Phenological shifts in the North Atlantic net primary production detected in the 21st century. Results from two Earth system models.

Jenny Hieronymus[1], Magnus Hieronymus[1], Matthias Gröger[2], Jörg Schwinger[3], Raffaele Bernadello[4], Etienne Tourigny[4], Valentina Sicardi[4], Itzel Ruvalcaba Baroni[1], Klaus Wyser[1]

[1]Swedish Meteorological and Hydrological Institute, SMHI, Norrköping, 601 76, Sweden
[2]Department of Physical Oceanography and Instrumentation, Leibniz Institute for Baltic Sea Research Warnemünde, Rostock, D-18119, Germany
[3]NORCE Climate & Environment, Bjerknes Centre for Climate Research, Bergen, 5007, Norway
[4]Barcelona Supercomputing Center, BSC, Barcelona, 08034, Spain

Correspondence to: Jenny Hieronymus (jenny.hieronymus@smhi.se)

Abstract. Changes in the day of peak Net Primary Production in the North Atlantic (25-65° N) are here identified in daily data from two fully coupled CMIP6 Earth System Models, EC-Earth3-CC and NorESM2-LM, for the period 1750-2100 (under SSP5-8.5 scenario). The majority of the region displays the largest change point in the day of peak NPP occurring after the year 2000 indicating a shift towards earlier peak NPP with the most change occurring in the northern parts of the domain. We compared the ESM data of NPP to estimates derived from the CAFE satellite based data and found significant differences between the Earth System Model simulations and the CAFE data. However, the two models well represent both the magnitude of peak NPP and the seasonal cycles. Furthermore, the occurrence of the first day with MLD shallower than 40 m shows positive correlation with the occurrence of the day of peak NPP for most of the domain and, similar to the day of peak NPP, displays the largest changepoints occurring around or after the year 2000. The early timing of the detected shifts in both models suggests that similar shifts could already have been initiated or could start in the near future. This highlights the need for long term monitoring campaigns in the North Atlantic.

1 Introduction

Net Primary Production (NPP) is the net rate of photosynthetic carbon fixation. In the ocean, NPP is performed by microscopic planktonic phototrophs with a turnover time of about one week. Though the individual plankton are small, the total marine NPP almost equals its terrestrial counterpart with an estimated size of marine NPP on the order of 50GtC/yr (eg. Kulk et al., 2020, Westberry et al., 2008, Silsbe et al., 2016, Carr et al., 2006). NPP constitutes the basis of the food chain and provides the energy for higher trophic levels. Changes in NPP thus affect the entire ecosystem and ultimately fisheries and human food supply (Stock et al., 2017). In addition, NPP is the first step in the biological carbon pump, a set of processes by which carbon is exported from the surface to the deep ocean through the sinking of organic matter.

Understanding how the NPP and the subsequent export of organic carbon from the euphotic zone will change in future climate is thus vital for evaluating future uptake of atmospheric carbon dioxide (Honjo et al., 2014).

The north Atlantic is a region of particular importance for carbon sequestration in the deep ocean (Goris et al. 2018; Baker et al., 2022). This region contributes about 0.55-1.94 GtC/yr (Sanders et al. 2014) to the global export production, estimated to be 4-12 GtC/yr (de Vries and Weber, 2017). Moreover, here, cold water increases CO_2 solubility and deep mixing and subduction in the subpolar portion of this area result in a net transport of carbon to depth, a combination of processes known as the solubility carbon pump.

NPP is affected by climate variability through precipitation, wind patterns, temperature and light and is thus projected to change with anthropogenic climate change (Laufkötter et al., 2015; Pearl et al., 1999; Myriokefalitakis et al., 2020). Though an increase in temperature may enhance the growth rate of phytoplankton and thereby the net primary production, global NPP is projected to decrease (Behrenfeld et al., 2006; Steinacher et al., 2010; Bopp et al., 2013) though the uncertainty displayed in state-of-the-art Earth system Models (ESMs) is very large (Kwiatkowski et al., 2020). A projected NPP decline is often explained as being caused by increased water column stability that decreases the amount of nutrients available for primary production (Behrenfeld et al., 2006; Steinacher et al., 2010) but processes such as retreat of sea ice and increased stratification in high latitudes reduces the light limitation leading to NPP increases (Kwiatkowski et al., 2020). Efforts have been made to estimate how NPP has already changed in the historical satellite record but the limited range of satellite time-series makes such endeavors difficult. Estimates range from -2.1% per decade over the period 1998-2015 (Gregg and Rousseaux, 2019) to no significant change (Kulk et al., 2020).

The seasonal cycle of phytoplankton blooms has been explained with various theories. An often cited theory is the Critical Depth Hypothesis (Sverdrup, 1953) which postulates that a bloom can occur when the mixed layer has shoaled to a critical depth where the light limited gross production outweighs respiration. It does not, however, give an explanation as to when a bloom starts and ends. A more recent theory, termed the Disturbance Recovery Theory, of the timing of blooms was given by Behrenfeld (2010) (see also Behrenfeld and Boss, 2018). The theory suggests a balance between the growth and the loss in terms of respiration, grazing and disturbances to the physical environment such as the depth of the mixed layer. Other theories include Smyth et al. (2014) that relates the seasonality to the shift between negative and positive net heat flux.

Changes in seasonality and the timing of algal blooms can occur along with climate change with cascading effects into higher trophic levels up to fish and marine mammals. Changes in the phenology, or the timing of recurring biological events, of phytoplankton blooms due to climate change have already been observed in the North Sea with the Continuous Plankton Recorder (CPR) since 1960, with data displaying a significantly earlier onset of the spring bloom (Chivers et al. 2020). A phenological change in phytoplankton blooms will affect zooplankton and larvae as the timing of available food resources

will change, an effect known as the match/mismatch hypothesis (Cushing, 1990, Durant et al., 2007). The suggested causes of phenological shifts range from bottom up controls, including thermal stratification occuring earlier in the year allowing for an earlier bloom initiation, to top down controls resulting from changes in zooplankton grazing pressure (Yamaguchi et al., 2022).

Henson et al. (2013) used historical simulations from six ESMs covering the years 1985-2009 and a high emission future scenario (RCP8.5) to study changes in the NPP phenology. They found a shift towards an earlier peak NPP for most areas around the globe. However, the monthly resolution of the CMIP5 data dampens the phenology signal considerably. In a more recent study, Henson et al. (2018) used higher frequency model output to investigate the effect of temporal resolution on results of phytoplankton phenology. They found that in order to detect long term trends in bloom timing, a maximum temporal resolution of not more than 20 days is required.

However, even though a 20-day resolution may be adequate to detect long term trends, it is certainly not enough for detecting the timing of a rapid change in phenology in the course of global warming. In this paper, we use a long time series of daily output from two ESMs that contributed to the 6th Coupled Model Intercomparison Project (Eyring et al., 2016), to investigate the evolution of oceanic net primary production and its phenology in a region 25-65° N in the North Atlantic in the period 1750-2100. Furthermore, we divide the region into biogeochemical provinces (Longhurst et al., 1995) in order to see how localities with similar biogeochemical functioning differ across the region. We then investigate the occurrence of change points in the time-series of the day of peak NPP for the different provinces using change point analysis. Furthermore, we investigate the cross-correlation between the day of mixed layer depth shallower than a certain limit (here arbitrarily taken to be 40 m) and the day of peak NPP. We also investigate the largest change points in the day of the mixed layer shoaling above the limit. The cross-correlation analysis is complementary to the change point analysis and highlights at which leads and lags, the timing of mixed layer shoaling and peak NPP are covariant in the different regions.

2 Method

Daily output of vertically integrated net primary production has been produced using NorESM2-LM and EC-Earth3-CC for 100 years pi-control, historical (1850-2014) and the very high emission scenario SSP5-8.5 (2015-2100, Kriegler et al., 2017). All runs are forced with prescribed atmospheric CO_2 concentrations (concentration driven) in accordance with Meinshausen et al., 2020. The models are described in section 2.1. Section 2.2 describes the observational data set and section 2.3 provides an overview of the change point analysis method used. The phenological indicator that we have used is the day of peak NPP which is calculated as a simple max of NPP. This is a well defined metric that is robust unless for bimodal distributions with two peaks of similar size, which was not found in our data. The metric has previously been used in eg. Nissen and Vogt (2021) and Henson et al. (2013).

The mixed layer depths used for the analysis are calculated differently in the two ESMs. In EC-Earth3-CC we have used the turbocline depth as a mixed layer depth proxy calculated with a turbulent mixing coefficient criterion of 5cm2/s while in NorESM2-LM the mixed layer depth has been calculated in accordance with the criterion of de Boyer Montégut et al. (2004) and with a density difference of 0.03 kg/m3.

2.1 Models

2.1.1 EC-Earth3-CC

EC-Earth3 is an ESM developed by a European consortium of institutes and universities (Döscher et al. 2022). It is available in several different configurations. For this work, we have used EC-Earth3-CC which consists of the Integrated Forecast System (IFS) CY36R4 of the European Centre for Medium-Range Weather Forecasts (ECMWF) for simulating physics of the atmosphere and land surface, NEMO3.6 (Madec, 2015) for ocean physics, LPJ-Guess (Smith et al., 2014) for terrestrial vegetation and PISCES (Aumont et al., 2015) for ocean biogeochemistry. In concentration driven form, PISCES is fed a spatially uniform atmospheric pCO_2 while a CO_2 mapping occurs within IFS to account for regional heterogeneities.

PISCES is a mixed Monod-quota model simulating two different phytoplankton functional types, diatoms and nanophytoplankton, two size classes of zoo-plankton, micro and meso, and nutrients nitrate, ammonium, phosphate, iron and silicate. Iron and silicate are modeled using quotas in phytoplankton and the other nutrients with fixed Redfield ratios. Phytoplankton growth depends on the external concentration in nutrients, light and temperature. PISCES is suited for a wide range of spatial and temporal scales, including quasi-steady state simulations on the global scale. PISCES further simulates the carbon system, as well as dissolved and particulate organic matter. Net primary production is the growth of phytoplankton thus the term excludes mortality, excretion and grazing. The integrated net primary productivity used for the analysis is integrated over the water column and also summed over the two different phytoplankton functional types.

PISCES has been used and validated in a number of settings (Ramirez-romero et al., 2020; Gutknecht et al., 2019; Kwiatkowski et al., 2018). Skyllas et al. (2019) validated EC-Earth3, in an offline ocean only NEMO-PISCES version, for a north-south (29-63^{o}N) transect in the North-west Atlantic using cruise data of temperature, salinity and nutrients and chlorophyll-a and found a good agreement with observations. Net primary production has not previously been validated for

EC-Earth3-CC although the air-sea CO_2 flux, which is strongly affected by net primary production, was compared to an observation based climatology in Döscher et al. (2022). Their results showed stronger uptake of CO_2 than observations in the North Atlantic, thought to be caused by too active convection in the Labrador Sea.

130

2.1.2 NorESM2-LM

The Norwegian Earth System Model NorESM2 (Seland et al., 2020, Tjiputra et al, 2020) is a fully coupled ESM, which is based on the Community Earth System Model version 2 (CESM2, Danabasoglu et al. 2020) but employs a different ocean component (the Bergen Layered Ocean Model, BLOM) and a modified atmosphere model (CAM6-Nor). The land surface
135 and terrestrial biogeochemistry is represented by the Community Land Model version 5 (CLM5). BLOM uses isopycnic coordinates in the vertical (below a bulk mixed layer represented by two non-isopycnic model layers on top) and it includes the iHAMOCC model to represent ocean biogeochemistry. BLOM is coupled to the sea-ice component CICE5, which is the same as in CESM2. The LM version of NorESM2 used in this study has an atmosphere-land resolution of 2o and a nominal ocean model resolution of 1o. iHAMOCC is derived from the HAMOCC model (Six and Maier-Reimer, 1996; Ilyina et al.
140 2013) and was adapted for the use with isopycnic coordinates by Assman et al.(2010). HAMOCC includes a relatively simple NPZD ecosystem model with one phytoplankton and one zooplankton compartment and an implicit representation of calcifying and silicifying organisms. The model simulates nitrogen and phosphorus nutrients as well as dissolved iron with phytoplankton nutrient uptake according to Redfield molar ratios. The growth of phytoplankton is further affected by light as well as temperature.

145

NorESM2-LM has been validated with regards to biogeochemical variables including net primary production in Tjiputra et al. (2020). The results show a seasonal cycle of marine NPP that is reasonably well captured in amplitude but with a too low annual mean.

150 **2.2 Satellite based data - The CAFE model**

Direct observational data records of net primary production are scarce and in order to validate the two ESMs with respect to NPP, we have chosen to use data from a satellite based approach. There are several different methods for deriving total water column NPP estimates from satellite data. Often they are either based on ocean color (Behrenfeld and Falkowski, 1997), carbon (Westberry et al., 2008) or absorption (Smyth et al., 2005).

In this work, we use data from the Carbon, Absorption and Fluorescence Euphotic-resolving (CAFE) model (Silsbe et al. 2016), freely available through the Ocean Productivity site (http://sites.science.oregonstate.edu/ocean.productivity/index.php). The model utilizes satellite derived properties and has been shown to compare well to in situ observations (Johnson and Bif, 2021). We here utilize the MODIS-aqua dataset from 2002 to 2021.

2.3 Longhurst Provinces

The seasonality of NPP depends on the physical characteristics of different ocean localities. In modelling the terrestrial environment, the division into regions of similar growth conditions such as boreal forest or savannah is well defined, while in the ocean, biological differences between regions exist but are more difficult to constrain (Sathyendranath et al., 1995). Division of the global ocean into biogeochemical provinces has been done in a number of references (Longhurst et al., 1995; Sathyendranath et al., 1995) with the object of determining the global or regional net primary production. Longhurst et al. (1995) made the static boundaries that we have used in this analysis (downloaded from: https://www.marineregions.org/). Although the boundaries in reality are shifting on seasonal and interannual time scales (Reygondeau et al., 2013), we have chosen to use the static boundaries as we then are able to compare the same localities in the two models and in the CAFE data. The North Atlantic is divided according to three main areas: Coastal, westerlies and polar which are further subdivided to give the regions shown in Fig. 1. Note that we have chosen not to include the entire Arctic basin causing the Arctic provinces to be cut off in the north. The boreal polar region (BPLR) is defined by the southward flowing Labrador current that continues northward along the Greenland coastline. The Atlantic Arctic region (ARCT) is defined by strong stratification caused by large inflow of meltwater while the Atlantic subarctic (SARC) is characterized by poleward flowing warm North Atlantic water. The west wind regions; Gulf Stream (GFST), North Atlantic Drift (NADR) and North West (NADW) and North East Atlantic sub polar gyre (NASE) are governed by westerly winds and a Sverdrup type circulation. We have also included the coastal region North west Atlantic continental shelf (NWCS).

2.4 Change point analysis

Change point detection is a method to identify abrupt change in a time-series. Formally, the problem is to find the best possible segmentation of a signal according to some chosen criterion. Depending on this criterion one can look for changes

in, for example, mean, variance, or a spectral characteristic of a given signal In climate science, the method has been used to detect shifts in a wide variety of quantities (Beaulieu et al., 2012) such as AMOC strength (Smeed et al., 2018), coastal organic C sequestration (Watanabe et al., 2018), and cod stock (Möllmann et al., 2021). We have used change point detection to identify rapid change in the calendar-day of peak NPP. The calculations have been performed using the Python package Ruptures (Truong et al. 2020).

Generally speaking, change point detection requires a search method, a cost function and a constraint on the number of change points to detect. Search methods can be either exact or approximate. Here we use a version of the former called optimal detection, as computational speed is not an issue. Moreover, we use primarily the kernel based cost function and a constraint where we directly pick the number of change points to find. Many methods of change point analysis, in fact, focus on finding a predetermined number of shifts in a predefined quantity, such as the time series mean or variance (Truong et al., 2020). Another option that does not require the number of change points to be determined beforehand, is to instead use a penalty that is related to the amplitude of the change of interest. A small penalty generates many change points, which may arise due to intra-annual variability or noise, while a large penalty instead only gives the largest, if any, changes in the time series. By choosing a large enough penalty, the number of change points can in this way be tuned. In the process of doing this research we have tested both approaches.

Furthermore, instead of predefining the type of time series change, we have chosen to primarily use a kernel based non-parametric cost function developed by Arlot et al. (2013), in the following called “the kernel based cost function”. This model can detect all sorts of changes in the probability distribution of the time-series; mean, variance and higher order changes such as skewness and kurtosis. The upside of this approach is that no large changes are missed. The downside is that the method does not provide information on which change point is related to what type of change. Therefore, we complement the method with analyses using the Least absolute deviation (L1) cost function that detects changes in the median and the Least squared deviation (L2) cost function that detects changes in the mean of the time-series. Both of these are also available through the ruptures package and the search method used is the same as for the kernel based cost function.

3 Results

3.1 ESMs vs CAFE

We have compared the daily ESM data with 8 day NPP estimates from the CAFE data (Silsbe et al, 2016). Seasonal mean NPP over the MODIS-aqua period 2003-2021, for March, April, May (MAM), June, July, August (JJA) and September, October, November (SON) seasons are shown in Fig. 2. Note that the 2003-2021 period modelled by the ESMs is not the same period as that in the observationally constrained CAFE model. The two ESMs are forced with greenhouse gas concentrations that are similar to those for the period, but the internal variability of the climate system as modelled by the two ESMs is not in sync with that in reality or with each other. The comparison that can be done is thus strictly climatological. Due to the smaller area seen by satellites in winter, the CAFE data contains missing data over the winter months. In order to correctly compare the seasonal cycles, the ESM data was bounded to the north by the maximum latitude present in the CAFE data (Fig. S1). This means that not all data points to calculate a mean over the entire SON are present. The SON mean over the northern part of the domain is thus biased towards September-October.

Figure 2 shows large spatial differences between CAFE, EC-Earth3-CC and NorESM2-LM data. Most notably, EC-Earth3-CC shows a very strong NPP in MAM over the Gulf stream region. The high resolution CAFE data show that the enhanced production occurs in the warm Gulf stream eddies while the low resolution of the ESMs gives a wider warm water transport. The NorESM2-LM results in the Gulf stream region are closer to the CAFE data, although the production in the northern part of the domain is underestimated in both ESMs.

Figure 3 shows the area mean seasonal cycle over the period 2003-2021 for the different provinces. Again, the data is bounded to the north by the extent of CAFE the data (Fig. S1). The size of the NPP peak is well captured by both ESMs with the notable exception of province GFST in EC-Earth3-CC. This strong GFST production in EC-Earth3-CC is clearly seen in Fig. 2. However, the CAFE data shows a flatter and wider peak which generates a higher mean NPP over the time period compared to both ESMs for all provinces except for GFST and NASW in EC-Earth3-CC (Tab. 1). It is also apparent from Fig. 3 that the timing of peak NPP differs between biogeochemical provinces and models (Tab. 1). In the CAFE data, the day of peak NPP occurs in day 164-178 (early to late June) in the northernmost provinces BPLR, ARCT, SARC and NADR while the subtropical gyres (NASW, NASE), GFST and NWCS generates an earlier peak NPP between day 114 (April 24) and day 130 (May 10). Similarly, in EC-Earth3-CC, the three arctic regions BPLR, ARCT and SARC display the latest peak NPP occurring from day 150 to day 166 (May 30 to June 15) while the peak NPP in NADR occurs earlier compared to CAFE (day 124, May 4). The earliest peak NPP occurs in NASE on day 83 (March 24). As in the CAFE data, the earliest peak NPP in NorESM2-LM occurs in NASW (April 26, day 116 compared to April 24 in CAFE) while the latest occurs in

NWCS (day 186, July 5). In NorESM, the three arctic regions display a day of peak NPP of 159 (June 8) for BPLR, 161 (June 10) for ARCT, and 176 (June 25) for SARC. NASE and GFST have a day of peak NPP of 138 (May 18) and 148 (May 28) respectively. In general, EC-Earth3-CC is closer to the CAFE data in size but NorESM2-LM is closer in timing (Tab. 1).

3.2 Historical and future NPP

The time-series of annual mean NPP for the different biogeochemical provinces from 100 years of pi-control, historical and SSP5-8.5 are shown in Fig. 4 for EC-Earth3-CC, NorESM2-LM along with the annual mean CAFE data for the period 2003-2021. The figure reveals a large interannual and multidecadal variability in EC-Earth3-CC compared to NorESM2-LM.

For most provinces, EC-Earth3-CC generates higher annual mean NPP than NorESM2-LM, with the exception of BPLR and NASE. For BPLR, the mean for the entire period (not shown) is 110 mgC/m2/day for EC-Earth3-CC and 112 mgC/m2/day for NorESM2-LM and for NASE 314mgC/m2/day for EC-Earth3-CC compared to 321mgC/m2/day for NorESM2-LM. The largest difference between the two models is seen in GFST, where EC-Earth3-CC generates a time-series mean of 640 mgC/m2/day compared to 282 mgC/m2/day in NorESM2-LM. The highest NPP in NorESM2-LM is instead found in NASE.

The standard deviation for the entire period is in EC-Earth3-CC between 23.8 mgC/m2/day and 71.6 mgC/m2/day depending on the province (not shown). The largest standard deviation is found in NASE and the lowest in NWCS. In contrast, the standard deviation in NorESM2-LM is between 9.17 mgC/m2/day and 22.0 mgC/m2/day with, similar to EC-Earth3-CC, the largest found in NASE and the lowest in NWCS.

In order to find how the NPP and the timing of peak NPP has changed over the time-series, we have compared the last 30 yr period of SSP5-8.5 (2070-2099) to the first 30 yr period of the historical simulation (1850-1879). The results are summarized in Tab. 2. EC-Earth3-CC shows an increased NPP for most provinces with the exception of NADR and NASE where the NPP is lower in the late period compared to the early historical period. In addition to those provinces, NorESM2-LM display decreased NPP also for NASW and SARC. The day of the year of peak NPP decreases for all regions except one in both EC-Earth3-CC (NASE) and NorESM2-LM (GFST).

To further find how the shift in phenology is distributed over the region, the spatial distribution of the day of peak NPP averaged over the 30 year period 1850-1879 for the two ESMs is shown in Fig 5. Also shown in the figure is the difference of the ESM results averaged over the period 1985-2014 and 2070-2099 from the early period 1850-1879. In the early period, 1850-1879, both ESMs displayed a pattern of later bloom in the northern parts compared to the rest of the domain. For EC-Earth3-CC, this is most notable in the Labrador Sea while in NorESM2-LM, the later bloom is also visible in the Gulfstream and the northwest continental shelf area (NWCS).

The period 1985-2014 shows small and scattered differences from the early period. In the late period, 2070-2099, most of the domain experienced an earlier peak NPP but with some notable exceptions. Parts of the Gulf stream region display a markedly later peak NPP in the NorESM2-LM data compared to the early period, 1850-1879. This corresponds to an expansion of the pattern of late peak NPP in the Gulf stream region seen in the early period.

In EC-Earth3-CC, the pattern of earlier peak NPP in the final 30 yrs of SSP5-8.5 is widespread over the domain although a notable feature is the much later bloom in NASE (27 days on average, Tab. 2). The size of NPP in this region was greatly reduced in 2070-2099 compared to 1850-1879 (-86.6 mgC/m2/day, Tab. 2) caused by a strong reduction in winter surface nitrate concentration (not shown). The NPP seasonality in this area shifts from a clear spring peak to an extended period of weak NPP (not shown) with peak NPP therefore occurring later in the year. Also shown in Fig. 5 is the deviation of the late period from the 1850-1879 mean divided by the standard deviation of the PI-control in each grid cell which gives a measure of the significance of the late period change. The results show large significance in the northern parts of the domain.

Averaging over the different provinces allows us to look at the mean change in the day of peak NPP as well as identify change points. Fig. 6 shows the time-series of the day of peak NPP averaged over the different provinces together with the largest (cf. Fig. S2 for the largest change points using the PELT search method). In EC-Earth3-CC, the largest change point occurs between 2002 (BPLR and ARCT) and 2066 (NASE) for all regions except NASW where the largest change point

occurs the year 1900. Note also that NASW is the province with the least change over the time period (Tab. 2). In NorESM2-LM, the largest change point is in general located later, between 2010 (NASW) and 2082 (NASE). When increasing to two change points, the pattern of most change occurring after the year 2000 is maintained with few change points occurring earlier (Fig. 6). Also shown in the figure is the largest change point found by the L1 and L2 cost function that indicates changes in the median and mean respectively. The results show that the L1, L2 and the kernel based cost function gives almost the same results for almost all provinces. The most discrepancy is found in NASW in EC-Earth3-CC which is also the region displaying the least change.

Fig. 7 shows the first day of the year at which the spatial mean Mixed Layer Depth (MLD) shoals to 40 m or shallower in each province. Similar to the results of the day of peak NPP (Fig. 6), the day of MLD shallower than 40 m occurs progressively earlier over SSP5-8.5 for most provinces and for both EC-Earth3-CC and NorESM2-LM. The largest change-point in the time-series (Tab. 3) occurs between 1997 (GFST) and 2067 (NASW) for EC-Earth3-CC and between 2025 (ARCT, NADR) and 2092 (NWCS) for NorESM2-LM. Increasing to two change points, the pattern is consistent with most points located after the year 2000 (Fig. 7). Note that the choice of 40 m is arbitrary. We have tested for other cut-off depths with similar results (Figs. S3-S5).

But how well do change points in the spatial mean day of peak NPP of the different provinces represent the separate grid points? The year during which the largest change point for every grid point occurs is shown in Fig. 8. The results broadly correspond to the results seen in the spatial mean time series with a major change point occurring after the year 2000. Few grid points display a change point earlier than that. Note that many of the grid cells displaying change points early in the time-series correspond to cells where the PELT search method could not find only one single change point (Fig. S6). This points to the fact that in these grid points, little significant change occurs (cf. Fig 5. bottom panel). Furthermore, EC-Earth3-CC displays an earlier major change point for most grid points as compared to NorESM2-LM. The northern part of the domain, i.e. regions where the euphotic zone is more vigorously coupled to the deep sea by vertical mixing like the Labrador

Sea, northern North Atlantic and the sub-polar gyre, shows the earliest change point in the EC-Earth3-CC results close to the year 2000. The south-eastern part of the domain displays the latest change point in both NorESM2-LM and EC-Earth3-CC.

To elucidate on the correlation between the day of MLD shallower than 40 m and the day of max NPP, the cross correlation between the area averaged time-series shown in Figs. 6 and 7 has been plotted in Fig. 9. The figure shows a notable correlation, well above the 95% confidence bound, between the two indices in most provinces for both ESMs. The maximum correlation occurs for zero lag in most provinces, indicating, as expected, that peaks in these variables tend to occur within the same year. Note that the strongest correlation for zero lag, at least in EC-Earth3-CC is seen in the west wind provinces, GFST, NADR, NADW and NASE that have a Sverdrup-like circulation. These are also open ocean regions where mixed layers can be expected to be less constrained by freshwater fluxes from land. Furthermore, a striking feature is the strong negative correlations found in the BPLR and ARCT provinces in EC-Earth3-CC. Looking at Fig. 8 we find that the day of MLD shallower than 40 m, at least in the BPLR province, occurs so early in the year that it can hardly affect the day of peak NPP. Thus, suggesting that the anti-correlation between these variables is owing to a lurking variable affecting both NPP and MLD. The similar correlation structure between BPLR and ARCT strongly suggests that the same is true about the ARCT province.

4. Discussion

The comparison between CAFE and the two ESMs showed that the size of peak NPP was well captured by the ESMs. The regional difference in NPP is larger in the ESMs compared to the CAFE data which is evident by the difference in annual mean between the provinces (Tab. 1). Peak NPP occurs latest in the year for the arctic provinces; BPLR, ARCT and SARC in EC-Earth3-CC which corresponds well to the CAFE data (although peak NPP in NADR occurs later than BPLR in CAFE).

The results showed that the most change in the day of peak NPP occurs after the beginning of the 21st century which is consistent with the results of Henson et al (2009), who found no long term trend in the subpolar north Atlantic towards earlier or later blooms in model data spanning 1959-2004. The earlier bloom displayed for most provinces (Tab. 2) under SSP5-8.5 is in agreement with Asch et al. (2019) who showed that blooms north of 40°N shifted earlier under RCP8.5 using 5 daily output from GFDL ESM2M including the biogeochemical model TOPAZ2.0. In contrast, Henson et al. (2018) reports, using an ocean only model (MEDUSA-2.0, NEMO), a start of bloom shifting later in the year under RCP8.5 in most parts of the North Atlantic. However, both studies relate to surface chlorophyll and not to NPP, which makes the comparison problematic. Although surface chlorophyll has the benefit of being easily validated to observations, it is not, in a simple manner connected to vertically integrated NPP. The chlorophyll maxima can be found below the surface (Cornec et al., 2021) and the relationship between the surface concentration and the subsurface profile differs between different localities (Sathyendranath et al., 1995). The seasonality of peak NPP is therefore not necessarily directly relatable to the seasonality of surface chlorophyll. Moreover, our temporal resolution is higher and both Henson et al. (2018) and Asch et al. (2019) use the start of bloom as well as length of bloom as a phenological indicator instead of the timing of the annual peak which further complicates the comparison.

The change-point analysis revealed that the largest change points in the day of peak NPP occur in most provinces after the year 2000 in both ESMs (Tab. 3, Fig. 6). A noteworthy observation is that very few change points occur in the latter part of the 21st century even though it is then that the warming in the SSP5-8.5 is the greatest. In EC-Earth3-CC many change points occur already in the historical simulation, while in NorESM2-LM they tend to occur in the middle of the 21st century, at a point when the very high emission scenario SSP5-8.5 has just started to diverge from the more moderate emission scenarios in terms of global mean surface temperature (IPCC, 2022). Change points occur earlier in EC-Earth3-CC than in NorESM2-LM in all provinces, and the same change points are generally found regardless of which cost function (L1,L2 or kernel based) is chosen in both models. This indicates that the change points found are affecting multiple statistical moments.

We compared the day of peak NPP with the day of MLD shallower than 40 m and the cross correlation did show the strongest correlation at zero lag. The fact that we saw significant correlations also with much longer lags likely reflected the low-frequency cycles of the Atlantic multi-decadal variability which affects many physical parameters, like SST and MLD, on multi-decadal time scales (e.g. Börgel et al. 2020). Both this type of low frequency variability and anthropogenic climate change could act as a lurking variable driving coherent changes in both NPP and MLD on a range of both positive and negative lags. Furthermore, we noted from Fig. 9 a strong anticorrelation between the MLD and NPP phenology for provinces BPLR and ARCT in EC-Earth3-CC. Given that both the provinces are far to the north and that SSP5-8.5 is a very strong warming scenario, we speculate that changes in sea-ice could be behind the observed correlation structure. This is supported by the fact that EC-Earth3 has been shown to overestimate sea ice concentrations in the Labrador Sea (Döscher et al., 2022). However, given that the timing of the MLD shallowing is unlikely to be important for the timing of peak NPP in these provinces we did not investigate further.

NPP and its timing is, of course, both in the models and in reality dependent on many other factors in addition to the MLD. Some examples are light availability, nutrient concentrations and temperature. The MLD can similarly both be affected by and affect some of these factors. In light of this, it is clear that MLD changes can both act as a driver of phenology changes in itself and act as a proxy for other drivers, which complicates the interpretation. The cross correlation analysis therefore does not point to the validity of a certain bloom timing theory such as Critical Depth Hypothesis or the Disturbance Recovery Theory (Behrenfeld, 2010) but it does highlight the covariance of the NPP and MLD phenology.

The biogeochemical modules included in the earth system models are by necessity simplistic, with PISCES simulating two phytoplankton functional types representing nanophytoplankton and diatoms and HAMOCC only one. However, even with reduced complexity interpretation is not straightforward. Compared to observations, community structure has been shown to affect the NPP and models containing a more dynamic phytoplankton community have a more non-linear response to climate change due to decreases in large cells and an increasing amount of regenerated production (Fu et al., 2016). Thus, more

complex biogeochemical models may generate different results. The simpler biogeochemical model included in NorESM2-LM may be the reason for the lower variability seen in the NPP results (Fig. 4).

For this work we have used fully coupled earth system models as opposed to ocean only models, that are often used for similar work (eg. Henson et al., 2018, 2009). The exchange of heat, momentum and freshwater is more accurately treated in coupled models than in ocean only models. This affects, for example, temperature and stratification. It has also been demonstrated that interactive coupling affects the variability of these variables (e.g. Bhatt et al. 1998, Barsugli and Battisti, 1998). The biogeochemical response is therefore expected to differ in the coupled vs. the uncoupled case. Because of more realistic physics with respect to uncoupled models, we believe using coupled models might constitute an important step forward in the larger effort of trying to understand what phenology changes might occur in the future.

5 Summary and conclusions

Using daily output for the period 1750-2100 under historical and SSP5-8.5 from two fully coupled ESMs, we have analyzed NPP and the day of peak NPP for an area covering 25-65°N in the North Atlantic. We have compared the vertically integrated NPP for the two ESMs with the satellite based CAFE model. Both ESMs show deviations from the CAFE data. At least for EC-Earth3-CC, this is especially true in the Gulf Stream region.

Furthermore, we separated the domain in biogeochemical provinces in accordance with Longhurst et al. (1995) in order to look at spatial averages. The seasonal cycle of the CAFE data displays a longer growing season in most provinces than the two ESMs. NorESM2-LM is typically better at capturing the timing of peak NPP and EC-Earth3-CC is closer in annual average NPP in most provinces. The multi-decadal variability is smaller in NorESM2-LM than in EC-Earth3-CC. The CAFE data seem to show larger interannual variability than both ESMs though this is difficult to determine given the relatively short (2003-2021) satellite record.

The ESM data show that shifts in the seasonal cycle occur mainly during the 21st century with reduced NPP as well as a peak NPP occurring earlier in the year for most of the biogeochemical provinces. The largest change towards an earlier day of peak NPP occurs in the northern parts for both ESMs. In EC-Earth3-CC the largest change occurs in the biogeochemical province BPLR and in NorESM2-LM in the arctic province ARCT. Moreover, the changes in the day of peak NPP are far outside the range of the natural variability diagnosed from the PI-control run in large parts of the domain. For future work it would be very interesting to see what a similar change point analysis would reveal in low and moderate emission scenarios. For all provinces, the change points occur earlier in EC-Earth3-CC than in NorESM2-LM.

Cross correlation analysis showed significant correlation between the day of MLD shallower than 40 m and the day of peak NPP in most regions. The peak correlation occurs at zero lag, but correlations are significant at many both positive and negative lags. The large range of correlated lags we ascribe to forced and unforced low frequency variability affecting both parameters.

Our results point to phenological shifts occurring in the early 21st century in the vertically integrated NPP in different parts of the North Atlantic (25-65^{o}N). Shifts in the phenology may have an impact on fishery yields through the mismatch of fish spawning and available resources. Furthermore, carbon sequestration in this highly productive region may be affected by changes in ecosystem structure in turn affecting the export production and the general efficiency of the biological pump.

Code availability: The EC-Earth3 code is available from the EC-Earth development portal for members of the consortium. All code related to CMIP6 forcing is implemented in the component models. Model codes developed at ECMWF, including the atmosphere model IFS, are intellectual property of ECMWF and its member states. Permission to access the EC-Earth3 source code can be requested from the EC-Earth community via the EC-Earth website (http://www.ec-earth.org/, EC-Earth consortium, 2019a) and may be granted if a corresponding software license agreement is signed with ECMWF. The repository tag for the version of EC-Earth that is used in this work is 3.3.1. Currently, only European users can be granted access due to license limitations of the atmosphere model. The component models NEMO, LPJ-GUESS, TM5, and PISM are not limited by their licenses.

The NorESM code can be accessed via zenodo: Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Gao, S., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moree, A., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: NorESM2 source code as used for CMIP6 simulations (includes additional experimental setups, extended model documentation, automated inputdata download, restructuring of BLOM/iHAMOCC input data), Zenodo [code], https://doi.org/10.5281/zenodo.3905091, 2020.

Data availability: The data used to produce the figures in this manuscript can be downloaded from Zenodo using the link: 10.5281/zenodo.7716480

The python package Ruptures can be downloaded from: https://centre-borelli.github.io/ruptures-docs/

Author contribution: JH performed the EC-Earth3-CC model run, made the analysis with contributions from MH, and drafted the manuscript. MG contributed to the research design. JS performed the NorESM2-LM run. ET, RB and VS made the EC-Earth3-CC setup and contributed to the EC-Earth3-CC model run. IRB contributed with disscussions and paper writing. KW assisted in setting up and running EC-Earth3-CC. All co-authors contributed to the writing of the paper.

Competing interests: The authors declare that they have no conflict of interest.

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

Acknowledgements: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820989 (project COMFORT, Our common future ocean in the Earth system – quantifying coupled cycles of carbon, oxygen, and nutrients for determining and achieving safe operating spaces with respect to tipping points). R. B. acknowledges support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No GA 708063 (NetNPPAO).

References

Arlot, S., Celisse, A. and Harchaoui, Z. (2012). A kernel multiple change-point algorithm via model selection. Preprint. Available at https://arxiv.org/abs/1202.3878v2.

Asch, R. G., Stock, C. A., & Sarmiento, J. L. (2019). Climate change impacts on mismatches between phytoplankton blooms and fish spawning phenology. Global Change Biology, 25(8), 2544–2559. https://doi.org/10.1111/gcb.14650

Assmann, K. M., Bentsen, M., Segschneider, J., and Heinze, C.: An isopycnic ocean carbon cycle model, Geosci. Model Dev., 3, 143–167, https://doi.org/10.5194/gmd-3-143-2010, 2010.

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., & Gehlen, M. (2015). PISCES-v2: An ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development*, *8*(8), 2465–2513. https://doi.org/10.5194/gmd-8-2465-2015

Barsugli, J. J., and D. S. Battisti (1998). The Basic Effects of Atmosphere–Ocean Thermal Coupling on Midlatitude Variability. J. Atmos. Sci., 55, 477–493, https://doi.org/10.1175/1520-0469(1998)055<0477:TBEOAO>2.0.CO;2

Bhatt, U. S., M. A. Alexander, D. S. Battisti, D. D. Houghton, and L. M. Keller (1998). Atmosphere–Ocean Interaction in the North Atlantic: Near-Surface Climate Variability. J. Climate, 11, 1615–1632, https://doi.org/10.1175/1520-0442(1998)011<1615:AOIITN>2.0.CO;2

Beaulieu Claudie, Chen Jie and Sarmiento Jorge L. (2012). Change-point analysis as a tool to detect abrupt climate variationsPhil. Trans. R. Soc. A. 370. 1228–1249. http://doi.org/10.1098/rsta.2011.0383

Behrenfeld, M. J., & Falkowski, P. G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*, *42*(1), 1–20. https://doi.org/10.4319/lo.1997.42.1.0001

Behrenfeld, M. J. (2010). Abandoning Sverdrup ' s Critical Depth Hypothesis on phytoplankton blooms. *Ecology*, *91*(4), 977–989.

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., & Boss, E. S. (2006). Climate-driven trends in contemporary ocean productivity. *Nature*, *444*(7120), 752–755. https://doi.org/10.1038/nature05317

Behrenfeld, M. J., & Boss, E. S. (2018). Student's tutorial on bloom hypotheses in the context of phytoplankton annual cycles. *Global Change Biology*, *24*(1), 55–77. https://doi.org/10.1111/gcb.13858

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., & Vichi, M. (2013). Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences*, *10*(10), 6225–6245. https://doi.org/10.5194/bg-10-6225-2013

Börgel, F., C. Frauen, T. Neumann, and H. E. M. Meier., The Atlantic Multidecadal Oscillation controls the impact of the North Atlantic Oscillation on North European climate, Environmental Research Letters, https://doi.org/10.1088/1748-9326/aba925. 2020

Chivers, W. J., Edwards, M., & Hays, G. C. (2020). Phenological shuffling of major marine phytoplankton groups over the last six decades. *Diversity and Distributions*, *26*(5), 536–548. https://doi.org/10.1111/ddi.13028

Cornec, M., Claustre, H., Mignot, A., Guidi, L., Lacour, L., Poteau, A., et al. (2021). Deep chlorophyll maxima in the global ocean: Occurrences, drivers and characteristics. Global Biogeochemical Cycles, 35, e2020GB006759. https://doi.org/10.1029/2020GB006759

Cushing, D. H. (1990). Plankton Production and Year-class Strength in Fish Populations: an Update of the Match/Mismatch Hypothesis. Advances in Marine Biology, 26, 249-293. https://doi.org/10.1016/S0065-2881(08)60202-3

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M.,

Lindsay, K., Lipscomb, W. H., Mills, M. J., … Strand, W. G. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), 1–35. https://doi.org/10.1029/2019MS001916

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, *109*(12), 1–20. https://doi.org/10.1029/2004JC002378

Durant, J. M., Hjermann, D., Ottersen, G., & Stenseth, N. C. (2007). Climate and the match or mismatch between predator requirements and resource availability. *Climate Research*, *33*(3), 271–283. https://doi.org/10.3354/cr033271

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geosci. Model Dev., 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

Fu, W., Randerson, J. T., & Keith Moore, J. (2016). Climate change impacts on net primary production (NPP) and export production (EP) regulated by increasing stratification and phytoplankton community structure in the CMIP5 models. *Biogeosciences*, *13*(18), 5151–5170. https://doi.org/10.5194/bg-13-5151-2016

Gregg, W. W., & Rousseaux, C. S. (2019). Global ocean primary production trends in the modern ocean color satellite record (1998-2015). *Environmental Research Letters*, *14*(12). https://doi.org/10.1088/1748-9326/ab4667

Goris, N., Tjiputra, J. F., Olsen, A., Schwinger, J., Lauvset, S. K., & Jeansson, E. (2018). Constraining projection-based estimates of the future North Atlantic carbon uptake. Journal of Climate, 31(10), 3959–3978. https://doi.org/10.1175/JCLI-D-17-0564.1

Gröger, M., Maier-Reimer, E., Mikolajewicz, U., Moll, A., and Sein, D.: NW European shelf under climate warming: implications for open ocean – shelf exchange, primary production, and carbon absorption, Biogeosciences, 10, 3767–3792, https://doi.org/10.5194/bg-10-3767-2013, 2013

Gutknecht, E., Reffray, G., Mignot, A., Dabrowski, T., & Sotillo, M. G. (2019). Modelling the marine ecosystem of Iberia-Biscay-Ireland (IBI) European waters for CMEMS operational applications. *Ocean Science*, *15*(6), 1489–1516. https://doi.org/10.5194/os-15-1489-2019

Henson, S. A., Dunne, J. P., & Sarmiento, J. L. (2009). Decadal variability in North Atlantic phytoplankton blooms. *Journal of Geophysical Research: Oceans*, *114*(4), 1–11. https://doi.org/10.1029/2008JC005139

Henson, S., Cole, H., Beaulieu, C., & Yool, A. (2013). The impact of global warming on seasonality of ocean primary production. *Biogeosciences*, *10*(6), 4357–4369. https://doi.org/10.5194/bg-10-4357-2013

Henson, S. A., Cole, H. S., Hopkins, J., Martin, A. P., & Yool, A. (2018). Detection of climate change-driven trends in phytoplankton phenology. *Global Change Biology*, *24*(1), e101–e111. https://doi.org/10.1111/gcb.13886

Honjo, S., Eglinton, T. I., Taylor, C. D., Ulmer, K. M., Sievert, S. M., Bracher, A., German, C. R., Edgcomb, V., Francois, R., Deboraiglesias-Rodriguez, M., Van Mooy, B., & Repeta, D. J. (2014). The role of the biological pump in the global carbon cycle understanding an imperative for ocean science. *Oceanography*, *27*(3), 10–16. https://doi.org/10.5670/oceanog.2014.78

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., & Núñez-Riboni, I. (2013). Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations. *Journal of Advances in Modeling Earth Systems*, *5*(2), 287–315. https://doi.org/10.1029/2012MS000178

IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. Cambridge University Press, Cambridge, UK and New York, NY, USA, 3056 pp., doi:10.1017/9781009325844.

Johnson, K. S., & Bif, M. B. (2021). Constraint on net primary productivity of the global ocean by Argo oxygen measurements. *Nature Geoscience*, *14*(10), 769–774. https://doi.org/10.1038/s41561-021-00807-z

Killick, R., Fearnhead, P., and Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. Journal of the American Statistical Association, 107(500):1590–1598,107(500):1590–1598.

Kulk, G., Platt, T., Dingle, J., Jackson, T., Jönsson, B. F., Bouman, H. A., Babin, M., Brewin, R. J. W., Doblin, M., Estrada, M., Figueiras, F. G., Furuya, K., González-Benítez, N., Gudfinnsson, H. G., Gudmundsson, K., Huang, B., Isada, T., Kovač, Ž., Lutz, V. A., … Sathyendranath, S. (2020). Primary production, an index of climate change in the ocean: Satellite-based estimates over two decades. *Remote Sensing*, *12*(5), 1–26. https://doi.org/10.3390/rs12050826

Kwiatkowski, L., Aumont, O., Bopp, L., & Ciais, P. (2018). The Impact of Variable Phytoplankton Stoichiometry on Projections of Primary Production, Food Quality, and Carbon Uptake in the Global Ocean. *Global Biogeochemical Cycles*, *32*(4), 516–528. https://doi.org/10.1002/2017GB005799

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J., Dunne, J., Gehlen, M., Ilyina, T., John, J., Lenton, A., Li, H., Lovenduski, N., Orr, J., Palmieri, J., Schwinger, J., Séférian, R., Stock, C., Tagliabue, A., … Ziehn, T. (2020). Twenty-first century ocean warming, acidification, deoxygenation, and upper ocean nutrient decline from CMIP6 model projections. *Biogeosciences*. 17, 3439–3470. https://bg.copernicus.org/articles/17/3439/2020/

Laufkötter C, Vogt M, Gruber N, Aita-Noguchi M, Aumont O, Bopp L, Buitenhuis E, Doney SC, Dunne J, Hashioka T, Hauck J. Drivers and uncertainties of future global marine primary production in marine ecosystem models. Biogeosciences. 2015 Dec 7;12(23):6955-84.

Longhurst, A., Sathyendranath, S., Platt, T., & Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *Journal of Plankton Research*, *17*(6), 1245–1271. https://doi.org/10.1093/plankt/17.6.1245

Madec, G. (2015). NEMO ocean engine, Note du Pole de modelisation de l'Institut Pierre-Simon Laplace No 27, ISSN No 1288-1619.

Myriokefalitakis, S., Gröger, M., Hieronymus, J., and Döscher, R.: An explicit estimate of the atmospheric nutrient impact on global oceanic productivity, Ocean Sci., 16, 1183–1205, https://doi.org/10.5194/os-16-1183-2020, 2020.

Möllmann, C., Cormon, X., Funk, S. et al. Tipping point realized in cod fishery. Sci Rep 11, 14259 (2021). https://doi.org/10.1038/s41598-021-93843-z

Nissen, C., & Vogt, M. (2021). Factors controlling the competition between Phaeocystis and diatoms in the Southern Ocean and implications for carbon export fluxes. *Biogeosciences*, *18*(1), 251–283. https://doi.org/10.5194/bg-18-251-2021

Paerl HW, Willey JD, Go M, Peierls BL, Pinckney JL, Fogel ML. Rainfall stimulation of primary production in western Atlantic Ocean waters: roles of different nitrogen sources and co-limiting nutrients. Marine Ecology Progress Series. 1999 Jan 18;176:205-14.

Ramirez-romero, E., Jordà, G., Amores, A., Kay, S., Segura-noguera, M., Macias, D. M., Maynou, F., Sabatés, A., & Alexander, M. A. (2020). *Assessment of the Skill of Coupled Physical – Biogeochemical Models in the NW Mediterranean*. *7*(July), 1–18. https://doi.org/10.3389/fmars.2020.00497

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., & Maury, O. (2013). Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles*, *27*(4), 1046–1058. https://doi.org/10.1002/gbc.20089

Sanders, R., Henson, S. A., Koski, M., De La Rocha, C. L., Painter, S. C., Poulton, A. J., Riley, J., Salihoglu, B., Visser, A., Yool, A., Bellerby, R., & Martin, A. P. (2014). The Biological Carbon Pump in the North Atlantic. *Progress in Oceanography*, *129*(PB), 200–218. https://doi.org/10.1016/j.pocean.2014.05.005

Sathyendranath, S., Longhurst, A., Caverhill, C. M., & Platt, T. (1995). Regionally and seasonally differentiated primary production in the North Atlantic. *Deep-Sea Research Part I*, *42*(10), 1773–1802. https://doi.org/10.1016/0967-0637(95)00059-F

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y. C., Kirkeväg, A., Schwinger, J., Tjiputra, J., Schanke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., … Schulz, M. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. In *Geoscientific Model Development* (Vol. 13, Issue 12). https://doi.org/10.5194/gmd-13-6165-2020

Silsbe, G. M., Behrenfeld, M. J., Halsey, K. H., Milligan, A. J., & Westberry, T. K. (2016). The CAFE model: A net production model for global ocean phytoplankton. *Global Biogeochemical Cycles*, *30*(12), 1756–1777. https://doi.org/10.1002/2016GB005521

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.

Skyllas, N., Bintanja, R., Buma, A. G. J., Brussaard, C. P. D., Gröger, M., Hieronymus, J., & van de Poll, W. H. (2019). Validation of stratification-driven phytoplankton biomass and nutrient concentrations in the northeast atlantic ocean as simulated by EC-earth. *Geosciences (Switzerland)*, *9*(10). https://doi.org/10.3390/geosciences9100450

Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., Moat, B. I., Frajka-Williams, E., Rayner, D., Meinen, C. S., Baringer, M. O., Bryden, H. L., & McCarthy, G. D. (2018). The North Atlantic Ocean Is in a State of Reduced Overturning. *Geophysical Research Letters*, *45*(3), 1527–1533. https://doi.org/10.1002/2017GL076350

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual- based dynamic vegetation model, Biogeosciences, 11, 2027– 2054, https://doi.org/10.5194/bg-11-2027-2014, 2014.

Smyth, T. J., Tilstone, G. H., & Groom, S. B. (2005). Integration of radiative transfer into satellite models of ocean primary production. *Journal of Geophysical Research: Oceans*, *110*(10), 1–11. https://doi.org/10.1029/2004JC002784

Steinacher, M., Joos, F., Frölicher, T. L., Bopp, L., Cadule, P., Cocco, V., Doney, S. C., Gehlen, M., Lindsay, K., Moore, J. K., Schneider, B., & Segschneider, J. (2010). Projected 21st century decrease in marine productivity: A multi-model analysis. *Biogeosciences*, *7*(3), 979–1005. https://doi.org/10.5194/bg-7-979-2010

Sein, D.V, Gröger, M., Cabos, W., Alvarez, F., Hagemann, S., de la Vara, A., Pinto, J.G., Izquierdo, A., Koldunov, N.V., Dvornikov, A. Y., Limareva, N., Martinez, B., Jacob, D. (2020), Regionally coupled atmosphere - ocean - marine biogeochemistry model ROM: 2. Studying the climate change signal in the North Atlantic and Europe J. Adv. Model. Earth Syst., DOI: 10.1029/2019MS001646

Stock, C. A., John, J. G., Rykaczewski, R. R., Asch, R. G., Cheung, W. W. L., Dunne, J. P., Friedland, K. D., Lam, V. W. Y., Sarmiento, J. L., & Watson, R. A. (2017). Reconciling fisheries catch and ocean productivity. *Proceedings of the*

National Academy of Sciences of the United States of America, *114*(8), E1441–E1449. https://doi.org/10.1073/pnas.1610238114

Sverdrup, H. U. (1953). On conditions for the vernal blooming of phytoplankton. *ICES Journal of Marine Science*, *18*(3), 287–295. https://doi.org/10.1093/icesjms/18.3.287

Tjiputra, J. F., Schwinger, J., Bentsen, M., L. Morée, A., Gao, S., Bethke, I., Heinze, C., Goris, N., Gupta, A., He, Y. C., Olivié, D., Seland, O., & Schulz, M. (2020). Ocean biogeochemistry in the Norwegian Earth System Model version 2 (NorESM2). *Geoscientific Model Development*, *13*(5), 2393–2431. https://doi.org/10.5194/gmd-13-2393-2020

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, *167*. https://doi.org/10.1016/j.sigpro.2019.107299

van Noije, T. P. C., Le Sager, P., Segers, A. J., van Velthoven, P. F. J., Krol, M. C., Hazeleger, W., Williams, A. G., and Chambers, S. D.: Simulation of tropospheric chemistry and aerosols with the climate model EC-Earth, Geosci. Model Dev., 7, 2435–2475, https://doi.org/10.5194/gmd-7-2435-2014, 2014.

Watanabe, K., Seike, K., Kajihara, R., Montani, S., & Kuwae, T. (2019). Relative sea-level change regulates organic carbon accumulation in coastal habitats. *Global Change Biology*, *25*(3), 1063–1077. https://doi.org/10.1111/gcb.14558

Westberry, T., Behrenfeld, M. J., Siegel, D. A., & Boss, E. (2008). Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Global Biogeochemical Cycles*, *22*(2), 1–18. https://doi.org/10.1029/2007GB003078

Yamaguchi, R., Rodgers, K. B., Timmermann, A., Stein, K., Schlunegger, S., Bianchi, D., Dunne, J. P., & Slater, R. D. (2022). Trophic level decoupling drives future changes in phytoplankton bloom phenology. *Nature Climate Change*, *12*(5), 469–476. https://doi.org/10.1038/s41558-022-01353-1

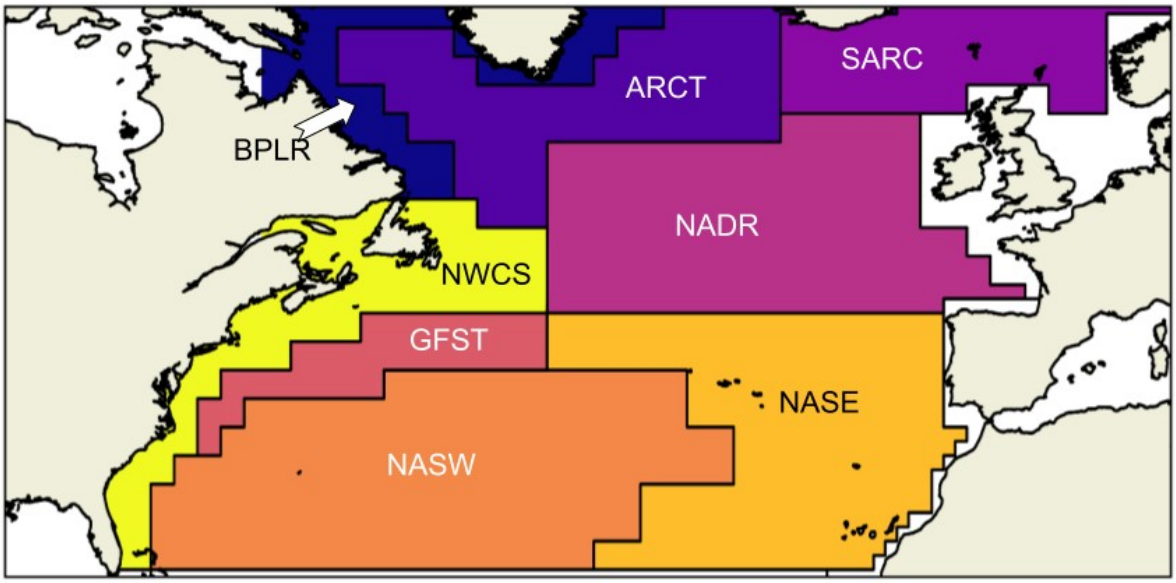
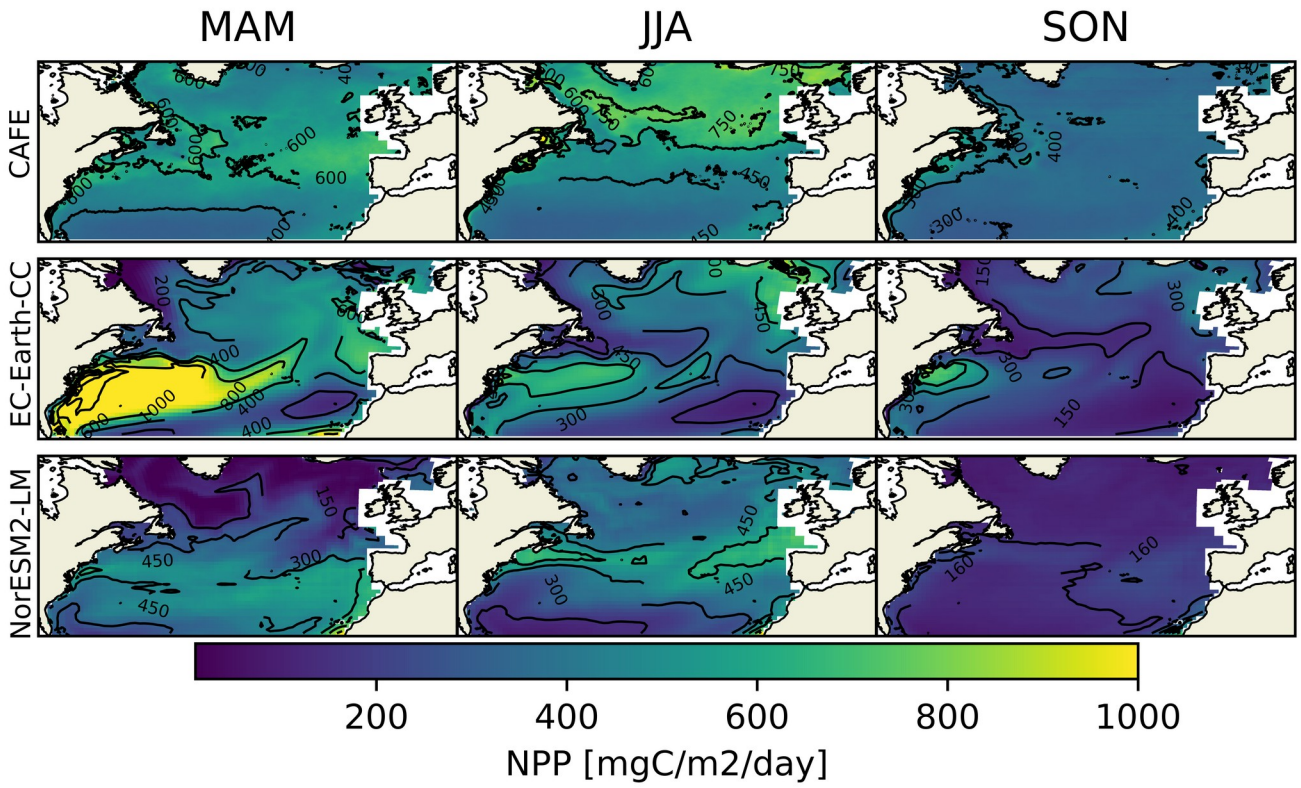


Figure 1: Study area and Longhurst provinces. BPLR: Boreal polar, ARCT: Atlantic arctic, SARC: Atlantic subarctic, NADR: North Atlantic drift, GFST: Gulf stream, NASW: North west Atlantic subtropical gyre, NASE: North east Atlantic subtropical gyre, NWCS North west Atlantic shelf.



655 **Figure 2: Seasonal mean vertically integrated NPP from the CAFE model (upper), EC-Earth3-CC (middle) and NorESM2-LM (bottom). The ESM data was masked by the maximum latitude present in the CAFE data to account for the smaller area seen by satellites in winter. This only affects the SON values which is biased towards September-October (Fig. S1).**

660

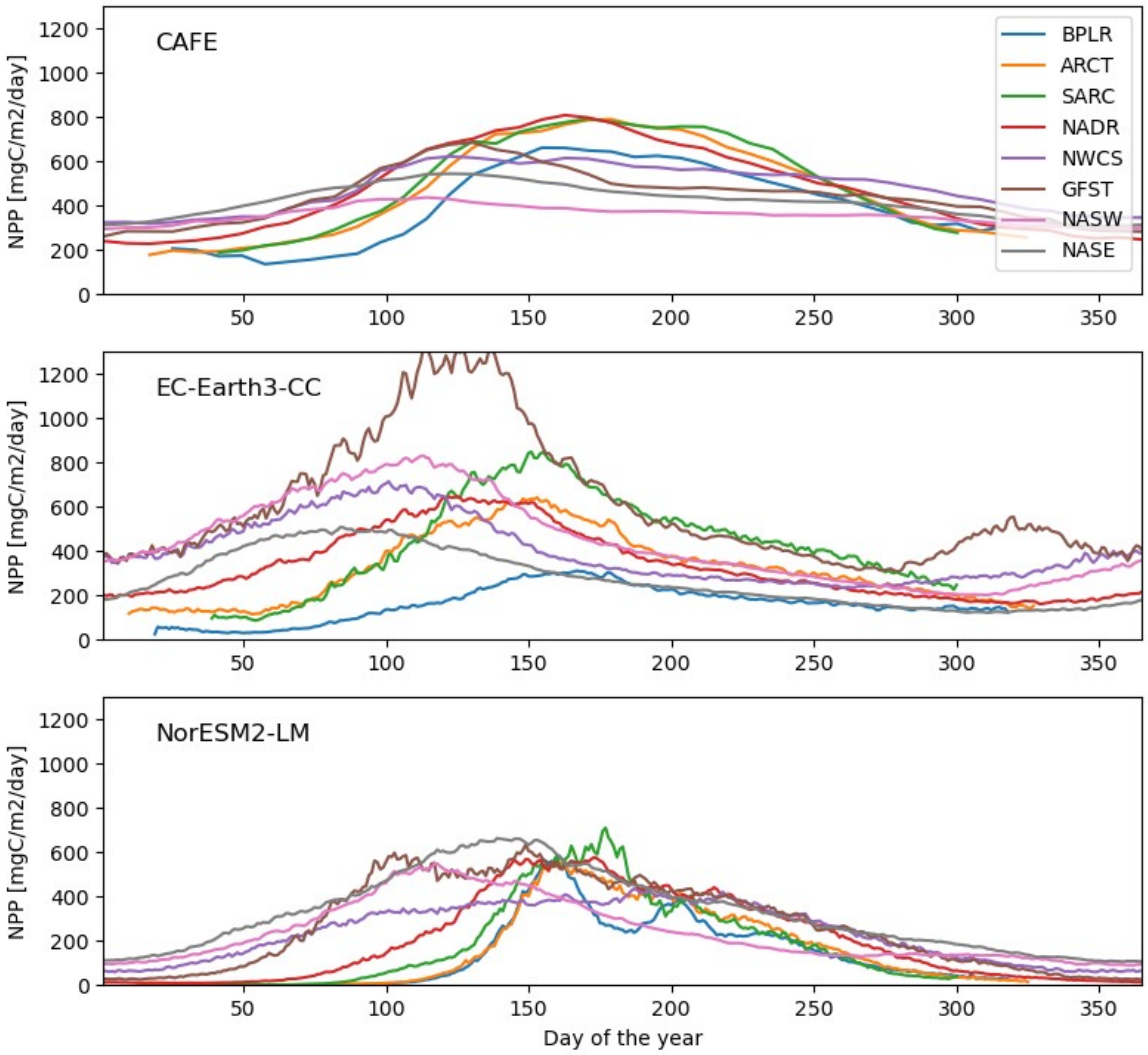


Figure 3: Seasonal cycle of vertically integrated NPP for CAFE, EC-Earth3-CC and NorESM2-LM averaged over 30-60 °N. The model data was masked by the maximum latitude present in the CAFE data to account for the smaller winter domain visible by satellites. A multi-year (2003-2021) average is shown.

	CAFE		EC-Earth3-CC		NorESM2-LM	
Province	**Mean NPP [mgC/m2/day]**	**Day of peak NPP**	**Mean NPP [mgC/m2/day]**	**Day of peak NPP**	**Mean NPP [mgC/m2/day]**	**Day of peak NPP**
BPLR	**405**	**155**	**161**	**166**	**141**	**159**
ARCT	**470**	**179**	**321**	**152**	**160**	**161**
SARC	**525**	**171**	**442**	**150**	**210**	**176**
NADR	**472**	**163**	**332**	**124**	**203**	**172**
NWCS	**477**	**122**	**396**	**100**	**239**	**186**
GFST	**441**	**130**	**608**	**126**	**276**	**148**
NASW	**358**	**114**	**442**	**112**	**238**	**116**
NASE	**41**	**122**	**273**	**83**	**326**	**138**

Table 1. Mean NPP and mean Day of peak NPP over the time period 2003-2021. The ESM data was masked to the real valued CAFE data.

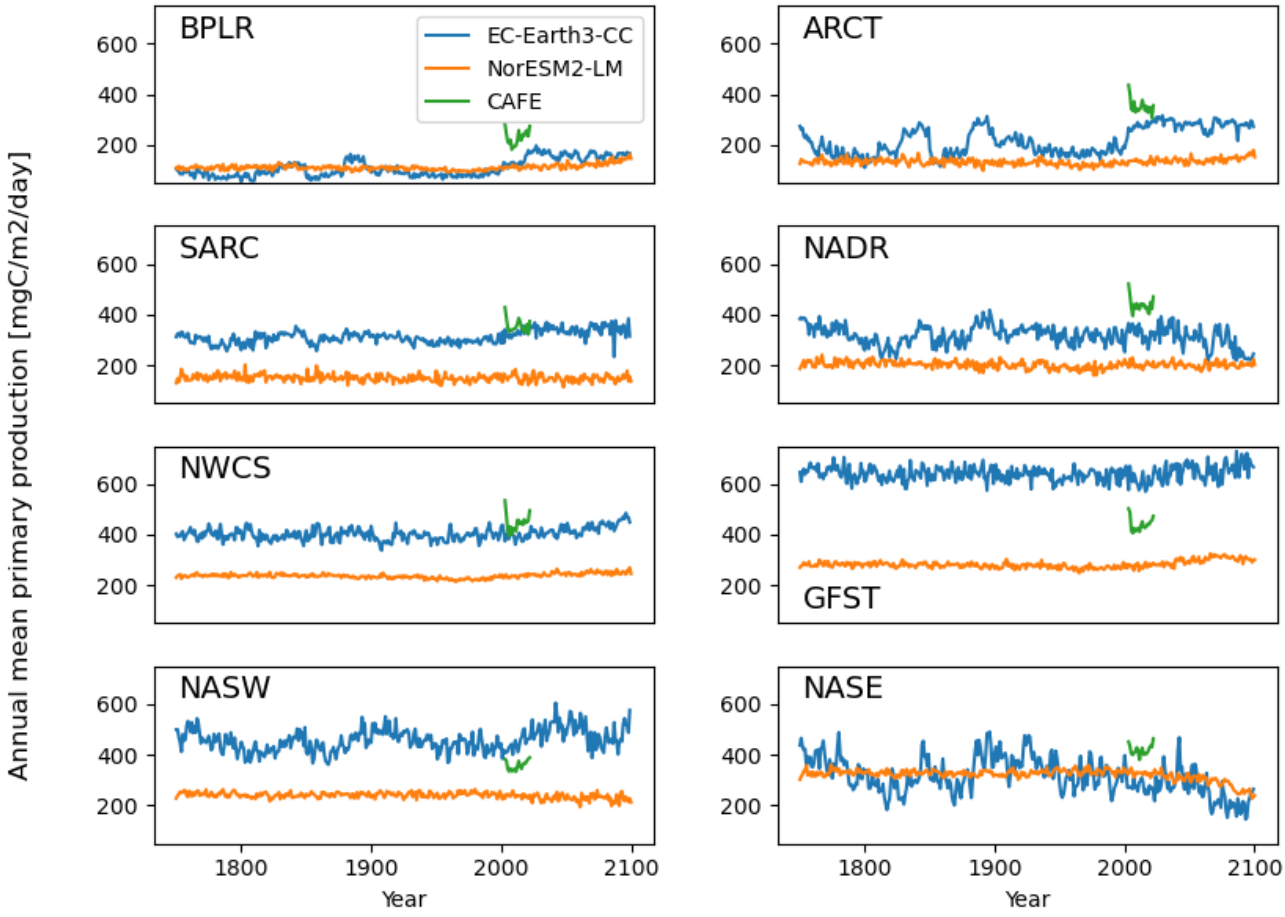


Figure 4: Time-series of annual mean vertically integrated NPP for the different biogeochemical provinces for EC-Earth3-CC (blue), NorESM2-LM (orange) and CAFE (green).

	EC-Earth3-CC 2070 to 2099 mean minus 1850 to 1889 mean		NorESM2-LM 2070 to 2099 mean minus 1850 to 1889 mean	
Province	**NPP**	**Day of peak NPP**	**NPP**	**Day of peak NPP**
BPLR	**79.4**	**-68.2**	**24.7**	**-12.09**
ARCT	**125**	**-25.8**	**15.4**	**-20.8**
SARC	**48.6**	**-8.84**	**-2.25**	**-18.0**
NADR	**-24.1**	**-7.71**	**-3.19**	**-10.1**
NWCS	**42.8**	**-12.7**	**12.2**	**-1.40**
GFST	**20.4**	**-5.73**	**28.0**	**13.3**
NASW	**47.4**	**-5.13**	**-14.1**	**-1.28**
NASE	**-86.6**	**27.0**	**-59.8**	**-12.8**

Table 2. Mean NPP over the period 2070-2099 minus mean NPP over the period 1850-1889 together with the difference in the day of peak NPP for the same periods.

675

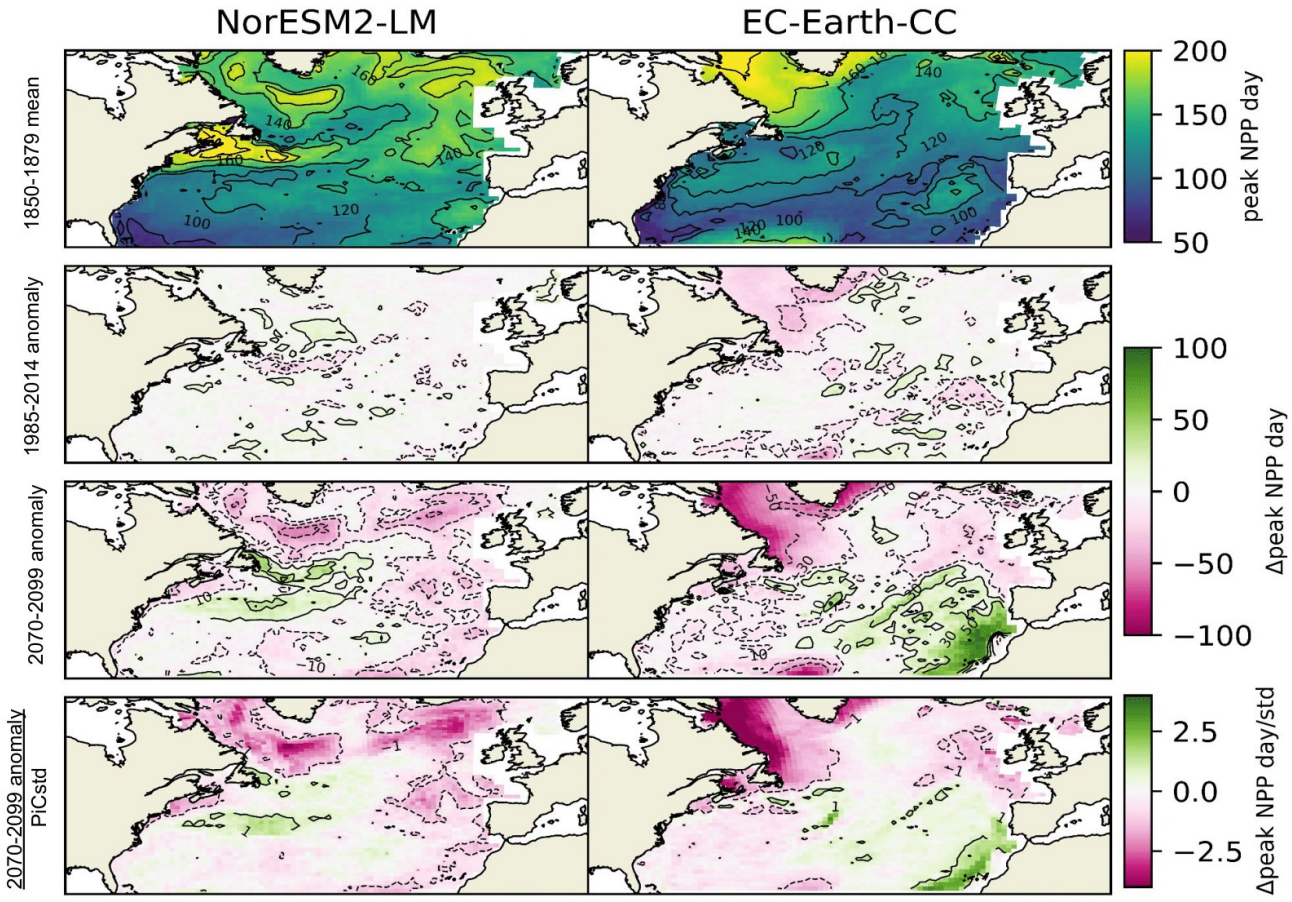


Figure 5: Mean day of peak NPP for NorESM2-LM (left) and EC-Earth3-CC (right) over the years 1850-1879 (top). The second panels show the mean over 1985-2014 minus the 1850-1879 mean. The third panels shows the mean over 2070-2099 minus the 1850-1879 mean (bottom). The bottom panels show the results from the third panels normalised by the yearly standard deviation of the day of peak NPP in the respective PI-control simulations, giving a view of how large the changes are compared to unforced variability.

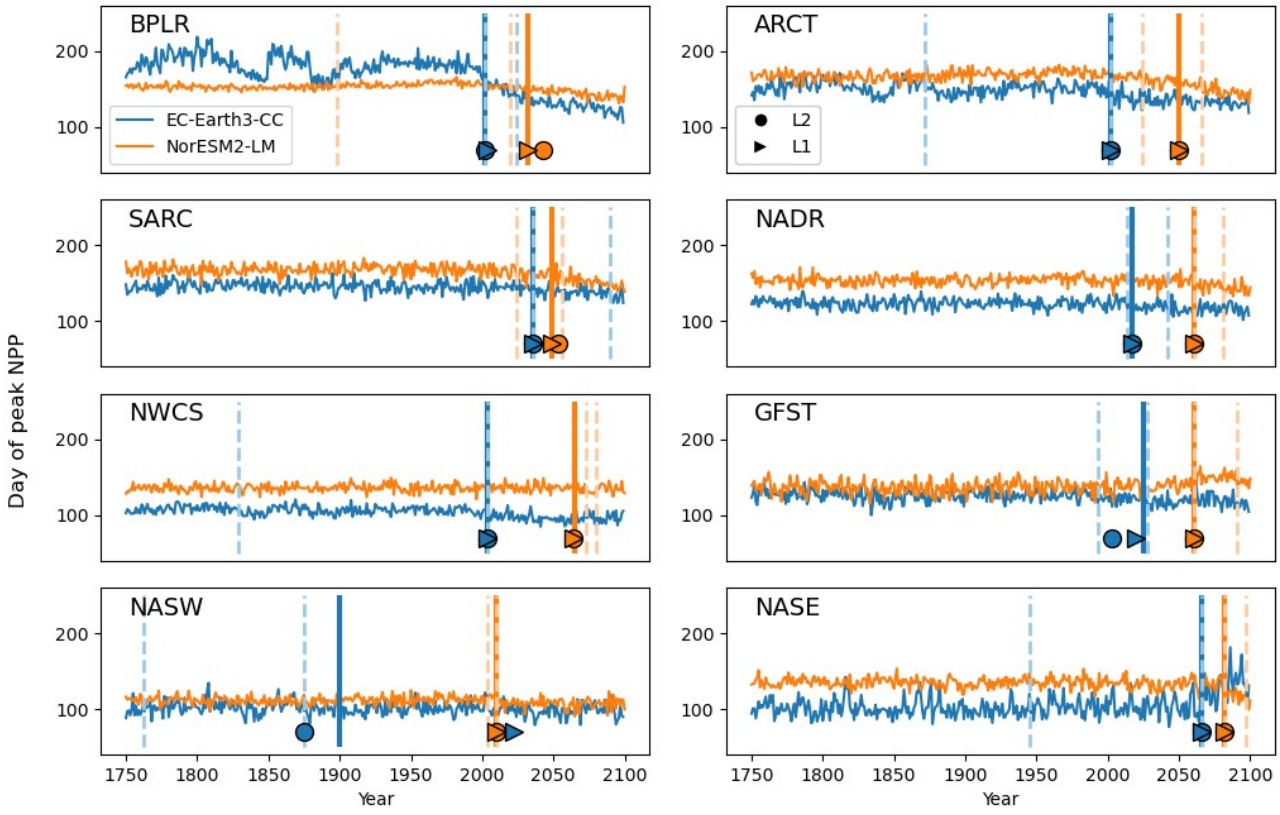
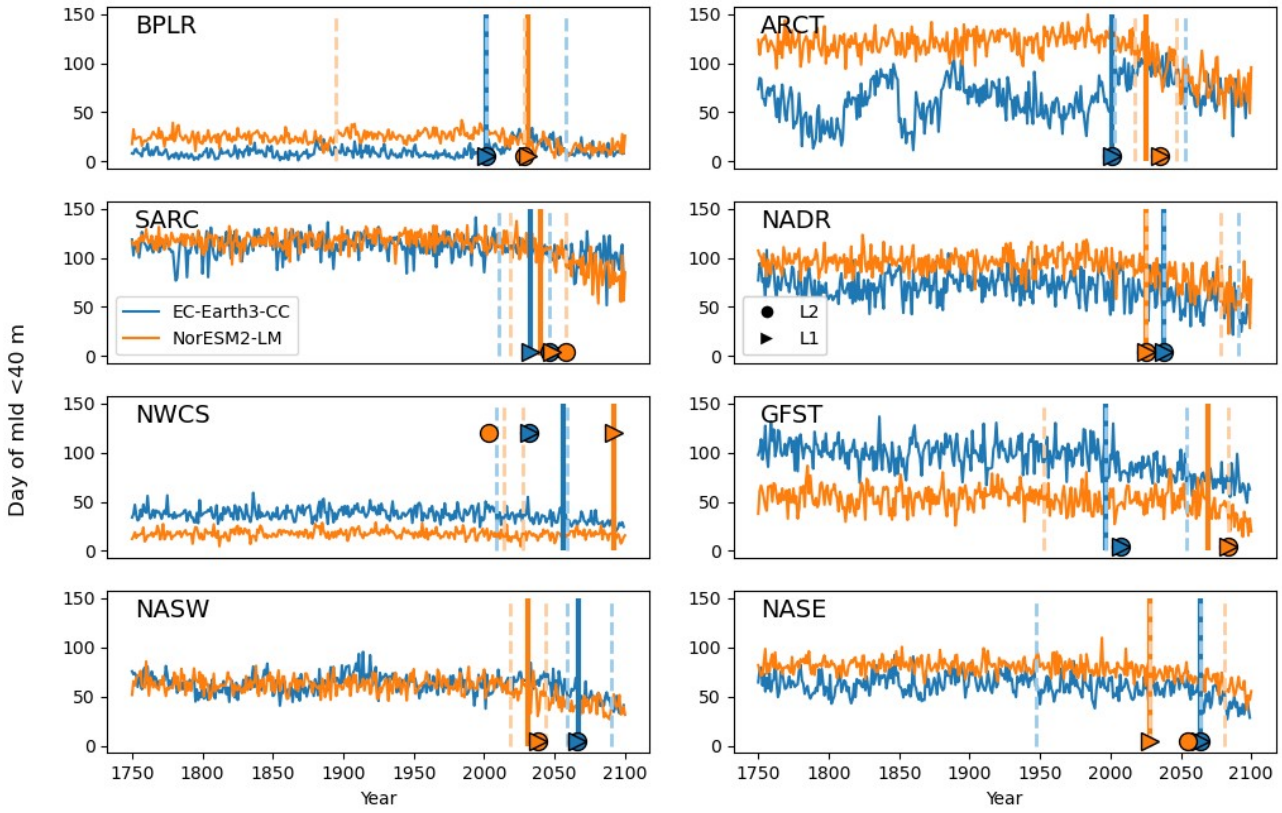


Figure 6: Day of peak NPP per region for EC-Earth3-CC (blue) and NorESM2-LM (orange). The major change points (calculated with a kernel based cost function) in the time-series are marked by the vertical lines. The largest change point is marked by solid lines and the two largest are marked with dashed lines. The centre of the circles represents the largest change point in the time series that corresponds to a change in the mean (L2) while the centre of the triangles represents the largest change point corresponding to a change in the median (L1).



.Figure 7: First day of the year when the mixed layer is 40m or less for EC-Earth3-CC (blue) and NorESM2-LM (orange). The major change points (calculated with a kernel based cost function) in the time-series are marked by the vertical lines. The largest change point is marked by solid lines and the two largest are marked with dashed lines. The centre of the circles represents the largest change point in the time series that corresponds to a change in the mean (L2) while the centre of the triangles represents the largest change point corresponding to a change in the median (L1)

	EC-Earth3-CC Largest Change point [year]		NorESM2-LM Largest Change point [year]	
Province	**Day of peak NPP**	**Day of MLD <40m**	**Day of peak NPP**	**Day of MLD <40m**
BPLR	**2002**	**2001**	**2032**	**2031**
ARCT	**2002**	**2001**	**2050**	**2025**
SARC	**2036**	**2033**	**2049**	**2040**
NADR	**2017**	**2038**	**2061**	**2025**
NWCS	**2004**	**2056**	**2065**	**2092**
GFST	**2025**	**1997**	**2061**	**2069**
NASW	**1900**	**2067**	**2010**	**2031**
NASE	**2066**	**2064**	**2082**	**2028**

Table 3. The table shows the largest change points of the day of peak NPP and the day of MLD<40m time series.

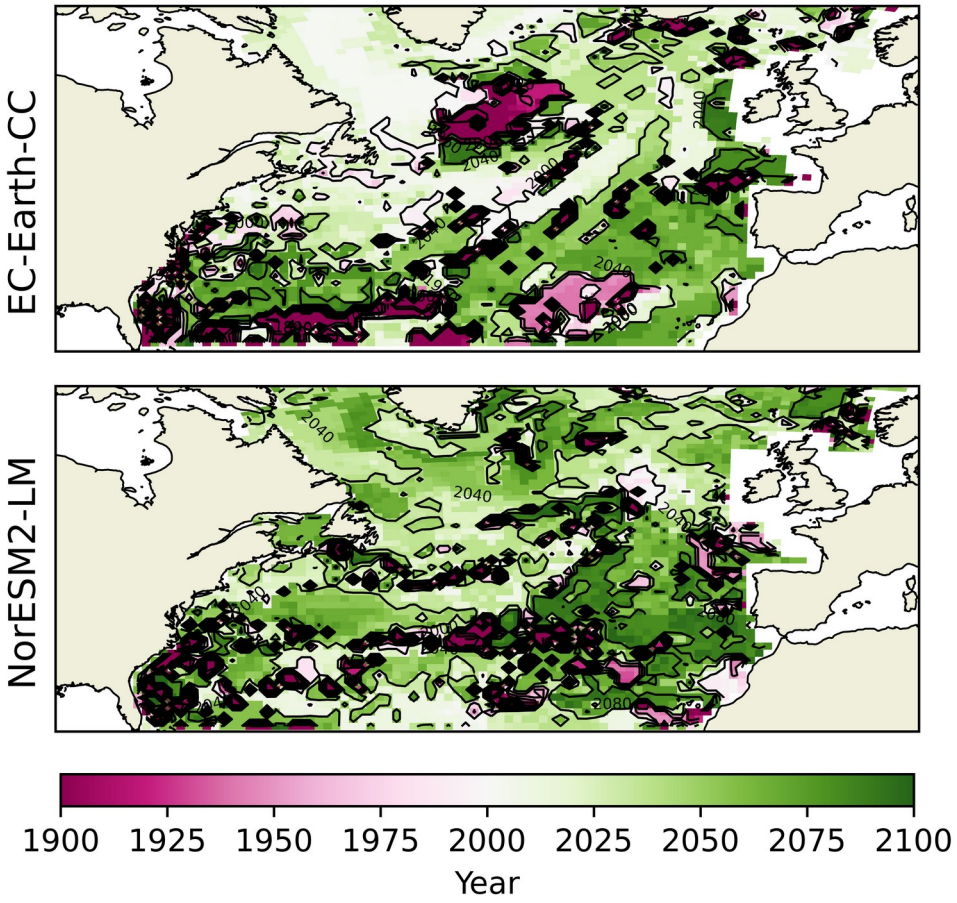


Figure 8: Year of change point of the day of maximum primary production for all grid spaces. The change point algorithm is here set to look for only one change point.

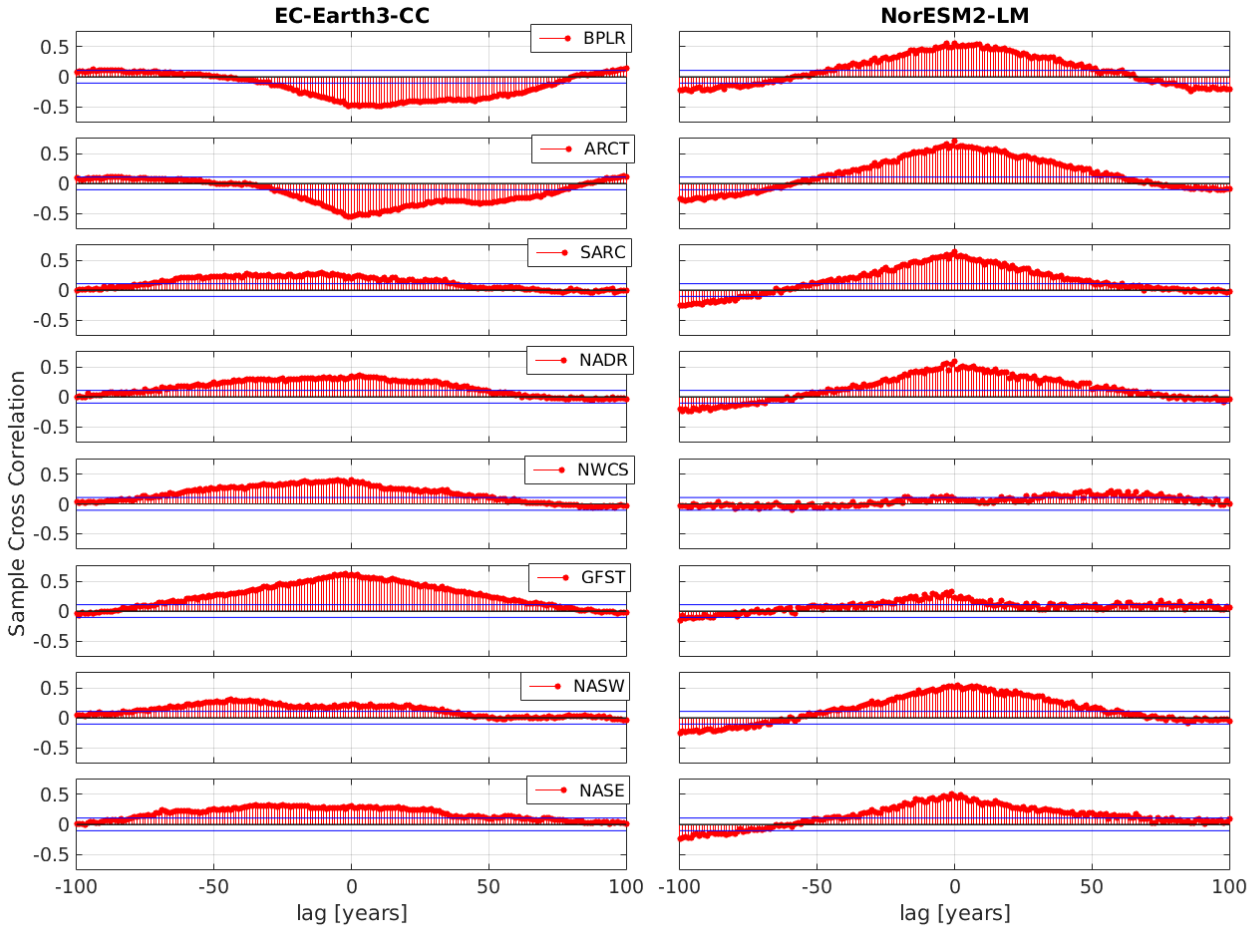


Figure 9: Cross correlation between the day of peak NPP and the first day of mixed layer depth (MLD) shallower than or equal to 40m. Negative lag means that the day of peak NPP proceeds the first day of MLD shallower than 40m, while the opposite holds for positive lag. The horizontal blue lines mark the 95% confidence bounds.