



# Greenland Sea primary production in 1998-2022: monitoring and parameterization using satellite and field data.

Aleksandra Cherkasheva<sup>1</sup>, Rustam Manurov<sup>1</sup>, Piotr Kowalczyk<sup>1</sup>, Alexandra N. Loginova<sup>1</sup>, Monika Zabłocka<sup>1</sup>, Astrid Bracher<sup>2,3</sup>

<sup>1</sup>Institute of Oceanology of Polish Academy of Sciences, Sopot, 81-712, Poland

<sup>2</sup>Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, 27570, Germany

<sup>3</sup>Institute of Environmental Physics, University of Bremen, Bremen, 28359, Germany

Correspondence to: Aleksandra Cherkasheva ([acherkasheva@iopan.pl](mailto:acherkasheva@iopan.pl))

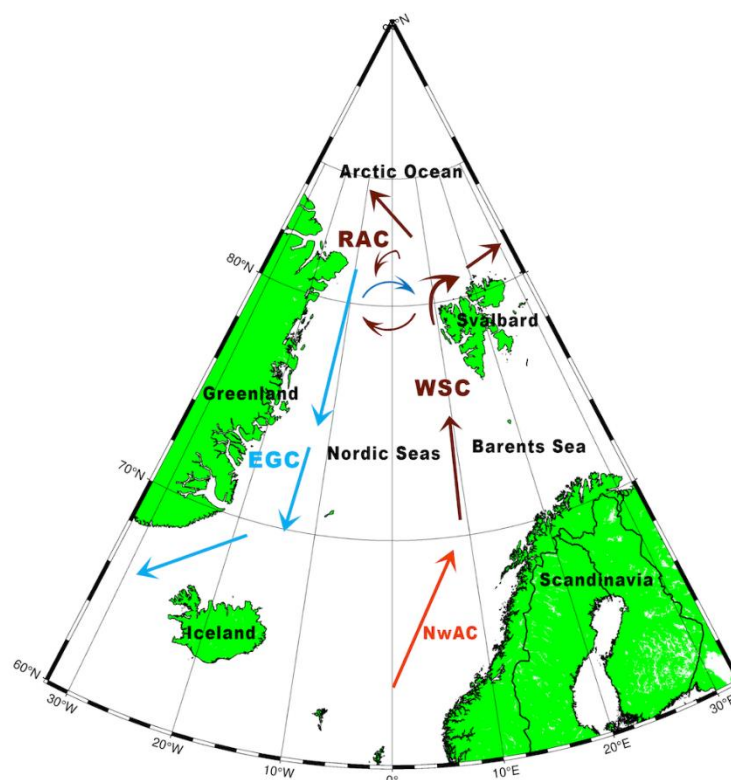
**Abstract.** Phytoplankton are responsible for releasing half of the world's oxygen and for removing large amounts of carbon dioxide from surface waters. Despite many studies on the topic conducted in the past decades, we are still far from a good understanding of ongoing rapid changes in the Arctic Ocean and how they will affect phytoplankton and the whole ecosystem. An example is the difference in net primary production modelling estimates, which differ twice globally and fifty times when only the Arctic region is considered. Here, we aim to improve the quality of Greenland Sea primary production estimates, by testing different versions of primary production model against in situ data and then calculating regional estimates and trends for 1998-2022 for those performing best. As a baseline, we chose the commonly used global primary production model and tested it with different combinations of empirical relationships and input data. Local empirical relationships were taken from measurements by the literature and derived from the unpublished Institute of Oceanology of Polish Academy of Sciences across the Fram Strait. For validation we took historical net primary production <sup>14</sup>C data from literature, and added to it our own gross primary production O<sub>2</sub> measurements to extend the limited validation dataset. Field data showed expected elevated values in the frontal zone together with differences between Arctic and Atlantic-dominated waters, and unexpected good agreement between primary production measured with <sup>14</sup>C and O<sub>2</sub> evolution methods. From all the model setups, those including local chlorophyll a profile and local absorption spectrum and using Level 2 photosynthetically active radiation data best reproduced in situ data. Our modelled regional annual primary production estimates are equal to 346 TgC/year for the Nordic Seas region and 342 TgC/year for the Greenland Sea sector of the Arctic defined as 45°W-15°E, 66°33'N-90°N. These values are higher than those previously reported. Monthly values show a seasonal cycle with less monthly variability than previously reported and with peak values observed in May. No significant increase or decrease in primary production was observed when studying regionally averaged trends. The accuracy of the selected here model setups to reproduce the field data in terms of Root Mean Square Difference is poorer than in the related global studies, but better than in the related Arctic studies.



## 1 Introduction

The Greenland Sea is one of the most productive regions of the Arctic Ocean (Arrigo and Van Dyken, 2011; Sakshaug, 2004), which means that changes in its ecosystem strongly affect the fisheries of the Arctic region. In addition to that, here deep convective mixing takes place (Rey et al., 2000), and large amounts of carbon are possibly transferred to the deep ocean, having a profound effect on the global carbon cycle (Hansell et al., 2009). Most of the atmospheric CO<sub>2</sub> uptake occurs as the Atlantic water is cooled on its way north along the Norwegian coast, and consequently, the Atlantic water contains a high anthropogenic CO<sub>2</sub> content (e.g., Olsen et al., 2006; Sabine et al., 2004; Vázquez-Rodríguez et al., 2009). According to recent work by Chierici et al. (2019), phytoplankton uptake of CO<sub>2</sub> played by far the most important role in the observed CO<sub>2</sub> change throughout the study area and explained up to 89% of the total CO<sub>2</sub> change.

However, correct quantification of phytoplankton uptake of CO<sub>2</sub> (or primary production) is challenging in the area due to several environmental factors. The first factor to consider is the dynamic circulation, characterised by warm and saline Atlantic waters in the eastern part, cold and fresh Arctic waters in the western part, and the large frontal zone with eddies in between these two water masses (Rudels and Quadfasel, 1991; Johannessen et al., 1987), as seen in Figure 1. The second factor is the presence of sea ice that, by melting, influences water stratification and circulation. Formation of sea ice constrains light exposure and nutrient supply of phytoplankton, and its melting enhances it, having a profound effect on phytoplankton blooms (e.g., Cherkasheva et al., 2014; McLaughlin and Carmack, 2010; Skogen et al., 2007; Slagstad et al., 2011; Von Appen et al. 2021). This effect is especially pronounced in the marginal ice zone, which is known to be the localisation of the large phytoplankton blooms in the area (Alexander and Neinauer, 1981; Cherkasheva et al., 2014; Perrette et al., 2011). Recently, it was found that such blooms can even be widespread under the ice, making up a large percentage of total biomass, which cannot be tracked with satellite data (Ardyna and Arrigo, 2020; Ardyna et al., 2020). Such under-ice blooms were documented, for example, in the area along the north coast of Svalbard and two degrees north of it (Assmy et al., 2017).



55 **Figure 1: Schematic image of circulation patterns in the European Arctic. EGC - East Greenland Current; WSC - West Spitsbergen Current; NwAC - Norwegian Atlantic Current; RAC - Return Atlantic Current. Modified from Kraft (2013).**

Another feature to be accounted for in the Arctic Ocean is often occurring deep subsurface chlorophyll maxima (SCMs) that significantly contribute to primary production (PP), but are not detected by ocean colour sensors. In the context of current sea-ice loss in the Arctic, the role of SCM layer on biogeochemical fluxes will potentially increase, and this remains to be quantified (Ardyna and Arrigo, 2020). Commonly used models usually assume either uniform chlorophyll-a (CHL) profile or global relationships between surface CHL and its profile (Antoine and Morel, 1996; Behrenfeld and Falkowski, 1997) with some exceptions (Ardyna et al., 2013; Cherkasheva et al., 2013).

Ocean colour PP models are also challenged in the Arctic Ocean by the limited availability of satellite and field data. The satellite data used as input to the models has large uncertainties and poor spatial coverage in the region due to specific conditions, such as low solar elevation, presence of sea ice (IOCCG, 2015), and extensive presence of clouds in summer (Eastman and Warren, 2010; Intrieri et al., 2002). The coverage of field data that could be used for validation is also limited. We have checked the availability of PP data for the satellite era with continuous ocean colour time series (1998-2022, after the SeaWiFS launch) in the largest field primary production database for the Arctic region ARCSS-PP (Matrai et al., 2013). The percentage of primary production data for the Greenland Sea sector of the Arctic in the ARCSS-PP database is only 0.3%



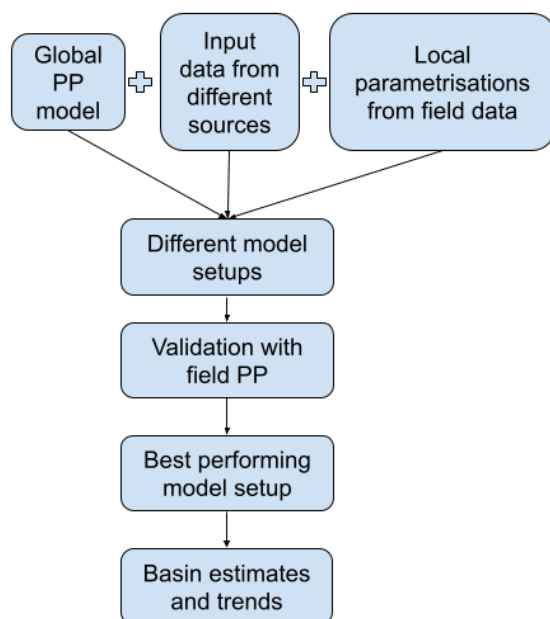
compared to the data from the whole Arctic region. While in terms of area (with land), the Greenland Sea Sector takes up 9% of the Arctic Ocean.

As a result of the above-mentioned factors, the quality of primary production modelling in the Arctic in general and in the Greenland Sea in particular still needs improvement. The most extensive assessment of the performance of primary production models is the series of studies called the Primary Production Algorithm Round Robin (PPARR), which uses the set of activities to compare PP models. According to one of the most cited PPARR studies, models estimating marine primary production range by a factor of two globally (Carr et al., 2006). If only the Arctic region is considered, the factor of difference increases to fifty (Carr et al., 2006). The most recent Arctic PPARR by Lee et al. (2015) showed that PP models need to be carefully tuned for the Arctic Ocean, because most of the models that performed relatively well were those that used Arctic-relevant parameters.

Taking into account these challenges, we have formulated the goals of the current article: 1) develop a model setup adapted for the Greenland Sea based on the global primary production model, 2) obtain more accurate regional primary production estimates, and 3) monitor the primary production variability for the period when ocean colour data are stably available (1998-2022).

## 2 Method and data

The overview of the procedure to choose the best performing model setup and then calculate basin estimates is presented in Figure 2. The global primary production model equation was taken as a base (Morel, 1991), and then the different setups with the variations of the input parameters (from different satellites and climatologies) and the variations of the local parameterizations (from field data) were calculated. The resulting different setups of a model were then validated against the in situ data to choose the best performing model setup. Finally, for the best performing model setup, the basin estimates and temporal trends were calculated.



**Figure 2: Scheme of the procedure applied to calculate Greenland Sea basin estimates and trends.**

## 95 2.1 Region

The spatial limits for the region of study were set at 45°W-15°E, 65°N-84°N for the results to be comparable with the studies of Hill et al. (2013) and Arrigo and van Dijken (2011) who calculated the basin primary production estimates for this part of the Arctic. All the data used were limited to the period of April-September 1998-2022.

## 2.2 Choice of a PP model

100 The choice of a PP model is not a straightforward task, as it was shown that no best model exists for all conditions (Saba et al., 2011). However, during one of the latest PPARR studies the Antoine and Morel (1996) model performed among the best models (in terms of lowest Root Mean Square Difference (RMSD) between in situ and modelled data) in eight out of ten regions that were studied (Saba et al., 2011). The conclusive recommendation of Saba et al. (2011) was that ‘in deeper waters, Antoine and Morel(1996) model might be an excellent choice’ and this encouraged our decision to use the Morel (1991)  
 105 equation, which is in the base of Antoine and Morel (1996) model. We are aware that when the PPARR was conducted for the Arctic region, the Antoine and Morel (1996) model did not perform best, as opposed to the models with Arctic-specific coefficients (Lee et al., 2015). However, it is a model of depth-resolved type, which correlated more with in situ primary production than other model types (Lee et al., 2015) and has high potential globally. We were not able to use Antoine and



Morel (1996) model as its components, i.e. look-up tables are not available online; therefore, here we applied a simplified version of Morel (1991) model adding to it the Arctic-specific coefficients.

## 2.3 Input data

### 2.3.1 Satellite and Reanalysis data

To calculate outputs for different model versions, seasonal basin estimates, and temporal trends, we have used the satellite data described in this section. For all satellite data we've taken monthly composites, as for the daily or 8-day averages the coverage was much poorer. For example, for the whole Arctic, Lee et al. (2015) found only 85 match-ups between satellite and in-situ daily data, only two of which are in the Greenland Sea sector (visual analysis of Figure 1 in Lee et al. (2015)).

The decision to test different satellite data products was also mainly based on the spatial availability of data for the test month of August 2022, since we faced the issue of most of the ocean colour data not being available for the latitudes higher than 78°N in our region of study. The test month of August 2022 was chosen based on the most recent expedition with available field data.

For CHL we have used Copernicus-GlobColour Level 4 CHL (SeaWiFS, MODIS, MERIS, VIIRS-SNPP & JPSS1, OLCI-S3A & S3B) monthly and interpolated data (resolution: 4 km) and Globcolour Level 3 CHL (MERIS, MODIS, VIIRS) for monthly data from Case1 waters (resolution: 4 km). Copernicus-GlobColour CHL for the test month showed 83% coverage, while Globcolour CHL showed 69% coverage.

For Photosynthetically Available Radiation (PAR) we have used Eumetsat OLCI Level 2 PAR daily data, which we combined into monthly composites (resolution: 1200 m at nadir) and Globcolour Level 3 MODIS/VIIRS merged PAR monthly product (resolution: 4 km). Eumetsat PAR for the test month showed 87% coverage, while Globcolour PAR showed 81% coverage. For OLCI Level 2 PAR data for all the months from April until September were available only for 2022, thus we took 2022 data and regarded it as climatology for further calculations. As an alternative for the satellite PAR data, we used reanalysis climatological estimates. Reanalysis estimates that incorporate observations and numerical simulations with data assimilation of monthly average downward solar radiation flux at ~1.9° resolution were obtained from NOAA/National Centers for Environmental Prediction and converted to PAR by multiplying by  $C=0.43$ , PAR to the shortwave radiation fraction (Olofsson et al., 2007).

### 2.3.2 Field data used for local coefficients

For local empirical coefficients we've used vertical CHL parameterisation for the Greenland Sea based on analysis of 1199 profiles from Cherkasheva et al. (2013) and particulate absorption data.



### 2.3.2.1 Particulate absorption data

For the particulate absorption database, we combined the measurements obtained during the cruise to the Fram Strait onboard of RV ‘Kronprinz Haakon’ in 2021 and data from Kowalczyk et al. (2019) for 2014-2016 from the same area.

140 Samples for particulate absorption analyses were taken from three depths at selected stations (5, 15 and 25 m) and filtered onto 0.7  $\mu\text{m}$  GF/F filters. The filters were stored at -80 deg. C freezer and analysed at home laboratory of the Institute of Oceanology of Polish Academy of Sciences (IOPAN). Filter papers with deposited particles were measured with Lambda 850, (Perkin Elmer, USA) in the spectral range 300 - 850 nm with 1 nm resolution, equipped with the integration sphere using the transmission-reflection method described by Tassan and Ferrari (2002), and Tassan and Ferrari (1995). The phytoplankton  
145 pigment absorption coefficient,  $a_{\text{ph}}(\lambda)$ , was calculated using the standard procedure described in Kowalczyk et al. (2019). The CHL specific phytoplankton pigments absorption coefficient at 443 nm,  $a_{\text{ph}}^*(443)$ , was calculated for a given sample as the ratio of  $a_{\text{ph}}(443)$  to Chla (Bricaud et al., 1995).

### 2.3.3 Field data used for validation

Due to the limited availability of publicly available field primary production data in the region (Section 1), for validation, we  
150 have used data obtained with two different methods: net primary production data obtained with the  $^{14}\text{C}$  method (Steemann Nielsen, 1952) (NPP\_C14) (see Section 2.3.3.1), and gross primary production obtained with optode  $\text{O}_2$  measurements (GPP\_O2) (see Section 2.3.3.2).

#### 2.3.3.1 Net primary production data with $^{14}\text{C}$ measurements.

For our region of study for 1998-2022, five points were available from the Matrai et al. (2013) database, and 19 points from  
155 RV Dana and RV Triton cruises (Richardson et al., 2005). We have also added nine data points from the two Spitsbergen fjords (Iversen and Seuthe, 2011, Piwosz et al., 2009; Wiktor and Wojciechowska, 2005). As a result, we got a total of 33 data points for the period of April-September 1999-2006. PP data were logarithmically transformed (base 10) (Campbell, 1995) before being analysed in this study.

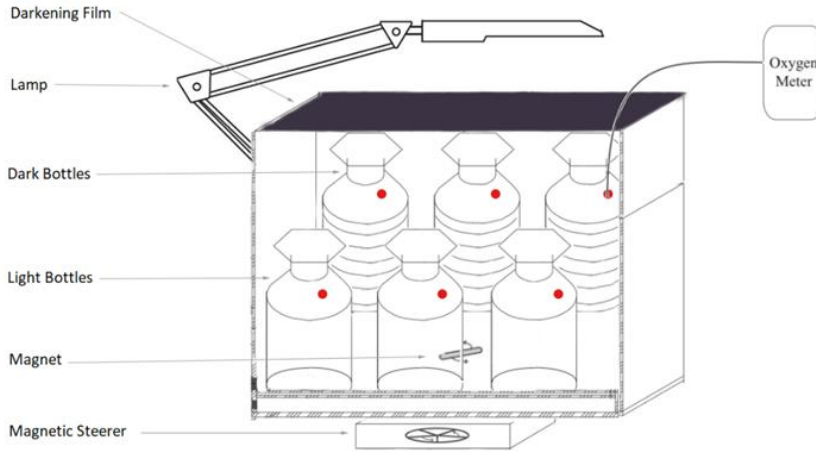
#### 2.3.3.2 Gross primary production with optode $\text{O}_2$ measurements.

160 These are data obtained and measured specially for this study during RV ‘Kronprinz Haakon’ 2021 Fram Strait cruise and RV ‘Maria S Merian’ 2022 Greenland Fjords cruise. This data set has 13 points for August 2021-2022.

The estimation of Gross Primary Production from the optode  $\text{O}_2$  measurements was derived with the method adapted from Campbell et al. (2016), and also described in Section 6 of the IOCCG Protocol Series (2022). The incubation bottles were overfilled with the seawater sample to prevent the formation of a headspace during closure of the glass stopper and placed into  
165 containers with seawater under continuous illumination and mixing. The incubation temperature in the cold room was set at +4°C, which was representative of the average conditions of the seawater in situ. The determination of the gross primary



production by oxygen modification was carried out with the help of oxygen measurements obtained using a Fixbox4 optical sensor (PreSens HMBH, Germany), which non-invasively utilises optical oxygen sensor spots installed in the 250 ml white glass bottles. Oxygen respiration was determined in foil-wrapped bottles of the same volume with the same optical sensor. The bottles were incubated in a thermostabilised luminostat at light levels representing surface, 15 m and 25 m Photosynthetic Available Radiation in the case of a 2021 cruise and surface, 40 m and 60 m in the case of a 2022 cruise (Figure 3). Each of the bottles had a duplicate in case of the 2021 cruise and a triplicate in case of the 2022 cruise. The light level reproducing surface light conditions was selected at the beginning of the cruise by choosing the light intensity of the lamp and the appropriate level of darkening film. The selection was made by visually comparing the measurement of downwelling irradiance spectra on deck with the TrioS downwelling irradiance sensor, and the measurements of the different combinations of light intensity and darkening film with the same downwelling irradiance spectra sensor in the cold room inside the incubation box. The light levels were assumed to be constant throughout the day, as both of the cruises took place in the beginning of August above 70°N in the period of the midnight sun at these latitudes. The oxygen concentration dynamics was determined every 6 hours for 24, 48 or 72 hours depending on the initial CHL concentration. Samples of the same Niskin bottles were filtered for CHL measurements in parallel with oxygen incubations. Optode sensors were calibrated for bottles in use prior to the cruise using 0% and 100% dissolved oxygen standards of nitrogen-saturated water and oxygen-saturated water, respectively. After measurements were done, the values were averaged for duplicates in case of 2021 cruise and triplicates in case of 2022 cruise. Then, the rates of change of oxygen in dark bottles (an estimate of community respiration, CR) and that in clear bottles (an estimate of net community production, NCP) were calculated by subtracting initial dissolved oxygen concentrations from the dissolved oxygen concentrations measured after incubation under light and dark conditions, respectively (Carritt and Carpenter, 1966; Carpenter, 1995). GPP was derived by summing NCP and CR (Carritt and Carpenter, 1966; Duarte et al., 2011). In case the measurements were taken for 48 or 72 hours, they were weighted by the incubation time to achieve a value for 24 hours. Then, to convert the O<sub>2</sub> production rates into <sup>14</sup>C incorporation rates, the specific photosynthetic quotient (PQ) value was used. Although no PQ value has been derived for the Arctic Ocean, a value of 1.25, proposed by Williams et al. (1979), has been widely applied in this region to convert O<sub>2</sub> molar stoichiometry units into C (i.e., Duarte and Agustí, 1998, Sanz-Martin et al., 2019, Vaquer-Sunyer et al., 2013). Therefore, we have used a PQ value of 1.25 and then integrated the data for the 2021 and 2022 cruises till the deepest depth available in 2021, which is 25m.



**Figure 3: Gross Primary Production Measurements Setup during the Kronprinz Haakon Fram Strait 2021 and Maria S Merian 2022 cruises. Three of such boxes were installed, one for each depth.**

## 2.4 Primary production model and versions of its setup

As a base model to calculate PP from satellite input data, we took the Morel (1991) simplified case model, which is a wavelength-integrated depth-resolved primary production model:

$$P = \left(\frac{12}{4.6}\right) CHL_{tot} \overline{PAR}(0^+) \overline{a^* \varphi_{\mu}} \quad (1)$$

For the baseline, we took values of spectrally-averaged constant CHL-specific absorption coefficient ( $a^*$ ), and average quantum yield valid for the euphotic layer ( $\overline{\varphi_{\mu}}$ ) from Morel (1991).  $CHL_{tot}$ , which is a total CHL integrated over the euphotic layer, was calculated from the surface CHL based on Morel and Berthon (1989) method. Following Morel and Berthon (1989), the model of Morel (1988) was used for the estimation of both  $Z_{eu}$  (euphotic layer depth) and  $CHL_{tot}$ .

After the baseline equation was set, there were several variations of the components of the model that we have tested when validating the model output against field PP data. The groups tested were:

1. Source of CHL data (see Section 2.3.1 above):
  - a) Hermes ACRI Globcolour Level 3 CHL (CHL\_L3)
  - b) Copernicus-GlobColour Level 4 CHL (CHL\_L4)
2. Source of photosynthetically active radiation data (see Section 2.3.1 above):
  - a) NOAA/NCEP Reanalysis PAR (PAR\_R)



- b) EUMETSAT Level 2 OLCI PAR (PAR\_L2)
  - c) Hermes ACRI Globcolour Level 3 MODIS/VIIRS merged PAR (PAR\_L3)
3. Shape of CHL vertical profile calculated using:
- a) Global relationship between surface CHL and CHL profile used in Antoine and Morel (1996) model which is the satellite data adapted version of Morel (1991) model. Either uniform or Morel and Berthon (1989) shape of the profile was assumed; the distinction between the two cases was made based on latitude (for high latitudes  $>70^{\circ}\text{N}$  the profile is assumed to be uniform) or on the mixed layer depth position. If the mixed layer depth was larger than 100 m or exceeded the euphotic depth, the profile was also assumed to be uniform (Antoine et al., 1996). The mixed layer depth values were taken from Boyer et al. (2018) climatology data (PROFILE\_GLOB)
  - b) Local Greenland Sea relationship between surface CHL and CHL profile approximated with a Gaussian fit based on the analysis of 1199 profiles (Cherkasheva et al., 2013) (PROFILE\_LOC)
4. The fraction of light spectrum. We have tested:
- a) Photosynthetically Available Radiation (PAR) used in the majority of primary production models (e.g., Lee et al., 2015) and Morel (1991) simplified model version (SPECTR\_PAR)
  - b) Photosynthetically Usable Radiation (PUR) used in Antoine and Morel (1996), which is a fraction of PAR absorbed by phytoplankton. To obtain PUR we multiplied PAR by a mean CHL-specific absorption spectrum computed from measurements for 14 phytoplankton species, grown in culture, and normalised with respect to a maximum value (Morel, 1991) (SPECTR\_PUR\_GLOB)
  - c) PUR accounted for the Greenland Sea species of phytoplankton. To obtain this version of PUR, the PAR values were multiplied by the climatology of the mean CHL-specific absorption calculated from the unpublished data from RV ‘Kronprinz Haakon’ 2021 Fram Strait cruise and data from Kowalczyk et al. (2019) for 2014-2016 (see Section 2.3.2.1). The mean CHL-specific absorption value was computed as the spectrally averaged percentage of absorption related to maximum value, which was assumed to be 100%. As the measurements did not cover all the area, for each point of the grid the closest available value was taken. (SPECTR\_PUR\_LOC)
5. Integration depth. We have compared integration of CHL data for two depths:
- a) Euphotic layer depth (Zeu) calculated using the Morel (1988) model following Morel and Berthon (1989), which was later confirmed by Morel and Maritorena (2001) (DEPTH\_ZEU)
  - b) Depth of the ‘extended’ productive layer (D) which is defined by Zeu multiplied by 1.5. D was introduced by Morel (1991) as in some cases Zeu does not cover the subsurface CHL maximum, thus giving false estimates of the integrated CHL profile (DEPTH\_D).



For comparison, we have also added to the further analysis 1) publicly available PP output of Global Ocean Colour (Copernicus-GlobColour), Bio-Geo-Chemical, L4 (monthly and interpolated) from Satellite Observations (1997-ongoing) based on Antoine and Morel (1996) algorithm, and 2) Behrenfeld et al. (1998) PP model based on Globcolour Level 3 CHL.

## 2.5 Model performance assessment

Field PP data were matched with modelled PP data using two methods: 1) commonly used method of matching field data location with satellite cell, cells with missing modelled PP data were excluded. 2) method to increase the number of collocations further on named as «interpolated»: field data location was matched with a satellite cell; if the modelled PP data was missing it was interpolated using pandas.DataFrame.interpolate linear interpolation in Python.

Model performance was assessed using *RMSD* for each participating model version, where *N* is the number of observations:

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (\log NPP_m(i) - \log NPP_d(i))^2} \quad (2)$$

Generally, the smaller a *RMSD* value becomes, the better a model performs. The *RMSD* consists of two components: 1) bias representing the difference between the means of in-situ and model data (*BIAS*), providing the measure of how well the mean is modelled;

$$BIAS = \overline{\log NPP_m} - \overline{\log NPP_d} \quad (3)$$

and 2) unbiased *RMSD* (*uRMSD*), providing the measure of how well variability is modelled.

$$uRMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N ((\log NPP_m(i) - \overline{\log NPP_m}) - (\log NPP_d(i) - \overline{\log NPP_d}))^2} \quad (4)$$

The target diagram (Joliff et al., 2009) was used to visualise *BIAS* (y-axis), *uRMSD* (x-axis) and *RMSD* (distance from a center) on a single plot. To plot a Target diagram, *BIAS* and *uRMSD* are normalized by the standard deviation of  $\log NPP_{insitu}$ . Normalized bias (*BIAS\**) is thus defined as:

$$BIAS^* = BIAS / \sigma_{insitu} \quad (5)$$

Normalized *uRMSD* (*uRMSD\**) is defined as:

$$uRMSD^* = \frac{uRMSD}{\sigma_{insitu}} \text{ (if } \sigma_{model} > \sigma_{insitu}) = -\frac{uRMSD}{\sigma_{insitu}} \text{ (if } \sigma_{model} < \sigma_{insitu}) \quad (6)$$

, where  $\sigma_{insitu}$  is the standard deviation of  $\log NPP_{insitu}$  and model is the standard deviation of  $\log NPP_{model}$ .

As an additional characteristic to assess the skill of the model, the Pearson correlation coefficient (*r*) was calculated. The closer *r* is to 1, the better a model version performs.

Skill statistics for all versions of the model normalized by the standard deviation are visually presented in the Target diagrams (Figure 7). The closer a model symbol is to the origin, the better a model performs.

As in Lee et al. (2015), model versions performing relatively better than the others were selected for further analysis using the two criteria: (1) bias was close to 0 ( $-0.1 < \text{bias} < 0.1$ ), and (2) Pearson's correlation coefficient (*r*) was greater than the model average (0.25).

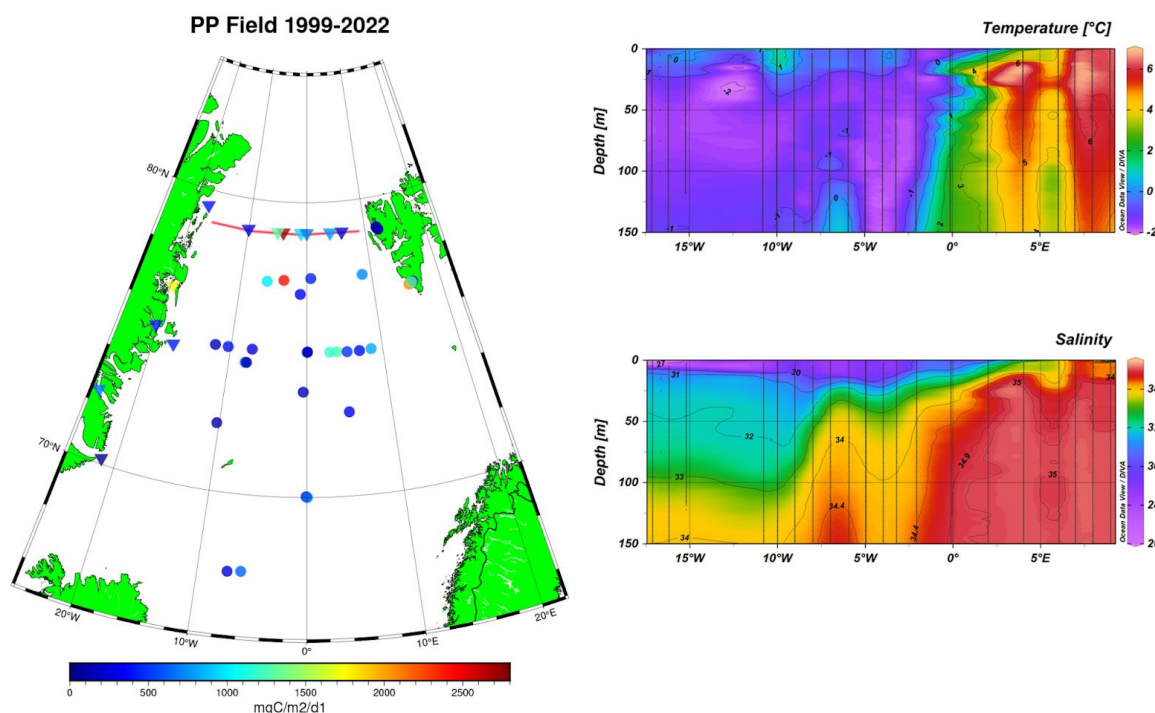


## 2.6 Calculation of trends and basin estimates

To compare our estimates to previous studies, we calculated basin primary production for two regions. The first region was set according to Hill et al. (2013) and covered the Nordic Seas region of the EASE grid representation of the Arctic Ocean starting at 65°N. Each cell in the grid was 100 x 100 km. As in the versions of the EASE grid available online now the Nordic Seas region borders have changed compared to Hill et al. (2013) version, we have digitized the map from Hill et al. (2013) using WebPlotDigitizer v4.6. Hill et al. (2013) paper includes Nordic Seas annual PP basin estimates accounting for SCM, and pan-Arctic monthly estimates accounting for SCM. However, monthly estimates for the Nordic Seas do not include SCM influence, thus we similarly to the method used in the paper assumed an underestimation of 75% for the calculations without SCM and corrected for that. The second region was set as in Arrigo and van Dijken (2011) and Arrigo and van Dijken (2015) at 45°W-15°E, 65°N-84°N. Trends were calculated using least squares linear fit.

## 3. Results and discussion

### 3.1 Field primary production data





**Figure 4: Left: Locations of field primary production data for 1999-2022 used for model validation. Circles indicate net primary production obtained with the  $^{14}\text{C}$  method for 1999-2006. Triangles indicate the gross primary production obtained with the dissolved  $\text{O}_2$  method and converted to  $\text{mgC}/\text{m}^2/\text{day}$  for 2021-2022. The red line shows the location of cross sections to the right. Right: Cross sections of temperature and salinity CTD measurements from the 2014 Fram Strait cruise across  $79^\circ\text{N}$ .**

290 Field primary production data were available for the period 1999-2006 for NPP\_C14 and for the years 2021-2022 for GPP\_O2. Looking at the spatial distribution of the field data in Figure 4, one can see the slightly larger values on the eastern Atlantic waters side as opposed to lower values on the polar waters side to the West.

In between the polar waters and the Atlantic Waters lies a frontal zone. The location of the frontal zone can clearly be seen in the temperature and salinity cross section across the Fram Strait at  $79^\circ\text{N}$  in 2014 (Figure 4, right side). The location of this

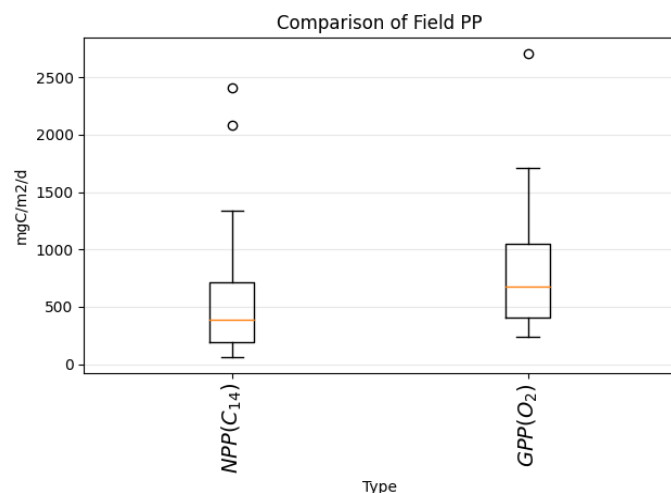
295 frontal zone is quite stable through the years, also confirmed, for example, by Granskog et al. (2012) and Gonçalves-Araujo et al. (2016). At the frontal zone (see area  $76^\circ\text{N}$ - $80^\circ\text{N}$ ,  $2^\circ\text{W}$ - $6^\circ\text{W}$  on the map in Figure 4) both the NPP and GPP are larger, with maximum values observed in the area for the both measurement methods.

These general patterns are in good agreement for both methods of data collection, although they were collected in different years and the type of measured primary production is different. The mean values of the data are 1.5 times higher for GPP\_O2

300 -  $884 \text{ mgC}/\text{m}^2/\text{day}$  for GPP\_O2 and  $589 \text{ mgC}/\text{m}^2/\text{day}$  for NPP\_C14. The range of values is similar for both methods, being just slightly higher for the GPP\_O2 values (Figure 5). In general, GPP\_O2 is supposed to have higher values than NPP\_C14 (e.g., Robinson et al. (2009), IOCCG Protocol Series (2022)). In the literature comparisons between the GPP\_O2 and NPP\_C14 show on average a twofold difference between these two estimates in the Arctic,  $21$ - $70 \text{ mgC}/\text{m}^3/\text{d}$  for the surface  $^{14}\text{C}$ -NPP, and  $55$ - $168 \text{ mgC}/\text{m}^3/\text{d}$  for the surface GPP\_O2 (Matrai et al., 2013, Sanz-Martin et al., 2019, Vaquer-Sunyer et al., 2013).

305 The fact that our GPP\_O2 values in general are not as high as expected could come from the difference in the integration depth between the two methods. All the GPP\_O2 data were integrated up to 25 m, while the NPP\_C14 data have different integration depths, varying from 30 m to 60 m. Unfortunately, it was not possible to have the same integration depths for both methods, as the majority of the  $^{14}\text{C}$  data from the literature had only integrated values. To sum up, the results show that similar spatial patterns for PP are obtained for the two datasets accomplished with different PP methods, collected in different years and

310 seasons. The standard deviation of the triplicate measurements that were then converted to GPP was  $38 \text{ mgC}/\text{m}^3/\text{day}$  for the difference between the first and last light bottle measurements, and  $50 \text{ mgC}/\text{m}^3/\text{day}$  for the difference between the first and last dark bottle measurements.



**Figure 5: Comparison of values range for field NPP measured with <sup>14</sup>C method (left bar, historical data, see Section 2.3.3.1 for details) and our GPP measured with dissolved O<sub>2</sub> evolution (right bar, see section 2.3.2 for details). The red lines indicate the median values.**

### 3.2 Sensitivity study

To test satellite based PP model sensitivity to changes in different model configurations, we have assessed which parameter affected the model output most. For this, we have calculated the RMSD difference between matchups of field PP and modelled PP within each of the groups of tested parameters as listed in Section 2.4. For groups assessing the source and the vertical distribution of CHL data, the difference was least. Minimum was observed for group 1, source of CHL data, with a difference in RMSD of 0.001, and for group 3, difference in the shape of a CHL profile, the difference in RMSD was also low, 0.011. Group 5, difference in integration depth, a showed also quite low RMSD difference of 0.032. The largest RMSD difference was observed for the groups assessing the light field - group 2, source of PAR data, showed 0.254 difference, and group 4, choice of PAR or PUR spectra, showed 0.195 difference.

This result is contradictory to most of the PP model sensitivity studies, where differences in CHL data generally have more influence on the final output than PAR data (e.g. Lee et al., 2015). This could be explained by the fact that in this study the CHL data differ only in processing algorithms, while the PAR data differ in both processing algorithms and sensors. Here, as for our current knowledge, take for the first time Level 2 PAR data from Eumetsat for PP modelling, which has larger values than Glocolour PAR Level 3 data. The difference within group 4, i.e. choosing the PAR spectrum, as in the simplified version of Morel (1991), or the PUR spectrum, as in the full version of Antoine and Morel (1996), is basically a choice between two physically different models, which explains a large difference.





345 **Figure 6: Bar plots illustrating the range of PP values for each model setup for 41 field points of both NPP\_C14 and GPP\_O2 with spatially interpolated chlorophyll data. Red line: median, bubbles: outliers, box: interquartile range. Top: model versions integrated to the euphotic layer depth, bottom: model versions integrated to the productive layer depth. Model versions have identification number [a,b,c,d,e]; [a]: 3 - Globcolour Level 3 CHL, 4 - Copernicus-GlobColour Level 4 CHL; [b]: 1 - NOAA/NCEP Reanalysis PAR, 2 - EUMETSAT Level 2 OLCI PAR, 3 - Globcolour Level 3 PAR; [c]: 0 - Global CHL profile following Morel and Berthon (1988), 1 - Local CHL profile following Cherlkasheva et al. (2013); [d]: 0 - no coefficient applied to PAR, 1 - PAR converted to local PUR, 2 - PAR converted to global PUR; [e]: 0 - profiles integrated till euphotic layer depth, 1 - profiles integrated till productive layer depth. Model versions reproducing field data best in terms of bias and correlation coefficient are highlighted (see Section 2.5), two models selected for further calculations are additionally outlined in blue. ‘A&M’ refers to Copernicus-GlobColour L4 PP data based on Antoine and Morel (1996). ‘PP Behrenfeld’ refers to our own calculations of PP based on Behrenfeld(1998) and Globcolour L3 CHL. The last bar corresponds to in situ primary production.**

Now we present all the six versions of the field dataset, and how they passed the performance test described in the last paragraph in Section 2.5. In the case of the NPP\_C14 datasets, many models passed the performance test and were able to reproduce the field data. These were ten combinations in the case of an interpolated dataset, and eleven combinations in the case of the noninterpolated dataset. In the case of the GPP\_O2 data set with 12 points, no model combinations passed the performance test. For the NPP\_C14+GPP\_O2 dataset, five combinations performed well for both interpolated and noninterpolated cases (see Table 1).

Field dataset name	Number of collocations	Best performing models
NPP	19-33	[4,2,1,1,0], [4,2,0,1,0], [4,2,1,2,0], [4,2,0,2,0], [3,2,1,1,0], [4,2,1,1,1], [4,2,1,2,1], [4,3,1,0,1], [4,3,0,0,1], [4,1,1,0,1], [4,1,0,0,1]
NPP_interpolated	32	[4,2,1,1,0], [4,2,0,1,0], [4,2,1,2,0], [4,2,0,2,0], [4,2,1,1,1], [4,2,1,2,1], [4,3,1,0,1], [4,3,0,0,1], [4,1,1,0,1], [4,1,0,0,1]
NPP+GPP	25-45	[4,2,1,1,1], [4,2,0,1,1], [3,3,1,1,1], [3,2,1,1,0], [3,3,0,1,0]
NPP+GPP_interpolated	41	[4,2,1,1,1], [4,2,0,1,1], [3,2,1,1,1], [3,2,1,1,0], [3,2,0,1,0]
GPP	6-12	-
GPP_interpolated	9	-

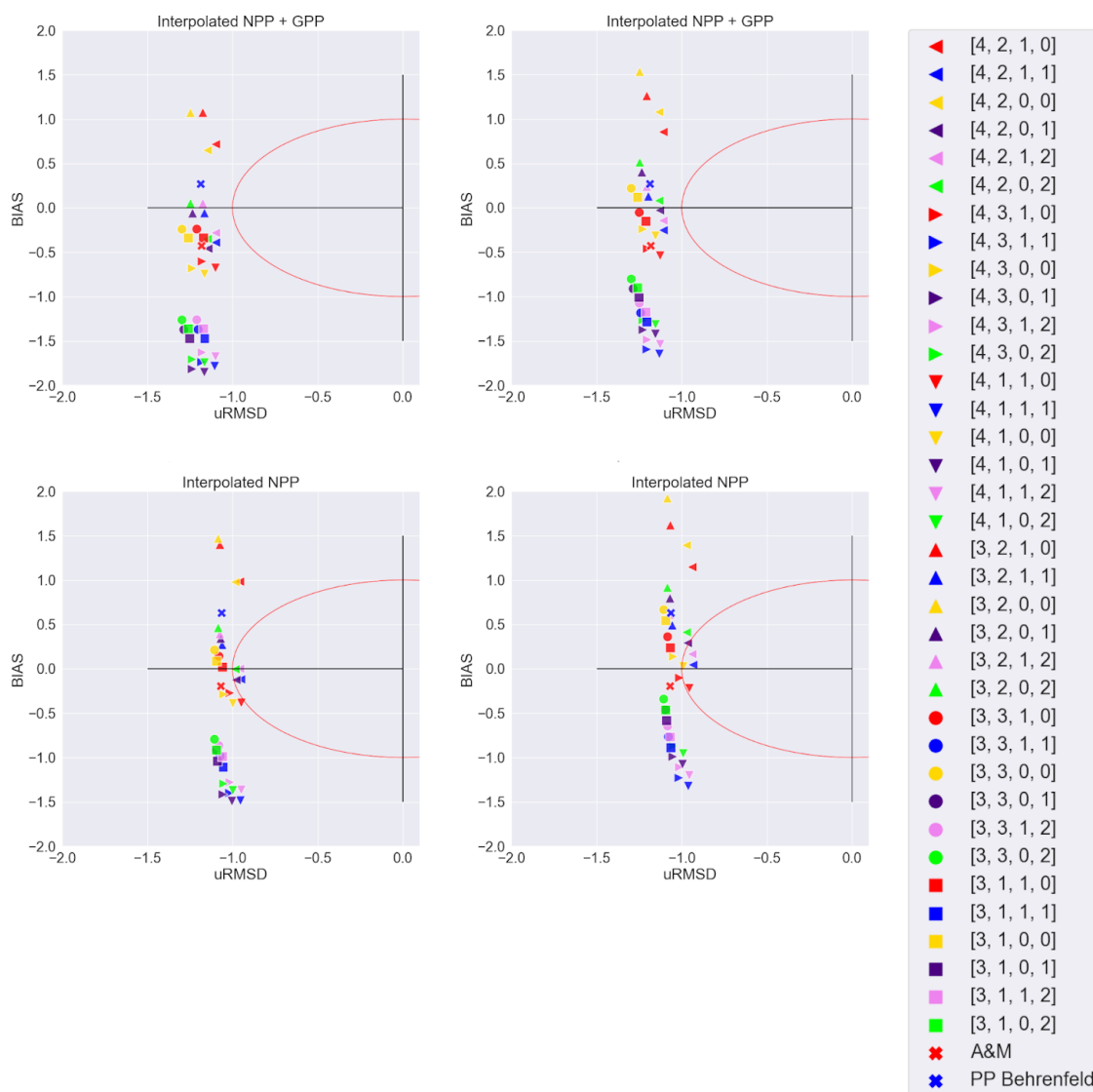
**Table 1: Name of field datasets with the selection of best performing models in terms of bias and correlation coefficient (see Section 2.5).**

365



For NPP\_C14 datasets, model setups with L4 CHL performed better than those with L3 CHL. For NPP\_C14+GPP\_O2 datasets, the best performing versions always had L2 PAR and local absorption spectrum. In Table 1 it is clear that the interpolated datasets do not always reach the maximum number of available data points, for example, 45 available points for NPP\_C14+GPP\_O2 dataset give only 41 matchups in its interpolated version. In the interpolated dataset, our goal was to  
370 obtain an equal number of collocations for all the model setups. Points that were not interpolated in L3 CHL fields as they were outlying the area with available data were excluded. Points that were not present in Copernicus-GlobColour L4 PP data based on Antoine and Morel (1996) were also excluded.

For further calculations we've chosen the model versions that passed the performance test to reproduce both the NPP\_C14 dataset and the NPP\_C14+GPP\_O2 dataset. These were the two versions: [4,2,1,1,1] and [3,2,1,1,0], see the blue highlight in  
375 Figure 6. Both of these versions contain Level 2 PAR, local CHL-a profile, and local absorption spectrum. In terms of values, one can see that both of these models have a range of values similar to field data in combination with several high outliers (Figure 6).



**Figure 7: Target diagrams illustrating relative model performance in reproducing field PP data. Right half of the diagrams is cut due to a lack of data in that area. Top: NPP\_C14+GPP\_O2 interpolated dataset; bottom: NPP\_C14 interpolated dataset; left: vertical profiles integrated to the depth of the euphotic layer; right: vertical profiles integrated to the depth of the productive layer. Legend has model identification number [a,b,c,d]; [a]: 3 - Globcolour Level 3 CHL, 4 - Copernicus-GlobColour Level 4 CHL; [b]: 1 - NOAA/NCEP Reanalysis PAR, 2 - EUMETSAT Level 2 OLCI PAR, 3 - Globcolour Level 3 PAR; [c]: 0 - Global CHL profile following Morel and Berthon (1988), 1 - Local CHL profile following Cherilkasheva et al. (2013); [d]: 0 - no coefficient applied to PAR, 1 - PAR converted to local**



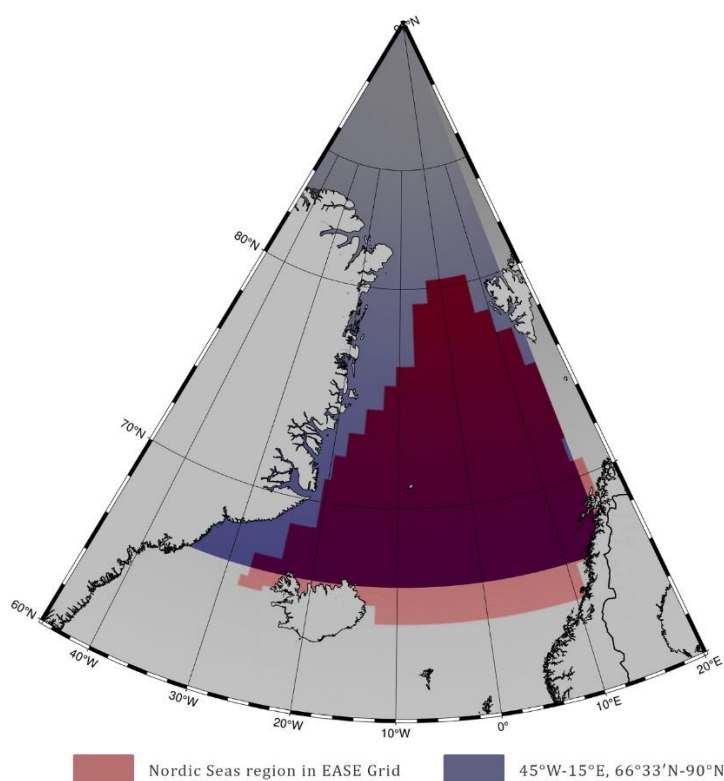
**PUR, 2 - PAR converted to global PUR; ‘A&M’ refers to Copernicus-GlobColour L4 PP data based on Antoine and Morel (1996). ‘PP Behrenfeld’ refers to our own calculations of PP based on Behrenfeld (1998) and Globcolour L3 CHL.**

390 Similar patterns are clear in the target diagrams, which have slightly different metrics and thus give results that are not identical to a performance test (Figure 7). The target diagram uses uRMSD and relative bias, as opposed to a performance test that uses the correlation coefficient and bias. The closer the model is to the target, the better the model performs. One can see that for NPP\_C14+GPP\_O2 interpolated dataset (41 points), the same five models as previously mentioned are performing best, and three more models with CHL L4, PAR L2, and local CHL profile were added to them. For NPP\_C14 interpolated dataset (32  
 395 points) the patterns are also similar to performance test, with the majority of well-performing models using CHL L4. It is also worth noting that the global Antoine and Morel model also performs quite well according to these metrics, but not reaching the inner part of the circle as other model setups.

As mentioned before, for further calculations we’ve chosen the two model versions that passed the performance test best. These two versions are also close to the target in Figure 7: [4,2,1,1,1] - blue east-oriented triangle on right images, and  
 400 [3,2,1,1,0] - blue north-oriented triangle on left images. Model setup [4,2,1,1,1] uses Copernicus-GlobColour Level 4 CHL, EUMETSAT Level 2 OLCI PAR, the local CHL profile following Cherkasheva et al. (2013), PAR converted to local PUR, and profiles integrated till productive layer depth. The setup [3,2,1,1,0] uses Globcolour Level 3 CHL, EUMETSAT Level 2 OLCI PAR, local CHL profile following Cherkasheva et al. (2013); PAR converted to local PUR and profiles integrated till euphotic layer.

### 405 **3.4 Basin estimates and trends**

For the two models selected in the previous section, we calculated the basin primary production estimates for the two regions plotted in Figure 8 for our results to be comparable with Hill et al. (2013) estimates (red region) and Arrigo and Van Dijken (2015) estimates (blue region).



410 **Figure 8: Two regions selected for the calculation of basin estimates for comparison with the of the PP model results**  
**from Hill et al. (2013) and Arrigo and Van Dijken (2015).**

The monthly basin estimates were present only in Hill et al. (2013) paper, and are shown corrected by us for the SCM presence  
 using the method applied by Hill et al. (2013). Annual Hill et al. (2013) estimates account for SCM and did not have to be  
 415 corrected.

The results of the basin estimates for the two models selected in our study are presented in Table 2. We have tested calculations  
 with both interpolating the missing pixels of data and not, but there was only 0.5 TgC/year (less than 1%) difference between  
 the two models selected in our study. Thus, here we show the noninterpolated values. The difference between the estimates of  
 the two models on average was minor, about 1-4% of the annual estimates for 1998-2022 depending on the choice of the  
 420 region. The annual average was 342-347 TgC/year. The results are slightly higher than Hill et al. (2013) results, which give  
 308 TgC/year. This difference could be attributed to the principal differences between the models used. Hill et al. (2013)  
 calculations are based on a model developed for the Chukchi Sea, which is using SeaWiFS CHL data only, without accounting  
 for PAR data (Hill and Zimmerman, 2010). In our case, CHL data are an integrated product of several sensors, PAR data is  
 used, and the model includes Greenland Sea parameterizations derived from particulate absorption and CHL data. Our model

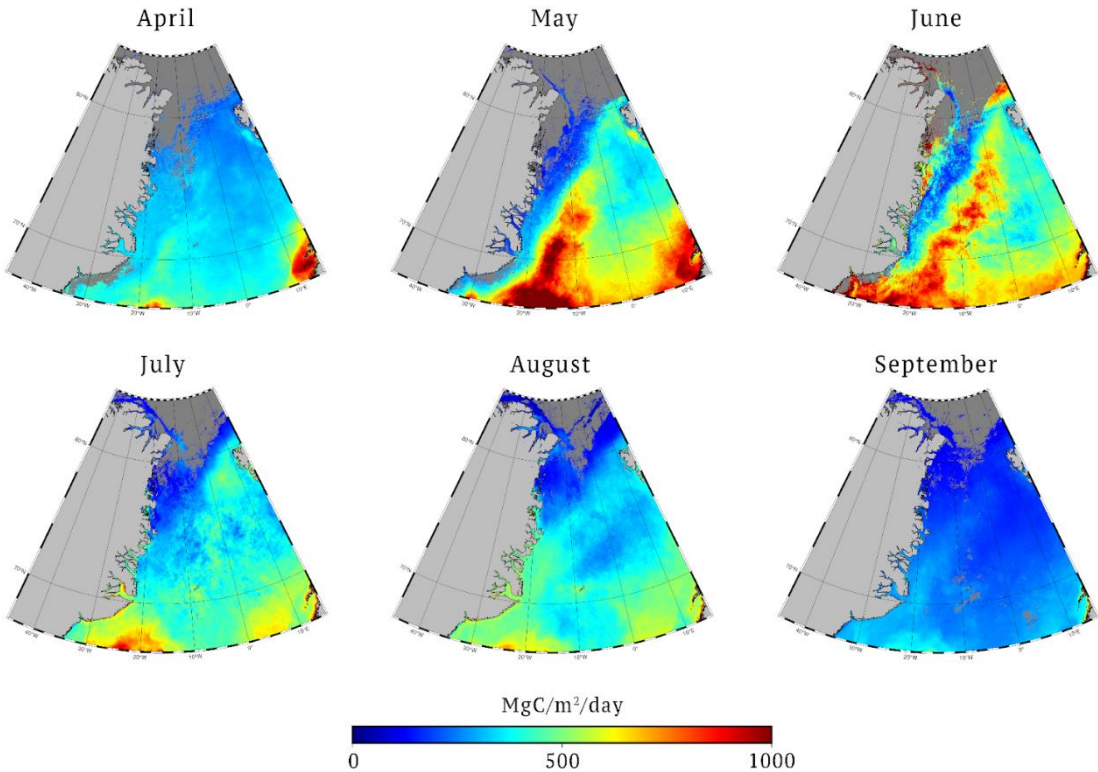


425 results calculated for the basin grid used in Arrigo and Van Dijken (2015), were with an annual average of 340 TgC/year  
 annual average significantly higher than the 136.3 TgC/year reported in that study. These results are difficult to compare as  
 Arrigo and Van Dijken (2015) did not yet account for the SCM, which has a potentially increasing impact on integrated primary  
 production in the Arctic with time (Ardyna and Arrigo, 2020). In terms of a seasonal evolution, our results are more uniform  
 throughout the year, ranging between 33 TgC/month to 78 TgC/month, while Hill et al. (2013) estimates show a wider range  
 430 of 13-123 TgC/month. This seasonal dynamics that we have observed is in line with the pattern that we have previously seen  
 when analysing CHL data in the area (Nöthig et al., 2015). In our models, the peak of the bloom is observed in May for the  
 Nordic Seas region similar to Hill et al. (2013). For the larger and further north region of Arrigo and Van Dijken (2015), the  
 peak value shifts to June, though the values in May are close as well. The spatial distribution of this bloom is seen in the  
 monthly maps (Figure 9). When compared to Ardyna et al. (2013) 228-230 TgC/year estimates, our results as in all other cases  
 435 give higher regional PP estimates of 333 TgC/year. This could be due to the fact that Ardyna et al. (2013) ten validation points  
 for the region are distributed in the Western part of the Greenland Sea which is less productive than the Eastern part. The data  
 set on which we based the selection of the model setup is, on the other hand, distributed in both the western and eastern parts  
 of the Greenland Sea (Figure 4). The other reason could be the different parametrisations of the CHL vertical profile. In general,  
 our larger estimates than those previously reported could have been explained by our additional use of GPP field data which  
 440 has higher values than NPP. However, as we have tested, the selected models are best at reproducing NPP data without GPP  
 data as well (Section 3.3). Another point worth noting is that although mentioned studies (Ardyna et al. (2013), Arrigo and  
 Van Dijken (2011), Hill et al. (2013)) accurately use local data from the Greenland Sea, in principle they are pan-Arctic studies  
 and have less Greenland Sea parametrisation parameters than used here.

#	Source	Region	Period	Annual (TgC/year)	Month (TgC/month)					
					Apr	May	Jun	Jul	Aug	Sep
1	Hill et al. (2013) with SCM	Nordic Seas region in EASE Grid	1998- 2007	308	33.4	122.5	42.3	27.5	13.4	44.5
2	Arrigo and Van Dijken (2015)	45°W-15°E, 66°33'N-90°N	1998- 2012	136.3	ND	ND	ND	ND	ND	ND
3	Ardyna et al. (2013)	Greenland-Norwegian Seas	1998	227.9	ND	ND	ND	ND	ND	ND
			2007	230.8	ND	ND	ND	ND	ND	ND
4	This study (related to #1)	Nordic Seas region in EASE Grid	1998- 2007	344.1	45.3	77.9	73.7	60.3	54.1	32.7

5	This study (related to #2)	45°W-15°E, 66°33'N-90°N	1998-2012	340.0						
6	This study (related to #3)	45°W-15°E, 66°33'N-90°N	1998	333.1						
			2007	333.7						
7	This study (own estimates)	Nordic Seas region in EASE Grid	1998-2022	346.6	44.4	77.1	76.0	60.6	55.2	33.2
8	This study (own estimates)	45°W-15°E, 66°33'N-90°N	1998-2022	342.1	42.8	69.7	74.1	62.3	58.3	35.0

445 **Table 2: Primary production basin estimates in the European Arctic from literature and calculated in this study using two models selected via performance tests in Section 3.3. The average between the two setups of models [4,2,1,1,1] and [3,2,1,1,0] is given. Monthly values from Hill et al (2013) are in italics as we have calculated them ourselves from Hill et al. (2013) averages without SCM using a method given in the source. ND - no data found in the source.**

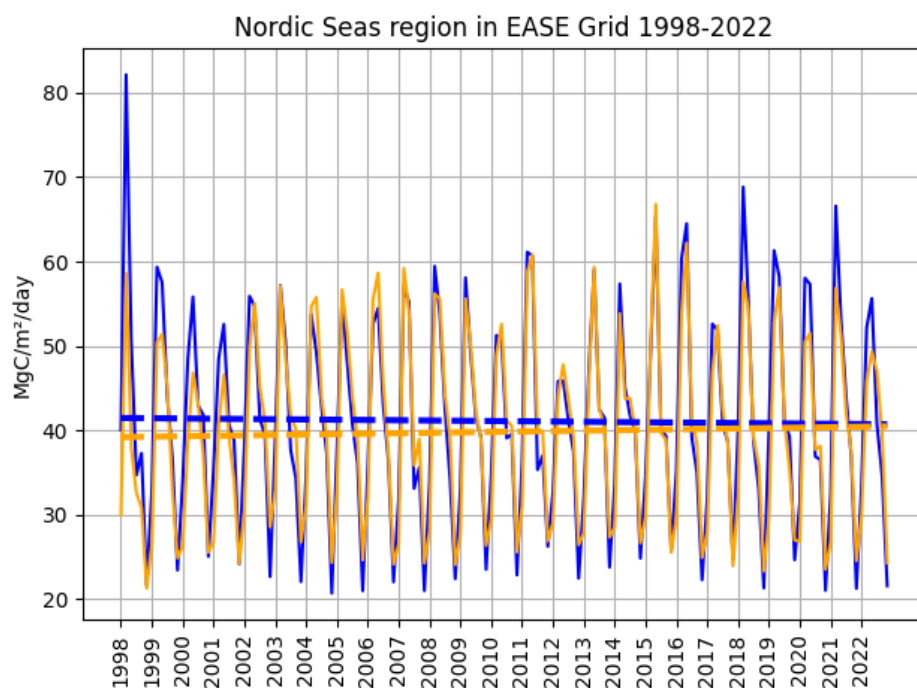


450



**Figure 9: Monthly primary production for 1998-2022 for one of the selected model setups [4,2,1,1,1], which uses Copernicus-GlobColour Level 4 CHL, EUMETSAT Level 2 OLCI PAR, local CHL profile following Cherkasheva et al. (2013), PAR converted to local PUR and profiles integrated till productive layer depth.**

455 When looking at the trends for the selected regions, no significant increase or decrease in primary production was found in the period 1998-2022 (see Figure 10).



**Figure 10: Time series and trend line for one of the two selected primary production model setups and the Nordic Seas region in the EASE grid for 1998-2022. Orange line: model setup [3,2,1,1,0], blue line: model setup [4,2,1,1,1]. The trends are significant ( $p < 0.01$ ).**

460

## 4 Concluding discussion

### 4.1 Field estimates

Although GPP measurements with oxygen sensors are not traditional, in our case they gave estimates comparable to historical NPP values derived with the  $^{14}\text{C}$  method for the Greenland Sea in terms of spatial patterns and 1.5 higher estimates in terms of average values.

465

The method should be applied with care, as the sensors are very sensitive to temperature changes and the way sampling bottles are filled, so an error could easily be introduced. We also recommend having triplicate samples for each measurement to



minimise errors. The average standard deviation of triplicate measurements was 38-50 mgC/m<sup>3</sup>/day for the difference between the initial and the last measurements. Unfortunately, we did not have simultaneous measurements of primary production with the other methods to be able to perform a direct intercomparison. In summary, we would recommend using oxygen sensors in a setup presented here to get a rough estimate of GPP, especially as an alternative in case of absence of <sup>14</sup>C and <sup>13</sup>C measurements of NPP, which are accordingly legally and logistically challenging.

## 4.2 Choice and performance of the model

According to the study by Lee et al. (2015) of the Arctic PP models, depth-resolved CHL models agreed better with in-situ data than any other type. This was a decisive factor for us in choosing a depth-resolved CHL model to work with. However, Lee et al. (2015) also mention that absorption-based models as well do have potential, since they exhibit lowest bias associated with weaker correlation when compared to field PP. Antoine et al. (2013) give a similar point of view recommending to use either locally tuned CHL algorithms for the Arctic, or not-CHL based algorithms (e.g., Hirawake et al., 2012; Mouw and Yoder, 2005). According to Antoine et al. (2013), combining the retrieval of nonwater absorption with locally tuned models for CDOM absorption (Matsuoka et al., 2013), might improve absorption-based models. Thus, the next step to achieve a goal of more accurate Greenland Sea PP estimates could be to test the absorption models focused precisely on the Greenland Sea and not on the Arctic as a whole, as previously done in Lee et al. (2015).

Our result clearly and expectedly showed that model setups with local CHL profile and local absorption spectrum perform better than global relationships. The fact that level 2 PAR is performing better than level 3 PAR is more challenging to explain. This could just be a mathematical feature of level 2 PAR having larger values, and thus these models follow closer the range of in-situ data. For the Lee et al. (2015) Arctic study the best performing cases were the models that used in situ CHL and satellite PAR, but they did not test PAR data from different sensors, and no conclusions could be made between the performance of L2 and L3 PAR.

The accuracy of the model ability here to reproduce the field data in terms of RMSD is poorer than reported for the global studies: the average RMSD for our selected model setups is 0.4 as opposed to RMSD=0.3 reported as an average for 21 models by Saba et al. (2011) tested in ten marine regions across the world. For the Arctic, however, the performance of our selected model setups is much better than the average shown in the Arctic intercomparison study by Lee et al. (2015), where the RMSD range is 0.61-0.67 for 32 models, and averages 0.65 for the depth resolved models as those used in this study.

## 5 Conclusion

We have collected an integrated field dataset of net primary production data obtained with the <sup>14</sup>C method (NPP\_C14), and gross primary production obtained with optode O<sub>2</sub> measurements (GPP\_O2). Tests with different setups of Morel (1991) primary production model were performed in a way that included both local and global empirical relationships and input data from various sources. The model versions that performed best when validated against the NPP\_C14 and GPP\_O2 data included the



local CHL profile and local absorption spectrum in their setup and used the Level 2 PAR data as input. In terms of choice of  
500 CHL input data and integration depth, there was no dependency.

When comparing the basin estimates for the Nordic Seas of our selected model versions with previous studies, it is clear that  
our estimates are higher than those previously reported. Annual estimates are on average 11% higher than reported by Hill et  
al. (2013) and 150% higher than reported by Arrigo and Van Dijken (2015). For the 1998-2022 period and Nordic Seas region,  
our annual estimates are 347 TgC/year. The seasonal cycle in our case has less monthly variation of 33-78 TgC/month than  
505 13-123 TgC/month previously reported by Hill et al. (2013) with the peak values similarly observed in May, and no significant  
increase or decrease in primary production was observed when studying regionally averaged trends.

The accuracy of the selected model setups to reproduce the field data in terms of RMSD is poorer than in the related global  
studies, but better than in the related Arctic studies.

As previously shown by Lee et al. (2015) and recommended by Antoine et al. (2013) the use of absorption-based models may  
510 improve the performance, especially in CDOM-dominated areas such as western part of the Greenland Sea (Granskog et al.,  
2012 and 2015, Goncalves-Araujo et al., 2016), and could be a next step toward improving primary production estimates in  
the Greenland Sea.

### Data availability

Primary production measured with optode O<sub>2</sub> measurements for 2021-2022 has been uploaded to PANGAEA Data Publisher  
515 for Earth & Environmental Science. Temporary link to be updated during the review process:  
<https://issues.pangaea.de/browse/PDI-36519>.

### Code availability

The Python codes developed for this manuscript are available at  
[https://github.com/9Di/environmental\\_data\\_algorithms/tree/main/Algorithms](https://github.com/9Di/environmental_data_algorithms/tree/main/Algorithms)

### 520 Author contribution

Conceptualisation: AC, AB

Methodology: AC, PK

Investigation: AC, PK, RM, MZ

Visualisation: RM, AC

525 Supervision: PK, AB

Writing—original draft: AC



Writing—review & editing: AB, AL, AC, PK

## Competing interests

The authors declare that they have no conflict of interest.

## 530 Acknowledgements

The research leading to these results has received funding from the Norwegian Financial Mechanism 2014-2021, NCN-POLS project MOPAR no. 2020/37/K/ST10/03254. The authors used data sets collected during the CDOM-HEAT project (2013-2016) supported by the Polish-Norwegian Research Programme operated by the National Centre for Research and Development under the Norwegian Financial Mechanism 2009–2014 in the framework of the project contract Pol-  
 535 Nor/197511/40/2013. Prof. Piotr Kowalczyk's participation in the Maria S. Merian 2022 expedition was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 869383 (ECOTIP). Prof. Astrid Bracher's contribution was supported in part by Helmholtz Impulse Fond DFG (German Research Foundation) Transregional Collaborative Research Centre ArctiC Application: Climate Relevant Atmospheric and Surface Processes, and Feedback Mechanisms (AC)3 (Project C03) and by the Helmholtz Infrastructure Initiative FRAM. The contribution of Dr.  
 540 Alexandra Loginova was supported by Norwegian Financial Mechanism 2014-2021: DOMUSE (2020/37/K/ST10/03018). Funding for field work on board RV Kronprins Haakon during the FS2021 cruise was received from the ARICE-NoTAC project awarded to Dr. R. Goncalves-Araujo (Danish Technical University). We thank Dr. Vasiliy Povazhnyy for the training on measuring GPP with oxygen sensors and consultations when preparing for the cruise and processing the data, and Dr. Elena Terzic for the help in field measurements onboard RV Maria S. Merian August 2022 cruise. Dr. Aleksandra Cherkasheva  
 545 thanks Prof. David Antoine and Bernard Gentili for the instructions on the Morel (1989) model setup which she received during her PhD work. The authors also thank the entire crews of RV Kronprins Haakon and RV Maria S. Merian for their time and effort.

## References

- Miller, B. B. and Carter, C.: The test article, *J. Sci. Res.*, 12, 135–147, doi:10.1234/56789, 2015.
- 550 Alexander, V. and H. J. Niebauer: Oceanography of the eastern Bering Sea ice-edge zone in spring, *Limnol. Oceanography*, 26(6), 1111–1125, 1981.
- Antoine, David & Morel, André: Oceanic primary production: 1. Adaptation of a spectral light-photosynthesis model in view of application to satellite chlorophyll observations. *Global Biogeochemical Cycles*. 10. 43-55. 10.1029/95GB02831, 1996.



- Antoine, D., André, J.-M., and Morel, A.: Oceanic primary production: 2. Estimation at global scale from satellite (Coastal  
 555 Zone Color Scanner) chlorophyll, *Global Biogeochemical Cycles*, 10( 1), 57– 69, doi:10.1029/95GB02832, 1996.
- Antoine, D., S. B. Hooker, S. Bélanger, A. Matsuoka, and M. Babin: Apparent optical properties of the Canadian Beaufort Sea  
 – Part 1: Observational overview and water column relationships, *Biogeosciences*, 10(7), 4493– 4509, doi:10.5194/bg-513 10-  
 4493-2013, 2013
- Ardyna, M., Mundy C. J., Mayot, N., Matthes, L.C., Oziel, L., Horvat, C., Leu, E., Assmy, P., Hill, V., Matrai, P. A., Gale,  
 560 M., Melnikov, I.A., Arrigo, K.R.: Under-Ice Phytoplankton Blooms: Shedding Light on the “Invisible” Part of Arctic Primary  
 Production . *Frontiers in Marine Science*, DOI=10.3389/fmars.2020.608032, 2020.
- Ardyna, M., M. Babin, M. Gosselin, E. Devred, S. Belanger, A. Matsuoka, and J.-E. Tremblay: Parameterization of vertical  
 chlorophyll a in the Arctic Ocean: Impact of the subsurface chlorophyll maximum on regional, seasonal, and annual primary  
 production estimates, *Biogeosciences*, 10, 4383–4404, 2013.
- 565 Ardyna, M., Arrigo, K.R.: Phytoplankton dynamics in a changing Arctic Ocean. *Nature Climate Change*, 10, 892–903, 2020.
- Arrigo, K.R., van Dijken, G.L.: Secular trends in Arctic Ocean net primary production. *Journal Geophysical Research*, 16.  
<http://dx.doi.org/10.1029/2011JC7273 C09011>, 2011.
- Arrigo, K. R., & van Dijken, G. L.: Continued increases in Arctic Ocean primary production. *Progress in Oceanography*, 136,  
 60–70. doi:10.1016/j.pocean.2015.05.002, 2015.
- 570 Assmy, P., Fernández-Méndez, M., Duarte, P., Meyer, A., Randelhoff, A., Mundy, C. J., et al.: Leads in Arctic pack ice enable  
 early phytoplankton blooms below snow-covered sea ice. *Scientific Reports* 7:40850, 2017.
- Behrenfeld, M. J., and P. G. Falkowski: Photosynthetic rates derived from satellite-based chlorophyll concentration.  
*Limnology and Oceanography*, 42/1, pp. 1-20, 1997.
- Behrenfeld, M. J., Prasil, O., Kolber, Z. S. et al.: Compensatory changes in photosystem II electron turnover rates protect  
 575 photosynthesis from photoinhibition. *Photosynth. Res.*, 58, 259–268. ll concentration, *Limnology Oceanography.*, 42(1), 1–20,  
 1998.
- Boyer, Tim P.; Garcia, Hernan E.; Locarnini, Ricardo A.; Zweng, Melissa M.; Mishonov, Alexey V.; Reagan, James R.;  
 Weathers, Katharine A.; Baranova, Olga K.; Seidov, Dan; Smolyar, Igor V.: World Ocean Atlas 2018. [Statistical mean of  
 Mixed Layer Depth on 1° grid for all decades]. NOAA National Centers for Environmental Information. Dataset.
- 580 <https://www.ncei.noaa.gov/archive/accession/NCEI-WOA18>. Accessed 17.01.2023, 2018.
- Bricaud, A., Babin, M., Morel, A., & Claustre, H.: Variability in chlorophyll-specific absorption coefficients of natural  
 phytoplankton: Analysis and parameterization. *Journal of Geophysical Research*, 100(7), 13,321–13,332, 1995.
- Campbell, J. W.: The lognormal distribution as a model for bio-optical variability in the sea, *Journal of Geophysical Research*,  
 100, 13237–13254, doi: /10.1029/95JC00458, 1995.
- 585 Campbell, K., Mundy, C.J., Landy, J.C., Delaforge, A., Michel, C., Rysgaard, S.: Community dynamics of bottom-ice algae  
 in Dease Strait of the Canadian Arctic, *Progress in Oceanography*, 149, 27-39, ISSN 0079-6611,  
<https://doi.org/10.1016/j.pocean.2016.10.005>, 2016.



- Carr, M.E.; Friedrichs, M.A.; Schmeltz, M.; Aita, M.N.; Antoine, D.; Arrigo, K.R.; Asanuma, I.; Aumont, O.; Barber, R.; Behrenfeld, M.; et al.: A comparison of global estimates of marine primary production from ocean color. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 53, 741–770, 2006.
- Carpenter, J.: *The Accuracy of the Winkler Method for Dissolved Oxygen Analysis*. Baltimore, MD: The Johns Hopkins University, 135–140, 1995.
- Carritt, D. E., and Carpenter, J. H.: Comparison and evaluation of currently employed modifications of the Winkler method for determining dissolved oxygen in seawater. *Journal of Marine Research* 24, 286–318, 1966.
- Cherkasheva, A., Nöthig, E.-M., Bauerfeind, E., Melsheimer, C., and Bracher, A.: From the chlorophyll a in the surface layer to its vertical profile: a Greenland Sea relationship for satellite applications. *Ocean Science*, 9, 431–445, doi:10.5194/os-9-431-2013, 2013
- Chierici, M., Vernet, M., Fransson, A., Børsheim, K. Y.: Net Community Production and Carbon Exchange From Winter to Summer in the Atlantic Water Inflow to the Arctic Ocean, *Frontiers in Marine Science*, vol. 6, 2019.
- Duarte, C. M., Agustí, S., and Regaudie-de-Gioux, A.: “The role of marine biota in the metabolism of the biosphere,” in *The Role of Marine Biota in the Functioning of the Biosphere*, ed. C. M. Duarte (Madrid: CSIC), 38–53, 2011.
- Duarte, C. M., and Agustí, S.: The CO<sub>2</sub> balance of unproductive aquatic ecosystems. *Science* 281, 234–236. doi: 10.1126/science.281.5374.234, 1998.
- Gonçalves-Araujo, R., Granskog, M. A., Bracher, A., Azetsu-Scott, K., Dodd, P. A., & Stedmon, C. A.: Using fluorescent dissolved organic matter to trace and distinguish the origin of Arctic surface waters. *Scientific Reports*, 6(1). doi:10.1038/srep33978, 2016.
- Goncalves-Araujo, R., Rabe, B., Peeken, I., Bracher, A.: High colored dissolved organic matter (CDOM) absorption in surface waters of the central-eastern Arctic Ocean: Implications for biogeochemistry and ocean color algorithms. *PLoS ONE* 13(1): e0190838. <https://doi.org/10.1371/journal.pone.0190838>, 2018.
- Granskog, M. A., C. A. Stedmon, P. A. Dodd, R. M. W. Amon, A. K. Pavlov, L. de Steur, and E. Hansen: Characteristics of colored dissolved organic matter (CDOM) in the Arctic outflow in the Fram Strait: Assessing the changes and fate of terrigenous CDOM in the Arctic Ocean, *J. Geophys. Res.*, 117, C12021, doi:10.1029/2012JC008075, 2012.
- Granskog, M. A., A. K. Pavlov, S. Sagan, P. Kowalczyk, A. Raczowska, and C. A. Stedmon: Effect of sea-ice melt on inherent optical properties and vertical distribution of solar radiant heating in Arctic surface waters, *J. Geophys. Res. Oceans*, 120, 7028–7039, doi:10.1002/2015JC011087, 2015.
- Eastman, R., and S. G. Warren: Interannual variations of Arctic cloud types in relation to sea ice, *Journal of Climate*, 23, 4216–4232, 2010.
- Intrieri, J., C. W. Fairall, M. D. Shupe, P. O. G. Persson, E. L. Andreas, P. S. Guest, and R. E. Moritz: An annual cycle of Arctic surface cloud forcing at SHEBA, *Journal of Geophysical Research*, 107(C10), 8039, doi:10.1029/2000JC000439, 2002.
- IOCCG: *Ocean Colour Remote Sensing in Polar Seas*, edited by M. Babin, et al., IOCCG Report Series, No. 16, International Ocean Colour Coordinating Group, pp. 1–130, Dartmouth, Canada, 2015.



- IOCCG Protocol Series: Aquatic Primary Productivity Field Protocols for Satellite Validation and Model Synthesis. Balch, W.M., Carranza, M., Cetinić, I., Chaves, J.E., Duhamel, S., Fassbender, A., Fernandez-Carrera, A., Ferrón, S., García-Martín, E., Goes, J., Gomes, H., Gundersen, K., Halsey, K., Hirawake, T., Isada, T., Juranek, L., Kulk, G., Langdon, C., Letelier, R.,
- 625 López-Sandoval, D., Mannino, A., Marra, J.F., Neale, P., Nicholson, D., Silsbe, G., Stanley, R.H., Vandermeulen, R.A. IOCCG Ocean Optics and Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation, Volume 7.0, edited by R.A. Vandermeulen, J. E. Chaves, IOCCG, Dartmouth, NS, Canada. doi:<http://dx.doi.org/10.25607/OBP-1835>, 2022.
- Iversen, K. R., & Seuthe, L.: Seasonal microbial processes in a high-latitude fjord (Kongsfjorden, Svalbard): I. Heterotrophic bacteria, picoplankton and nanoflagellates. *Polar biology*, 34(5), 731–749, 2011.
- 630 Johannessen, J.A., Johannessen, O.M., Svendsen, E., Schuchman, R., Manley, T., Cambell, W.J., Josberger, E.G., Sandven, S., Gascard, J.C., Olaussen, T., Davidson, K., Vanleer, J.: Mesoscale eddies in the Fram Strait marginal ice zone the 1983 and 1984 Marginal Ice Zone Experiments. *Journal of Geophysical Research*, 92, 6754–6772, 1987.
- Hansell, D.A., C.A. Carlson, D.J. Repeta, and R. Schlitzer: Dissolved organic matter in the ocean: A controversy stimulates new insights. *Oceanography* 22(4):202–211, <https://doi.org/10.5670/oceanog.2009.109>, 2009.
- 635 Hill, V.J., Matrai, P.A., Olson, E., Suttles, S., Steele, M., Codispoti, L.A., Zimmerman, R.C.: Synthesis of integrated primary production in the Arctic Ocean: II. In situ and remotely sensed estimates, *Progress in Oceanography*, 110, 107–125, doi: 10.1016/j.pocean.2012.11.005, 2013.
- Hill, V.J., Zimmerman, R.C.: Assessing the accuracy of remotely sensed primary production estimates for the Arctic Ocean, using passive and active sensors. *Deep Sea Research I* 57, 1243–1254, 2010.
- 640 Hirawake, T., K. Shinmyo, A. Fujiwara, and S. I. Saitoh: Satellite remote sensing of primary productivity in the Bering and Chukchi Seas using an absorption-based approach, *ICES Journal Marine Science*, 69, 1194–1204, 2012.
- Kowalczyk, P., Sagan, S., Makarewicz, A., Meler, J., Borzycka, K., Zabłocka, M., et al.: Bio-optical properties of surface waters in the Atlantic Water inflow region off Spitsbergen (Arctic Ocean). *Journal of Geophysical Research: Oceans*, 124, 1964–1987. <https://doi.org/10.1029/2018JC014529>, 2019.
- 645 Kraft, A.: Arctic pelagic amphipods - community patterns and life-cycle history in a warming Arctic Ocean. PhD Thesis, FB2, University of Bremen, Germany, 2013.
- Lee, Y. J., et al.: An assessment of phytoplankton primary productivity in the Arctic Ocean from satellite ocean color/in situ chlorophyll-a based models, *Journal of Geophysical Research Oceans*, 120, 6508–6541, doi:10.1002/2015JC011018, 2015.
- Matrai, P. A., E. Olson, S. Suttles, V. Hill, L. A. Codispoti, B. Light, and M. Steele: Synthesis of primary production in the
- 650 Arctic Ocean: I. Surface waters, 1954–2007, *Progress in Oceanography*, 110, 93–106, 2013.
- McLaughlin, F.A., and E. C. Carmack: Deepening of the nutricline and chlorophyll maximum in the Canada Basin interior, 2003–2009, *Geophysical Research Letters*, 37, L24602, doi: 10.1029/2010GL045459, 2010.
- Meier, W. N. and J. S. Stewart: Arctic and Antarctic Regional Masks for Sea Ice and Related Data Products, Version 1 [Data Set]. Boulder, Colorado USA. National Snow and Ice Data Center. <https://doi.org/10.5067/CYW3O8ZUNIWC>. Date Accessed
- 655 03-08-2023, 2023.



- Morel, A.: Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). *Journal of Geophysical Research*, 93(C9), 10749. doi:10.1029/jc093ic09p10749, 1988.
- Morel, A.: Light and marine photosynthesis: a spectral model with geochemical and climatological implications, *Progress in Oceanography*, 26, 3, 1991, 263-306, ISSN 0079-6611, doi: 10.1016/0079-6611(91)90004-6, 1991.
- 660 Morel, A. and Berthon, J. F.: Surface Pigments, Algal Biomass Profiles, and Potential Production of the Euphotic Layer: Relationships Reinvestigated in View of Remote-Sensing Applications. *Limnology and Oceanography*, 34, 1545–1562, 1989.
- Morel, A., & Maritorena, S.: Bio-optical properties of oceanic waters: A reappraisal. *Journal of Geophysical Research: Oceans*, 106(C4), 7163–7180. doi:10.1029/2000jc000319, 2001.
- Mouw, C. and Yoder, J. A.: Primary production calculations in the Mid-Atlantic Bight, including effects of phytoplankton community size structure, *Limnology and Oceanography*, 50, 1232–1243, 2005.
- 665 Nöthig, E.-M., Bracher, A., Engel, A., Metfies, K., Niehoff, B., Peeken, I., Bauerfeind, E., Cherkasheva, A., Gäbler-Schwarz, S., Hardge, K., Kiliass, E., Kraft, A., Kidane, Y. M., Lalande, C., Piontek, J., Thomisch, K. & Wurst, M.: Summertime plankton ecology in Fram Strait—a compilation of long- and short-term observations, *Polar Research*, 34:1, DOI: 10.3402/polar.v34.23349, 2015.
- 670 Olsen, A., Omar, A. M., Bellerby, R. G. J., Johannessen, T., Ninnemann, U., Brown, K. R., et al.: Magnitude and origin of the anthropogenic CO<sub>2</sub> increase and <sup>13</sup>C Suess effect in the Nordic seas since 1981. *Global Biogeochemical Cycles* 20:GB3027. doi: 10.1029/2005GB00266, 2006.
- Olofsson, P., P. E. Van Laake, and L. Eklundh: Estimation of absorbed PAR across Scandinavia from satellite measurements: Part I: Incident PAR. *Remote Sensing of Environment*, 110, 252–261, 2007.
- 675 Piwosz, K., Walkusz, W., Hapter, R., Wieczorek, P., Hop, H., & Wiktor, J.: Comparison of productivity and phytoplankton in a warm (Kongsfjorden) and a cold (Hornsund) Spitsbergen fjord in mid-summer 2002. *Polar Biology*, 32(4), 549-559, 2009.
- Perrette, M., A. Yool, G. D. Quartly, and E. E. Popova: Near-ubiquity of ice-edge blooms in the Arctic, *Biogeosciences*, 8(2), 515–524, 2011.
- Rey, F., Noji, T.T., Miller, L.: Seasonal phytoplankton development and new production in the central Greenland Sea. *Sarsia* 85, 329–344, 2000.
- 680 Richardson, K., Markager, S, Buch, Lassen, M.F., Kristensen, A.S.: Seasonal distribution of primary production, phytoplankton biomass and size distribution in the Greenland Sea, *Deep Sea Research Part I: Oceanographic Research Papers*, 52, 6, 2005, 979-999, DOI: 10.1016/j.dsr.2004.12.005, 2005.
- Robinson, C., Tilstone, G., Rees, A., Smyth, T., Fishwick, J., Tarran, G., Luz, B., Barkan, E., Efrat, D.: Comparison of in vitro and in situ plankton production determinations. *Aquatic Microbial Ecology*. 54. 13-34. 10.3354/ame01250, 2009.
- 685 Rudels, B., and D. Quadfasel: Convection and deep water formation in the Arctic Ocean -Greenland Sea system. *Journal of Marine Systems*, 2, 435-450, 1991.
- Saba, V. S., Friedrichs, M. A. M., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M. J., Ciotti, A. M., Dowell, M., Hoepffner, N., Hyde, K. J. W., Ishizaka, J., Kameda, T., Marra, J., Mélin, F., Morel, A., O'Reilly, J., Scardi, M., Smith Jr., W.



- 690 O., Smyth, T. J., Tang, S., Uitz, J., Waters, K., and Westberry, T. K.: An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe. *Biogeosciences*, 8, 489–503, <https://doi.org/10.5194/bg-8-489-2011>, 2011.
- Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., et al.: The oceanic sink for anthropogenic CO<sub>2</sub>. *Science* 305, 367–371. doi: 10.1126/science.1097403, 2004.
- 695 Sakshaug, E.: Primary and Secondary Production in the Arctic seas. In: Stein, R., Macdonald, R. (Eds.), *The organic carbon cycle in the Arctic Ocean*. Springer, Berlin, pp. 57–81, 2004.
- Sanz-Martín, M., Vernet, M., Cape, M.R., Mesa, E., Delgado-Huertas, A., Reigstad, M., Wassmann, P., and Duarte, C.M.: Relationship Between Carbon- and Oxygen-Based Primary Productivity in the Arctic Ocean, Svalbard Archipelago. *Frontiers in Marine Science*. 6:468. doi: 10.3389/fmars.2019.00468, 2019.
- 700 Skogen, M.D., Budgell, W.P., Rey F.: Interannual variability in Nordic seas primary production. *ICES Journal of Marine Science*, 64:889-898, 2007.
- Slagstad, D., I.H. Ellingsen, and P. Wassmann: Evaluating primary and secondary production in an Arctic Ocean void of summer sea ice: An experimental simulation approach. *Progress in Oceanography*, 90:117–131, <http://dx.doi.org/10.1016/j.pocean.2011.02.009>, 2011.
- 705 Steemann Nielsen, E.: The use of radio-active carbon (C<sup>14</sup>) for measuring organic production in the sea. *Journal du Conseil International pour l'Exploration du Mer*, 18 (1952), pp. 117-140, 1952.
- Tassan, S., & Ferrari, G. M.: An alternative approach to absorption measurements of aquatic particles retained on filters. *Limnology and Oceanography*, 40, 1358–1368, 1995.
- Tassan, S., & Ferrari, G. M.: A sensitivity analysis of the 'Transmittance-Reflectance' method for measuring light absorption by aquatic particles. *Journal of Plankton Research*, 24(8), 757–774. <https://doi.org/10.1093/plankt/24.8.757>, 2002.
- 710 Vázquez-Rodríguez, M., Touratier, F., Lo Monaco, C., Waugh, D. W., Padin, X. A., Bellerby, R. G. J., et al.: Anthropogenic carbon distributions in the Atlantic Ocean: data-based estimates from the Arctic to the Antarctic. *Biogeosciences* 6, 439–451. doi: 10.5194/bg-6-439-2009, 2009.
- Vaquer-Sunyer, R., Duarte, C. M., Regaudie-De-Gioux, A., Holding, J. M., García-Corral, L. S., Reigstad, M., et al.: Seasonal patterns in Arctic planktonic metabolism (Fram Strait - Svalbard region). *Biogeosciences* 10, 1451–1469. doi: 10.5194/bg-10-1451-2013, 2013.
- 715 Von Appen W.J., Waite A., Bergmann M., Bienhold C., Boebel O., Bracher A., Cisewski B., Hagemann J., Hoppema M., Iversen M., Konrad C., Krumpen T., Lochthofen N., Metfies K., Niehoff B., Nöthig E.-M., Purser A., Salter I., Schaber M., Scholz D., Soltwedel T., Torres-Valdes S., Wekerle C., Wenzhöfer F., Wietz M., Boetius A.: Sea ice derived meltwater stratification delays export and favors pelagic secondary production: results from continuous observations" *Nature Communication* 12, 7309 (2021). <https://doi.org/10.1038/s41467-021-26943-z>, 2021.
- 720 Wiktor, J., & Wojciechowska, K.: Differences in taxonomic composition of summer phytoplankton in two fjords of West Spitsbergen, Svalbard. *Polish Polar Research*, vol. 26, no. 4, pp. 259–268, 2005.



Williams, P. J. B., Raine, R., and Bryan, J.: Agreement between the  $^{14}\text{C}$  and oxygen methods of measuring phytoplankton  
725 production: reassessment of the photosynthetic quotient. *Oceanologica Acta* 2, 411–416, 1979.