Arctic Ocean CO$_2$ uptake: an improved multi-year estimate of the air–sea CO$_2$ flux incorporating chlorophyll-a concentrations

Sayaka Yasunaka$^{1,2}$, Eko Siswanto$^1$, Are Olsen$^3$, Mario Hoppema$^4$, Eiji Watanabe$^2$, Agneta Fransson$^5$, Melissa Chierici$^6$, Akihiko Murata$^{1,2}$, Siv K. Lauvset$^{4,7}$, Rik Wanninkhof$^8$, Taro Takahashi$^9$, Naohiro Kosugi$^{10}$, Abdirahman M. Omar$^7$, Steven van Heuven$^{11}$, and Jeremy T. Mathis$^{12}$

$^1$Research and Development Center for Global Change, Japan Agency for Marine-Earth Science and Technology, Yokosuka, Japan

$^2$Institute of Arctic Climate and Environment Research, Japan Agency for Marine-Earth Science and Technology, Yokosuka, Japan

$^3$Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway

$^4$Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Climate Sciences Department, Bremerhaven, Germany

$^5$Norwegian Polar Institute, Fram Centre, Norway

$^6$Institute of Marine Research, Tromsø, Norway
Abstract

We estimated monthly air–sea CO₂ fluxes in the Arctic Ocean and its adjacent seas north of 60° N from 1997 to 2014, after mapping partial pressure of CO₂ in the surface water (\(pCO_{2w}\)) using a self-organizing map (SOM) technique incorporating chlorophyll-a concentration (Chl-a), sea surface temperature, sea surface salinity, sea ice concentration, atmospheric CO₂ mixing ratio, and geographical position. The overall relationship between \(pCO_{2w}\) and Chl-a is negative in most regions when Chl-a \(\leq\) 1 mg m\(^{-3}\), whereas there is no significant relationship when Chl-a > 1 mg m\(^{-3}\). In the Kara Sea and the East Siberian Sea and the Bering Strait, however, the relationship is
typically positive in summer. The addition of Chl-a as a parameter in the SOM process enabled us to improve the estimate of $pCO_2w$ via better representation of its decline in spring, which resulted from biologically mediated $pCO_2w$ reduction. Mainly as a result of the inclusion of Chl-a, the uncertainty in the CO$_2$ flux estimate was reduced, and a net annual Arctic Ocean CO$_2$ uptake of $180 \pm 130$ TgC y$^{-1}$ was determined to be significant.

1. Introduction

The Arctic Ocean and its adjacent seas (Fig. 1) are thought to act as a sink for atmospheric CO$_2$ because of the high solubility of CO$_2$ in its low-temperature waters, combined with its extensive primary production during the summer season (Bates and Mathis, 2009). The Arctic Ocean and its adjacent seas consist of complicated subregions that include continental shelves, central basins, and sea-ice-covered areas. Therefore, the surface partial pressure of CO$_2$ ($pCO_2w$) distribution is not only affected by ocean heat loss and gain, and biological production and respiration, but also by sea-ice formation and melting, river discharge, and shelf–basin interactions (cf. Bates and Mathis, 2009). CO$_2$ measurements are sparse in this very heterogeneous area (Fig. 2), and hence air–sea CO$_2$ flux estimates in the Arctic have shown poor agreement.
As global warming progresses, melting of sea ice will increase the area of open water and enhance the potential for uptake of atmospheric CO₂ (e.g., Bates et al., 2006; Gao et al., 2012). However, other processes will occur simultaneously, some of which could suppress CO₂ uptake. For example, increasing seawater temperatures, declining buffer capacity due to the freshening of Arctic surface water by increased river runoff and ice-melting, and increased vertical mixing supplying high-CO₂ water to the surface (Bates and Mathis, 2009; Cai et al., 2010; Chierici et al., 2011; Else et al., 2013; Bates et al. 2014; Fransson et al., 2017), will all result in a tendency for reduced uptake. The combined effect of all these processes on ocean CO₂ uptake has not yet been clarified for the Arctic.

Yasunaka et al. (2016) successfully prepared monthly maps of air–sea CO₂ fluxes from 1997 to 2013 for the Arctic north of 60° N by applying, for the first time, a self-organizing map (SOM) technique to map \( pCO₂ \) in the Arctic Ocean. The advantage of the SOM technique is its ability to empirically determine relationships among variables without making any a priori assumptions (e.g., what types of regression functions are applicable, and for which sub-regions the same regression function can be adopted). The SOM technique has been shown to reproduce the
distribution of $pCO_{2w}$ from unevenly distributed observations better than multiple regression methods (Lefèvre et al., 2005; Telszewski et al., 2009). The uncertainty of the CO$_2$ flux estimated by Yasunaka et al. (2016), however, was large ($\pm 3.4$–$4.6 \text{ mmol m}^{-2} \text{ d}^{-1}$), and the estimated CO$_2$ uptake in the Arctic Ocean was smaller than the uncertainty ($180 \pm 210 \text{ TgC y}^{-1}$). One possible reason for the large uncertainties is that the effect of biological processes on explaining the variability was not included among the parameters used in the SOM process, which could lead to an underestimation of the seasonal amplitude of $pCO_{2w}$.

Several studies have estimated chlorophyll-a concentrations (Chl-a) in the Arctic from satellite remote sensing reflectance (Rrs) (e.g. Arrigo and Dijken, 2004; Cota et al., 2004). Perrette et al. (2011) showed that satellite-derived Chl-a successfully captured a phytoplankton bloom in the ice-edge region. Changes in the seasonal cycle of Chl-a have also been observed and are likely a consequence of the recent sea-ice loss in the Arctic (Ardyna et al., 2014). Measurements in several areas of the Arctic show that the relationship between $pCO_{2w}$ and Chl-a is typically negative, as expected (Gao et al., 2012; Ulfsbo et al., 2014). Exceptions do occur, however, in the coastal regions, where the relationship is positive (Mucci et al., 2010). Chierici et al. (2009) produced $pCO_{2w}$ algorithms for the subpolar North Atlantic and found that the inclusion of Chl-a
improved the fit substantially during the period from May to October.

In the present study, we examined the relationship between $pCO_2w$ and Chl-a in the Arctic Ocean and its adjacent seas, and presented monthly air–sea CO2 flux maps for regions north of 60° N using a SOM technique similar to that of Yasunaka et al. (2016), and with Chl-a added to the SOM process.

2. Data  

2.1. $pCO_2w$ measurements

We used $pCO_2w$ observations (converted from the fugacity of CO2 values; a correction of <1 %) from the Surface Ocean CO2 Atlas version 4 (SOCATv4; Bakker et al., 2016; http://www.socat.info/; 1,983,799 data from >60° N), and from the Global Surface $pCO_2$ Database Version 2014 (LDEOv2014; Takahashi et al., 2015; http://cdiac.ornl.gov/oceans/LDEO_Underway_Database/; 302,150 data from >60° N).

Duplicate data points were eliminated; the SOCAT versions of these duplicates were used, except for the data obtained from onboard the USCGC Healy as these have been reanalyzed by Takahashi et al. (2015). Altogether 200,409 duplicates were removed. We also used shipboard $pCO_2w$ data obtained during cruises of the R/V Mirai of the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) that have not yet been
included in SOCATv4 or LDEOv2014 (cruises MR09_03, MR10_05, MR12_E03, and MR13_06; http://www.godac.jamstec.go.jp/darwin/e; 95,725 data from >60° N). In total, we used 2,181,265 $\text{pCO}_2$ data, 33 % more than used by Yasunaka et al. (2016).

To improve the data coverage, especially for the ice-covered regions, we also used 2166 $\text{pCO}_2$ values calculated from dissolved inorganic carbon and total alkalinity data extracted from the Global Ocean Data Analysis Project version 2 (GLODAPv2; Key et al., 2015; Olsen et al., 2016; http://cdiac.ornl.gov/oceans/GLODAPv2/). 90% of these data were obtained at cruises without available underway $\text{pCO}_2$ data. We extracted values of samples obtained from water depths shallower than 10 m, or the shallowest values from the upper 30 m of each cast if there were no values from above 10 m. We used the CO2SYS program (Lewis and Wallace, 1998; van Heuven et al., 2009) and the dissociation constants reported by Lueker et al. (2000) and Dickson (1990) for the calculation. This resulted in 94 % more calculated $\text{pCO}_2$ values than used by Yasunaka et al. (2016).

The availability of $\text{pCO}_2$ data (directly measured and calculated) varied spatially and temporally (Fig. 2). Most of the available data are from the subpolar North Atlantic, the Greenland Sea, the Norwegian Sea, the Barents Sea, and the Chukchi Sea while much less data are available for the Kara Sea, the Laptev Sea, the East Siberian Sea, and the
Eurasian Basin. The number of $pCO_2w$ data strongly increases after 2005, but there is also a substantial number of data from before 2004.

2.2. Other data

To calculate Chl-a, we used merged Rrs data from multi-ocean color sensors processed and distributed by the GlobColour Project (Maritorena et al., 2010; http://hermes.acri.fr/index.php?class=archive). For compatibility with the spatio-temporal resolution of the gridded $pCO_2w$ data (see Sect. 3.3), we selected monthly mean Rrs data with a spatial resolution of $1^\circ \times 1^\circ$.

Sea surface temperature (SST) data were extracted from the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation SST Version 2 (Reynolds et al., 2002; http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html). These data are provided at a resolution of $1^\circ$ (latitude) $\times 1^\circ$ (longitude) $\times$ 1 month. The sea surface salinity (SSS) data were retrieved from the Polar Science Center Hydrographic Climatology version 3.0, which also has a resolution of $1^\circ \times 1^\circ \times 1$ month (Steele et al., 2001; http://psc.apl.washington.edu/nonwp_projects/PHC/Climatology.html). The sea ice concentration (SIC) data were obtained from the NOAA/National Snow and Ice Data Center Climate Data Record of Passive Microwave Sea Ice Concentration version
9, which has a resolution of 25 km × 25 km × 1 month (Meier et al., 2013; http://nsidc.org/data/G02202). These data were averaged into 1° × 1° × 1 month grid-cells. Zonal mean data for the atmospheric CO₂ mixing ratio (xCO₂a) were retrieved from the NOAA Greenhouse Gas Marine Boundary Layer Reference data product (Conway et al., 1994; http://www.esrl.noaa.gov/gmd/ccgg/mbl/index.html) and were interpolated into 1° × 1° × 1 month grid-cells. We used monthly sea level pressure data and the 6-hourly 10-m wind speeds from the U.S. National Centers for Environmental Prediction–Department of Energy Reanalysis 2 (NCEP 2) (Kanamitsu et al., 2002; http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html). We also used the 6-hourly 10-m wind speeds from the U.S. National Centers for Atmospheric Prediction and the National Center for Atmospheric Research Reanalysis 1 (NCEP1) (Kalnay et al., 1996; https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html) to optimize the gas transfer velocity.

Surface nitrate measurements were extracted from GLODAPv2 (Key et al., 2015; Olsen et al., 2016) and the World Ocean Database 2013 (Boyer et al., 2013).

3. Methods
3.1. Calculation of chlorophyll-a concentrations

We calculated Chl-a from Rrs by using the Arctic algorithm developed by Cota et al. (2004). Several assessments have shown that the Arctic algorithm by Cota et al. (2004) has a large uncertainty (e.g., Matsuoka et al., 2007; Lewis et al., 2016), so we evaluated the sensitivity of our results to this choice by using two other algorithms for Chl-a: the standard algorithm, of O’Reilly et al. (1998), and the coastal algorithm, of Tassan (1994).

To ensure that we were working with Rrs relatively unaffected by suspended sediments and colored dissolved organic matter, we masked the Chl-a data following the method of Siswanto et al. (2013). We plotted the Rrs spectral slope between 412 and 555 nm (Rrs555-412 slope; sr⁻¹ nm⁻¹) against logarithmically transformed Chl-a. We then defined a boundary line separating phytoplankton-dominated grid-cells (Rrs555-412 slope < boundary value) from potentially non-phytoplankton-dominated grid-cells (Rrs555-412 slope ≥ boundary value):

\[
\text{Rrs}_{555-412} \text{ slope} = -0.000003 \{\log(\text{Chl-a})\}^2 + 0.00002 \{\log(\text{Chl-a})\} + 0.00006. \quad (1)
\]

Grid-cells were considered invalid and masked out if 1) \(\text{Rrs}_{555-412} \text{ slope} \geq \text{boundary value} \),
and 2) Rrs at 555nm (Rrs\textsubscript{555}) > 0.01 sr\textsuperscript{-1} (or normalized water-leaving radiance > 2 mW cm\textsuperscript{-2} \mu m\textsuperscript{-1} sr\textsuperscript{-1}; see Siswanto et al., 2011; Moore et al., 2012).

The criteria described in the previous paragraph could mask out grid-cells having coccolithophore blooms, which are sometimes observed in the Arctic Ocean (e.g., Smyth et al., 2004), as they also have Rrs\textsubscript{555} > 0.01 sr\textsuperscript{-1} (Moore et al., 2012). Unlike waters dominated by non-phytoplankton particles, whose Rrs spectral shape peaks at 555 nm, the Rrs spectral shape of waters with coccolithophore blooms peaks at 490 or 510 nm (see Iida et al., 2002; Moore et al., 2012). Therefore, grid-cells with Rrs spectral peaks at 490 or 510 nm (already classified using the criteria of Rrs at 490nm (Rrs\textsubscript{490}) > Rs at 443nm (Rrs\textsubscript{443}) and Rrs at 510nm (Rrs\textsubscript{510}) > Rrs\textsubscript{555}) were considered as coccolithophore grid-cells, and were re-included.

3.2. Chlorophyll-a interpolation

Chl-a values calculated from Rrs are often missing because of cloud cover, low angle of sunlight, or sea ice. For the period and area analyzed here, data are missing for 86 % of the space and time grid-cells. Because pCO\textsubscript{2w} mapping requires a complete Chl-a field without missing values, we interpolated Chl-a data as follows. 1) Chl-a was set to 0.01 mg m\textsuperscript{-3} (minimum value of Chl-a) in high-latitude regions in winter when there was no
light (north of 80° N in December and January, and north of 88° N in November and
February). 2) Whenever SIC was greater than 99 %, Chl-a was set to 0.01 mg m⁻³ (full
ice coverage, thus minimum Chl-a). We chose the strict criterion of SIC > 99 % because
weak but significant primary production has been found to occur under the sea ice in
regions with SIC around 90 % (Gosselin et al., 1997; Ulfbo et al., 2014; Assmy et al.,
2017). 3) The remaining grid-cells with missing data were filled, wherever possible,
using Chl-a averaged over ±1° latitude and ±1° longitude; this mainly compensated for
missing Chl-a values as a result of cloud cover. 4) Parts of the remaining missing Chl-a
values, mainly for the pre-satellite period of January–August 1997, were set to the
monthly climatology Chl-a values based on the 18-year monthly mean from 1997 to
2014. 5) The final remaining missing Chl-a data, mainly for the marginal sea-ice zone,
were generated by linear interpolation using surrounding data. With each interpolation
step the number of the grid-cells with missing data decreased; 23 % of the grid-cells
without Chl-a data were filled by the first step, and in the same manner the subsequent
steps provided data for 12, 8, 5, and 52 %, respectively.

3.3. Gridding procedure for pCO₂ data

The individual pCO₂ data were gridded to 1° × 1° × 1 month grid-cells covering the
years from 1997 to 2014 using the same procedure as Yasunaka et al. (2016). As reference values, we calculated the long-term mean (i.e., the climatology) and its standard deviation for a window size of ±5° of latitude, ±30° of longitude, and ±2 months (regardless of the year) for each 1° × 1° × 1 month grid-cell. We then eliminated data in each grid-cell that differed by more than three standard deviations from the climatology. We next recalculated the climatology and its standard deviation using a smaller window size of ±2° of latitude, ±10° of longitude, and ±1 month (regardless of the year) for each 1° × 1° × 1 month grid-cell. We again eliminated data that differed from the climatology by more than three standard deviations. This procedure identified in total about 0.5 % of the data as extreme or erroneous values. These excluded values are randomly distributed in time and space. The remaining observations were binned into 1° × 1° × 1 month grid-cells for each year from 1997 to 2014. Although some studies have used $pCO_2w$ normalized to a certain year, based on the assumption of a constant rate of increase for $pCO_2w$ (e.g., Takahashi et al., 2009), we used “non-normalized” $pCO_2w$ values from all years; therefore, $pCO_2w$ can increase both non-linearly in time and non-uniformly in space. We believe that this better represents the real variability and trends of $pCO_2w$. 
3.4. $p$CO$_2$ estimation using a self-organizing map

We estimated $p$CO$_2w$ by the SOM technique used by Yasunaka et al. (2016), with the exception that Chl-a was added as a parameter to train the SOM (a “training parameter”) in addition to SST, SSS, SIC, $x$CO$_2a$, and geographical position $X (= \sin[\text{latitude}] \times \cos[\text{longitude}])$ and $Y (= \sin[\text{latitude}] \times \sin[\text{longitude}])$. Chl-a, SST, SSS, and SIC are closely related to processes causing variation in $p$CO$_2w$, such as primary production, warming/cooling, mixing, and freshwater input, whereas $x$CO$_2a$ is related to seasonal changes from terrestrial CO$_2$ uptake and release, and the anthropogenic CO$_2$ increase. Thus, Chl-a, SST, SSS, and SIC represent spatio-temporal variability at seasonal to interannual time-scales, and $x$CO$_2a$ acts as a temporal variable representing the seasonal cycle and the long-term trend. The use of geographical position as a training parameter can prevent a systematic spatial bias (Yasunaka et al., 2014).

Briefly, the SOM technique was implemented as follows: first, the approximately one million 1° × 1° × 1 month grid-cells in the analysis region and period were assigned to 5000 groups, which are called “neurons”, of the SOM by using the training parameters. Second, each neuron was labeled, whenever possible, with the $p$CO$_2w$ value observed at the grid-cell at which Chl-a, SST, SSS, SIC, $x$CO$_2a$, and X and Y values were most similar to those of the neuron. Finally, each grid-cell in the analysis region and period
was assigned the $p_{\text{CO}_2w}$ value of the neuron whose Chl-a, SST, SSS, SIC, $x_{\text{CO}_2a}$, and X and Y values were most similar to those at that grid-cell. If the most similar neuron was not labeled with a $p_{\text{CO}_2w}$ value, then the $p_{\text{CO}_2w}$ value of the neuron that was most similar and labeled was used. A detailed procedure can be found in Telzewski et al. (2009) and Nakaoka et al. (2013).

3.5. Calculation of air–sea CO2 fluxes

We calculated monthly air–sea CO2 flux ($F$) values from the $p_{\text{CO}_2w}$ values estimated in Sect. 3.4 by using the bulk formula:

$$F = kL(p_{\text{CO}_2w} - p_{\text{CO}_2a}),$$  \hspace{1cm} (2)

where $k$ is the gas transfer velocity and $L$ is the solubility of CO2. The solubility of CO2 ($L$) was calculated as a function of SST and SSS (Weiss, 1974). We converted the interpolated NOAA marine boundary layer $x_{\text{CO}_2a}$ data (Sect. 2.2) to $p_{\text{CO}_2a}$ by using monthly sea-level pressure data and the water-vapor saturation pressure calculated from monthly SST and SSS (Murray, 1967).

The gas transfer velocity $k$ was calculated by using the formula of Sweeney et al.
(2007):

\[ k = 0.19 \left( \frac{Sc}{660} \right)^{-0.5} <W_{\text{NCEP2}}^2>, \] (3)

where Sc is the Schmidt number of CO\textsubscript{2} in seawater at a given SST and was calculated according to Wanninkhof (2014), “\(< >\)” denotes the monthly mean, and \(<W_{\text{NCEP2}}^2>\) is the monthly mean of the second moment of the NCEP2 6-hourly wind speed. The coefficient 0.19, which is the global average of \(0.27<W_{\text{NCEP1}}^2>/<W_{\text{NCEP2}}^2>\), is based on the one determined by Sweeney et al. (2007) but optimized for NCEP2 winds following the same method as Schuster et al. (2013) and Wanninkhof et al. (2013).

The suppression of gas exchange by sea ice was accounted for by correcting the air–sea CO\textsubscript{2} fluxes using the parameterization presented by Loose et al. (2009); the flux is proportional to \((1-\text{SIC})^{0.4}\). The SIC effect used in this study is consistent with the linear SIC effect derived from another approach by Takahashi et al. (2009) and Butterworth and Miller (2016) within the uncertainty of 30%. Following Bates et al. (2006), in the regions with SIC > 99 %, we used SIC = 99 % to allow for non-negligible rates of air–sea CO\textsubscript{2} exchange through leads, fractures, and brine channels (Semiletov et al., 2004; Fransson et al., 2017).
4. Uncertainty

4.1. Uncertainty in chlorophyll-a concentration data

Figure 3 shows original and interpolated Chl-a for 2012 as an example. Most interpolated Chl-a data are low concentrations because of high SIC and lack of daylight. Overall, the interpolated Chl-a data seem to fit well with the original data.

To evaluate our choice of Chl-a algorithm (i.e. the Arctic algorithm of Cota et al., 2004), we compared its calculated Chl-a values with those determined by using the standard algorithm of O’Reilly et al. (1998) and the coastal algorithm of Tassan (1994). The root mean square difference (RMSD) and correlation coefficient (r) between the original (i.e. non-interpolated) Chl-a values determined with the Arctic algorithm and the standard algorithm are 0.80 mg m\(^{-3}\) and 0.90, respectively, and between those determined with the Arctic algorithm and the coastal algorithm are 0.81 mg m\(^{-3}\) and 0.87, respectively (Table 1). For the interpolated Chl-a values, they are 0.37 mg m\(^{-3}\) and 0.92, and 0.48 mg m\(^{-3}\) and 0.86, respectively. The lower RMSD results from the fact that most of the interpolated Chl-a values represent low concentrations. The \(pCO_{2w}\) and \(CO_2\) fluxes determined using Chl-a from any of these algorithms as input to the SOM are consistent within uncertainties of the \(pCO_{2w}\) and \(CO_2\) flux estimates (see Sects. 4.2
To validate our Chl-a interpolation, we repeated the interpolation after randomly eliminating 10% of the satellite Chl-a values. We then used the eliminated original Chl-a data as independent data for validation. The RMSD and correlation coefficient between the interpolated and the independent original Chl-a data are 0.90 mg m\(^{-3}\) and 0.80, respectively. It means the interpolated Chl-a generally reproduced the Chl-a concentrations, and therefore is a meaningful parameter in the SOM process.

4.2. Uncertainty of \(pCO_2w\) mapping

Both observed and estimated \(pCO_2w\) tended to be high in the subpolar North Atlantic, the Laptev Sea, and the Canada Basin, and low in the Greenland Sea and the Barents Sea (Figs. 4a and 4b; Note that the spatial change depicted in Figs. 4a and 4b include differences in the observed seasons depending on the regions). However, the east-west contrast in the Bering Sea and the contrast between the Canada Basin and the Chukchi Sea are weaker in our estimates than those in the observations, and mean bias and RMSD are relatively large in those areas (Figs. 5a and 5b). The temporal changes in the observed and estimated \(pCO_2w\) are in phase, although the variability of the estimated values is somewhat suppressed compared to that of the observed data (Fig. 4c; Note that...
the temporal change depicted in Fig. 4c includes changes in the positions of the
observation points). Mean bias and RMSD show no long-term change (Fig. 5c).

The correlation coefficient between estimated and observed $p_{\text{CO}_2w}$ values is 0.82,
and the RMSD is 30 μatm, which is 9 % of the average and 58 % of the standard
deviation of the observed $p_{\text{CO}_2w}$ values (a performance level of 58 % is categorized as
“good” by Maréchal [2004]). The differences between the estimated and observed
values stem not only from the estimation error but also from the error of the gridded
observation data. The uncertainty of the $p_{\text{CO}_2w}$ measurements is 2–5 μatm (Bakker et
al., 2014), the uncertainty of the $p_{\text{CO}_2w}$ values calculated from dissolved inorganic
carbon and total alkalinity can be up to 14 μatm (Lueker et al., 2000), and the sampling
error of the gridded $p_{\text{CO}_2w}$ observation data was deduced from the standard errors of
monthly observed $p_{\text{CO}_2w}$ in the 1° × 1° grid-cells to be 7 μatm (Yasunaka et al., 2016).

To validate our estimated $p_{\text{CO}_2w}$ values for periods and regions without any observed
data, we repeated the mapping experiments after systematically excluding some of the
observed $p_{\text{CO}_2w}$ data when labeling the neurons; four experiments were carried out, by
excluding data (1) for 1997–2004, (2) for January to April, (3) from north of 80°N, and
(4) from the Laptev Sea (90°E – 150°E), where there are only a few $p_{\text{CO}_2w}$ observations.
The RMSDs between the estimates and the excluded observations are 54 μatm on
average, which is 1.8 times the RMSDs of the estimates based on all observations.

Hence the uncertainty in $p_{CO2w}$ might be as large as 16% in regions and periods without data.

### 4.3. Uncertainty of CO$_2$ flux estimates

Signorini and McClain (2009) estimated the uncertainty of the CO$_2$ flux resulting from uncertainties in the gas exchange parameterization to be 36%, and that resulting from uncertainties in the wind data to be 11%. The uncertainty for SIC is 5% (Cavalieri et al., 1984; Gloersen et al., 1993; Peng et al., 2013). The standard error of the sea-ice effect on gas exchange was estimated to about 30% by Loose et al. (2009). The uncertainty of $p_{CO2a}$ is about 0.5 μatm (http://www.esrl.noaa.gov/gmd/ccgg/mbl/mbl.html), and that of $p_{CO2w}$ was 30 μatm (Sect. 4.2); therefore, we estimated the uncertainty of $Δp_{CO2}$ ($=p_{CO2w}−p_{CO2a}$) to be 34% (average $Δp_{CO2}$ in the analysis domain and period was −89 μatm). The overall uncertainty of the estimated CO$_2$ fluxes is 59% ($[(0.36^2 + 0.11^2 + 0.05^2 + 0.3^2 + 0.34^2)^{1/2}]$) in the sea-ice covered region and 51% ($[(0.36^2 + 0.11^2 + 0.05^2 + 0.34^2)^{1/2}]$) in the ice-free region. In regions and periods without any observed data, where the uncertainty in $p_{CO2w}$ is 54 μatm and the uncertainty of the $Δp_{CO2}$ estimates can be as
high as 61 %, the uncertainty is 78 % ($\sqrt{0.36^2 + 0.11^2 + 0.05^2 + 0.3^2 + 0.61^2}$) in sea-ice covered regions, and 72 % ($\sqrt{0.36^2 + 0.11^2 + 0.05^2 + 0.61^2}$) in ice-free regions.

The average of the estimated CO$_2$ flux in the analysis domain and period is 4.8 mmol m$^{-2}$ d$^{-1}$; hence the uncertainty of the CO$_2$ flux estimate corresponds to 2.8 mmol m$^{-2}$ d$^{-1}$ in sea-ice covered regions and 2.4 mmol m$^{-2}$ d$^{-1}$ in ice-free regions. In regions and periods without observed data, the uncertainty corresponds to 3.7 mmol m$^{-2}$ d$^{-1}$ in the sea-ice covered region and 3.5 mmol m$^{-2}$ d$^{-1}$ in ice-free regions.

5. Results and discussion

5.1. Relationship between $p$CO$_2$ and chlorophyll-a

We compared the observed $p$CO$_2$$_{w}$ and the original non-interpolated Chl-a in spring (March–May) and summer (July–September) (Fig. 6a and b). In spring, when much of the Arctic Ocean is ice-covered, Chl-a is high in the Barents Sea and the Bering Strait ($>1$ mg m$^{-3}$). In summer, when the ice cover is less extensive, Chl-a is high in the Chukchi Sea, the Kara Sea, the Laptev Sea, and the East Siberian Sea ($>1$ mg m$^{-3}$) and especially high in the coastal region ($>2$ mg m$^{-3}$). $p$CO$_2$$_{w}$ is high in the Norwegian Sea in spring; and also high in the Kara Sea, the Laptev Sea and the Canada Basin during summer ($>300 \mu$atm). On the other hand, it is lower in the Chukchi Sea, Bering Strait
The overall relationship between $pCO_2w$ and Chl-a is negative where Chl-a $\leq 1$ mg m$^{-3}$ (70% of all the data; correlation coefficient $r = -0.36$, $P < 0.01$), but there is no significant relationship where Chl-a $> 1$ mg m$^{-3}$ (Fig. 7), as also shown by Olsen et al. (2008) in the subpolar North Atlantic. That is probably because high Chl-a usually appeared in the coastal region (Fig. 6b) which are affected by the river runoff.

To determine the spatial variability of the relationship between $pCO_2w$ and Chl-a, we calculated the correlation coefficients between $pCO_2w$ and Chl-a in a window of $\pm 5^\circ$ of latitude, and $\pm 30^\circ$ of longitude for each monthly $1^\circ \times 1^\circ$ grid-cell (Fig. 6d). We found negative relationships between $pCO_2w$ and Chl-a in the Greenland/Norwegian Seas and the Canada Basin. In the Greenland/Norwegian Seas, the relationship between $pCO_2w$ and Chl-a is strongly negative ($r < -0.4$) in spring and weakly negative ($-0.4 < r < 0$) in summer. Chl-a there is higher in summer than in spring, whereas nutrient concentrations are high in spring and low in summer (Fig. 6c). These correlations suggest that primary production draws down the $pCO_2w$ in spring, whereas in summer the primary production mostly depends on regenerated nutrients (Harrison and Cota, 1991) and the net CO$_2$ consumption is small, as also reported for the subpolar North Atlantic (Olsen et al., 2008). In the eastern Barents Sea, the Kara Sea and the East Siberian Sea, and the
Bering Strait, the relationships are positive because of water with high $pCO_{2w}$ and Chl-a in the coastal region subjected to river discharge (Murata, 2006; Semiletov et al., 2007; Anderson et al., 2009; Manizza et al., 2011). In the Chukchi Sea, the relationship is weak ($-0.2 < r < 0.2$), probably because the relationship is on smaller spatial and temporal scales than those represented by the window size used here, as shown by Mucci et al. (2010). Calcifying plankton bloom would also weaken the correlation since it will affect the $pCO_{2w}$ in the different way (Shutler et al., 2013; Fransson et al., 2017).

### 5.2. Spatiotemporal CO$_2$ flux variability

The 18-year annual mean CO$_2$ flux distribution shows that all areas of the Arctic Ocean and its adjacent seas were net CO$_2$ sinks over the time period that we investigated (Fig. 8). The annual CO$_2$ influx to the ocean is strong in the Greenland/Norwegian Seas ($9 \pm 3$ mmol m$^{-2}$ d$^{-1}$; 18-year annual mean $\pm$ uncertainty averaged over the area shown in Fig. 1), the Barents Sea ($10 \pm 3$ mmol m$^{-2}$ d$^{-1}$), and the Chukchi Sea ($5 \pm 3$ mmol m$^{-2}$ d$^{-1}$). In contrast, influx is weak and not statistically significantly different from zero in the Eurasian Basin and the Canada Basin, or in the Laptev Sea and the East Siberian Sea. Our annual CO$_2$ flux estimates are consistent with those reported by Yasunaka et al. (2016) and other previous studies (Bates and Mathis,
2009, and references therein). The estimated 18-year average CO$_2$ influx to the Arctic Ocean is 5 ± 3 mmol m$^{-2}$ d$^{-1}$, equivalent to an uptake of 180 ± 130 TgC y$^{-1}$ for the ocean area north of 65° N, excluding the Greenland/Norwegian Seas and Baffin Bay (10.7 × 10$^6$ km$^2$; see Fig. 1). This is within the range of other estimates (81–199 TgC y$^{-1}$; Bates and Mathis, 2009).

The CO$_2$ influxes to the Greenland/Norwegian Seas, the Barents Sea, and the Chukchi Sea are strongest in October (>10 mmol m$^{-2}$ d$^{-1}$; Fig. 9), when the winds strengthen with the approach of winter and the SIC and $p$CO$_2$w are still as low as in the summer. The total CO$_2$ influx to the Arctic Ocean is also strongest in October. The CO$_2$ influx shows a secondary maximum in February in the Greenland/Norwegian Seas and the Barents Sea because the strongest winds occur in that month (not shown).

The CO$_2$ influx has been increasing in the Greenland Sea and northern Barents Sea, and decreasing in the Chukchi Sea and southern Barents Sea (Fig. 10). The CO$_2$ flux trend corresponds well with the ΔpCO$_2$ trend, which in turn corresponds well with the SST trend. High $p$CO$_2$w (>500 μatm) has been sometimes observed in the Chukchi Sea after 2010 (Hauri et al. 2013). The increasing CO$_2$ influx in the northern Barents Sea also corresponds with the sea-ice retreat.
5.3. Impact of incorporating chlorophyll-a data in the SOM

To determine the impact of including Chl-a data in the SOM process, the analyses were repeated without Chl-a data. The RMSD of the resulting estimated $p_{CO_2w}$ values is 33 $\mu$atm, 3 $\mu$atm larger than the uncertainty of the estimates generated by including Chl-a in the SOM. Chl-a data thus improved the $p_{CO_2w}$ estimate (namely, a 10 % reduction of RMSD), even though 40 % of the Chl-a data labeled with $p_{CO_2w}$ observations were interpolated Chl-a values. This improvement resulted mainly from an improved representation of the seasonal cycle of $p_{CO_2w}$ by the SOM technique, which reproduced the observed cycle better when Chl-a was included (upper panels of Fig. 11). The seasonal cycles of $p_{CO_2w}$ estimates derived with the inclusion of Chl-a has a larger amplitude than the uncertainties, whereas the uncertainties are larger than the seasonal amplitude of $p_{CO_2w}$ derived without Chl-a (lower panels of Fig. 11). Note that the much larger seasonal amplitude in observed $p_{CO_2w}$ in each region (upper panels of Fig. 11) is due to the seasonal changes in the extent of the observed area: area average of observed $p_{CO_2w}$ in winter was obtained mostly from the data in the lower latitudes, where $p_{CO_2w}$ tends to be higher than that in the higher latitudes. The winter CO$_2$ influx in the Greenland/Norwegian Seas estimated including Chl-a is about 3 mmol m$^{-2}$ d$^{-1}$ less than those calculated without using Chl-a (Fig. 9), but this difference is smaller than
the uncertainties. The pattern in trends is similar to that for the estimates without using Chl-a (see Fig. 10 in Yasunaka et al., 2016).

The inclusion of Chl-a data also reduced the uncertainty of the estimated annual air-sea CO$_2$ flux integrated over the entire Arctic Ocean. Compared to the flux estimate determined by Yasunaka et al. (2016) of 180 ± 210 TgC y$^{-1}$, the CO$_2$ uptake in the Arctic Ocean estimated here is significant within its uncertainty (180 ± 130 TgC y$^{-1}$).

This improvement is the result of 1) the inclusion of Chl-a data in the SOM process (which reduced the uncertainty by 23 %); 2) the separate uncertainty estimates for ice-free and ice-covered regions (8 %); and 3) the addition of new observational $p$CO$_{2w}$ data (7 %).

6. Conclusions

By applying an SOM technique with the inclusion of Chl-a data to estimate $p$CO$_{2w}$, we produced monthly maps of air-sea CO$_2$ fluxes from 1997 to 2014 for the Arctic Ocean and its adjacent seas north of 60° N. The overall relationship between $p$CO$_{2w}$ and Chl-a is negative, consistent with it being determined by primary production, but the relationship depended on the season and the region. Adding Chl-a to the SOM process improved representation of the seasonal cycle of $p$CO$_{2w}$. A major goal of the
carbon-cycle research community in recent years has been to reduce the uncertainty in estimates of carbon reservoirs and fluxes. Our results contribute to this in that CO₂ uptake in the Arctic Ocean is demonstrated with high significance. The resulting estimate of the annual Arctic Ocean CO₂ uptake of 180 TgC y⁻¹ is significant with an uncertainty of only ± 130 TgC y⁻¹. This is a substantial improvement over earlier estimates, and is due mainly to the incorporation of Chl-a data.

The addition of new observational data from SOCATv4 and GLODAPv2 reduced the overall uncertainty in the mapped pCO₂w: a 33 % increase in the number of observations induced a 7 % reduction in the uncertainty. However, there are still too few observations in the Kara Sea, the Laptev Sea, the East Siberian Sea and the Eurasian Basin to determine seasonal and interannual variations there. To improve our understanding of the variability in air–sea CO₂ fluxes in the Arctic, it is of critical importance to obtain additional ocean CO₂ measurements to fill these data gaps.

The monthly CO₂ flux data presented in this paper will be available at the JAMSTEC website (http://www.jamstec.go.jp/res/ress/yasunaka/co2flux_v2).

Acknowledgments
We thank the many researchers and funding agencies responsible for the collection of data and quality control for their contributions to SOCAT and GLODAPv2. We are grateful for the use of the CO2SYS program obtained from the Carbon Dioxide Information Analysis Center (http://cdiac.ornl.gov/oceans/co2rprt.html), and SOM Toolbox Version 2 developed by the Laboratory of Information and Computer Science at Helsinki University of Technology (http://www.cis.hut.fi/projects/somtoolbox). We thank ACRI-ST, France, for developing, validating, and distributing the GlobColour data used in this work. This work was financially supported by the Arctic Challenge for Sustainability (ArCS) Project funded by the Ministry of Education, Culture, Sports, Science and Technology, Japan. Are Olsen was supported by grants from the Norwegian Research Council (Subpolar North Atlantic Climate States [SNACS] 229752 and the Norwegian component of the Integrated Carbon Observation System [ICOS-Norway] 245927). Mario Hoppema was partly supported by the German Federal Ministry of Education and Research (grant no. 01LK1224I; ICOS-D). Siv K. Lauvset acknowledges support from the Norwegian Research Council (VENTILATE, 229791) and the EU H2020 project AtlantOS (grant agreement no. 633211). Rik Wanninkhof and Taro Takahashi acknowledge support from the Office of Oceanic and Atmospheric Research (OAR) of the United States NOAA, including resources from the Ocean Observation
and Monitoring Division of the Climate Program Office (fund reference100007298).

References

Anderson, L. G., Jutterström, S., Hjalmarsson, S., Wåhlström, I., and Semiletov, I. P.:


Ardyna, M., Babin, M., Gosselin, M., Devred, E., Rainville, L., and Tremblay, J.-É.:


Johnsen, G., King, J. A., Koch, B. P., Koenig, Z., Kwasniewski, S., Laney, S. R.,

Nicolaus, M., Pavlov, A. K., Polashenski, C. M., Provost, C., Rösel, A., Sandbu, M.,

Spreen, G., Smedsrud, L. H., Sundfjord, A., Taskjelle, T., Tatarek, A., Wiktor, J.,

Bakker, D. C. E., Pfeil, B., Smith, K., Harasawa, S., Landa, C., Nakaoka, S., Nojiri, Y.,

Metzl, N., O'Brien, K. M., Olsen, A., Schuster, U., Tilbrook, B., Wanninkhof, R.,


Burger, E., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Cosca, C., Currie, K.,

Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Greenwood, N., Gregor, L.,

Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humpt,

L., Monteiro, P., Munro, D., Murata, A., Newberger, T., Omar, A. M., Ono, T.,

Paterson, K., Pierrot, D., Robbins, L. L., Sabine, C. L., Saito, S., Salisbury, J.,

Schneider, B., Schlitzer, R., Sieger, R., Skjelvan, I., Steinhoff, T., Sullivan, K. F., St,

L., Monteiro, P., Munro, D., Murata, A., Newberger, T., Omar, A. M., Ono, T.,

Paterson, K., Pierrot, D., Robbins, L. L., Sabine, C. L., Saito, S., Salisbury, J.,

Schneider, B., Schlitzer, R., Sieger, R., Skjelvan, I., Steinhoff, T., Sullivan, K. F.,

Sutherland, S. C., Sutton, A. J., Sweeney, C., Tadokoro, K., Takahashi, T., Telszewski,


Dickson, A. G.: Standard potential of the reaction: AgCl(s) + 1/2H₂(g) = Ag(s) + HCl(aq), and the standard acidity constant of the ion HSO₄⁻ in synthetic seawater from 273.15 to 318.15 K. J. Chemical Thermodynamics, 22, 113–127, 1990.


Gao, Z., Chen, L., Sun, H., Chen, B., and Cai, W.-J.: Distributions and air–sea fluxes of


Murata, A.: Increased surface seawater pCO₂ in the eastern Bering Sea shelf: An effect


blooms of *Karenia mikimotoi* using MODIS measurements: a case study of
Seto-Inland Sea, Japan, Remote Sensing of Environment, 129, 185–196,

Smyth, T. J., Tyrrell, T., Tarrant, B.: Time series of coccolithophore activity in the
Barents Sea, from twenty years of satellite imagery, Geophys. Res. Lett., 31, L11302,

Steele, M., Morley, R., and Ermold, W.: PHC: A global ocean hydrography with a high

Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G., Sarmiento, J. L.,
and Wanninkhof, R.: Constraining global air-sea gas exchange for CO$_2$ with recent
bomb $^{14}$C measurements, Global Biogeochem. Cycles, 21, GB2015,

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman,
D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C. E.,
Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y.,
Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook,
Baar, H. J. W.: Climatological mean and decadal changes in surface ocean pCO$_2$, and


Figure 1: Map of the Arctic Ocean and its adjacent seas. Gray contour lines show the 1000, 2000, 3000, and 4000 m isobaths. Blue lines show the 17-year annual mean position of the ice edge (SIC = 15%). Area for the mapping is north of 60° N (heavy black circle). Sectors selected for regional analysis are the Arctic Ocean (dashed magenta line), the Greenland/Norwegian Seas (green 1), the Barents Sea (green 2), and the Chukchi Sea (green 3).
Figure 2: (a) The number of ocean surface CO$_2$ data in the grid boxes ($1^\circ \times 1^\circ$) used in this study. Data are from SOCATv4, LDEOv2014, GLODAPv2, and collected by R/V Mirai of JAMSTEC between 1997 and 2014. (b) Monthly time series of the number of CO$_2$ data in the analysis area (north of 60° N).
Figure 3: (a) Original and (b) interpolated Chl-a [mg m$^{-3}$] in July 2012 (upper panels), and along 75°N in 2012 (lower panels). Black lines denote SIC of 50% and 90%. Gray areas in (a) indicate missing Chl-a data.
Figure 4: (a) Observed $p_{CO_{2w}}$ averaged over the whole analysis period [μatm]. (b) Estimated $p_{CO_{2w}}$ averaged over the grid boxes in which observed $p_{CO_{2w}}$ values were available [μatm]. (c) Monthly time series of observed $p_{CO_{2w}}$ averaged over the entire analysis area (black), and estimated $p_{CO_{2w}}$ averaged over the grid boxes in which observed $p_{CO_{2w}}$ values were available (green) [μatm].
Figure 5: (a) Bias (estimate–observation) and (b) root-mean-square-difference between observed and estimated $p_{CO_{2}}$w averaged over the whole analysis period [$\mu$atm]. (c) Bias (estimate–observation; black) and root-mean-square-difference (green) between observed and estimated $p_{CO_{2}}$w averaged over the entire analysis area [$\mu$atm].
Figure 6: (a) Observed $pCO_{2w}$ [$\mu$atm], (b) non-interpolated Chl-a [mg m$^{-3}$] and (c) surface nitrate concentration [$\mu$mol l$^{-1}$] in March–May (left), and July–September (right) from 1997 to 2014. (d) Spatial correlation (correlation coefficient, $r$) between $pCO_{2w}$ and Chl-a in a window size of ±1 month, ±5° of latitude, and ±30° of longitude in March–May (left), and July–September (right). Darker hatched areas represent values in grids where correlations are insignificant ($P > 0.05$).
Figure 7: Observed $pCO_{2w}$ [μatm] vs. satellite Chl-a [mg m$^{-3}$] in the Arctic Ocean and its adjacent seas (north of 60° N) from 1997 to 2014. Colors indicate the number of data pairs in a 0.1 mg m$^{-3}$ × 5 μatm bin when Chl-a ≤ 5 mg m$^{-3}$, or in a 1 mg m$^{-3}$ × 5 μatm bin when Chl-a > 5 mg m$^{-3}$.
Figure 8: Eighteen-year annual means of CO$_2$ flux [mmol m$^{-2}$ day$^{-1}$] (negative values indicate flux into the ocean). Darker hatched areas represent show values in grids where fluxes were smaller than the uncertainty, estimated as described in the text.
Figure 9: Eighteen-year monthly mean CO₂ flux [mmol m⁻² day⁻¹] averaged over (a) the Greenland/Norwegian Seas, (b) the Barents Sea, (c) the Chukchi Sea, and (d) the Arctic Ocean. Black lines with triangles show estimates without Chl-a by Yasunaka et al. (2016); magenta lines with open circles show estimates with Chl-a. Error bars show the uncertainty, estimated as described in the text.
Figure 10: Trends in (a) CO$_2$ flux [mmol m$^{-2}$ day$^{-1}$ decade$^{-1}$], (b) $\Delta p$CO$_2$ [µatm decade$^{-1}$], and (c) SIC [% decade$^{-1}$]. Darker hatched areas represent values in grids where trend values were less than the uncertainty, estimated as described in the text.
Figure 11: Eighteen-year averaged $pCO_{2w}$ seasonal variations [μatm] in (a) the Greenland/Norwegian Seas, (b) the Barents Sea, and (c) the Chukchi Sea. Black lines with triangles show estimates without Chl-a; magenta lines with open circles show estimates with Chl-a; green lines with closed circles show observed values. The upper panels show $pCO_{2w}$ averaged over the grid boxes in which observed $pCO_{2w}$ values were available, and the lower panels show $pCO_{2w}$ averaged for all grid cells with each region. Error bars show the uncertainty, estimated as described in the text.
Table 1: RMSD [mg m\(^{-3}\)] and correlation (r) between Chl-a values

<table>
<thead>
<tr>
<th></th>
<th>standard algorithm</th>
<th>coastal algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSD</td>
<td>r</td>
</tr>
<tr>
<td>Chl-a from Arctic algorithm</td>
<td>0.80</td>
<td>0.90</td>
</tr>
<tr>
<td>Interpolated Chl-a from Arctic algorithm</td>
<td>0.37</td>
<td>0.92</td>
</tr>
</tbody>
</table>