ESSDD

Interactive
comment

Interactive comment on "A synthetic satellite dataset of *E. huxleyi* spatio-temporal distributions and their impacts on Arctic and Subarctic marine environments (1998–2016)" *by* Dmitry Kondrik et al.

A.J. Poulton (Referee)

a.poulton@hw.ac.uk

Received and published: 5 November 2018

In this study, Kondrik et al. have compiled satellite observations of coccolithophore blooms in the high-latitude northern hemisphere and combined them with various algorithms, published by the authors, to estimate coccolith concentrations and the impact of coccolithophores on the air-sea CO2 fluxes. The dataset is of considerable interest, with coccolithophore blooms in the high-latitude polar seas generally understudied and often poorly sampled in situ. The 18-year time-series of observations represents an exciting opportunity to examine temporal trends over a relatively long period and I am

sure the dataset will be used extensively. The manuscript is well written and I only have minor comments/suggestions for further clarity.

pg 1, Ln 1 - How do you know its E. huxleyi rather than other coccolithophores? Would it not be safer to say coccolithophores? Though many factors make E. huxleyi the most likely source of satellite-detectable reflectance, other coccolithophores can bloom and some can be a significant presence within blooms. Also, its more typical to give the full species name (i.e. Emiliania huxleyi).

pg 1, Ln 7 – What do the authors mean by 'activity' in the context of the first line of the paragraph? Distribution and impact on the air-sea flux of CO2 is what is presented.

pg 1, Ln 16 – 'Ongoing climate change is a background of numerous emerging hot topics' seems a rather cryptic opening line for the paper and it's not obviously clear what the authors mean.

pg 1, Ln 20 – 'most widespread in the world's oceans': please clarify this statement, do you mean 'the' most widespread coccolithophore?

pg 1, Ln 25 – Rivero-Calle et al. (2015) show increases in occurrence across the North Atlantic rather than a polewards expansion. Other authors have discussed polar expansion ranges (e.g. Smyth et al., 2004; Winter et al., 2014) and are more relevant to the current study.

pg 2, Ln 6 – Please rephrase 'solely satellite remote sensing approach means..'.

pg 2, Ln 21 – Please explain 'viz. North', do you mean the North Atlantic?

pg 4, Ln 30-32 – Please note that the use of a fixed carbon mass per coccolith (m) has limitations and that coccolith content between different morphotypes of E. huxleyi can be considerable (e.g., Poulton et al., 2011; Müller et al., 2015) and may lead to over- or underestimation depending on which morphotype(s) is present in the bloom. This directly influences the scaling up of coccolith mass to PIC content in this study, and is an important factor when considering bloom PIC production (see e.g. Poulton et al.,

2013; Balch et al., 2014).

Figure 2c - What are the units for panel c? Tons per unit area/pixel? Would it not make more sense to express in similar volumetric units as in panel b (i.e. m-3)? It is also not clear how the authors get to 30 tons of PIC; e.g. 250-400 x109 coccoliths m-3 equates to ∼50 to 80 mg C m-3 or ∼4 to 7 mmol C m-3 based on a coccolith mass of 0.2 pg C.

References

Balch, W.M., Drapeau, D.T., Bowler, B.C., Lyczkowski, E.R., Lubelczyk, L.C., Painter, S.C., and Poulton, A.J.: Surface biological, chemical, and optical properties of the Patagonian Shelf coccolithophore bloom, the brightest waters of the Great Calcite Belt, Limnol. Oceanogr., 59, 1715-1732, doi: 10.4319/lo.2014.59.5.1715.

Müller, M.N., Trull, T.W., and Hallegraeff, G.M.: Differing responses of three Southern Ocean Emiliania huxleyi ecotypes to changing seawater carbonate chemistry, Mar. Ecol. Prog. Ser., 531, 81-90, doi: 10.3354/meps11309, 2015.

Poulton, A.J., Young, J.R., Bates, and Balch, W.M.: Biometry of detached Emiliania huxleyi coccoliths along the Patagonian Shelf, Mar. Ecol. Prog. Ser., 443, 1-17, doi: 10.3354/meps09445, 2011.

Poulton, A.J., Painter, S.C., Young, J.R., Bates, N.R., Bowler, B., Drapeau, D., Lyczsckowski, E., and Balch, W.M.: The 2008 Emiliania huxleyi bloom along the Patagonian Shelf: Ecology, biogeochemistry, and cellular calcification, Global Biogeochem. Cycl., 27, 1-11, doi: 10.1002/2013GB004641, 2013.

Smyth, T.J., Tyrrell, T., and Tarrant, B.: Time series of coccolithophore activity in the Barents Sea, from twenty years of satellite imagery, Geophys. Res. Lett., 31, L11302, doi: 10.1029/2004GL019735, 2004.

Winter, A., Hendericks, J., Beaufort, L., Rickaby, R.E.M., and Brown, C.W.: Poleward expansion of the coccolithophore Emiliania huxleyi, J. Plankton Res., 36, 316-325, doi: 10.1093/plankt/fbt110, 2014.

ESSDD

Interactive
comment

Interactive comment on "A synthetic satellite dataset of *E. huxleyi* spatio-temporal distributions and their impacts on Arctic and Subarctic marine environments (1998–2016)" *by* Dmitry Kondrik et al.

Dmitry Kondrik et al.

ekazakov@niersc.spb.ru

Received and published: 7 November 2018

Dear Dr. Poulton,

Thank you for your thoughtful comments and recommendations. We are especially appreciative of the list of references.

Below are our answers.

Pg.1. Ln. 1: a) We will certainly change E. huxleyi for Emiliania huxleyi.

b) For all target seas we collected published reports from in situ/shipborne/laboratory studies explicitly indicating that the coccolithophore blooms were produced by E. huxleyi (see the attached specific list of references) with two exceptions for the Norwegian and Iceland seas, where along with E. huxleyi, Coccolithus pelagicus composes the coccolithophore community. However, as in situ determinations showed in the overwhelming cases the concentrations of cells of Coccolithus pelagicus were marginal (see e.g. Dylmer et al., 2015). This is the reason why we prefer leaving E. huxleyi instead of coccolithophores. A large number of papers on calcifying alga blooms in our targeted seas define the bloom-producing species as E. huxleyi.

Pg. 1, Ln. 7: By "activity" we meant the release of $CaCO_3$ in water and a decrease of uptake of dissolved CO_2 by E. huxleyi cells (e.g. Kondrik et al., 2018). In the revised version of the paper we will specify the actual meaning of the employed word "activity".

Pg. 1., Ln.16: It appeared to us that the issue of consequences of ongoing climate change–driven consequences is presently a commonplace, not requiring any further specialization. Indeed, the consequences are multifaceted, with numerous forward and feedback interactions and relate to many spheres of knowledge. So we choose to extend this phrase a little bit and provide this sentence with a reference that reasonably overarches the main dimensions of this phenomenon.

Pg. 1, Ln. 20: Yes, we will change for "the most widespread coccolithophore".

Pg. 1. Ln. 25; You are right, and we will add the reference "Winter et al., 2014".

Pg. 2., Ln 6:We agree that this phrase is kind of awkward and we will reword it as follows: "solely satellite remote sensing approach is..."

Pg. 2. Ln. 21: the following change will be made: the North, Labrador (with adjacent North Atlantic open waters), Norwegian, Barents, Greenland and Bering seas.

Pg. 4, Lns 30-32+ Figure 2c: The total content of PIC, Mpic, was determined for each 8-day time-period through multiplication of the carbon mass per coccolith, m, the coc-

colith concentration, Ccc, MLD and the bloom area, S. The value of m was equalled to 0.2pg. While most historical reports support this estimation, it is likely that the employment of this conversion might lead to either (i) some underestimation of PIC since it nevertheless neglects rare, relatively large, suspended calcite particles (PIC concentration per coccolith is ∼0.26 pg by Balch et al.(1991) and 0.5-0.6 pg by Holliganet al.(1983)) or (ii) some underestimation as there are in situ data indicating that many coccoliths in E. huxleyi blooms are either fragmented due to wave action (Holliganet al. 1993b) or just of a smaller size (PIC concentration is 0.13 pg) (Fernandez et al. 1993, Fritz 1999). Thus on balance, the selected value of m, in all probability, is a reasonably good estimate which is supported by the historical literature (Balch et al. 2005). The respective details are provided in section 2. Accordingly, the numbers in Figure 2c are indeed in tons as they reflect the content of PIC in a pixel-size column with the vertical extent equal to the respective MLD that was ascribed to each pixel within the bloom area. The respective methodology is described in detail in Kondrik et al., 2017 and will be given in the text.

Again, we express our gratitude to the referee for his very valuable comments.

Publications explicitly indicating the kind of coccolithophore species forming bloom in the target seas:

Barents Sea (Olson & Strom, 2002)

Bering Sea (Sukhanova and Flint, 1998)

North Sea (Holligan et al., 1993b; Buitenhuis et al., 1996)

Norwegian Sea (Baumann et al., 2000)

Labrador Sea (Okada & McIntyre, 1979)

North Atlantic (Holligan et al., 1993a)

Greenland Sea (Dylmer et al., 2015)

References

Balch, W. M., P. M. Holligan, S. G. Ackleson, and K. J. Voss. 1991. "Biological and Optical Properties of Mesoscale Coccolithophore Blooms in the Gulf of Maine." Limnology and Oceanography 36: 629–643. doi:10.4319/lo.1991.36.4.0629

Balch, W. M., H. Gordon, B. C. Bowler, D. T. Drapeau, and E. S. Booth. 2005. "Calcium Carbonate Measurements in the Surface Global Ocean Based on Moderate-Resolution Imaging Spectrometer Data." Journal of Geophysical Research110: C07001. doi:10.1029/2004JC002560

Baumann, Karl-Heinz &Andruleit, HA &Samtleben, C. (2000). Coccolithophores in the Nordic Seas: Comparison of living communities with surface sediment assemblages. Deep-sea Research Part Ii-topical Studies in Oceanography - DEEP-SEA RES PT II-TOP ST OCE. 47. 1743-1772. 10.1016/S0967-0645(00)00005-9.

Buitenhuis, E., J. van Bleijswijk, D. Bakker, and M. Veldhuis, Trends in inorganic and organic carbon in a bloom of Emilianiahuxleyi in the North Sea, Mar. Ecol. Prog. Ser., 143, 271-282, 1996

Dylmer, C. V., Giraudeau, J., Hanquiez, V., &Husum, K. (2015). The coccolithophoresEmilianiahuxleyi and Coccolithuspelagicus : Extant populations from the Norwegian–Iceland Seas and Fram Strait. Deep Sea Research Part I: Oceanographic Research Papers. https://doi.org/10.1016/j.dsr.2014.11.012

Fernandez, E., P. Boyd, P. M. Holligan, and D. S. Harbour. 1993. "Production of Organic and Inorganic Carbon within a Large Scale Coccolithophore Bloom in the North Atlantic Ocean." Marine Ecology Progress Series 97: 271–285. doi:10.3354/meps097271

Fritz, J. J. 1999. "Carbon Fixation and Coccolith Detachment in the Coccolithophore E. Huxleyi in Nitrate-Limited Cyclostats." Marine Biology 133: 509–518. doi:10.1007/s002270050491

Holligan, P. M., M. Viollier, C. Dupouy, and J. Aiken. 1983. "Satellite Studies on the

Distribution of Chlorophyll and Dinoflagellate Blooms in the Western English Channel." Continental Shelf Research 2 (2–3): 81–96. doi:10.1016/0278-4343(83)90009-2

Holligan, P. M., E. Fernández, J. Aiken, W. M. Balch, P. Boyd, P. H. Burkill, M. Finch, et al. 1993a. "A Biogeochemical Study of the Coccolithophore, Emilianiahuxleyi, in the North Atlantic." Global Biogeochemical Cycles 7: 879–900. doi:10.1029/93GB01731.

Holligan PM, Groom SB, Harbour DS (1993b) What controls the distribution of the coccolithophore, Emilianiahuxleyi, in the North Sea? Fish Oceanogr 2(3/4):17

Kondrik Dmitry, Dmitry Pozdnyakov & Lasse Pettersson (2017) Particulate inorganic carbon production within E. huxleyi blooms in subpolar and polar seas: a satellite time series study (1998–2013), International Journal of Remote Sensing, 38:22, 6179-6205, DOI: 10.1080/01431161.2017.1350304

Kondrik, D. V., Pozdnyakov, D. V., and Johannessen, O. M.: Satellite evidence that E. huxleyi phytoplankton blooms weaken marine carbon sinks, Geophysical Research Letters, 45, 2, 846-854, doi:10.1002/2017GL076240, 2018

Okada, H., and A. McIntyre. 1979. "Seasonal Distribution of Modern Coccolithophores in the Western North Atlantic Ocean." Marine Biology 54: 319–328. doi:10.1007/BF00395438

Olson, M. B., and S. L. Strom. 2002. Phytoplankton growth, microzooplankton herbivory and community structure in the southeast Bering Sea: Insight into the formation and temporal persistence of an Emilianiahuxleyi bloom. Deep-Sea Res. II 49: 5969–5990

Sukhanova, I. N., and M. V. Flint, Anomalous blooming of coccolithophorids over the eastern Bering Sea shelf, Oceanology, 38, 502 – 505, 1998

Winter, A., Hendericks, J., Beaufort, L., Rickaby, R.E.M., and Brown, C.W.: Poleward expansion of the coccolithophore Emiliania huxleyi, J. Plankton Res., 36, 316-325, doi: 10.1093/plankt/fbt110, 2014

Open Access

Earth System
Science
Data

Discussions

ESSDD

Interactive
comment

Interactive comment on "A synthetic satellite dataset of *E. huxleyi* spatio-temporal distributions and their impacts on Arctic and Subarctic marine environments (1998–2016)" *by* Dmitry Kondrik et al.

NEUKERMANS (Referee)

griet.neukermans@obs-vlfr.fr

Kondrik and collaborators present a 19-year satellite time series of Emiliania huxleyi bloom area, calcite content, and associated increase in in-water pCO_2 in four selected areas of the high-latitude northern hemisphere. The dataset is only partly unique, in the sense that a 19-year global remote sensing dataset of E. huxleyi bloom extent, coccolith concentration, and PIC content can also be easily obtained elsewhere. Therefore uniqueness only applies to pCO_2. This dataset could be useful, but I request a few substantial modifications that I believe are necessary to improve understanding and

C1

quality of the dataset: (1) some flaws in the dataset (pointed out below, 1a and 1b) will need to be fixed, (2) error estimates for remotely sensed quantities must be provided, and (3) in its present form, the study/data is not correctly positioned within the state-of-the-art literature and other available datasets.

(1a) It appears from Fig. 4 that the E. huxleyi bloom dataset includes false positives, a problem that is particularly evident in the Bering Sea (1998-2001) where the authors have detected blooms initiating in winter and lasting about 10 months as previously reported from ocean colour remote sensing data (Iida et al., 2002). However, ship-borne measurements have identified resuspended diatom frustules as the cause of these bright waters in winter-spring instead of E. huxleyi blooms (Broerse et al., 2003). This invalidates the authorial E. huxleyi bloom detection algorithm and all derived products in the Bering Sea from late fall to spring. I further fail to see how the algorithms used by the authors (Kondrik et al. 2017; Kondrik et al. 2018) to detect E. huxleyi blooms present an advance to NASA's standard method of E. huxleyi bloom classification (Brown and Yoder, 1994), and many other subsequent bloom detection methods (Iglesias-Rodriguez et al., 2002; Iida et al., 2002; Iida et al., 2012; Moore et al., 2012). (1b) The remote sensing algorithm for pCO2 estimation is a simple linear regression between observations of Delta_pCO2 and remote sensing reflectance Rrs in a blue waveband. This relationship is strictly empirical and does not appear to have theoretical grounds; I believe the user should be aware of this. Not surprisingly, there is an enormous spread along this regression line such that for a given reflectance value the estimated Delta_pCO2 has a confidence interval with a width of 50 ppm and even wider for denser blooms. Furthermore, the residuals of the regression are clearly unevenly distributed, with a strong tendency to underestimate Delta_pCO2 at higher reflectances. This relationship should be explicitly stated, which is not presently the case, including all relevant regression statistics, and especially a figure showing the observations and the fitted line so that the user can better grasp the errors of the algorithm. (2) Whereas the statistics of the validation of the retrieved coccolith concentration are given in section 2.2, the accompanying figure is missing.

No uncertainty assessment is given for pCO2 (see previous comment). (3) A 19-year global remote sensing dataset of PIC concentration merging all ocean colour satellite missions can be obtained here: http://www.globcolour.info/ in temporal resolutions ranging from daily to monthly, spatial resolution ranging from 4km to 100km, and various geographical projections. From PIC concentration, coccolith concentration can be derived using a fixed mass per coccolith (as you do too), and PIC content can also be easily derived by combining with a climatology for Mixed layer depth available here http://www.ifremer.fr/cerweb/deboyer/mld/Surface_Mixed_Layer_Depth.php. I therefore suggest you remove all statements of uniqueness of your PIC dataset (e.g., page 2, lines 24-26). The statements on page 2 lines 11-16, "Prior to the publication of Kondrik et al. (2018), no attempts have been undertaken to retrieve from space... No concatenated time series data are available to date on the associated bloom intensity..." are thus simply incorrect. I also suggest you appropriately reference the work of (Shutler et al., 2013) entitled "Coccolithophore surface distributions in the North Atlantic and their modulation of the air-sea flux of CO2 from 10 years of satellite Earth observation data Âż, which is very similar to your work on remote sensing of pCO2 in Ehux blooms, but is mentioned nowhere. Page2 Line 8-10: "Until recently, only few satellite studies were published on the typical locations of E. huxleyi blooms and associated concentrations of PIC in surface waters within the bloom area". It appears to me you missed a vast body of literature: (Balch et al., 1991; Balch et al., 1996; Gordon et al., 2001; Smyth et al., 2004; Signorini and McClain, 2009; Moore et al., 2012; Hopkins et al., 2015; Balch et al., 2016; Neukermans et al., 2018) etc.

Further comments : Title : add "blooms" after "E. huxleyi" Abstract : delete "detailed information on E. huxleyi impacts within the bloom area on marine environments", as this suggests that you are detailing ecological impacts

P1, L16 : "Ongoing climate change is a background of numerous emerging hot topics." is a rather meaningless opening sentence. P1 L25 : Rivero-Calle is not the right reference for poleward expansion of coccolithophores, instead use (Winter et al., 2014;

Neukermans et al., 2018). "gradually propagating in the poleward direction" ; the poleward expansion is not gradual, as expansion rates exhibit stark jumps as demonstrated in (Neukermans et al., 2018). P2, L1-4 : a lot of statements for only one reference. P2, L23 : replace 1918-2016 by 1998-2016 P2, L20 : remove "original" P3 L1 : spell out OC CCI P6 L1 : "in the cause of satellite processing" ?, rephrase P7 L10-15 and L24-28 : same paragraph appears twice. P7 L31 :"1,105,6800 km2" commas are in the wrong place

References : Balch WM, Bates NR, Lam PJ, Twining BS, Rosengard SZ, Bowler BC, Drapeau DT, Garley R, Lubelczyk LC, Mitchell C, et al. 2016. Factors regulating the Great Calcite Belt in the Southern Ocean and its biogeochemical significance. Global Biogeochem Cycles 30(8): 1124–1144. Wiley-Blackwell. doi: 10.1002/2016GB005414 Balch WM, Holligan PM, Ackleson SG, Voss KJ. 1991. Biological and optical properties of mesoscale coccolithophore blooms in the Gulf of Maine. Limnol Oceanogr 36(4): 629–643. doi: 10.4319/lo.1991.36.4.0629 Balch WM, Kilpatrick KA, Holligan P, Harbour D, Fernandez E. 1996. The 1991 coccolithophore bloom in the central North Atlantic. 2. Relating optics to coccolith concentration. Limnol Oceanogr 41(8): 1684–1696. doi: 10.4319/lo.1996.41.8.1684 Broerse AT., Tyrrell T, Young J., Poulton A., Merico A, Balch W., Miller P. 2003. The cause of bright waters in the Bering Sea in winter. Cont Shelf Res 23(16): 1579–1596. doi: 10.1016/j.csr.2003.07.001 Brown CW, Yoder JA. 1994. Coccolithophorid blooms in the global ocean. J Geophys Res 99(C4): 7467. doi: 10.1029/93JC02156 Gordon HR, Boynton GC, Balch WM, Groom SB, Harbour DS, Smyth TJ. 2001. Retrieval of coccolithophore calcite concentration from SeaWiFS Imagery. Geophys Res Lett 28(8): 1587–1590. doi: 10.1029/2000GL012025 Hopkins J, Henson SA, Painter SC, Tyrrell T, Poulton AJ. 2015. Phenological characteristics of global coccolithophore blooms. Global Biogeochem Cycles 29(2): 239–253. doi: 10.1002/2014GB004919 Iglesias-Rodriguez MD, Brown CW, Doney SC, Kleypas JA, Kolber D, Kolber Z, Hayes PK, Falkowski PG. 2002. Representing key phytoplankton functional groups in ocean carbon cycle models: Coccolithophorids. Global Biogeochem Cycles 16(4): 47-1-47–20. doi: 10.1029/2001GB001454 Iida T,

Mizobata K, Saitoh S-I. 2012. Interannual variability of coccolithophore Emiliania huxleyi blooms in response to changes in water column stability in the eastern Bering Sea. Cont Shelf Res 34: 7–17. Pergamon. doi: 10.1016/J.CSR.2011.11.007 Iida T, Saitoh SI, Miyamura T, Toratani M, Fukushima H, Shiga N. 2002. Temporal and spatial variability of coccolithophore blooms in the eastern Bering Sea, 1998-2001. Prog Oceanogr 55(1–2): 165–175. Pergamon. doi: 10.1016/S0079-6611(02)00076-9 Moore TS, Dowell MD, Franz BA. 2012. Detection of coccolithophore blooms in ocean color satellite imagery: A generalized approach for use with multiple sensors. Remote Sens Environ 117: 249–263. doi: 10.1016/j.rse.2011.10.001 Neukermans G, Oziel L, Babin M. 2018. Increased intrusion of warming Atlantic water leads to rapid expansion of temperate phytoplankton in the Arctic. Glob Chang Biol 24(6): 2545–2553. Wiley/Blackwell (10.1111). doi: 10.1111/gcb.14075 Shutler JD, Land PE, Brown CW, Findlay HS, Donlon CJ, Medland M, Snooke R, Blackford JC. 2013. Coccolithophore surface distributions in the North Atlantic and their modulation of the air-sea flux of CO_2 from 10 years of satellite Earth observation data. Biogeosciences 10(4): 2699–2709. doi: 10.5194/bg-10-2699-2013 Signorini SR, McClain CR. 2009. Environmental factors controlling the Barents Sea spring-summer phytoplankton blooms. Geophys Res Lett 36(10): L10604. doi: 10.1029/2009GL037695 Smyth TJ, Tyrell T, Tarrant B. 2004. Time series of coccolithophore activity in the Barents Sea, from twenty years of satellite imagery. Geophys Res Lett 31(11): L11302. doi: 10.1029/2004GL019735 Winter A, Henderiks J, Beaufort L, Rickaby REM, Brown CW. 2014. Poleward expansion of the coccolithophore Emiliania huxleyi. J Plankton Res 36(2): 316–325. Oxford University Press. doi: 10.1093/plankt/fbt110

ESSDD

Interactive
comment

Interactive comment on "A synthetic satellite dataset of *E. huxleyi* spatio-temporal distributions and their impacts on Arctic and Subarctic marine environments (1998–2016)" *by* Dmitry Kondrik et al.

Dmitry Kondrik et al.

ekazakov@niersc.spb.ru

Received and published: 24 December 2018

1. Regarding the status of our database.

With all respect for the reviewer, we can't agree with the reviewer's opinion that if any dataset(s) including the parameter(s) listed in our paper already exist(s) then our dataset can not be qualified as unique. The uniqueness of our dataset resides in that that

(A) it combines a spatially and temporarily collocated setof parameters (not solely e.g.

coccolith concentration)inherent in /related to the E. huxleyi blooms phenomenon in a number of polar and subpolar marine regions

(B)over the satellite measurement period of nearly 20 years (1998-2016), it is

(C) based on merged data from several satellites of the modern era (such as SeaWiFS, MODIS, MERIS, VIIRS), and

(D) designed specifically for the user convenience. Thus importantly, the user does not need to compose such a comprehensive database but use the already collected and user-friendly organized data source. Incidentally, this is explicitly corroborated by the reviewer himself/herself: even a spaceborne database on coccolith concentration per se is not available and needs to be retrieved from satellite datasets of PIC.

Summing up:

Given that our E-huxleyi-focused a ready-made database is yet unparalleled in terms of its combined areal+temporal coverage (6 seas in 3 oceans, 19 years, respectively), and the number of concatenated variables/parameters, we insist that, to date, it is veritably unique.

Other critical remarks relating to the issue of our database are commented on below.

2. Regarding the presence or absence of E. huxleyi blooms in the Bering Sea.

We considered this issue in detail in our work (Kondrik et al., 2017a), and it would obviously be improper to give here the respective entire excerpt from the above paper. In capsule:

A. Broerse et al.(2003) recognized that the area in which they took water samples, was on the very edge of the "bright patch". They write: "From the 7 February 2001 satellite image (Fig. 1(5)), it is not clear whether the sampling transect actually reached the edge of bright water patch". It is also worth pointing out that along with the diatom frustules,Broerse et al. also found coccoliths in their samples.

B. The ability of this alga to vegetate under conditions of very low levels of down-welling PAR irradiance is documented by Okada and McIntyre (1979): they have shown through their around-the-year shipborne measurements in the Labrador Sea at a latitudinal location (e.g. Station 'Bravo,' 56.5 ° N) similar to the location of the turquoise area in the Bering Sea that E. huxleyi cells indeed vegetated over a very long time period including not only summer time but also the winter period.

C. The appearance of turquoise areas in pelagic marine waters is a very strong argument in favor of attributing them to E. huxleyi blooms as no other hydrocoles possess such optical properties, which would render the truly turquoise color of their communities when observed from above. As Shutler et al. (2010) point out, this is a unique characteristic within phytoplankton species. Optically, diatom frustules are not identical to coccoliths. So that they would not produce the same remote sensing reflectance spectrum as coccoliths do.

An additional, albeit unnecessary argument: the phenomenon of huge blooms of E. huxleyi with extraordinarily high concentrations of coccoliths lasted only a few years and since 2001 have never re-occurred while diatoms blooms and associated release of frustules arethe annual event in the Bering Sea.

D. Finally, (although this argument is certainly optional, it only makes us additionally confident of our interpretation and robustness of our E. huxleyi bloom identification algorithm) we revealed the driving mechanism of the phenomenon of E. huxleyi blooms of exceptional intensity during 1998-2001, but this is the subject of a new paper, and we can't disclose it before its publication (expected in 2019).

In light of the above, the reviewer's assertion that our algorithm is invalidated because of the "false positives" in the Bering Sea could not be accepted.

3. Regarding the contested adequacy of our retrieval algorithms.

3a. On the advantage of our coccolith concentration retrieval algorithm.

We are not going to discuss here the advantages and disadvantages of E. huxleyi bloom detection algorithms suggested by other workers: it deserves a separate paper. Iida et al. (2002) have done it in detail with respect to e.g. the Brown and Yoder (1994) algorithm and pointed to some problems with it. Incidentally, Brown and Yoder themselves acknowledged certain limitationsof a world-wide application of their algorithm. Moore et al. (2012) commented on the feasibility of the algorithms in question developed by other teams that the reviewer specified in the his/her list of references.

The advantages of our algorithm were discussed in Kondrik et al. (2017a), and we hope that the reviewer does not expect us to dwell upon them. They can be epitomized as follows: our algorithm

(i) was developed on the basisof a nearly 20 year merged and skillfully harmonized OC CCI data provided by SeaWiFS, MODIS, MERIS, and VIIRS sensors;a comparative analysis of the OC CCI,GlobColour products, as well as the products from the MEaSUREs was conducted to prove the preference of the OC CCI data.

(ii) is based on extensive statistical analysis of satellite spectrometric [Rrs(lambda)] data fromsix marine environments specifically at high northern latitudes inthe North Atlantic, Arctic and North Pacific Oceans;

(iii) employsseveral criteria in conjunction, viz.: (a) location of maxima at the wavelengths typical of E. huxleyi bloom in Rrs spectra; (b) Rrs absolute value ranges at six wavelengths obtained through a dedicated/ large-size statistical sets of spaceborne data from the six seas; (c) consistency with the results of independent application of the BOREALI hydro-optical algorithm (Korosov et al., 2009; Kondrik et al., 2017a), which through retrieving inter alia the concentration of both coccoliths and chlorophyll-apermits to obtain the spatial distribution of the E. huxleyi bloom. This triple checking assured a higher reliability of the algorithm.

3b. Delta pCO2 retrieval algorithm

Again we believe that it would be improper to give here the respective entire excerpt from the paper on pCO2published in a refereed journal (Kondrik et al., 2018a). In a nutshell:

(i) the algorithm has the accuracy of delta pCO2 retrieval that is characterized by the following statistical parameters r2 = 0.54, pâĽ0.001, and RMSE = 23.4μatm;

(ii) the ensemble of blue data points in fig. 1 (Kondrik e al., 2018a) that looks like an "enormous spread" were obtained using climatological data and added solely to increase the statistical significance of the regression dependence established through using only in situ data that we could find for our study regions (red dots, their number is 187). Most of these points are within the declared error of 23.4 uatm; the indicated red points do not have the problem of Delta_pCO2 values overestimation indicated by the reviewer. It is also necessary to emphasize that a) "confidence interval" the reviewer refer tois in fact the "prediction limit" while the "confidence limit" has a much smaller variation (about 10 uatm). Also, it is important to be aware that the variation is given in uatm(units of partial pressure), but not in ppmas the reviewer writes.

(iii) all corrections for water temperature were duly conducted using the concurrently collected radiometric and IR satellite data.

(iv) the developed delta pCO2 regression dependence has a truly physical basis. Indeed, the increment of pCO2 in surface water within the E. huxleyi bloom is intimately related to the intracellular production of CO2 through the reaction of calcification and associated generation of coccoliths. The latter are very efficient reflectors of sun light coming into water (just because they don't absorb light but only reflect it). Therefore, the greater the amount of CO2 released through calcification, the more intense the optical signal coming out from the bloom area, especially at the wavelength of Rrs maximum – the parameter in our algorithm that is related to delta pCO2. Incidentally, returning to point 2C in our argumentations above, this is an important difference between coccoliths and diatomic frustules as the latter are not solely reflectors but also

absorbers.

4. The graphical illustration of validation of the retrievals of coccolith concentrations is available in our easily accessible papers published elsewhere, we doubt that the inclusion of those illustrations would be justified.

5. We acknowledge the reviewer's critical remarks in C3 –C4. All necessary changes are entered, the respective references [e.g. Shutler et al. (2010, 2013; Winter et al., 2014] are added to the reference list.

We certainly appreciate the list of references provided by the reviewer although, actually, we were aware of nearly all listed publications. The reason why they were not used is explained in point 1of our answers. As to the worksby Shutler et al. (2010,2013), it is indeed our flaw. We are earnestly grateful to the reviewer for this valuable critical remark.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2018-101, 2018.

Topical Editor Decision: Publish subject to minor revisions (review by editor) (07 Jan 2019)
by David Carlson
Comments to the Author:
The data description seems thorough and well-organised and I suspect this product will serve many users. Please, however, can the authors attend to a few points:

1) I feel somewhat surprised to see that these authors have ignored two coccolithophore data sets recently published in ESSD - Loveday & Smyth, https://doi.org/10.5194/essd-10-2043-2018 which seems quite relevant to the remote sensing aspects albeit using AVHRR rather than ocean colour, and Daniels et al. https://doi.org/10.5194/essd-10-1859-2018 which addresses the issue of calcification rates. In one of their responses the authors mentioned that the impacts of these blooms on pCO2 represented the unique contribution of this data. But these data also both draw on and contribute to the other two data sets?
Such a comparison may also prove useful for validation (e.g. see https://doi.org/10.5194/essd-10-2275-2018).

2) Please also attend to these comments from a third reviewer:

"I found the paper interesting and useful. The authors have made every attempt to validate their products and build on previous work, and to provide a theoretical basis for the algorithms where possible. They have brought together a large body of different types of data to generate their products, and to validate them. They have taken the trouble to re-process some of the OC-CCI data, when it appeared that the masking applied may not have been appropriate for their particular application.

While this paper was under review, another paper on the same topic, but using AVHRR paper, has been published in ESSD (Loveday and Smyth, ESSD 2018). It would be good to refer to this publication. That other attempts have been made recently to address a similar problem in no way deters from the value of the paper under review: the users should be given the opportunity to accept the product that best suits their requirement, and to evaluate the products themselves.

The writing style leaves some problems with the grammar, and some instances where the statements lean towards the hyperbole (especially in the introduction). I do not know if ESSD editors and copy editors can help the authors deal with them?

Assuming that such minor problems with the language can be fixed, I recommend the paper for publication."

Pending appropriate responses, I may also ask one of the reviewers who volunteered to read a revised version to take a quick final look.

Thank you for considering ESSD.

Dear Editor,

In accordance with the suggestions made by all three reviewers, we have revised the text and send you the resultant version of our paper.

We also noted brief descriptions and links to global multiyear databases from Loveday & Smith 2018 and PIC from NASA Ocean Color in the article.

With our cordial regards

Eduard Kazakov in the name of the paper's authors.

A synthetic satellite dataset of *E.~~.~~miliania huxleyi* blooms spatio-temporal distributions and their impacts on Arctic and Subarctic marine environments (1998-2016)

Dmitry Kondrik[1], Eduard Kazakov[1], Dmitry Pozdnyakov[1]

[1]Nansen International Environmental and Remote Sensing Centre, Saint Petersburg, 199034, Russian Federation

Correspondence to: Eduard Kazakov (ekazakov@niersc.spb.ru)

Abstract. A 19-year (1998-2016) continuous dataset of coccolithophore *E.~~.~~miliania huxleyi* distributions and activity, i.e. the release of $CaCO_3$ in water and the decrease of uptake of dissolved CO_2 by *Emiliania huxleyi* cells (e.g. Kondrik et al., 2018a), in Arctic and Subarctic seas is presented. The dataset is based on optical remote sensing data (mostly OC CCI data) with assimilation of different relevant in-situ observations, preprocessed with authorial algorithms. Alongside with bloom locations, we also provide both detailed information on *E. huxleyi* impacts ~~within the bloom area on marine environments~~on carbon balance and the subdatasets of quantified coccolith concentrations, particulate inorganic carbon content and CO_2 partial pressure in water driven by coccolithophores. All data are presented on a regular 4x4 km grid at a temporal resolution of 8 days. The paper describes the theoretical and methodological basis for all processing and modeling steps. The data are available on Zenodo: https://doi.org/10.5281/zenodo.1402033.

1 Introduction

~~Ongoing climate change is a background of numerous emerging hot topics. Among them,~~Among the topics related to the ongoing climate change, there are alterations of both biodiversity in marine environments and the carbon balance in the atmosphere-ocean system (Rost et al., 2008). In some specific cases both processes are interrelated being spurred up by one and the same agent(s). Along with other marine inhabitants, coccolithophores are such entities, and more specifically, the algal species named *Emiliania huxleyi* – a unicellular planktonic organism that is the most widespread coccolithophore in the world's oceans. Being simultaneously a calcifying and photosynthetic primary producer of, respectively, inorganic and organic carbon, *E.~~.~~miliania huxleyi*, in the course of its life cycle, enhances both the concentration of calcite and carbon dioxide partial pressure in ocean surface water. At least within *E.~~.~~miliania huxleyi* bloom areas, both processes are capable of changing the carbon balance, and hence affect both CO_2 fluxes between the atmosphere and surface ocean and the aquatic biogeochemistry. Being a spatially huge phenomenon invariably occurring in both hemispheres, and ~~gradually~~ steadily propagating in the poleward direction (Winter et al., 2014) due to CO_2 accumulation in the atmosphere ~~(Rivero Calle et al., 2015)~~ and ensuing climate warming (Johannessen, 2008), *E.~~.~~miliania huxleyi* blooms are believed to be highly relevant to understanding the comprehensive nature of the changes unfolding on our planet.

Historically, the initial building up of knowledge on coccolithophores in general and *E. miliania huxleyi*, specifically, was broadly based on in situ approaches effected in the course of both shipborne and laboratory activities. Extensive data were obtained on *E. miliania huxleyi* cell morphometry, internal structure, intracellular dark – and photoreactions, factors controlling/affecting the cell growth, as well as intrinsic optical properties, such as sun light total and spectral absorption, scattering/backscattering (Balch et al., 1996a). In addition, regression relationships were established between *E. miliania huxleyi*-driven changes in both inherent hydro-optical parameters and CO_2 partial pressure in surface water within the bloom area (Holligan et al. 1993).

However, as this phenomenon extends over marine areas in excess of hundreds of thousand square kilometres (Balch et al., 2016; Kondrik et al., 2018a), and is spatially and temporally highly dynamic, solely satellite remote sensing approach ~~means~~ ~~are~~is able to comply with the challenge of studying it.

Until recently, only few satellite studies were performed and published on the typical locations of *E. miliania huxleyi* blooms and associated concentrations of particulate inorganic carbon in surface ocean within the bloom area (e.g. Gordon et al., 2001; Balch et al., 2016).

Prior to the publication by Kondrik et al. (2018a), ~~no attempts~~ to the best of our knowledge, only a couple of studies (Shutler et al., 2010; 2013) have been undertaken to either retrieve from spaceborne data both the total content of inorganic carbon produced by a *E. miliania huxleyi* bloom (PIC) and increase in CO_2 partial pressure ($\Delta p CO_2$) in surface water within the bloom area or else reveal intraannual and interannual variations over long time periods in the location and intensity of *E. miliania huxleyi* blooms. No concatenated time series data of a nearly 20 year duration are available to date on the associated quantifications of bloom surface, bloom intensity, $\Delta p CO_2$ for all ~~E. huxleyi~~*Emiliania huxleyi* blooms occurring within extensive latitudinal belts and encompassing waters of different oceans i.e. marine tracts significantly distanced longitudinally. Meanwhile, the above specified information is an indispensable step towards a further pan-global inventory of the effects produced by *E. huxleyi* blooms on both marine chemistry and ecology, and CO_2 exchange fluxes between the atmosphere and ocean as such fluxes condition the status of the world's oceans as a sink of CO_2.

In addition to the studies cited above, it is also worthwhile to mention a few sources of multi-year satellite data on coccolithophore blooms that can be useful for the potential users to widen their multifaceted databases in their studies.

The NASA OCEANCOLOR portal https://oceancolor.gsfc.nasa.gov/atbd/pic/ offers the extensive data on particulate inorganic carbon retrieved from MODIS with the Balch et al. methodology (2005). Downloadable from https://oceandata.sci.gsfc.nasa.gov/, these data are, respectively, at a 4 km and 1 day spatial and temporal resolution and cover the time period starting from 2000 onward.

Also, reported by Loveday and Smyth (2018) a forty year time series (1978-2018) from AVHRR observation data on coccolithophore blooms is available at https://doi.org/10.1594/PANGAEA.892175. Employing specially developed coccolithophore bloom area masks that were developed from remote sensing reflectance spectra, these data are monthly global-wide and available at a spatial resolution of 0.1° (~ 10 km). Although these data do not encompass any additional

parameters such as particulate inorganic carbon or CO2 partial pressure in surface water within the bloom, nevertheless they can be valuable due to an exceptionally long observation period.

~~Based on the employed spaceborne ocean colour information,~~ Unlike the publications mentioned above the present paper reports on extensive concatenated original datasets generated for subpolar and polar seas of the Northern Hemisphere, viz. the

5   North, Labrador (with adjacent North Atlantic open waters), Norwegian, Barents, Greenland and Bering seas. ~~T~~Based on the employed spaceborne ocean colour information, the obtained datasets are processed into a nearly two decadal (19~~1~~98-2016) time series for each of the target seas/marine areas, and encapsulate information about~~.~~

~~The collected data base of~~ PIC and $\Delta p\mathrm{CO_2}$ values in surface water within the bloom area together with intraannual and interannual variations in the location and intensity of ~~*E. huxleyi*~~*Emiliania huxleyi* blooms over such a variety of seas and across

10   a nearly 20-year time period is presently unique.

Conjoined with a wealth of presently available supplementary data from satellite and shipborne missions on the environmental conditions under which target ~~*E. huxleyi*~~*Emiliania huxleyi* blooms emerged and developed, the synthetic dataset we are reporting herein opens the way to detailed analysis of forward and feedback mechanisms governing the temporal and spatial dynamics of this phenomenon. Further utilization of the results of such analysis in regional and global climatic models promises

15   to predict future directions of development of the phenomenon in question (Rost et al., 2008).

2 Methodology and dataset content

Based on the facility of available satellite OC CCI (Ocean Colour Climate Change Initiative) and SeaWiFS data in the visible part of the spectrum, the following products have been generated to achieve the goals specified in the previous section, viz.: 1. ~~*E. huxleyi*~~*Emiliania huxleyi* bloom extent; 2. Concentration of coccoliths within the bloom; 3. Total content of particulate

20   inorganic carbon (PIC) produced by the bloom; 4. Increase in CO_2 partial pressure in marine surface waters due to the blooming phenomenon.

2.1 Bloom area quantification

Quantification of ~~*E. huxleyi*~~*Emiliania huxleyi* bloom areas was performed in two stages. Firstly, RGB (red-green-blue) images were generated based on the weighted remote sensing reflectance, R_{rs}, which is the upwelling spectral radiance just above the

25   water–air interface normalized to the downwelling spectral irradiance at the same level (Bukata et al., 1995). R_{rs} values in the channels centered at 670, 555, and 443 nm were employed. Analysis of the spaceborne radiometric data collected by Kondrik et al. (2017a, b) from the 5 target seas, yielded statistically robust specific ranges of $R_{rs}(\lambda)$ highlighting ~~*E. huxleyi*~~*Emiliania huxleyi* blooms as turquoise areas; the areas of blooms of other (noncalcifying) algae were reflected in the images as green. Areas with scarce noncalcifying algae abundance showed up as blue or dark blue. The land mask was overlaid so that land

30   areas were coloured light yellow.

In the second stage of quantification of ~~E. huxleyi~~*Emiliania huxleyi* bloom extent, an additional criterion was imposed on the revealed turquoise areas: R_{rs} values should be maximal at 490 nm and/or 510 nm, while at other wavelengths they need to be in excess of 0.001 (412 nm), 0.008 (443 nm), 0.01 (490 nm), 0.008 (510 nm), 0.008 (555 nm), and ~0 (670 nm). Such a selection provided the highest accuracy of bloom delineation. With the known pixel size, the bloom area can be confidently

5   quantified. An example of ~~E. huxleyi~~*Emiliania huxleyi* bloom extent masking is shown in Figure1.
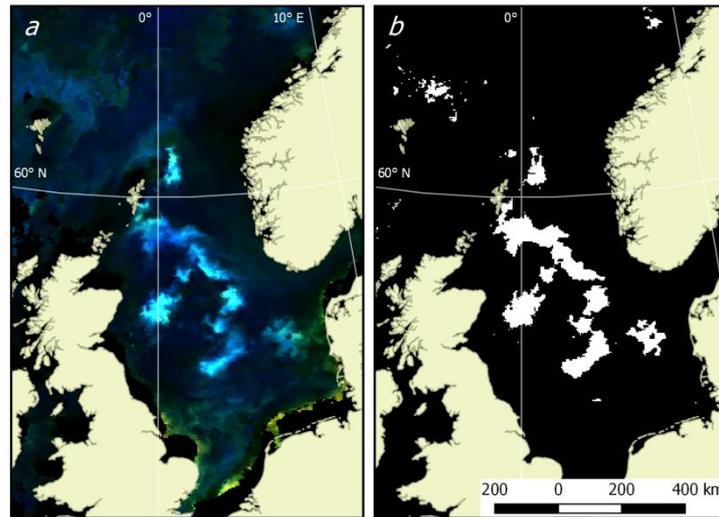
Figure 1: Example of the bloom masking algorithm performance. *a*= source of the OC CCI RGB imagery for the North Sea (2016.06.09, with land mask); *b* = calculated bloom mask (white pixels stand for bloom detected, black pixels are areas void of bloom).

10   **2.2 Determination of the coccolith concentration**

Determination of the coccolith concentration within the bloom was performed with the BOREALI algorithm (Bio-Optical REtrieval ALgorIthm, Korosov et al. 2009), based on the Levenberg–Marquardt (L-M) finite difference technique (Press et al. 1992). The L-M technique solves the inverse problem, i.e. in our case allows to retrieve the concentrations of water constituents from spectral subsurface remote-sensing reflectance, $R_{rsw}(\lambda)$, which is the upwelling spectral radiance just beneath the water–

15   air interface normalized to the downwelling spectral irradiance at the same level (Jerome et al., 1996). A hydro-optical model accommodating spectral specific absorption and backscattering coefficients of ~~E. huxleyi~~*Emiliania huxleyi* cells and coccoliths as well as pure water per se, non-calcifying alga and dissolved organic matter was developed and employed to run the BOREALI (Kondrik et al., 2017a).

The results of validation of coccolith concentration retrievals with BOREALI were assessed through the following statistical

20   measures: coefficient of correlation, r, linear regression equation, $f(x)$, coefficient of determination, R^2, root mean square deviation/error, RMSE, systematic error, BIAS, and MAE. BIAS and MAE were then also normalized to the absolute values of coccoliths concentrations determined by using each model: $r = 0.88$; $f(x) = 0.6159x + 6.9197$; $R^2 = 0.77$; RMSE = 3.55×10^9 coccoliths·m^{-3}; BIAS = 25.30%; MAE = 32.30%.

In addition, ascertained by both RGB and R_{rs} approaches, ~~E. huxleyi~~*Emiliania huxleyi* bloom areas were further checked up using the results of coccolith concentration retrievals. This was done through the application of a threshold. A threshold of 90 $\times 10^9$ coccoliths· m^{-3} was chosen because, firstly, it assures the best correspondence between the bloom surfaces, determined by our radiometric and BOREALI algorithms. Secondly, this threshold is very close to the average value of coccolith concentrations in developed ~~E. huxleyi~~*Emiliania huxleyi* blooms reported from the world's oceans (for references, see Balch et al. 1996b; Balch et al. 2005). The numerical assessments of bloom surfaces delineated/quantified by above independent ways converged precisely.

2.3 Coccolith content, particulate inorganic carbon and CO_2 partial pressure increment determination

Determination of the coccolith content (CC) was performed through establishing mixed layer depth (MLD) within the bloom area. The climatology of Montegut et al. (2004) was applied. The identified areas of ~~E. huxleyi~~*Emiliania huxleyi* blooms with retrieved concentrations of coccoliths were overlapped by the respective climatological MLD fields, and for each pixel, the value of MLD was further used for calculating CC. Further, CC values were used to quantify the total content of particulate inorganic carbon (PIC). It was done for each 8-day time period (corresponding to the temporal resolution of the spaceborne radiometric data employed) through multiplying the carbon mass per coccolith, m, and CC followed by summarizing the results of multiplication within all pixels of respective bloom extent. The value of m was equalled to 0.2 pg (Balch et al., 2005). The moment, at which the PIC assessment could be ideally performed in each bloom, corresponded to the situation when two conditions were fulfilled: (a) the bloom attained its largest surface, and (b) the spectral curvature of remote sensing reflectance, $R_{rs}(\lambda)$, exhibited a maximum at about 490 nm as the location of R_{rs} maximum at about 490 nm is an indication that the bloom is prevalently composed of coccoliths (Kondrik et al., 2017a).

Remote determinations of ~~E. huxleyi~~*Emiliania huxleyi*-driven pCO_2 increment (ΔpCO_2) consisted in establishing a relationship between ~~E. huxleyi~~*Emiliania huxleyi*-driven changes in pCO_2, that is, ΔpCO_2, in bloom pixels, and the respective values of R_{rs} (490). Such a relationship (Kondrik et al., 2018a) with the following statistical characteristics: coefficient of determination, r^2 = 0.54, p $\ll$0.001, and RMSE = 23.4 µatm was used to quantify the spatial variations of ΔpCO_2 in the target seas followed by recalculating ΔpCO_2 for the water temperatures (retrieved from spaceborne data) that actually occurred during respective ~~E. huxleyi~~*Emiliania huxleyi* bloom events (Copin-Montegut, 1988).
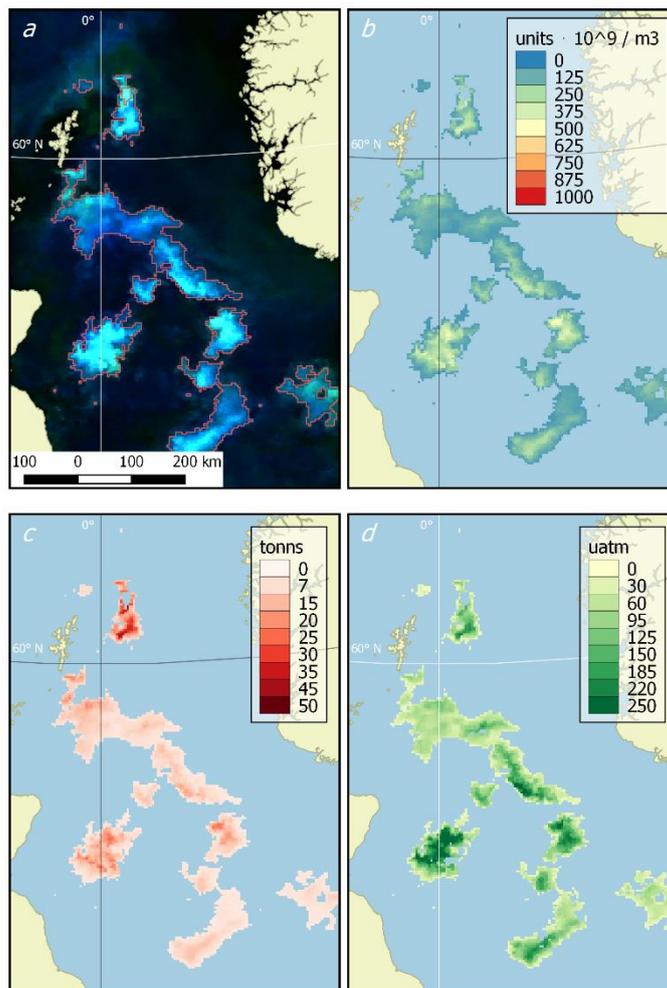
Figure 2: Example of dataset products (the North Sea, 2016.06.09). *a* = source OC CCI RGB imagery with the bloom mask contoured in red, *b* = coccolith concentration ($10^9 \cdot m^{-3}$), *c* = content of particulate inorganic carbon (tonns), *d* = increase in CO_2 partial pressure in water (µatm).

2.4 Additional technical workflow

In ~~the cause~~<ins>process</ins> of satellite data processing, several preceding procedures were performed.

1. Reprojection of satellite images. Given the high latitudinal location of the target seas, it was relevant to use an equal-area polar projection. Therefore, the NASA 'Ease-Grid' was employed. The system of coordinates of the WGS-84 (World Geodetic System 1984) is at the basis of 'Ease-Grid'.

2. Correction of Automatic Cloud Masking in the images from SeaWiFS in 1998–2001. In all images of the OC CCI product obtained in 1998–2001 (when only the SeaWiFS sensor was operational), all putative bloom areas proved to be masked. The errors of automatic cloud masking most probably resulted from very high values of brightness stemming from bloom areas

(comparable with cloud-produced signals), which may have led to possible mistakes in the masking algorithm. The problem was overcome via manual processing of the data of a lower level, i.e. directly from the SeaWiFS level 2 product (http://oceancolor.gsfc.nasa.gov/cgi/browse.pl?sen=am) for the period of 1998–2001 in all studied areas. As a result, in the RGB-images the areas masked as clouds in OC CCI images proved to exhibit large bloom areas with the brightness of signals

5   typical of ~~E. huxleyi~~*Emiliania huxleyi*. This approach was legitimate as OC CCI data obtained by different sensors have been brought to the SeaWiFS standard channels, and the entire data time series (1998-2016) was radiometrically uniform.

3. Filling Missing Pixels Masked as Ragged Clouds. In the case of ragged clouds, some pixels of RGB images are not informative. A special algorithm for filling such gaps included averaging of $R_{rs}(\lambda)$ values from neighboring pixels and from temporarily previous and following images of the same pixel. The use of this algorithm in each of the cloud-masked images

10   of the areas studied over 19 years and included in the OC CCI product helped increase the analysed area, sometimes to a significant extent. Calculated from 1998 to 2016 as arithmetic means for the Barents, Bering, North, Norwegian and Greenland seas, the quantitative estimates of such an increase attained for each 8-day-averaged image reached, respectively, ~107, 370, 31, 15, and 13 times. Thus, obtained were images with significantly larger cloud-free areas assuring a more accurate estimation of the borders of bloom areas, and their displacement, as well as of bloom areas per se.

15   Examples of products visualizations (for the North Sea) are shown in Figure 2.

3 Data sources

Data on R_{rs} in six channels (centered at 412, 443, 490, 510, and 670 nm) are from the OC CCI product (Ocean Colour Climate Change Initiative dataset, Version 3.0, European Space Agency, available online at http:// www.esa-oceancolour-cci.org/).

For the bio-optical retrieval algorithm validation, we employed the PANGAEA database (www.pangaea.de) of the

20   concentration of coccoliths within the target coccolithophore blooms in the North Atlantic including the North and Norwegian Seas (Charalampopoulou et al. 2008, 2011).

The bio-optical in situ database spanning between 1997 and 2012 (16 years) was employed for ocean-colour satellite applications as having a global coverage (Valente et al., 2016). The data were acquired from several sources: MOBY (Marine Optical Buoy), BOUSSOLE (BOUée pour l'acquiSition d'une Série Optique à Long termE), AERONET-OC (Aerosol Robotic

25   NETwork-Ocean Color), SeaBASS (SeaWiFS Bio-optical Archive and Storage System), NOMAD (NASA bio-Optical Marine Algorithm Dataset), MERMAID (MERIS Match-up In situ Database), AMT (Atlantic Meridional Transect), ICES (International Council for the Exploration of the Sea), HOT (Hawaii Ocean Time-series), and GeP&CO (Geochemistry, Phytoplankton, and Color of the Ocean). This database comprises a large number of variables, including the spectral remote sensing reflectance, R_{rs}, and chlorophyll-a concentration.

30   Data on mixed layer depth (MLD) were derived from the Montegut climatology (Montegut et al. 2004).

Data on bathymetry inherent in the target seas were taken from the website http://www.ngdc.noaa.gov/mgg/bathymetry/arctic/arctic.html (Jakobsson et al. 2012).

7

The GLobal Ocean Data Analysis Project (GLODAP) database (Key et al., 2015; Olsen et al., 2016), http://cdiac. ornl.gov/oceans/GLODAPv2/ was employed for pairing in situ NO_3 values at those points for which in situ pCO_2 values were available. In the cases when the desired NO_3 matching values were unavailable in the GLODAP database, the respective data were employed from the World Ocean Atlas 2013 (WOA13, NOAA, Garcia et al., 2014; https://www.nodc.noaa.gov/OC5/woa13/).

The SOCAT v4 database (The Surface Ocean CO_2 Atlas, Bakker et al., 2016; http://www.socat.info/access.html) comprises more than 6 million pCO_2 measurements performed at latitudes north of 40°N. The data employed by us from SOCAT V4 database met the following requirements: (*1*) measurements conducted during 1998–2016 and within a 10 m top layer (if there were data from several depths, the measurements from the shallowest depth were used); (*2*) pCO_2 data should necessarily have both corresponding seawater salinity data and valid R_{rs} spectra; (*3*) a daily mean pCO_2 value was employed provided there were several in situ measurements; (*4*) pCO_2 measurements conducted at a distance not less than 8 km offshore (to avoid the impact of adjacency effect on R_{rs} satellite data); (*5*) pCO_2 measurements were within the location and timing of ~~E. huxleyi~~*Emiliania huxleyi* blooming; and (*6*) data used from SOCAT v4 database overlap the data from either the GLODAP database or the WOA13 climatology database (depending upon which one was used for comparison).

~~The GLobal Ocean Data Analysis Project (GLODAP) database (Key et al., 2015; Olsen et al., 2016), http://cdiac. ornl.gov/oceans/GLODAPv2/ was employed for pairing in situ NO_3 values at those points for which in situ pCO_2 values were available. In the cases when the desired NO3 matching values were unavailable in the GLODAP database we resorted to the respective data from the World Ocean Atlas 2013 (WOA13, NOAA, Garcia et al., 2014; https://www.nodc.noaa.gov/OC5/woa13/).~~

4 Data spatio-temporal domain

The published dataset covers a time period of 19 years, from 1998 to 2016, with a time resolution of 8 days (a total of 874 time periods), and a spatial domain with the total area of ~~100,000~~~~105,6800~~11,056.800 km^2 at a resolution of 4x4 km, divided into 4 regions described in Table 1 and shown in Figure 3.

All data a represented in the Lambert Azimuthal Equal area projection with the parameters corresponding to the widespread NSIDC EASE-Grid North (EPSG: 3973) coordinate system.

The selection of 4 regions in this work resides in several reasons. They include all seas where coccolithophore blooms usually occur in subpolar and polar regions of the Northern Hemisphere (North, Norwegian, Greenland, Barents, Bering and Labrador seas). The exclusion from our dataset of blooms occurring in the northern parts of Atlantic Ocean (see, e.g. Holligan et al. 1993) was dictated by some technical restrictions: the hydro-optical model employed for obtaining coccolith concentration values was based prevalently on the data from high-latitude areas, and thus should be at first validated for geographically different marine environments such as open parts of the Atlantic Ocean.
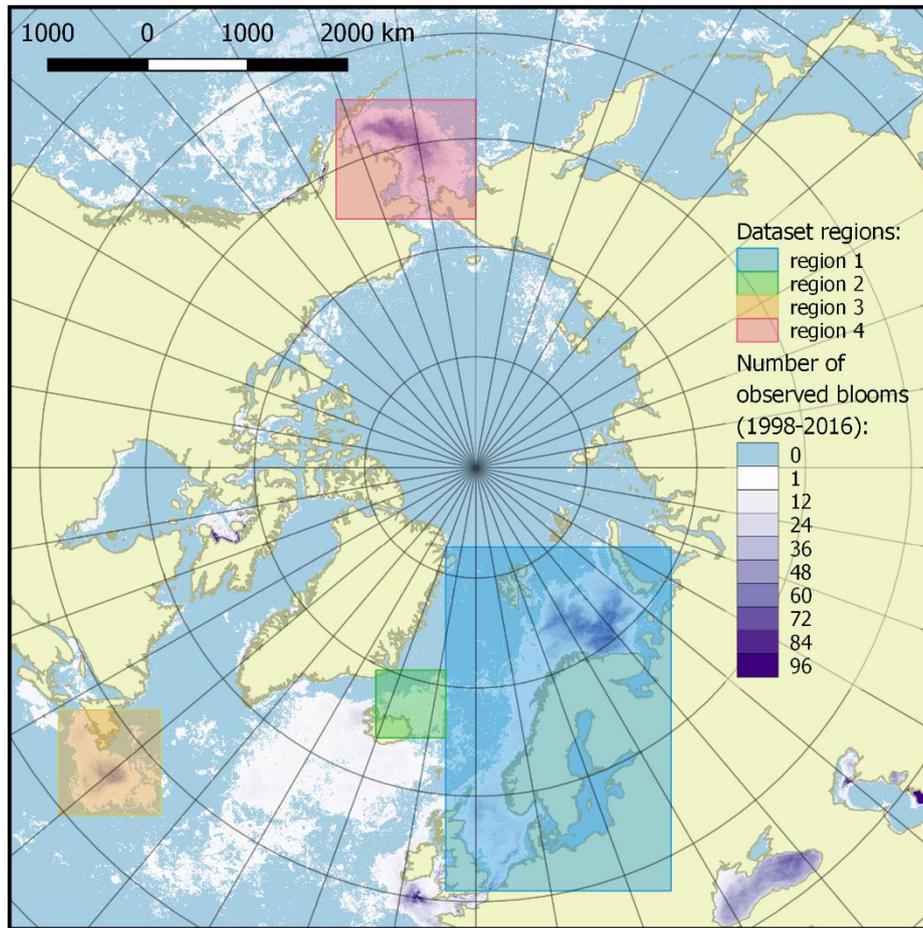
Figure 3: Dataset of target spatial regions. Regions are shown as coloured boxes, and the colourbar indicates the number of bloom observations in each pixel over the time period 1998 - 2016.

5 Dataset overview

5   The 19-year period data covers 4 blooming regions differing in nature. This allows to evaluate the bloom-related processes at different scales and time intervals in order to reveal both interannual dynamics and seasonal variations of parameters relevant to the bloom phenomenon. ~~E. huxleyi~~*Emiliania huxleyi* blooms in the Arctic and Subarctic seas are characterized by significant instability: the difference in intensity of blooming in different years can reach tens of times. Figure 4 and Table 2 collectively illustrate for the above four marine regions the temporal dynamics in bloom intensity (i.e. blooming area). For example, in the

10   Bering Sea (region 4), the most extensive blooms were observed exclusively from 1998 to 2001, but later on, their intensity decreased drastically. In region 1, mainly in the Barents, Norwegian and Northseas, the blooming activity over the years we are reporting on was very irregular, with a peak in 2016.

9
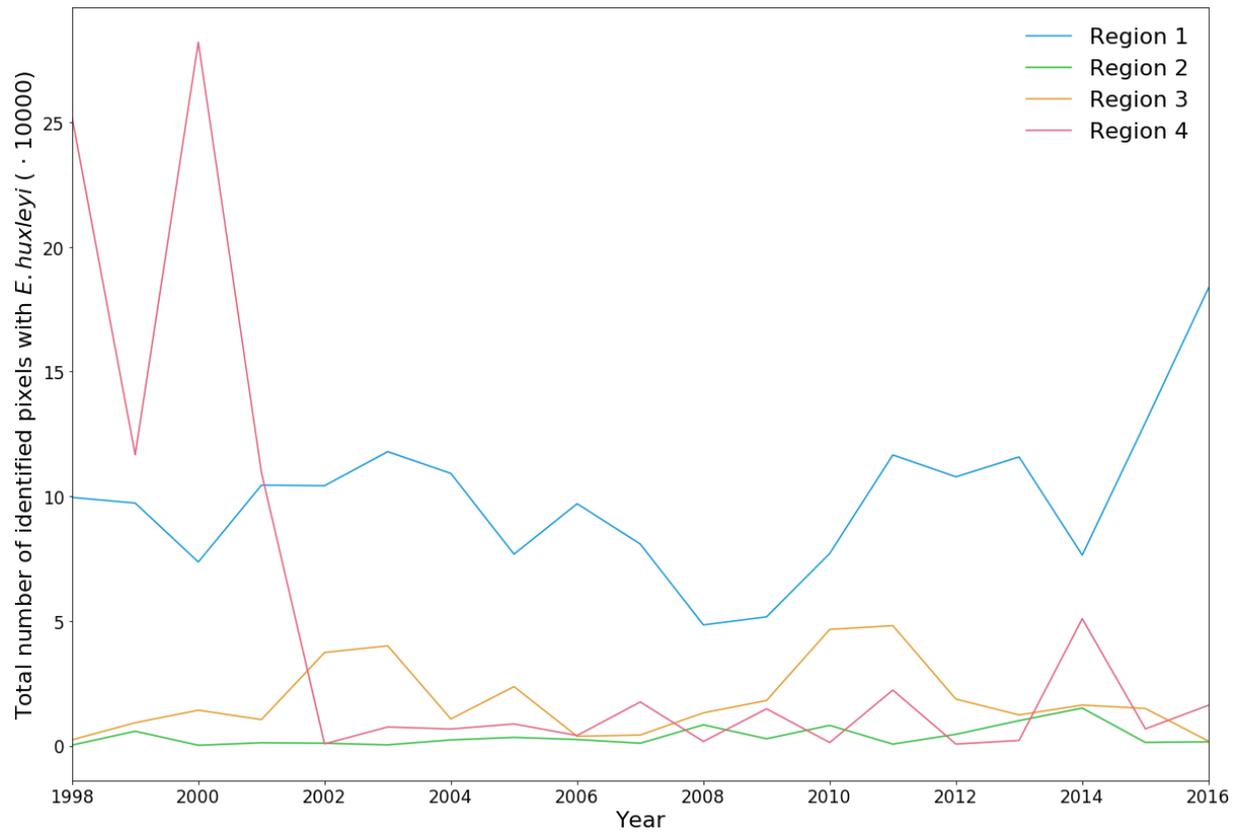
Figure 4: Total number dynamics of identified pixels with ~~E. huxleyi~~*Emiliania huxleyi* for each blooming season in the period 1998-2016 within the four regions specified in Figure 3.

With the data collected, it's possible to highlight the patterns of development of the regularly occurring blooms. They can be

5 characterized with the beginning/end of blooming periods, and the overall dynamics of coccolith concentration during the

blooms. Such patterns can be established based on the published dataset. Figure 5 shows an example of bloom development

in the Greenland Sea (region 2) in the period June 26 - August 13, 2014. However, these periods are generally unstable, which

is clearly seen in Figure 6, which displays the blooming area configuration in July, 20 for different years for the same area.

10
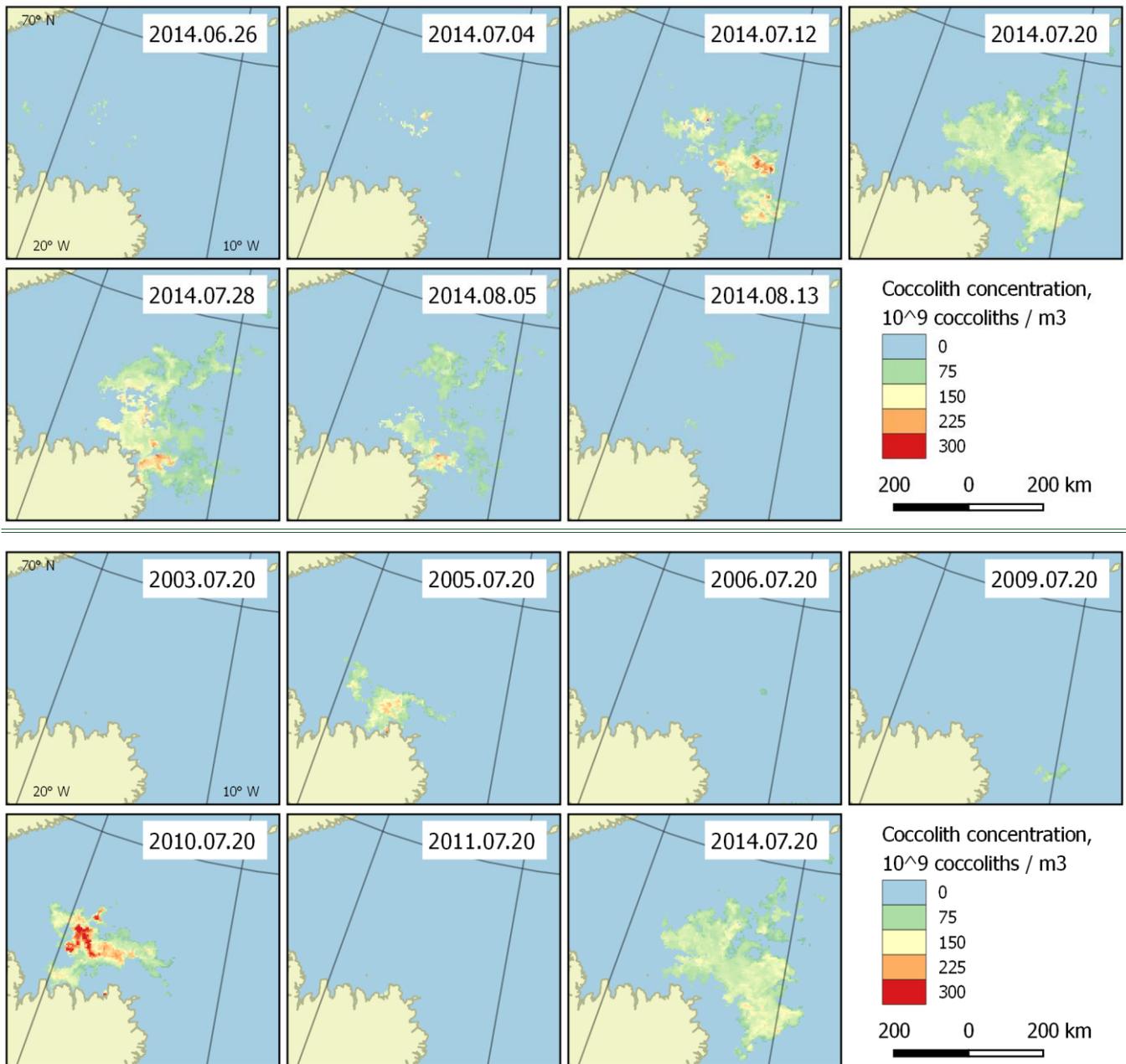
Figure 5: Bloom development in the Greenland Sea (region 2) in June-August 2014. The peak falls on July 20.
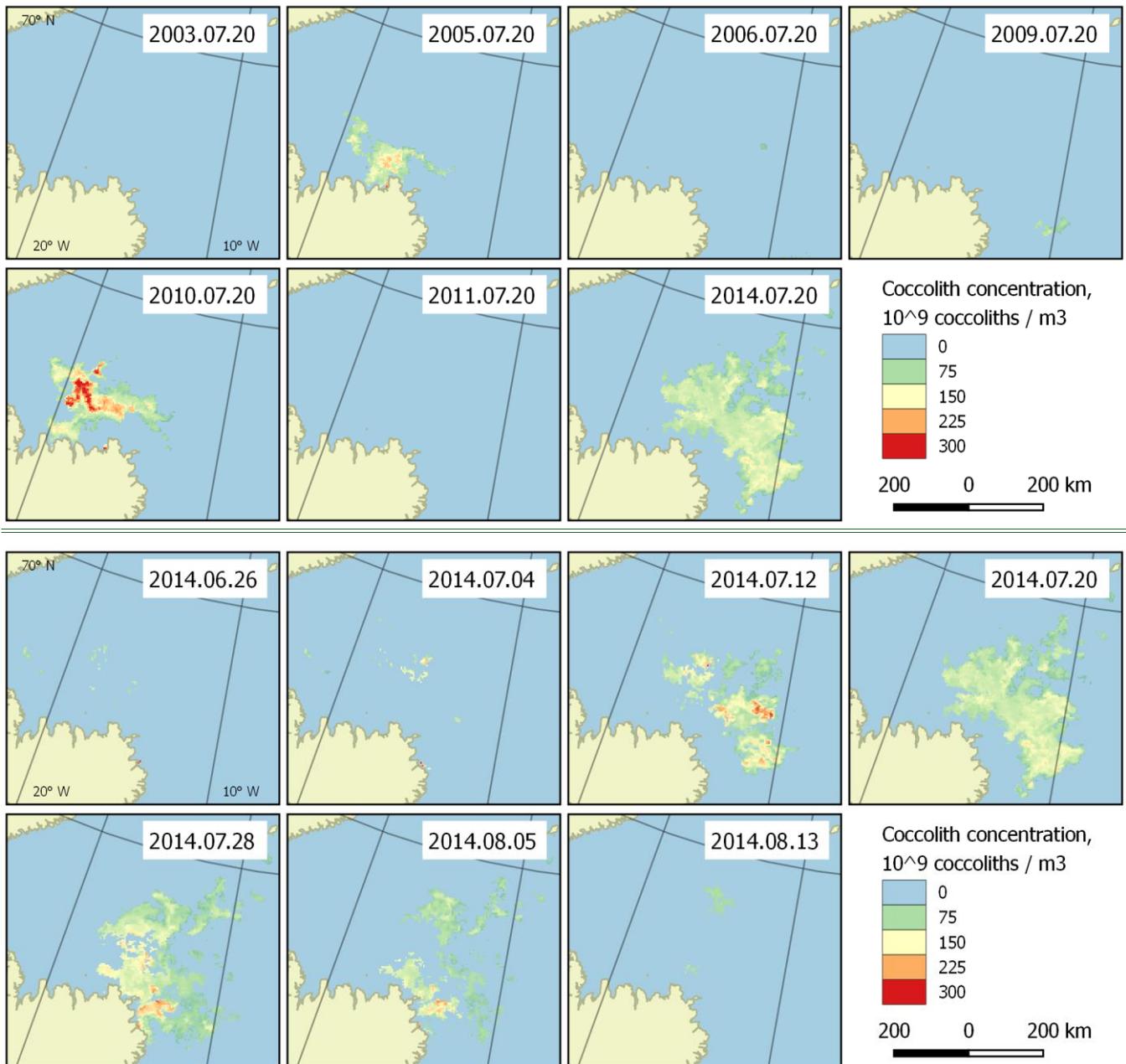
Figure 6: Bloom intensity in the Greenland Sea (region 2) on July 20 in different years. Its instability is obvious.

Technically, each dataset contains 4 subdatasets: bloom status, coccolith concentration, particulate organic carbon content and
5   CO_2 partial pressure in water driven by coccolithophores. The last three categories contain the parameter values directly calculated. The first subdataset contains information about the quality and content of data. This information is organised as a set of flags attributed to data on reliable observations of blooming presence or absence, or inaccurate data (usually due to

clouds) as well as data on coastal land. Figure 7 provides both an example of a status matrix and the matrix containing coccolith concentration values.
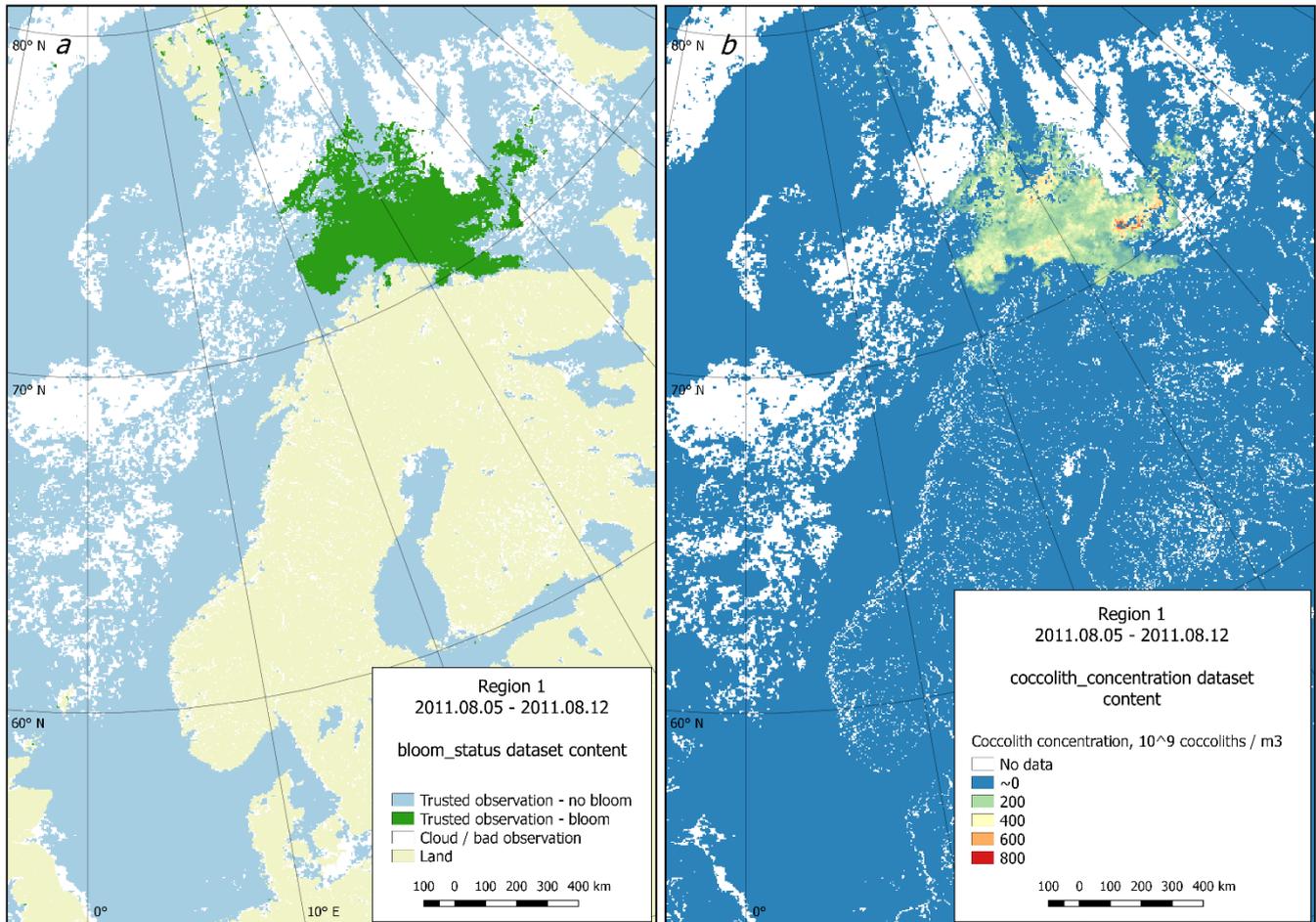
Figure 7: Dataset content example (region 1, 2011.08.05). a - bloom status subdataset visualization, b - coccolith concentration subdataset visualization.

6 Data availability

Dataset is available on Zenodo (Kondrik et al. 2018b; https://doi.org/10.5281/zenodo.1402033). Data granules are divided into directories by regions and years, each child directory contains files with 8-day periods data on the bloom status, coccolith concentration, PIC, ΔpCO$_2$. Data are stored in NetCDF4 format with GDAL-support, that allows to use the data immediately with any NetCDF-based or GIS software. Tips about how to read the data and QGIS styles for fast visualizations are also provided.

7 Conclusions

We have composed a detailed 19-year dataset of ~~E. huxleyi~~*Emiliania huxleyi* blooms in the Arctic and Subarctic seas, including the information about their influence on the carbon cycle in the ocean. These data are based mostly on satellite remote sensing observations, but also on available shipborne measurements and results of processing with authorial algorithms. We hope that

5 the publication of these data, on the one hand, will promote further studies aimed at elucidating ~~E. huxleyi~~*Emiliania huxleyi* bloom driving mechanisms and their forcing factors and, on the other hand, will facilitate understanding the patterns of this phenomenon distribution and its impact on the ocean and the atmosphere.

Author contributions

Dmitry Pozdnyakov is responsible for theoretical background and methodology development. Dmitry Kondrik also contributed

10 to theoretical background research, and responsible for data processing algorithms development and programming. Eduard Kazakov conceived the dataset structure and contributed to data processing algorithms programming, data analysis and visualizations. All authors equally contributed to the writing of the manuscript and data quality control.

Competing interests

The authors declare that they have no competing interests.

References

Balch, W. M., Kilpatrick, K. A., and Trees, C. C.: The 1991 coccolithophore bloom in the central North Atlantic. 1. Optical properties and factors affecting their distribution, Limnol. Oceanogr., 41, 8, 1669-1683, 1996a.

25 Balch, W.M., Kilpatrick, K.A., Holligan, P.M., Harbour, D., and Fernandez, E.: The 1991 coccolithophore bloom in the central North Atlantic. 2. Relating optics to coccolith concentration, Limnol. Oceanogr., 41, 8, 1684-1696, 1996b.

Balch, W. M., Gordon, H., Bowler, B. C., Drapeau, D. T., and Booth, E. S: Calcium Carbonate Measurements in the Surface Global Ocean Based on Moderate-Resolution Imaging Spectrometer Data, J. Geophys. Res., 110, C07001, doi:10.1029/2004JC002560, 2005.

Balch, W. M., Bates, N. R., Lam, P. J., Twining, B. S., Rosengard, S. Z., Bowler, B. C., Drapeau, D. T., Garley, R., Lubelczyk, L. C., Mitchell, C., and Rauschenberg, S.: Factors regulating the Great Calcite Belt in the Southern Ocean and its biogeochemical significance, Global Biogeochem. Cy., 30, 8, 1124-1144, doi:10.1002/2016GB005414, 2016.

Bakker, D. C. E., Pfeil, B., O'Brien, K. M., Currie, K. I., Jones, S. D., Landa, C. S., et al.: Surface Ocean CO_2 Atlas (SOCAT) V4, doi:10.1594/PANGAEA.866856, 2016.

Bukata, R. P., Jerome, J. H., Kondratyev, K. Ya., and Pozdnyakov, D. V.: Optical Properties and Remote Sensing of Inland and Coastal Waters, CRC Press, Boca Raton, FL, 362 pp, 1995.

Charalampopoulou, A., Poulton, A. J., Tyrrell, T. and Lucas, M.: Surface Seawater Carbonate Chemistry, Nutrients and Phytoplankton Community Composition on a Transect between North Sea and Arctic Ocean, 2008, Supplement to. doi: 10.1594/PANGAEA.763990, 2008.

Charalampopoulou, A., Poulton, A. J., Tyrrell, T., and Lucas, M: Irradiance and pH Affect Coccolithophore Community Composition on a Transect between the North Sea and the Arctic Ocean, Marine Ecology Progress Series 431, 25–43, doi:10.3354/meps09140, 2011.

Copin-Montegut, C.: A new formula for the effect of temperature on the partial pressure of CO_2 in seawater, Marine Chemistry, 25, 1, 29-37, doi:10.1016/0304-4203(88)90012-6, 1988.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., et al.: World Ocean Atlas 2013, volume 4: Dissolved inorganic nutrients (phosphate, nitrate, silicate), in S. Levitus & A. Mishonov (Eds.), NOAA Atlas NESDIS 76, 25 pp, 2014.

Gordon, H. R., Boynton, G. C., Balch, W. M., et al.: Retrieval of Coccolithophore Calcite Concentration from SeaWiFS Imagery, Geophys. Res. Lett., 28, 8, 1587-1590, 2001.

Holligan, P. M., Fernández, E., Aiken, J., Balch, W. M., Boyd, P., Burkill, P. H., et al. A biogeochemical study of the coccolithophore, *Emiliania huxleyi*, in the North Atlantic, Global Biogeochem. Cy., 7, 4, 879–900. doi:10.1029/93GB01731, 1993.

Jakobsson, M., Mayer, L. A., Coakley, B., Dowdeswell, J. A., Forbes, S., Fridman, B., Hodnesdal, H. et al: The International Bathymetric Chart of the Arctic Ocean (IBCAO) Version 3.0., Geophys. Res. Lett., 39, L12609, doi:10.1029/2012GL052219, 2012.

Jerome, J. H., Bukata, R. P., and Miller, J. R.: Remote sensing reflectance and its relationship to optical properties of natural water, International Journal of Remote Sensing, 17,1, 43–52, 1996.

Johannessen, O. M.: Decreasing Arctic ice mirrors increasing CO_2 on decadal time scale, Atmospheric and Oceanic Science Letters, 1, 1, 51-56, 2008.

Key, R. M., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al.: Global Ocean Data Analysis Project, version 2 (GLODAPv2), ORNL/CDIAC-162, ND-P093. Carbon Dioxide Information Analysis Center. Oak Ridge, TN: Oak Ridge National Laboratory, US Department of Energy, doi:10.3334/CDIAC/OTG.NDP093_GLODAPv, 2015.

Kondrik, D. V., Pozdnyakov, D.V., and Pettersson, L.H.: Particulate inorganic carbon production within *E. huxleyi* blooms in subpolar and polar seas: a satellite time series study (1998-2013), International Journal of Remote Sensing, 38, 22, 6179-6205, doi:10.1080/01431161.2017.1350304, 2017a.

Kondrik, D., Pozdnyakov, D. V., and Pettersson, L.H.: Tendencies in Coccolithophorid Blooms in Some Marine Environments of the Northern Hemisphere According to the Data of Satellite Observations in 1998-2013, Izvestya, Atmospheric and Oceanic Physics, 53, 9, 955-964, 2017b.

Kondrik, D. V., Pozdnyakov, D. V., and Johannessen, O. M.: Satellite evidence that *E. huxleyi* phytoplankton blooms weaken marine carbon sinks, Geophysical Research Letters, 45, 2, 846-854, doi:10.1002/2017GL076240, 2018a.

Kondrik, D., Kazakov, E., and Pozdnyakov, D.: Dataset of E. huxleyi blooms: spatio-temporal distribution and their impact on high-latitudinal marine environments (1998-2016) (Version 1) [Dataset], Zenodo, doi:10.5281/zenodo.1402033, 2018b.

Korosov, A. A., Pozdnyakov, D. V., Folkestad, A., Pettersson, L. H., Sorensen, K., and Shuchman, R.: Semi-empirical algorithm for the retrieval of ecology-relevant water constituents in various aquatic environments, Algorithms, 2, 1, 470-497, doi:10.3390/a201047, 2009.

Loveday, B. R. and Smyth, T.: A 40-year global data set of visible-channel remote-sensing reflectances and coccolithophore bloom occurrence derived from the Advanced Very High Resolution Radiometer catalogue, Earth Syst. Sci. Data, 10, 2043-2054, https://doi.org/10.5194/essd-10-2043-2018, 2018.

Montegut, C., Madec, G., Fisher, A., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology, Journal of Geophysical Research, 109, C12003, 2004.

Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al.: The Global Ocean Data Analysis Project version 2 (GLODAPv2)—An internally consistent data product for the world ocean. Earth System Science Data, 8, 2, 297–323, doi:10.5194/essd-8-297-2016, 2016.

Press, W., Teukolsky, S., Vettering, W., and Flannery, B.: Numerical recipes in C: The art of scientific computing (2nd ed.), New York: Cambridge University Press, 1992.

Rivo Calle, S., Gnanadesikan, A., Del Castillo, C. E., Balch, W., and Guikema, S. D.: Multidecadal increase in North Atlantic coccolithophores and the potential role of rising CO₂, Science Express, doi:10.1126/science.aaa8026, 2015.

Rost, B., Zondervan, I., and Wolf-Gladrow, D.: Sensitivity of phytoplankton to future changes in ocean carbonate chemistry: current knowledge, contradictions and research directions, Marine Ecology Progress Series, 373, 227-237, 2008.

Shutler, J., Grant, M., Miller, P., Rushton, E., and Anderson, K. Coccolithophore bloom detection in the north Atlantic using SeaWiFS: Algorithm description, application and sensitivity analysis. Remote Sensing of Environment, 114, 1008-1016, 2010, doi:10.1016/j.rse.2009.12.024.

Shutler, J., Land, P., Brown, C., Findlay, H., Donlon, C., Medland, M., Snooke, and Blackford, J. Coccolithophore surface distributions in the North Atlantic and their modulation of the air-sea flux of CO2 from 10 years of satellite Earth observation data, Biogeosciences, 10, 2699-2709, 2013, doi:10.5194/bg-10-2699-2013.

Valente, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Taberner, M., Antoine, D., et al.: A Compilation of Global Bio-Optical in Situ Data for Ocean-Colour Satellite Applications, Earth System Science Data, 8, 235–252, doi:10.5194/essd-8-235-2016, 2016.

Winter, A., Hendericks, J., Beaufort, L., Rickaby, R.E.M., and Brown, C.W.: Poleward expansion of the coccolithophore Emiliania huxleyi, J. Plankton Res., 36, 316-325, 2014, doi:10.1093/plankt/fbt110.

5

Table 1. Spatial regions description

Region number	Extent coordinates (NSIDC EASE-Grid North, EPSG:3973)		Region Area, km²	Contained waters
1	Xmin	-300000.00	7 819 600	The Barents, Norwegian, North seas, the Northern part of the Greenland Sea
	Ymin	-4260000.00		
	Xmax	1960000.00		
	Ymax	-800000.00		
2	Xmin	-1000000.00	476 000	Southern part of the Greenland sea, Western part of the Norwegian Sea
	Ymin	-2720000.00		
	Xmax	-300000.00		
	Ymax	-2040000.00		
3	Xmin	-4180000.00	1 081 200	Southern part of the Labrador Sea, the North Atlantic Ocean part to the south of the Labrador Sea
	Ymin	-3500000.00		
	Xmax	-3160000.00		
	Ymax	-2440000.00		
4	Xmin	-1400000.00	1 680 000	The Bering Sea
	Ymin	2500000.00		
	Xmax	0.00		
	Ymax	3700000.00		

Table 2. Total number of identified pixels with *E. huxleyi* for each blooming season in the period 1998-2016 within the four regions.

Year	Total number of pixels with *E. huxleyi*				Year	Total number of pixels with *E. huxleyi*			
	Region 1	Region 2	Region 3	Region 4		Region 1	Region 2	Region 3	Region 4
1998	99538	214	2336	252003	2008	48399	8319	13131	1656
1999	97259	5754	9168	116622	2009	51620	2745	18102	14749
2000	73642	138	14205	282046	2010	77050	8110	46591	1232
2001	104425	1142	10432	109541	2011	116555	603	48101	22259
2002	104237	949	37335	694	2012	107791	4532	18630	618
2003	117877	312	40018	7466	2013	115764	10011	12302	2079
2004	109156	2275	10686	6657	2014	76396	15047	16245	50900
2005	76768	3300	23651	8679	2016	129569	1265	14890	6705
2006	97004	2444	3729	4061	2017	183546	1536	1779	16184
2007	80835	955	4237	17505					