial conditions (decadal mean 1860-1869). Dashed lines indicate the transects through the Atlantic and Pacific analyzed in this study. The Taylor diagram **(b)** shows the correspondence between model results and the satellite-based estimates (Taylor, 2001). In this diagram the polar coordinates represent the correlation coefficient R (polar angle) and the normalized standard deviation $\sigma_{\mathrm{model}}/\sigma_{\mathrm{obs}}$ (radius). Panels **d**, **f**, and **h** show the projected changes by the end of the 21st century under SRES A2 for the three models. The changes are shown on an exponential scale and represent the difference between 2090-2099 and 1860-1869 (decadal means).

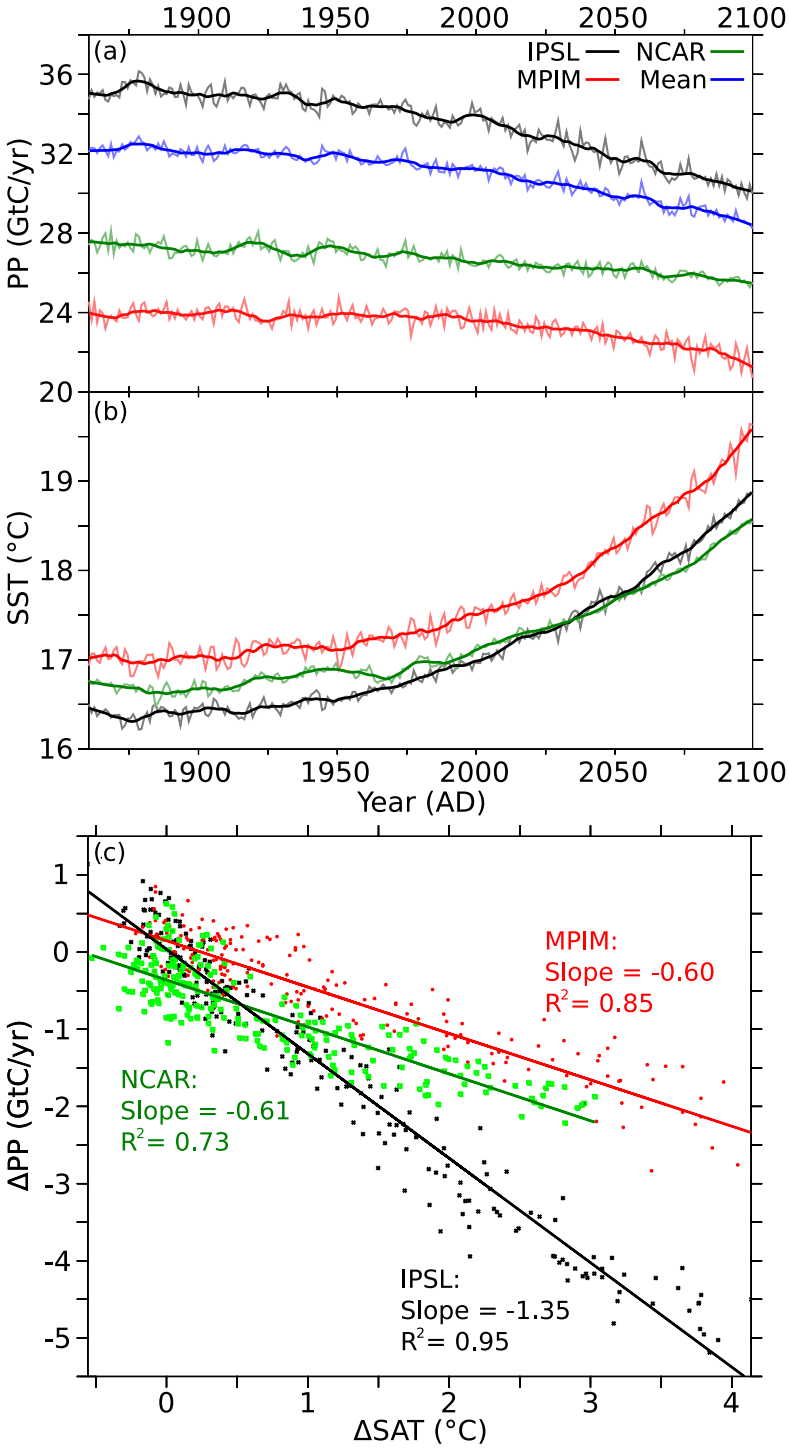

Fig. 2. **(a)** Global annual mean primary production (PP, $\frac{\mathrm{GtC}}{\mathrm{yr}}$) and **(b)** sea surface temperature (SST, °C) simulated by the IPSL (black), MPIM (red), and the NCAR (green) models for the period 1860-2100 under SRES A2. The blue curve indicates the weighted mean PP derived from the regional skill scores of the three models. **(c)** ΔPP as a function of changes in global mean surface air temperature (SAT) for the same models and time period.

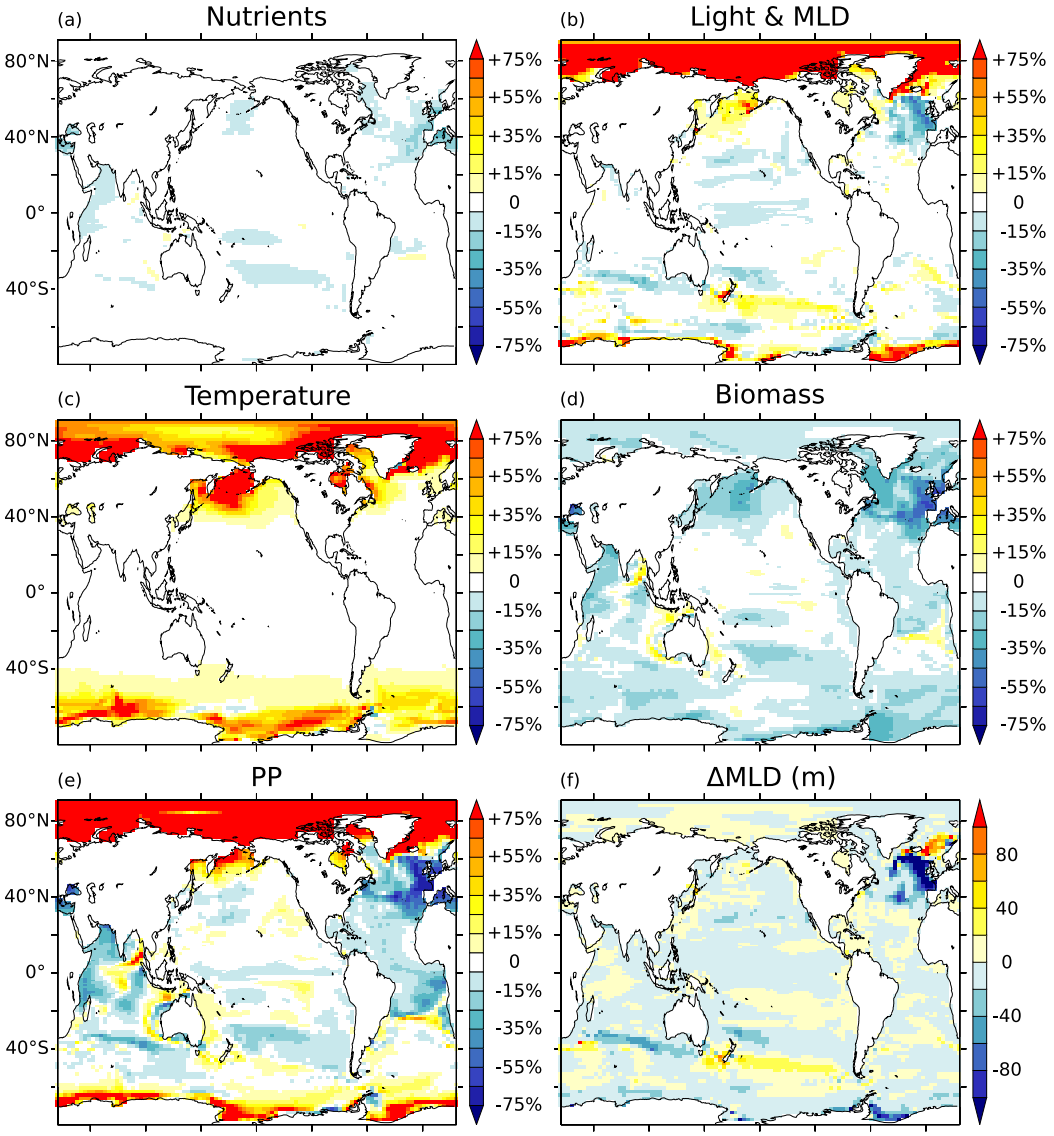

Fig. 3. Long-term changes in PP limitation by nutrients **(a)**, light **(b)**, and temperature **(c)** simulated by the NCAR model. In the NCAR model, these factors, together with changes in the biomass proxy **(d)**, determine the changes in PP **(e)**. Panels (a-e) show relative changes in percent from preindustrial (average 1860-1869) to projected future conditions under SRES A2 (average 2090-2099). Positive values indicate changes that enhance PP, negative values indicate changes that tend to reduce PP. All values are averaged over the compensation depth (75 m), where all of the production is restricted to occur. The light limitation factor (b) also accounts for changes in mixed layer depth **(f)**.

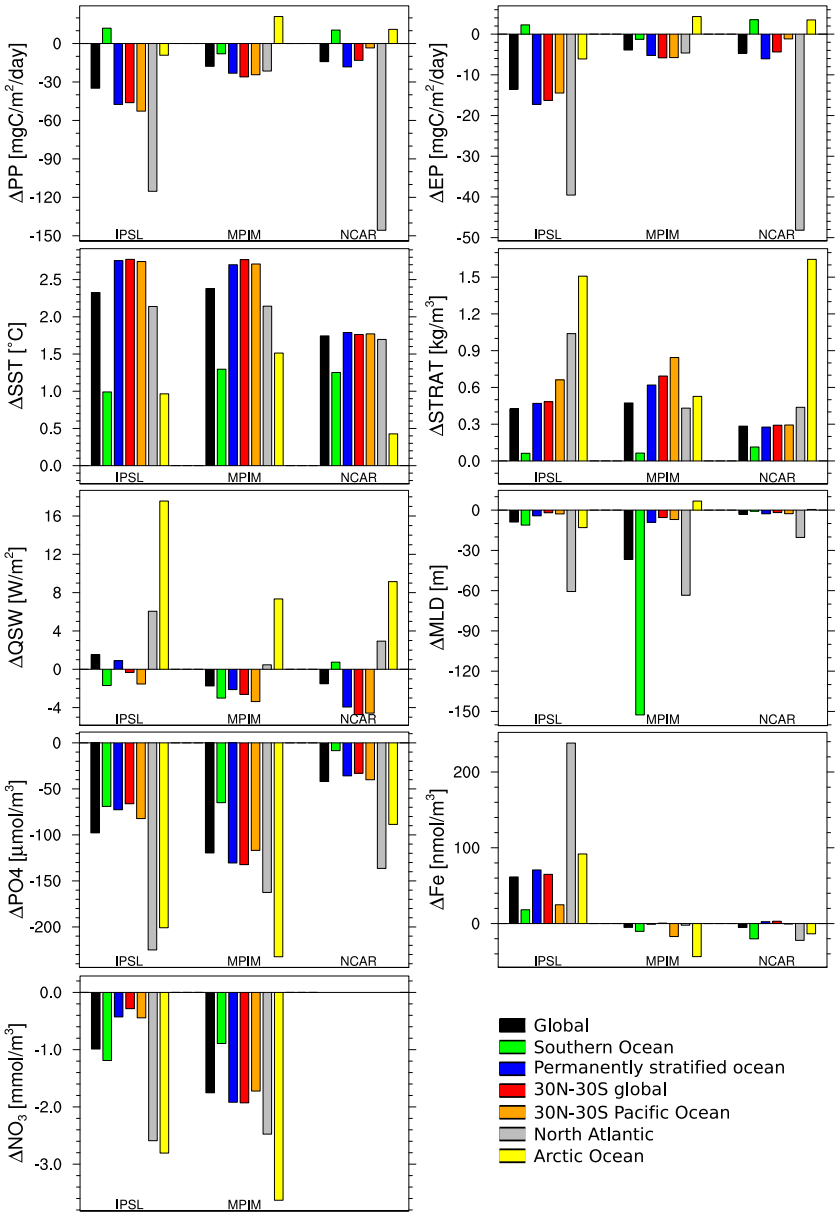

Fig. 4. Long-term trends of PP, EP and related properties simulated by the IPSL, MPIM, and NCAR models under SRES A2. Each panel shows the projected changes of one property with respect to preindustrial conditions (average 2090-2099 minus average 1860-1869) for the following regions: Global oceans (black), Southern Ocean (green; south of 45°S), permanently stratified, low-latitude oceans (blue; annual mean SST $>$ 15°C), low-latitude oceans (red; 30°N-30°S), low-latitude Pacific (black, 30°N-30°S), North Atlantic (gray; 30°N-80°N), and Arctic Ocean (yellow). The properties are vertically integrated PP, POC export (EP), surface temperature (SST, averaged over top 75 m), stratification (STRAT), short wave heat flux (QSW) at the surface, mixed layer depth (MLD), and surface nutrient concentrations (PO_4, Fe, NO_3; averaged over top 75 m). NO_3 is not available for the NCAR model.

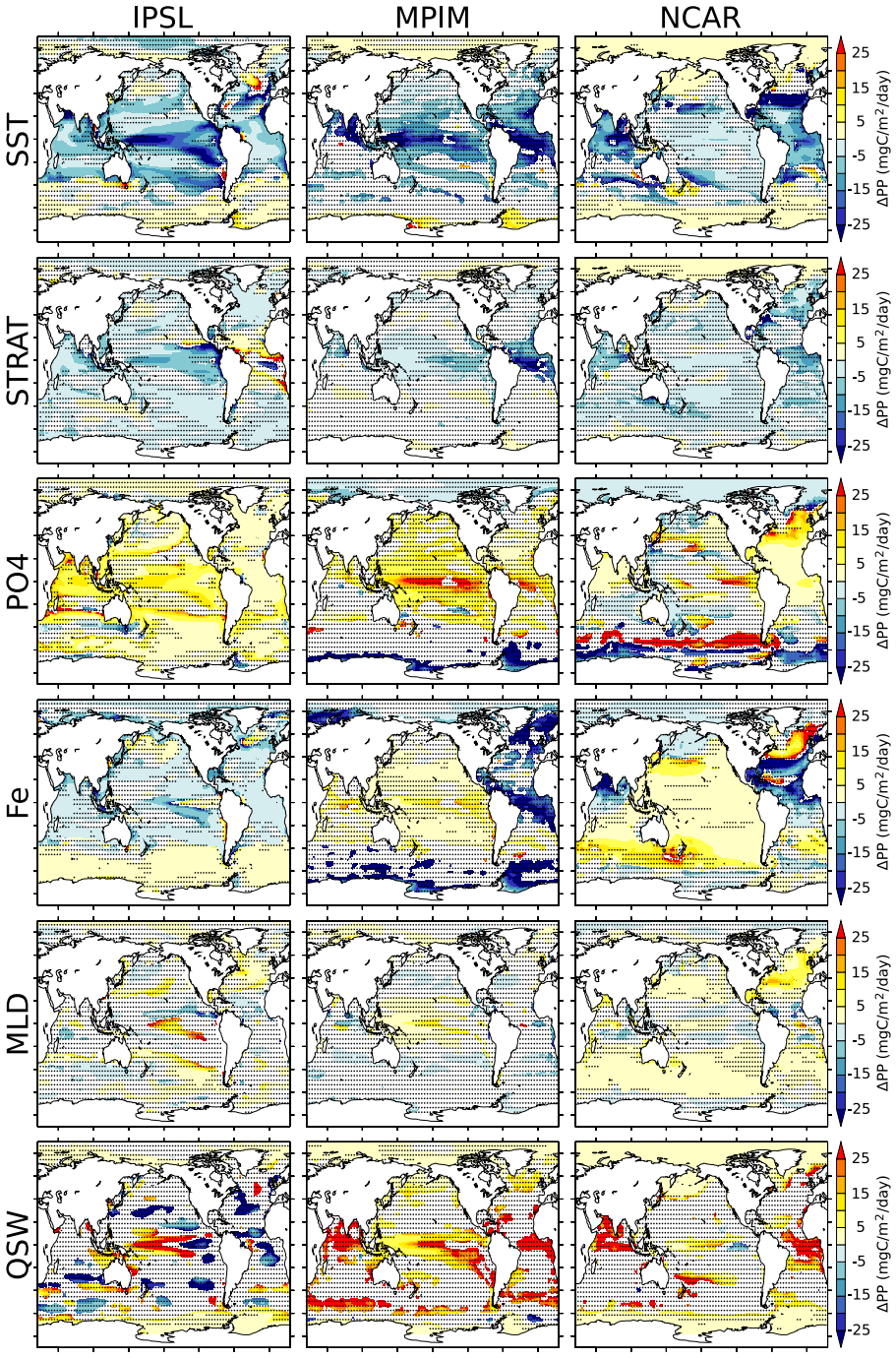

Fig. 5. Linear correlation between vertically integrated annual mean PP and surface temperature (SST), stratification index (STRAT), PO_4, Fe, mixed layer depth (MLD), and light (QSW) for IPSL (left), MPIM (middle), and NCAR (right). The regression has been calculated for simulated annual mean values in each grid cell from 1860 to 2100 (SRES A2). SST, PO_4 and Fe are averaged over top 75 m depth. Normalized regression slopes (ΔPP $[\frac{\mathrm{mgC}}{\mathrm{m^2\ day}}]$ per relative change of SST, STRAT, PO4, Fe, MLD, and QSW in percent) are shown where $R^2 > 0.1$. Areas where $R^2 < 0.5$ are shaded.

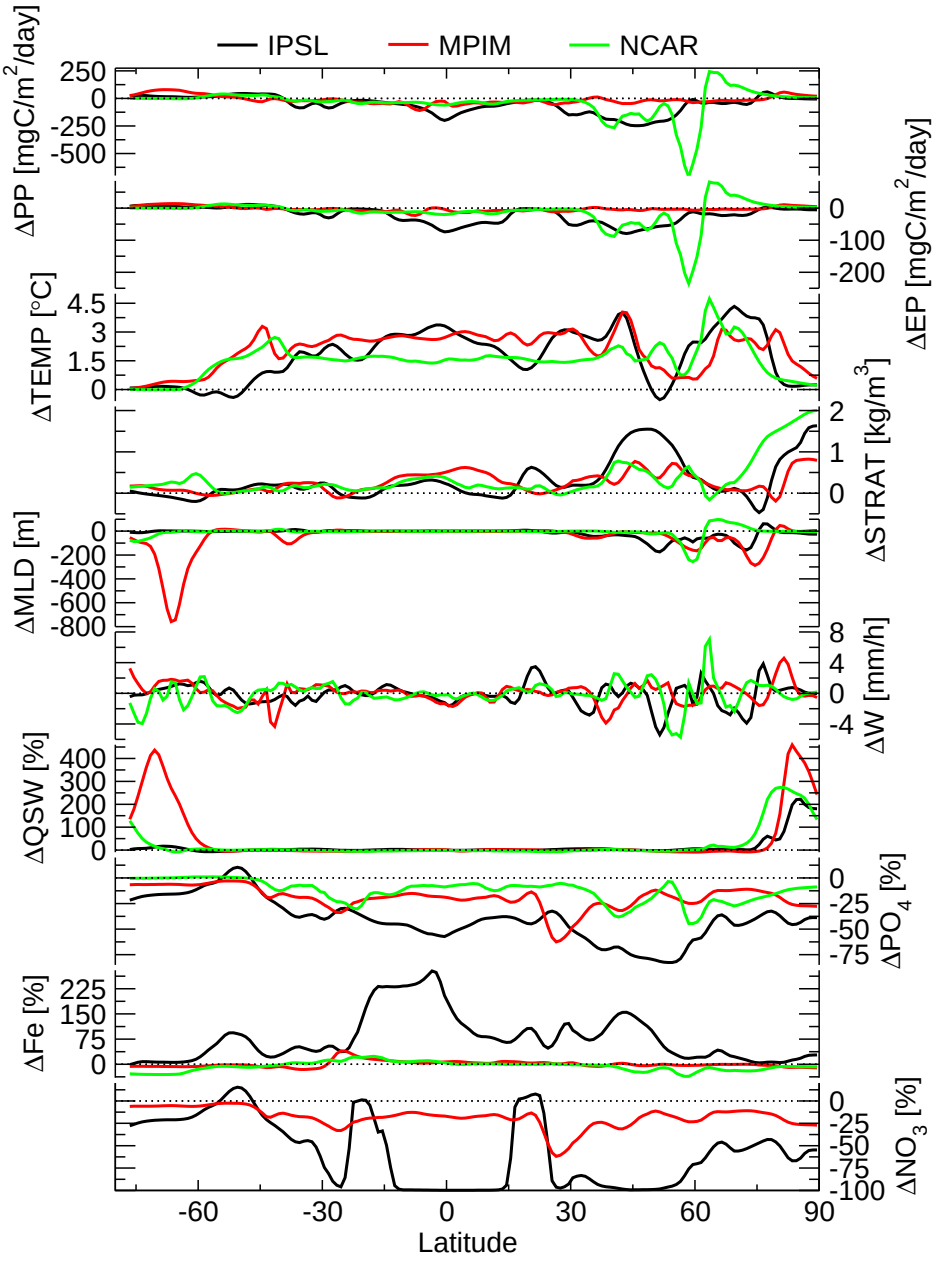

Fig. 6. Long-term trends of PP, EP and related properties simulated by the IPSL (black), MPIM (red), and NCAR (green) models under SRES A2 for a transect through the Atlantic that covers major productivity features (dashed lines in Fig. 1c,e,g). Changes in vertical velocity (ΔW) are shown in addition to the parameters displayed in Fig. 4. The large relative changes in NO_3 projected by the IPSL model at low and mid latitudes result from small absolute changes at locations where NO_3 is almost depleted.

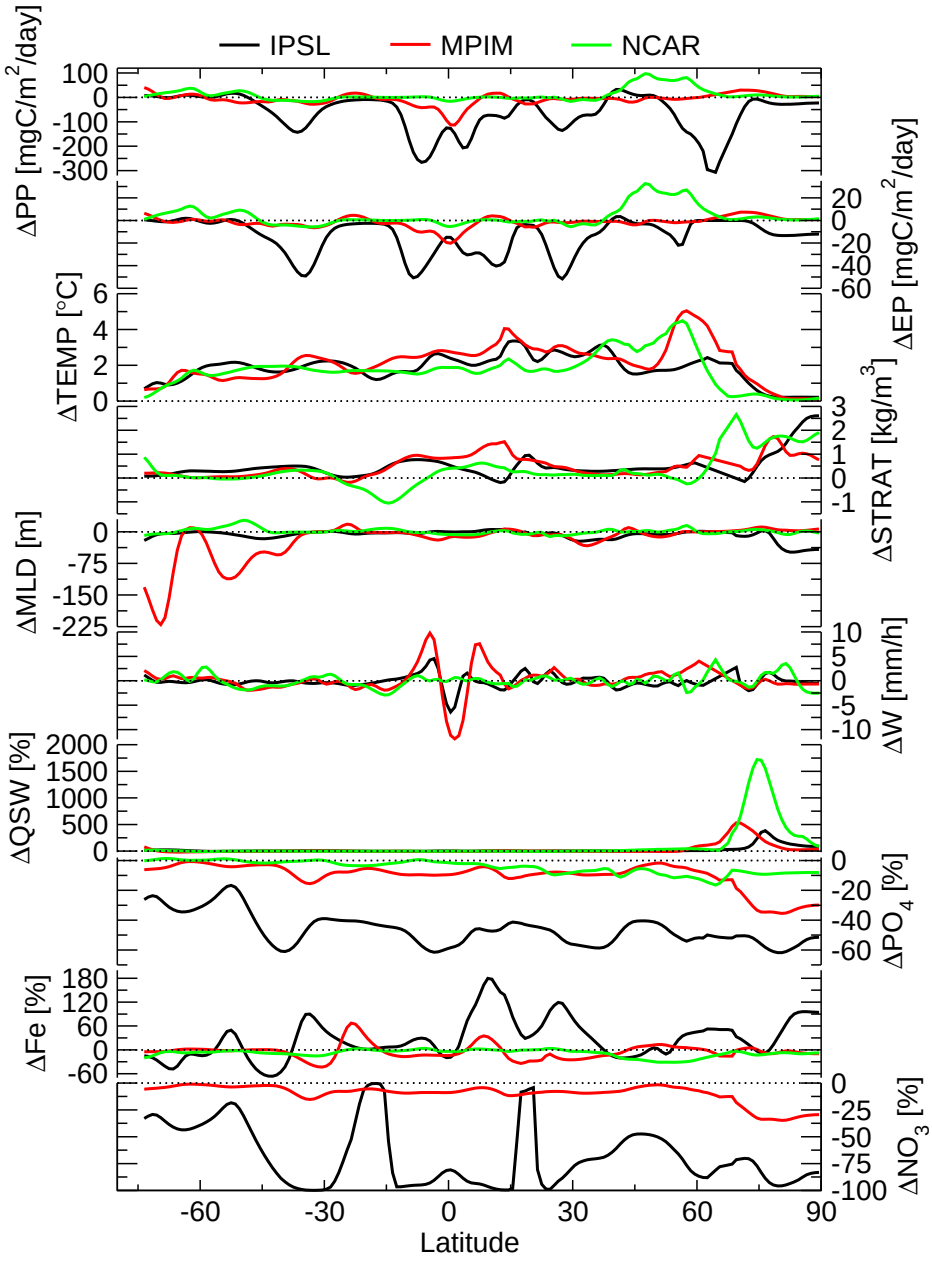

Fig. 7. Same as Fig. 6 but for a transect through the Pacific (dashed lines in Fig. 1c,e,g).

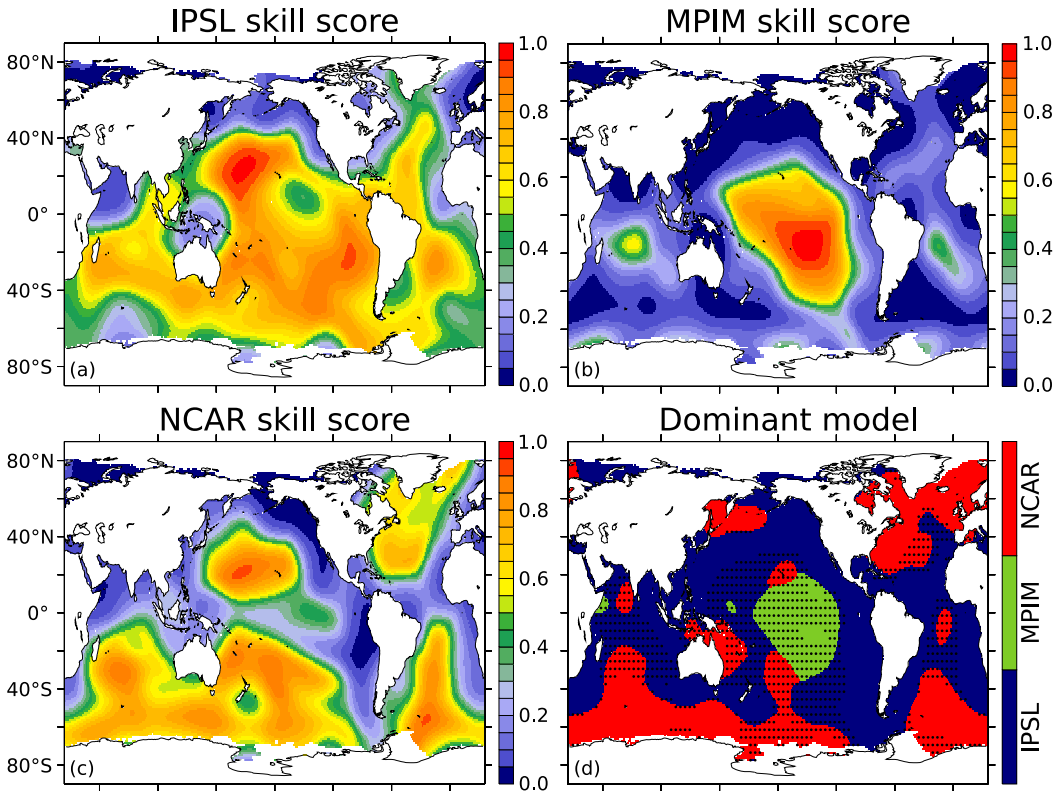

Fig. 8. Regional skill scores showing the ability of the IPSL **(a)**, MPIM **(b)**, and NCAR **(c)** models to reproduce the satellite-based estimates of PP. Panel **(d)** shows which model has the highest skill score at a specific point and therefore dominates the skill-score weighted multi-model mean shown in Fig. 9. The dotted areas indicate regions where the contribution of the model with the highest skill score to the multi-model mean is less than 50%.

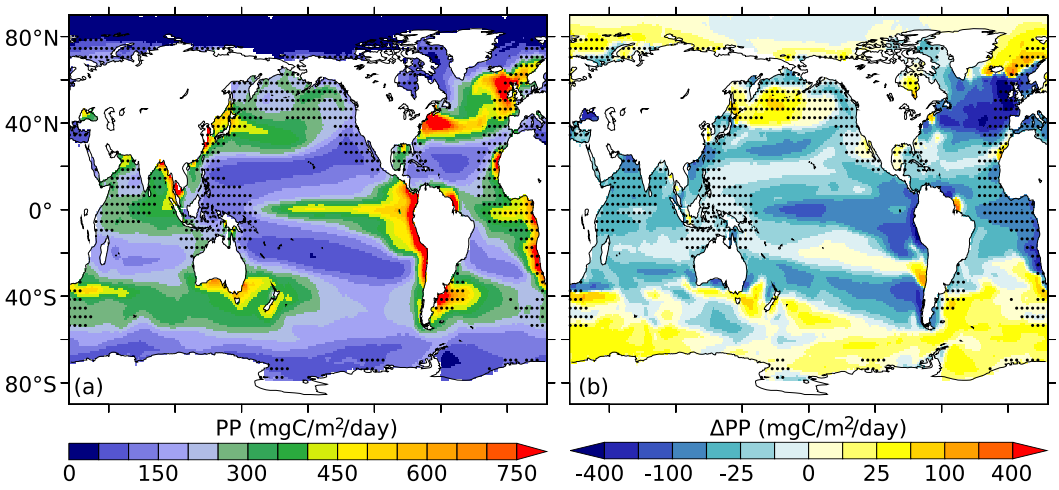

Fig. 9. (a) Multi-model mean of vertically integrated annual mean PP under preindustrial conditions (decadal mean 1860-1869) and **(b)** projected changes by the end of the 21st century under SRES A2. The changes are shown on an exponential scale and represent the difference between 2090-2099 and 1860-1869 (decadal means). The multi-model means have been computed by using the regional skill scores shown in Fig. 8 as weights. The dotted areas indicate that none of the regional skill scores is higher than 0.5. Where no observation-based data is available to calculate skill scores (e.g. in the Arctic) the arithmetic mean of the model results is shown.

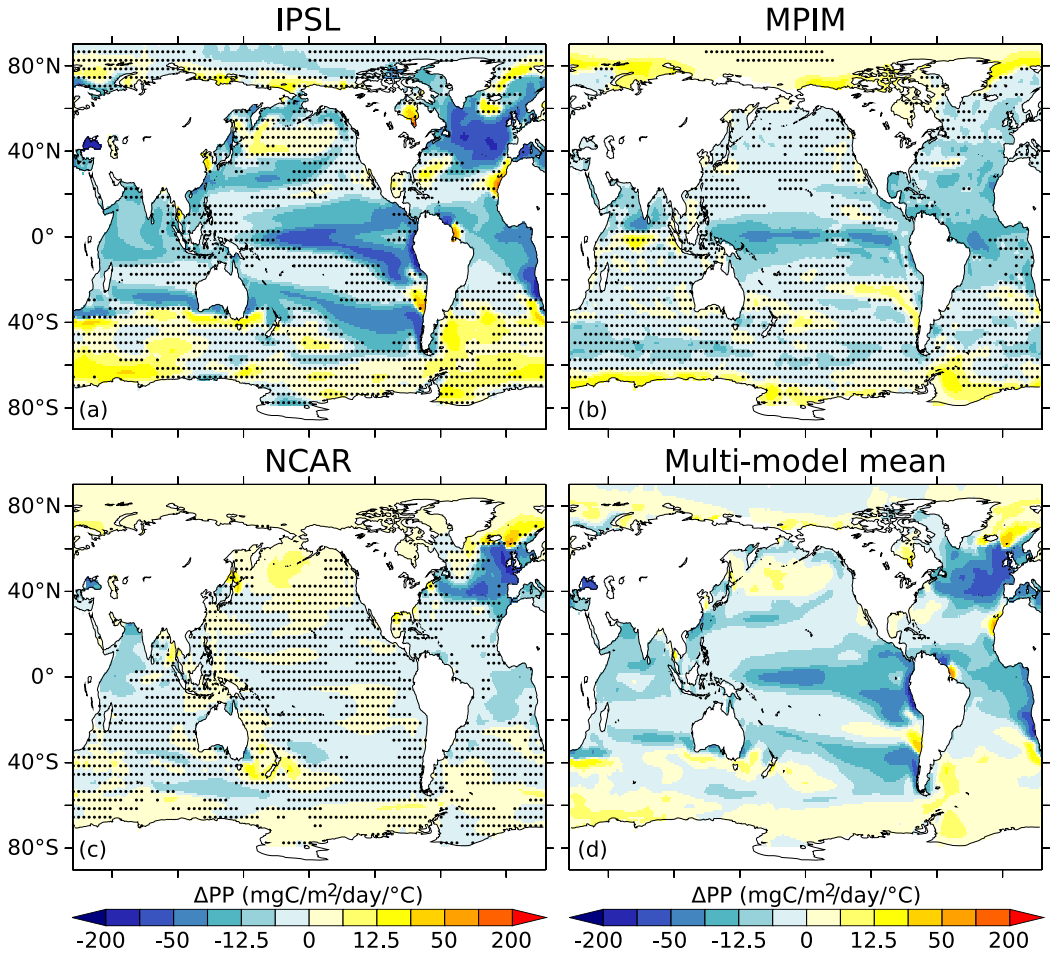

Fig. 10. Projected changes in vertically integrated annual mean primary production for a nominal increase in global mean surface air temperature (SAT_{glob}) of 1°C. The panels show the slope $\Delta PP/\Delta SAT_{glob}$ at each grid cell for the IPSL **(a)**, MPIM **(b)**, and NCAR **(c)** models. The multi-model mean **(d)** is the weighted mean (based on regional skill scores) of the individual slopes. The changes are shown on an exponential scale and are calculated from a linear regression of annual mean values over the period 1860-2099. Areas where $R^2 < 0.1$ are shaded in panels a-c.

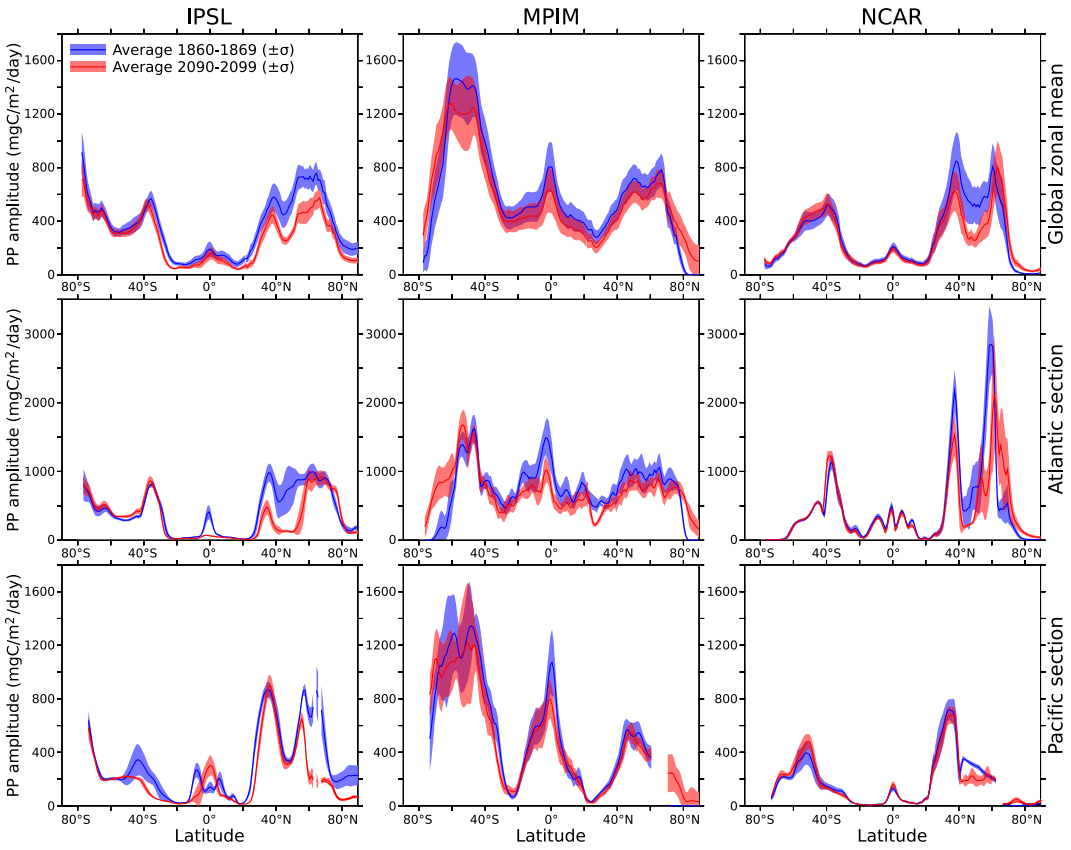

Fig. 11. Seasonal PP amplitude (maximum - minimum) zonally averaged (top), and for specific sections in the Atlantic (middle) and Pacific (bottom) as simulated by the three models IPSL (left), MPIM (middle), and NCAR (right) for preindustrial conditions (blue; decade 1860-1869) and projected by the end of the century (red; decade 2090-2099). Lines indicate the decadal mean and shadings the interannual variability ($\pm\sigma$). Please note that the scale of the vertical axis is different for the Atlantic section.

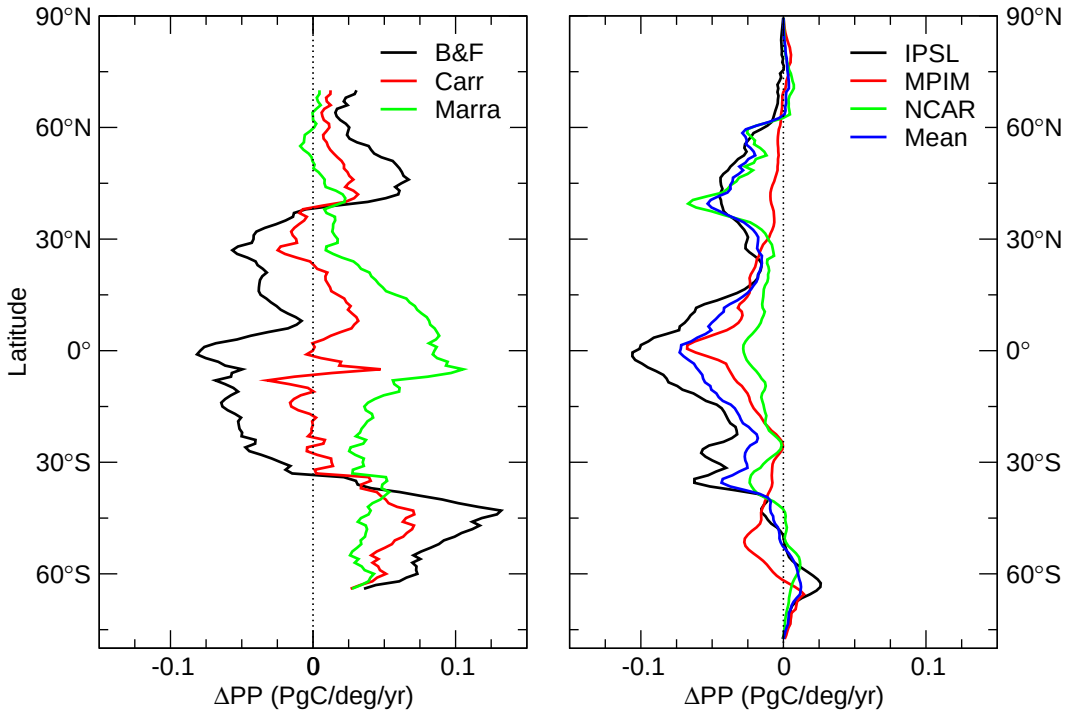

Fig. 12. Changes in zonally integrated PP under global warming as found with an empirical approach (left; cf. Fig. 11 in Sarmiento et al., 2004) and simulated with the mechanistic models IPSL, MPIM, and NCAR (right). In the left panel the productivity is calculated with the three different primary production algorithms of Behrenfeld and Falkowski (1997a, B&F), Carr (2002), and Marra et al. (2003). The multi-model mean shown in the right panel (blue) has been calculated using regional skill scores.