



SIMBA

Système Intégré de Modélisation de la
Baleine noire de l'Atlantique

(Integrated modelling system for the North Atlantic Right Whale)

Phenology Algorithm Theoretical Based Document (ATBD)

Milestone #3


Date: March 31, 2022

Customer: CSA

Project no.: 9F040-190633/004



VERSION AND SIGNATURE

Prepared by	Version	Date	Modifications
Christopher Illori	1.0	02/2022	Initial version
Yanqun Pan, Thomas Jaegler	2.0	03/2022	Revised version
Verification by			Modifications
Christiane Dufresne	2.0	03/2022	Revision and edition
Simon Bélanger	2.1	03/2022	Final revision
Final Authorization			Signature
Simon Bélanger	2.1	03/2022	



APPLICABLE AND REFERENCE DOCUMENTS

ID	Description	Reference
PMTP	Project Management and Technical Plan	M2-T1.3.3
ATBD-Chla	Technical document: Chla	M3-T2.1.4
ATBD-Pheno	Technical document: phenology metrics	M3-T2.3.5
OCR-R	Ocean Colour Radiometry Processing Report	M3-T2.7.1



TABLE OF CONTENTS

VERSION AND SIGNATURE	I
APPLICABLE AND REFERENCE DOCUMENTS	II
TABLE OF CONTENTS.....	III
ACRONYMS	IV
LIST OF FIGURES	II
LIST OF TABLES	III
1. INTRODUCTION	4
1.1. CONTEXT	4
1.2. SCOPE OF THE DOCUMENT	5
1.3. BACKGROUND	5
2. METHODOLOGY.....	9
2.1. DESCRIPTION OF DATA.....	9
2.2. DATA PROCESSING.....	9
2.3. METHODS FOR ESTIMATING BLOOM.....	10
3. PRELIMINARY RESULTS	13
3.1. BLOOM THRESHOLD ($CHLA_{THRES15\%}$).....	13
3.2. BLOOM METRICS VARIABILITY: 2015-2020	14
4. FUTURE WORK.....	19
5. SUMMARY.....	25
6. REFERENCES	26



ACRONYMS

AOI	Area of interest
Chla	Chlorophyll-a
C-TEP	Coastal Thematic Exploitation Platform
CMEMS	Copernicus Marine Environment Monitoring Service
CSA	Canadian Space Agency
DFO	Department Fisheries and Oceans
DOY	Day of year
GoM	Gulf of Maine
GoSL	Gulf of St. Lawrence
GoSLM	Gulf of St. Lawrence and Maine
ESA	European Space Agency
IOCCG	International Ocean-Colour Coordinating Group
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
NARW	North Atlantic Right Whale
NASA	National Aeronautics and Space Administration
NRT	Near Real Time
OC	Ocean Color
OCR	Ocean Color Radiometry
OC-CCI	Ocean Color Climate Change Initiative
OLCI	Ocean and Land Color Imager
PFT	Phytoplankton Functional Type
SDM	species distribution modeling
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
TC	Transport Canada
VIIRS	Visible Infrared Imaging Radiometer Suite



LIST OF FIGURES

Figure 1. Idealized model of phenological events and linkages in the Gulf of Maine (GoM). Yellow highlighted bars, circles, and arrows indicate match points where one phenological event relies on another. Width of yellow and gray colored bars indicates the duration of events. Phytoplankton blooms, zooplankton diapause strategy (specific to <i>Calanus</i> copepods), non-diapause strategy (other zooplankton including euphausiids), and river outflow event timing are depicted based on magnitude measurements. The bottom row depicts the NAWR migration and growth strategy in relation with its favorite prey, <i>Calanus</i> spp. (modified from Staudinger et al., 2019).	6
Figure 2. Schematic of bloom phenology illustrating the metrics to be determined for each pixel (modified from Ferreira et al., 2021).	10
Figure 3. Top: mean of the 24 yearly median values computed between 1998 and 2021; bottom: coefficient of variation for me year median ($100 \cdot \text{std}/\text{mean}$).	14
Figure 4. Annual bloom frequency (number of blooms) for 2015-2020	15
Figure 5. Bloom initiation date (in DOY) for the first bloom (B^1_{init}) between 2015-2020	16
Figure 6. Bloom duration (in days) for the first bloom (B^1_{Dur}) between 2015-2020	17
Figure 7. Biomass cumulated ($\text{mg} \cdot \text{m}^{-3}$) of the first bloom (B^1_{Area}) between 2015-2020	18
Figure 8. Oceanographic sub-regions of the Gulf of St. Lawrence (from Galbraith et al., 2015)	20
Figure 9. Example of monitoring grid for the management of the NAWR established by DFO in the Gulf of St Lawrence.	21
Figure 10. Example of phenoregions	22
Figure 11. DOY of maximum primary production between March and September for MODIS time-series (2002-2020).	24



LIST OF TABLES

Table 1. Phenology metrics used in this work. *NRT* column indicates whether or not the metric will be calculable in near real time mode for the prediction of NARW SDM. *n* stans for n^{th} bloom..... 12

1. INTRODUCTION

1.1. Context

This document is produced as part of the project entitled “SIMBA - Système Intégré de Modélisation de la Baleine noire de l’Atlantique (Integrated modelling system for the North Atlantic Right Whale)” which is within the portfolio of projects of the smartWhales initiative, supported by the Canadian space agency (CSA), Fisheries and Oceans Canada (DFO) and Transport Canada (TC). It describes the steps to extract phenology metrics over the Gulf of St. Lawrence and Maine (GoSLM) using Chla 8-day products compiled within the framework of the GlobColour project (www.globcolour.info). The GlobColour provides a time series of daily, weekly and monthly observations of OCR products from different satellite missions that allow the characterization of ocean surface.

In SIMBA, Ocean Color Radiometry (OCR) is a major source of information for the species distribution modeling (SDM). The work package 2 (WP2) deals with OCR data processing including both archive and near real-time (NRT). Briefly, we propose:

- 1) to develop of a regional chlorophyll-a (Chla) algorithm for the GoSL, Scotian shelf and GoM (GoSLM);
- 2) to extract phytoplankton bloom metrics from Chla time series,
- 3) to compute of phytoplankton primary production,
- 4) to identify of Phytoplankton Functional Types (PFT; Xi et al., 2020), and
- 5) to implement a novel presence/absence algorithm for direct detection of surface Calanus swarms.

Within the SIMBA project, a regional algorithm (based on principal component analyses of the ocean reflectance - PCA) and validation have already been performed for Chla retrieval in the GoSLM (see *Applicable and Reference Documents*). Results showed that the developed regional algorithm performed slightly better than other algorithms, such as OCX (O’Reilly et al., 1998), OC5 (Gohin et al 2011) and GSM (Garver, Siegel and Maritorena) (Maritorena & Siegel, 2005).

Additionally to the PCA-based algorithm output, the Chla dataset product serves as input for this phenology algorithm. The latter will be used to extract phytoplankton bloom metrics and will be implemented in the Coastal Thematic Exploitation Platform (C-TEP) to produce weekly phenology metrics from NRT Chla products. The C-TEP is a dedicated cloud computing platform for the management and analysis of large satellite ocean-related data. The weekly Bloom metrics produced from NRT Chla products will be provided to the modelling team of WP3 as inputs into the SDM algorithm.

1.2. Scope of the document

This ATBD describes the workflow to compute a suite of phenology metrics for the GoSLM. First, we give a background for the work. Next, we describe data and the phenology extraction algorithm including the relevant indices that are derived for the GoSLM. Lastly, we present and discuss the results briefly. Our main objective is to derive metrics which best predict the seasonal variability of surface phytoplankton concentration in the GoSLM. Hence we have produced 10 metrics from which metrics that best predict the distribution of the North Atlantic Right Whale (NARW) will be chosen for SDM prediction by modellers in WP3. We expect that the phenology metric will be significant to predict the arrival of the NARW in the Gulf of St. Lawrence, which is the ultimate goal of SIMBA. All metrics were extracted from time series of daily regionally-tuned Chla concentration products from the GlobColour processor and CMEMS that shall be produced in NRT later in the project. The phenology metric algorithm will be implemented in the C-TEP for NRT data processing.

1.3. Background

Phytoplankton bloom intensity and timing (phenology) provide an understanding about nutrient availability and predation by migratory marine species, whose reproduction relies on marine primary production (Viser et al. 2011). Seasonal presence of marine mammals coincides more or less with nutrient concentration and phytoplankton abundance (Biggs et al., 2020). Thus, knowledge about the variability of phytoplankton over the season can provide information about their occurrence and timing. Satellite-derived phenology data sets can be used to determine the timing and magnitude of recurring biological and seasonal growth of phytoplankton (e.g., Racault et al., 2012), which will be theoretically tied to zooplankton bloom, and ultimately to the

distribution and migration of the NARW in the GoSLM (the intermediate temporal habitat for the NARW). While the timing of whale migration is not directly linked to the timing of the phytoplankton bloom, whales usually capitalize on secondary production of zooplankton. For instance, the phenological state of the Calanus species (*finmarchicus*, *hyperboreus* or *glacialis*), which the NARW forages on, is likely linked to that of phytoplankton (Figure 1). Thus, knowledge about the regional and inter-annual variability of phytoplankton blooms over a long term can improve understanding about the overall system productivity (Cushing, 1990) and processes such as arrival time, migration timing and foraging pattern of long-distance migrants.

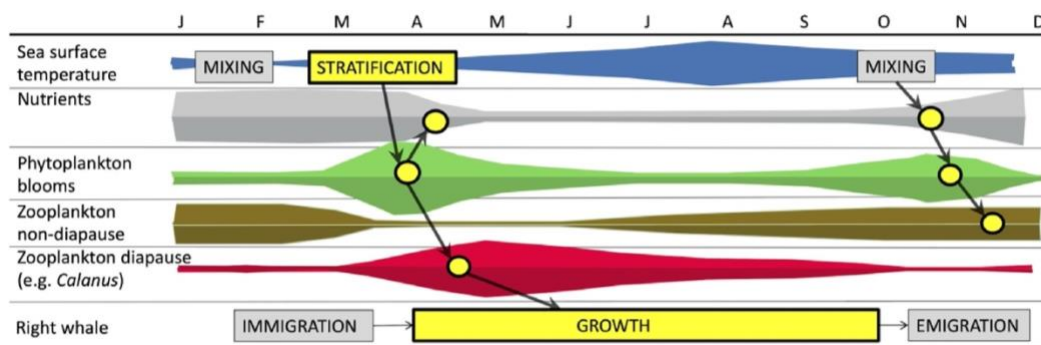


Figure 1. Idealized model of phenological events and linkages in the Gulf of Maine (GoM). Yellow highlighted bars, circles, and arrows indicate match points where one phenological event relies on another. Width of yellow and gray colored bars indicates the duration of events. Phytoplankton blooms, zooplankton diapause strategy (specific to Calanus copepods), non-diapause strategy (other zooplankton including euphausiids), and river outflow event timing are depicted based on magnitude measurements. The bottom row depicts the NAWR migration and growth strategy in relation with its favorite prey, Calanus spp. (modified from Staudinger et al., 2019).

In the North Atlantic, spring phytoplankton bloom is one of the major biological events that have been documented from space for many years (Siegel et al. 2002). Part of this area, notably the GoSLM, serves as an intermediate temporary habitat for the NARW. Thus, the timing of recurring biological and seasonal growth of phytoplankton will be central to the migration and presence of the NARW, which may rely on the abundance of Calanus finmarchicus and Calanus hyperboreus -

the preferred preys of the NARW (Pendleton et al. 2020). To effectively track the movement of the NARW and limit their exposure to increased risk of vessel collisions, it is important to assess the distribution of its preferred prey. For example, *Calanus finmarchicus* is ubiquitous over the biogeographic range of NARW but has recently shown a pronounced decline in abundance due to rapid warming in the GoM (Record et al. 2019), while *Calanus hyperboreus* is uniquely thriving in the GoSL (Khan et al. 2016) but has recently and unexpectedly been in short supply in the Magdalen Shallows that now attracts much of the NARW from spring to fall (Blais et al. 2021; Gavrilchuk et al. 2021).

Chla concentration is a proxy for phytoplankton biomass, a potential environmental driver impacting the NARW behaviour. In situ data provide a valuable source of information on Chla concentration, but inherently lack the spatio-temporal coverage needed in the context of NARW and phytoplankton/zooplankton's distribution. Remote sensing, however limited to the surface, provides the only way to observe and monitor phenological dynamics at both global and regional scales. Long-term records of species abundance observed from moderate resolution satellite observations can provide up-to-date measurements of ocean surface properties, and are therefore well-suited for monitoring seasonal-to-decadal trends in global to regional phenological shift of phytoplankton dynamics. A variety of studies have examined several aspects of phytoplankton dynamics in oceans using OCR products, all of which have different objectives. Using a 10-year time series of satellite imagery, Platt et al. (2009) carried out a spatiotemporal analysis of phytoplankton dynamics by examining their independence on time and latitude. In another study by Gittings et al. (2018), the effects of warming on the phytoplankton phenology and abundance were investigated. Likewise, Racault et al. (2012) examined global phytoplankton dynamics and retrieved phenology metrics using time series of OC data over 9 years. At a much lower scale, Racault et al. (2015) used the ESA's OC-CCI data to obtain phenology metrics for characterizing phytoplankton seasonality in the entire Red Sea basin.

Most phenology extraction methods based on the use of OC data usually focus on the timing of the initiation of the spring bloom and/or the timing of its peak. Platt and Sathyendranath (2008), suggested several metrics but the most commonly used are still the timing of the maximum Chla concentration during the spring bloom (e.g., Yoder and Kennelly, 2003) and the timing of initiation of the spring bloom (e.g., Cole et al., 2012; Brody et al., 2013). Although there are no universally accepted metrics, studies usually use a threshold of the annual mean/median (Siegel et al., 2002)



or cumulative distribution (Brody et al., 2013) of Chla concentration to detect the key phenological state. For example, Racault et al. (2015) used four parameters including **initiation**, **peak**, **termination** and **duration** to study the bloom patterns in the Red Sea basin. Another important thing to consider is the temporal resolution. Some studies have used satellite data with a daily or weekly resolution, but most use monthly or fortnightly composites to avoid gaps (e.g., Salgado-Hernanz et al. 2019). One advantage of using daily products is the increased probability of having per pixel data for each date in a cycle. Therefore, our first attempt was to use archived daily Chla 4-km resolution products to design and test the phenology algorithm using archived dataset. Future work will be done using the regional product being produced by our partners.

2. METHODOLOGY

2.1. Description of data

We used time series of daily Chla products (4.6 x 4.6 km spatial resolution) for the period 1998-2021 readily available to develop the phenology algorithm. Test data used for the development of the algorithm include Level 4 daily interpolated Chla products from the CMEMS (<https://marine.copernicus.eu/>) produced using the GlobColour processor (<https://hermes.acri.fr/>). We have used these archived products for the development of the phenology metrics and efforts are underway to produce the same metrics directly from Chla products for the SIMBA project which have recently been produced following the implementation of the PCA algorithm in the GlobColour processor. The results presented here will be compared against data from the PCA-based algorithm for performance assessment in the coming weeks. Within the SIMBA project, the main input for producing phenology metrics in the C-TEP will be the PCA-based Chla daily interpolated products which have been produced from merged product generated with data from OLCI, SeaWiFS, MERIS, MODISA, VIIRSJ and VIIRSN sensors. It is important to note, however, that the Level 4 interpolated regional products (refer to as L4-REG in MS2 report) may suffer from high noise levels. Data aggregation (downscaling) will be likely needed, as discussed in section 4 below (also see OCR processing report).

2.2. Data processing

We first build a three-dimensional array by combining daily maps of Chla ($CHL_{x,y,t}$). Any invalid pixels (e.g., pixels with -999) were assigned an 'na' value to differentiate them from other valid pixels needed for calculating metrics. This was to ensure a robust per pixel retrieval. The algorithm currently computes metrics from daily products but will be adapted to L4-REG after aggregation to minimize the impact of noise inherent in daily full resolution data. This noise is added on top of the natural variability and complexity of surface waters in the GoSLM.

2.3. Methods for estimating bloom

The algorithm computes eight bloom metrics (Table 1). The choice of the method was made with the aim of determining the metrics in NRT (or almost) during the predictive modeling of the NAWR habitat. In addition to the metric that can be estimated in NRT, two additional metrics could be computed after the completion of the whole seasonal cycle, i.e. the bloom frequency and the total biomass accumulated during the seasonal cycle. These metrics may be included in the hindcast NARW species distribution model (SDM). The metrics that will be selected in both hindcast and forecast SDM, if any, remain to be determined by the modelling team. The metrics generated for the modeling team are listed in Table 1 and illustrated in Figure 2. As depicted in Figure 2, more than one bloom can occur during the seasonal cycle at our latitude. Therefore, the metrics of each bloom (1^{st} , 2^{nd} , ..., n^{th}) are calculated, which is different from most methods found in the scientific literature.

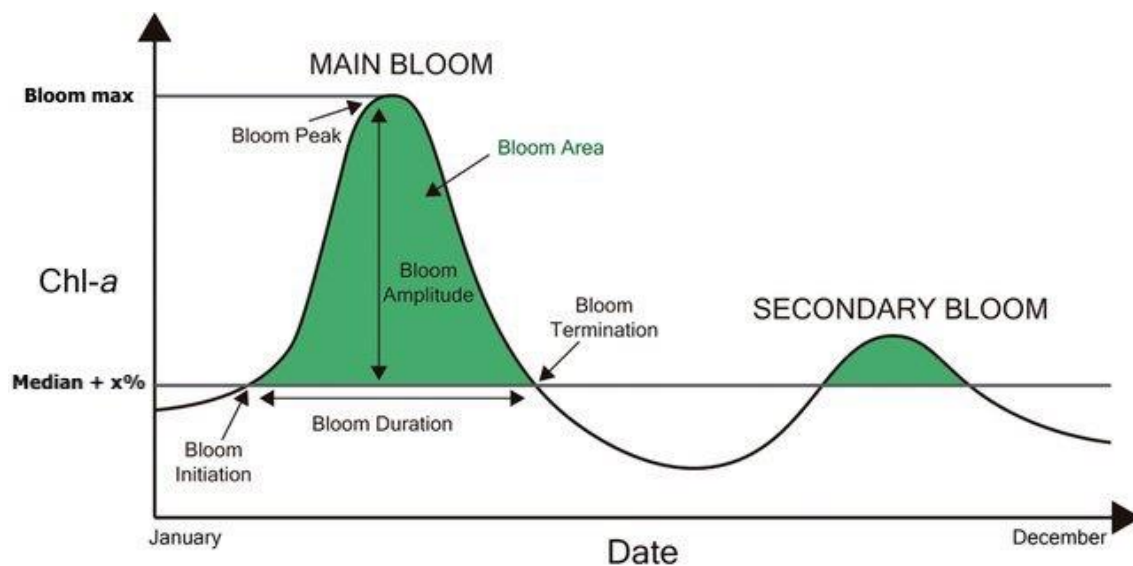


Figure 2. Schematic of bloom phenology illustrating the metrics to be determined for each pixel (modified from Ferreira et al., 2021).

We used a threshold-based approach (Cole et al. 2012; Racault et al. 2012; Fleming and Kaitala, 2006; Siegel et al. 2002) to extract all phenology metrics. This method uses a threshold above the annual median to determine the spring bloom start (B_{init}), which is likely the most important metric in the context of the NARW SDM. The median value varies spatially, i.e. each pixel has a value. One advantage of using a threshold-based method is that it is insensitive to the percentage used (Siegel et al., 2002); though values outside 5-15% may lead to error (Racault et. al, 2015). Siegel et al. (2002) tested different values ranging from 1% to 30% and found very little difference. This approach is better suited for spring bloom detection in NRT (Brody et al. 2013; Greve et al. 2005), and is ideal for the SIMBA project due to the nature of the NARW migration pattern. We tested different threshold criteria and found almost no sensitivity in the estimation of most metrics and very little sensitivity in B_{Area} and B_{Dur} when thresholds ranged within median plus 5 to 15% ($Chla_{thresX\%}$). For B_{Dur} , the mean extended by 4 and 5 days for the entire region when the threshold was decreased from 15% to 10% and from 10% to 5%, respectively. Results shown here are based on a 15% threshold above the median computed using a climatology value for the median ($Chla_{thres15\%}$; see below). We applied the algorithm to the whole time series of Chla data currently available in the CMEMS database (1998-2021) to produce the metrics and a complete coverage of the annual cycle in the GoSLM. A **bloom** is considered to have occurred if the **CHL value is maintained above the median + 15% threshold for a minimum of 15 days** (or 16 days when using binned data a 8-day) (Brody et. al, 2013; Cole et. al, 2012). We applied these criteria to each pixel and for each year, yielding one map per year per parameter. The metrics used to characterize bloom for the entire GoSLM are detailed in Table 1.

Table 1. Phenology metrics used in this work. *NRT* column indicates whether or not the metric will be calculable in near real time mode for the prediction of NARW SDM. *n* stans for n^{th} bloom

Metric	Description	Unit	NRT*
B_{init}^n	Day of year of initiation of each bloom	-	Yes
B_{term}^n	Day of year of termination of each bloom	-	Yes
B_{Dur}^n	Duration of of each bloom ($B_{term} - B_{init}$)	Days	Yes
B_{peak}^n	Day of year of Chla maximum of each bloom	-	Yes
B_{Amp}^n	Difference between Chla maximum and the threshold of each bloom	$\text{mg}\cdot\text{m}^{-3}$	Yes
B_{Area}^n	Biomass cumulated of each bloom	$\text{mg}\cdot\text{m}^{-3}$	Yes
B_{Freq}	Bloom frequency during the seasonal cycle	$\text{mg}\cdot\text{m}^{-3}$	No
Y_{area}	Total biomass accumulated during the seasonal cycle	$\text{mg}\cdot\text{m}^{-3}$	No

* NRT cannot be achieved within 1 day, but at best 15 days after the acquisition (see OCR report for more explanation; ref M3-T2.7.1).

The bloom amplitude, B_{Amp} , was estimated by taking the difference between Chla maximum and the threshold. B_{init} and B_{term} are defined as the day of year (DOY) of initiation and termination of each bloom, respectively. The later is the day when Chla biomass fell below $\text{Chla}_{\text{thresX\%}}$. B_{Dur} and B_{Area} were calculated as the duration and cumulated biomass of the blooms, respectively, while B_{peak} represents the DOY of Chla maximum for each bloom. Bloom duration (B_{Dur}) is the time between bloom initiation and bloom termination ($B_{term} - B_{init}$), and bloom amplitude (B_{Amp}) was estimated by calculating the difference between the maximum and the threshold of Chla concentration.

3. PRELIMINARY RESULTS

For the purpose of the ATBD, we have analyzed all the data from 1998-2021, but we present results mainly from the recent years (2015-2020). While we have computed metrics from CMEMS products, the phenology analysis for the SIMBA project (currently underway) will be based on L4-REG products, which are being produced by the ACRI team (refer to *OCR data processing report* listed in the *Applicable and Reference Documents* section). We first present the threshold variability assessed using 24 years of OCR data.

3.1. Bloom threshold ($Chla_{thres15\%}$)

The choice of the threshold to detect phytoplankton bloom is determinant. As mentioned above, we assume a bloom when the Chla is greater than 15% of the median value (Cole et al. 2012). When bloom phenology is computed on archived data, one can use dynamic threshold that varies from year to year (yearly $Chla_{thres15\%}$). This is done in most published work using threshold-based approaches (e.g., Siegel et al., 2002; Cole et al., 2012; Ferreira et al., 2021). In addition, $Chla_{thres15\%}$ is slightly higher compared to most phenology study (median + 5%; e.g., Siegel et al., 2002; Cole et al., 2012; Brody et al., 2013; Ferreira et al., 2021) to reduce the number of small secondary blooms. We first examine the variability in the yearly median Chla. Figure 3 presents the mean of all yearly medians and its coefficient of variation (i.e. standard deviation divided by the mean). As expected, the most productive areas are coastal areas in the nearshore regions. Higher threshold is found in the St Lawrence estuary (sub-region names within the GoSL could be found in Figure 8), along the New-Brunswick coast and in the southern Gulf of Maine near Cape Cod and the Great South Channel (Figure 3, top panel). More importantly the interannual variability of the yearly median is usually less than 20% in the GoMSL. Much larger variability is found in the frontal zone between the cold waters from the Labrador and GoMSL and the warm waters from the Gulf Stream, where the coefficient of variation can reach as much as 50%. The variation in yearly median for year to year is likely related to the interannual variability in nutrients availability and phytoplankton productivity, but also linked to the data availability that can create biases (e.g. more/less clear sky during a given period of the year).

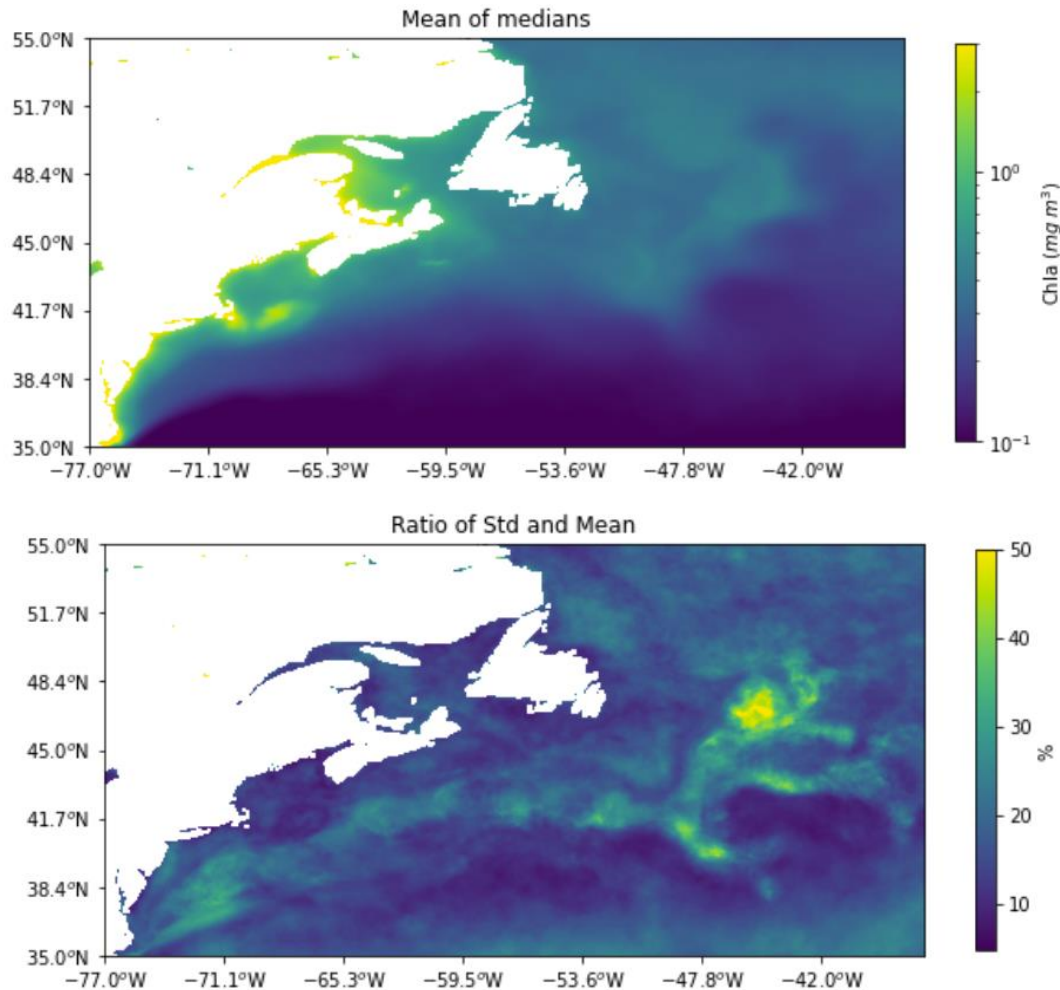


Figure 3. Top: mean of the 24 yearly median values computed between 1998 and 2021; bottom: coefficient of variation for me year median ($100 \cdot \text{std}/\text{mean}$).

3.2. Bloom metrics variability: 2015-2020

We first examine the number of blooms detected by the algorithm for the period 2015 to 2020 (Figure 4). While one or two blooms per year characterize a large part of the area of interest (shades of blue in Figure 4), several blooms are not unusual in the eastern part of the AOI, especially on the Newfoundland-Labrador shelves. In contrast, no bloom was detected in some areas. In 2015, for example, no bloom was detected in most parts of the GoM, while 2 to 4 blooms

were detected in the GoSL. In 2016, the situation was reversed, with many pixels without any bloom over the Magdalen shallows and several blooms in the GoM. In fact, more than 4 blooms also frequently occur in the GoSL (2017, 2018, 2020), which generally displays more frequent bloom events than the GoM (Figure 4). Except in 2016, 3 or 4 blooms occurred in GoM, while GoSL hosts as many as 6 or 7 events per year. Blooms are more frequent in the southern part of GoSL, except in 2019, when more blooms occurred in the northern GoSL. Interestingly, the location of high bloom frequency varies within each Gulf. For example, high bloom frequency occurred east and west of Anticosti Island in 2019 and 2020, respectively.

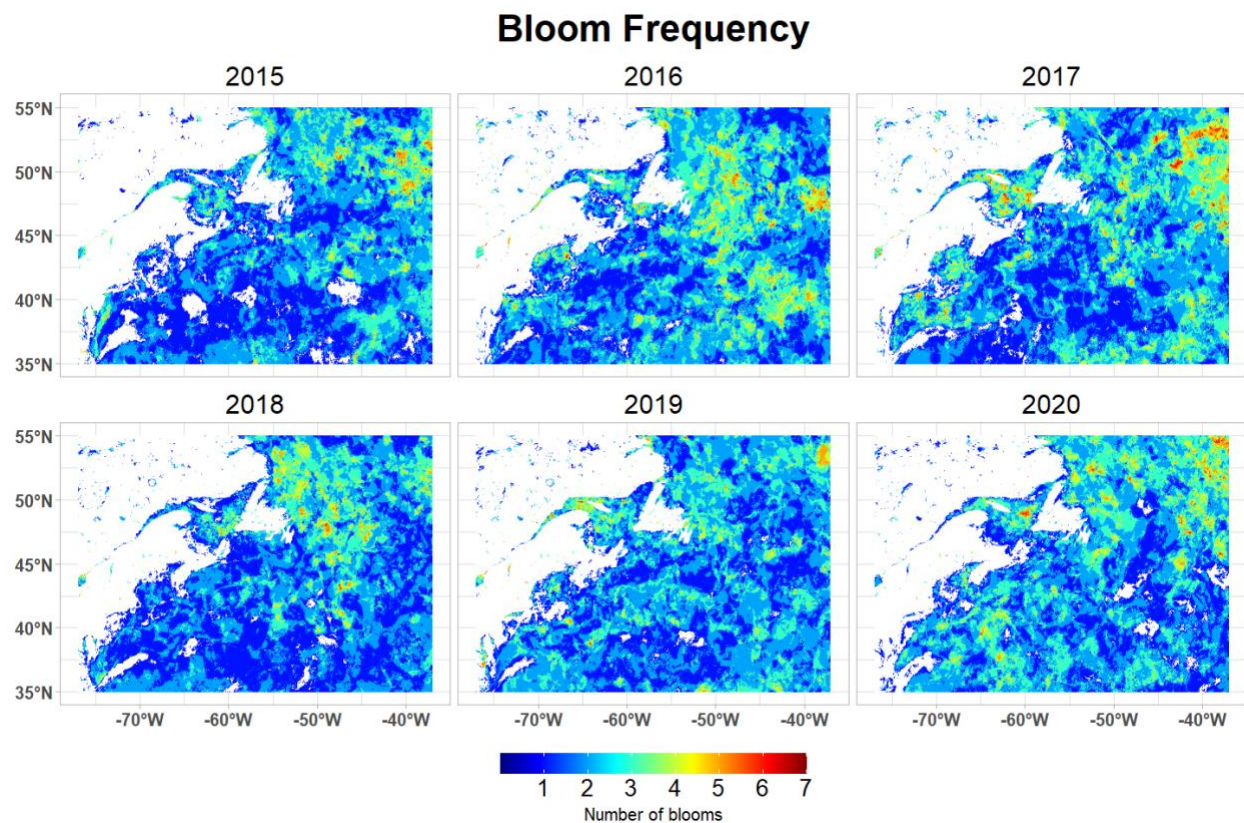


Figure 4. Annual bloom frequency (number of blooms) for 2015-2020

Figure 5 to Figure 7 show additional bloom metrics for the period 2015-2020: the day of year for the initiation (B_{init}), the duration (B_{Dur}) and the area (B_{Area}) of the first bloom, which is expected to play a key role in terms of NARW arrival in the GoSL. The initiation of the first bloom shows large variability in both time and space (Figure 5). In the GoMSL coastal area, the bloom is first detected over the Scotian Shelf (e.g., 2015, 2016, 2018). Both GoSL and GoM generally host the first bloom in spring (April = 90 to 120) and our results show no clear differences between the two areas. In some years, the algorithm identified B^1_{init} in late summer or beginning of fall (orange and red on Figure 5). This result suggests that no spring bloom was detected and only late summer or fall blooms were detected, which is particularly evident in 2015 in the GoSL, in 2016 in the western GoSL and in 2019 in the GoM. The large spatial variability in B^1_{init} in the AOI may be introduced by noise in the daily data set.

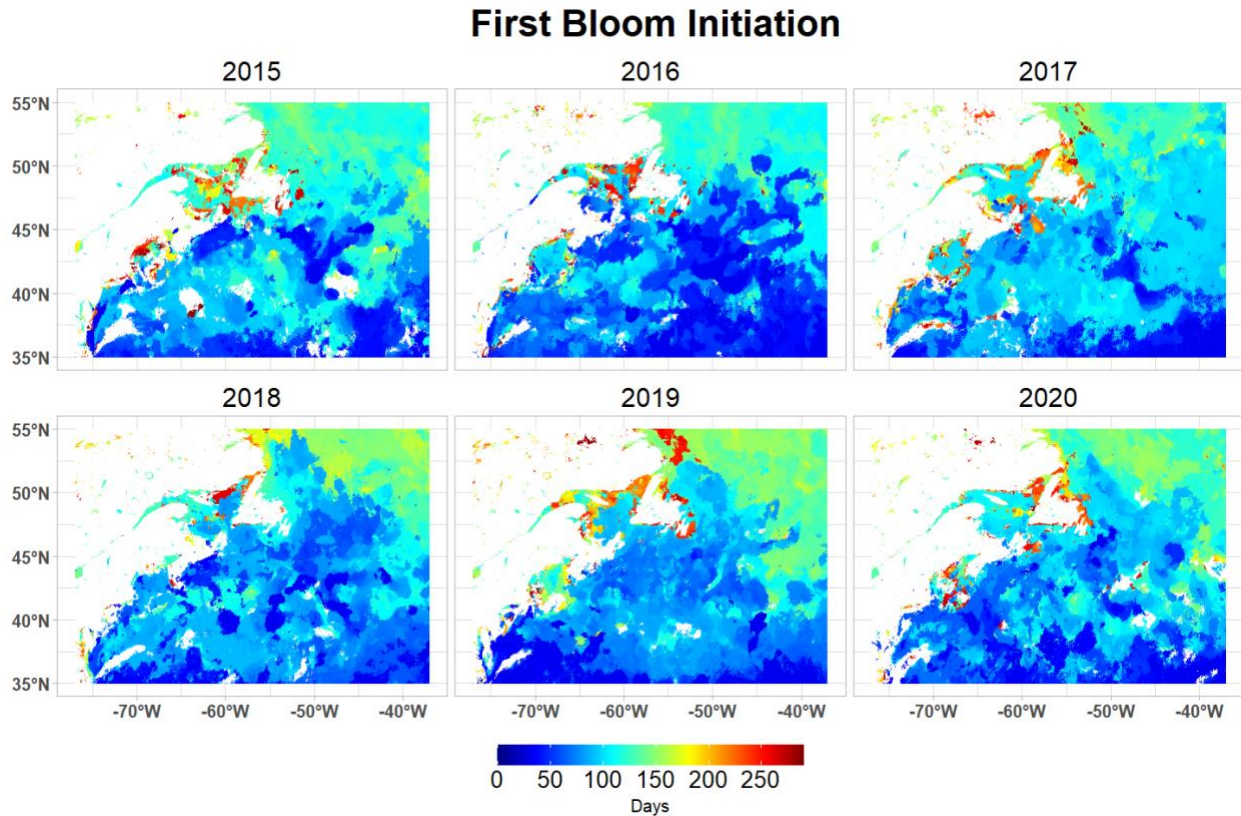


Figure 5. Bloom initiation date (in DOY) for the first bloom (B^1_{init}) between 2015-2020

The duration of the first bloom (B^1_{Dur}) varies substantially over the area, ranging from a couple of weeks (because of our 15-day criterion to define a bloom, see section 2.3) to several months (Figure 6). The first bloom generally lasts longer than the second one, but some inshore areas within GSL (e.g. Anticosti Island) exhibit longer 2nd bloom (not shown). In the GoMSL, most of the first blooms last less than 30 days, except in GoM in 2016 where the bloom lasted about 2 months after starting in April (DOY 90 to 120; Figure 5). The B^1_{Dur} is more variable (noisy) in the eastern part of the domain, outside the main AOI for SIMBA.

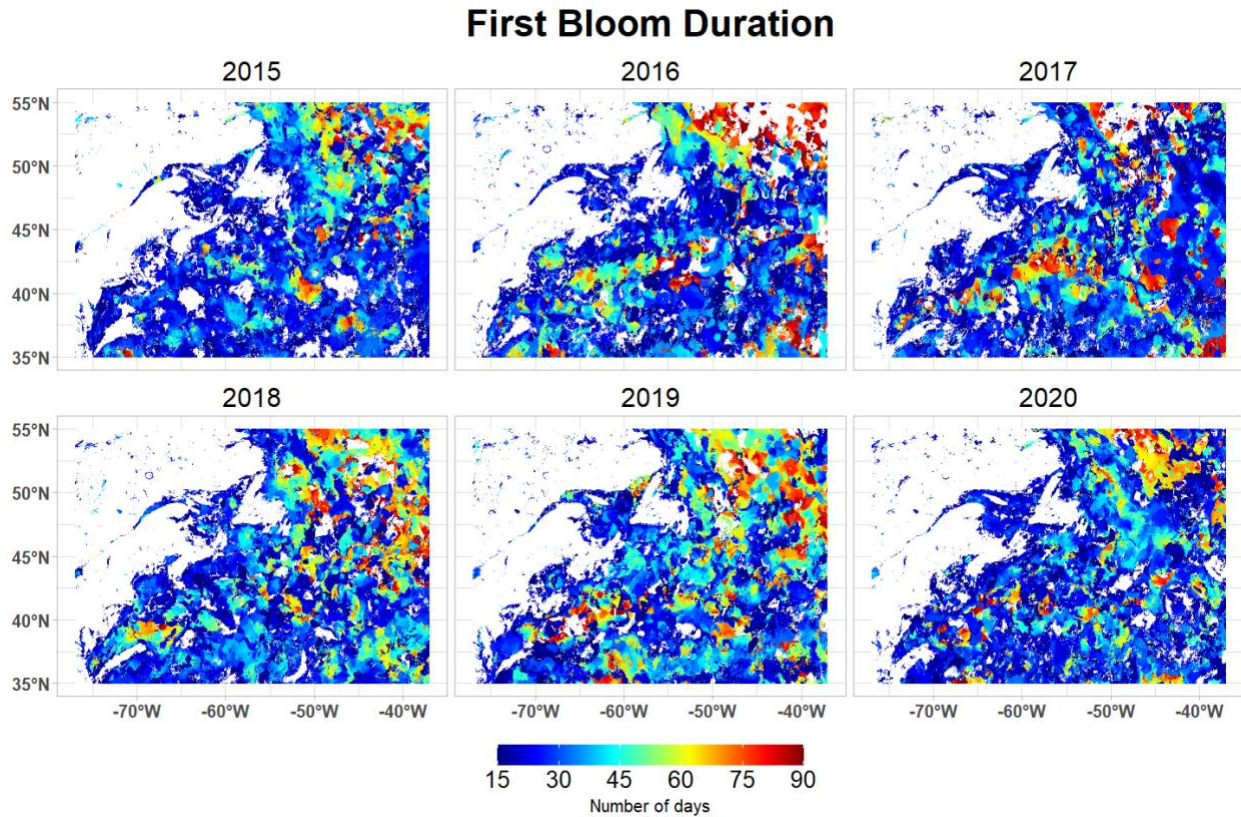


Figure 6. Bloom duration (in days) for the first bloom (B^1_{Dur}) between 2015-2020

Figure 7 shows the first bloom area, defined as the cumulated biomass over the bloom duration (mg m^{-3}), i.e. a good proxy of the first biomass production during the season. The highest biomass

accumulations are located in coastal areas, especially in the St. Lawrence Estuary, along the coasts of the GoM and in the Northumberland Strait. Offshore regions might also exhibit high biomass, such as the north-eastern part of our AOI (2016, 2017) and along the Newfoundland shelf-break (2018, 2019). It is interesting to note the strong bloom observed in 2017 in the Magdalen Shallows, where NARW were first observed in the GoSL. That year, the first bloom was weak in the GoM.

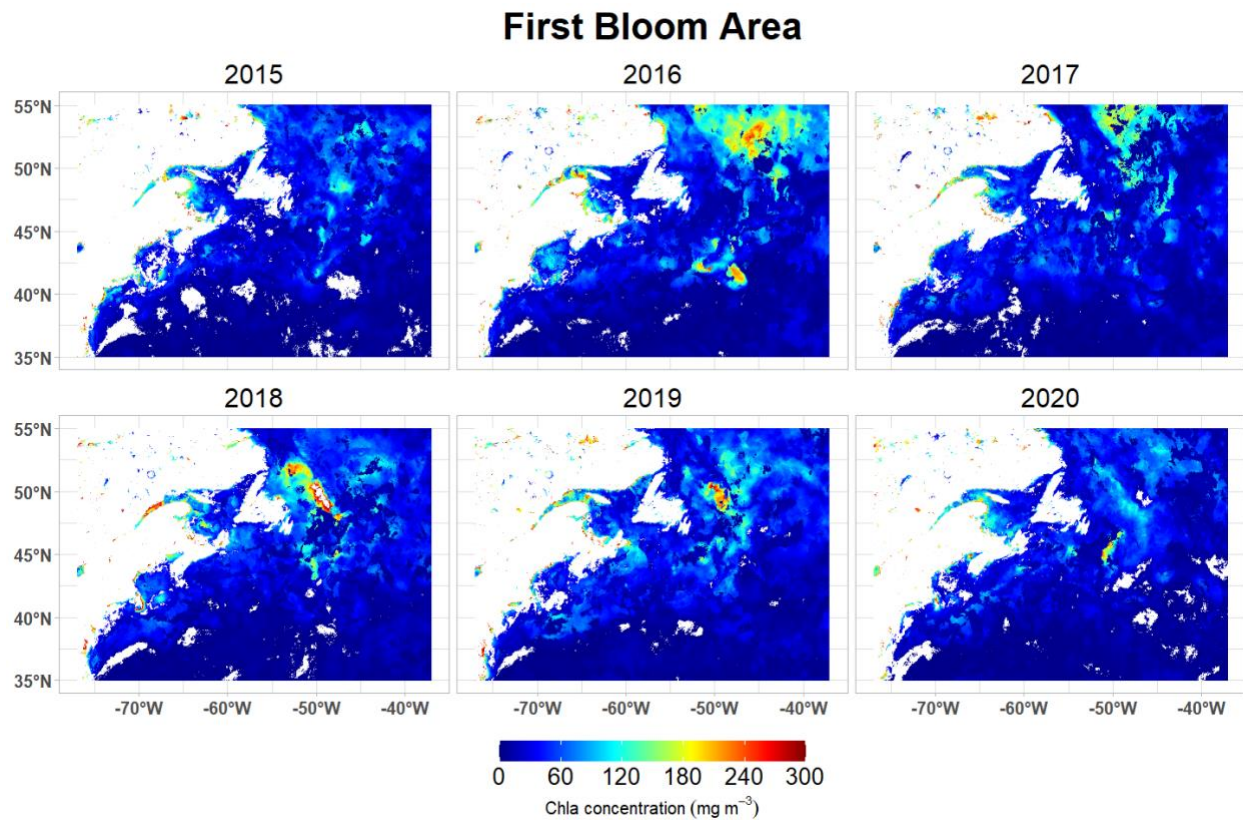


Figure 7. Biomass cumulated (mg.m^{-3}) of the first bloom (B^1_{Area}) between 2015-2020

4. FUTURE WORK

The framework for extracting the phenology metric has been developed, but more work is required to make the metric suitable for the modelers. During Phase 3, iterations are expected between the modeling team and the remote sensing to fine tune the method. Foreseen work is described below:

Yearly median Chla

The stability of the yearly median Chla on which the threshold is based must be assessed. Phenology metric computed using yearly median will be compared to those computed using the climatology of the median (Figure 3, top panel). In addition, trends in the yearly median between 1998 and 2021 will be calculated. If trends are significant, we will need to account for it for the final metric calculation.

Implications of NRT metrics on NARW SDM

The method described above can only provide bloom metrics in **quasi** NRT. The bloom initiation, for example, can only be confirmed after 16 days of its start date, while bloom peak, termination, duration and area will be output only when the bloom ended. The implications for the NARW SDM predictions will need to be assessed by the modeling team.

Spatial data aggregation

Spatial aggregation (downscaling) of the data may be necessary to yield smooth phenology metrics in space. As mentioned above, the algorithm will first use as input the L4-REG PCA-Chla product, which is a binned on daily product at 1-Km spatial resolution interpolated to fill the gaps (see the *OCR processing report*). At this high spatial resolution, noisy data will introduce large uncertainty in the bloom metrics output (e.g., Cole et al. ,2012). In such cases, further data aggregation will be needed. Here, the L4 daily 4-km resolution data were used and the noise was obvious in the metrics. More work is needed to find the best compromise to aggregate the Chla data in both space and time.

Three approaches will be considered and discussed with the modeling team:

1. L4-REG could be resampled to reduce the spatial resolution from 1-km to 2-km, 4-km, 8-km, 15-km, 25-km in order to remove as much noise as possible.
2. The study area has already been subdivided into sub-regions based on ecological features. For example, DFO has established a framework for the monitoring of the NARW in the Gulf of St. Lawrence (Figure 8 and Figure 9). Phenology metrics could be computed for these subregions specifically. Discussions with DFO scientists and modelers is required to determine whether such a regionalization would be relevant.

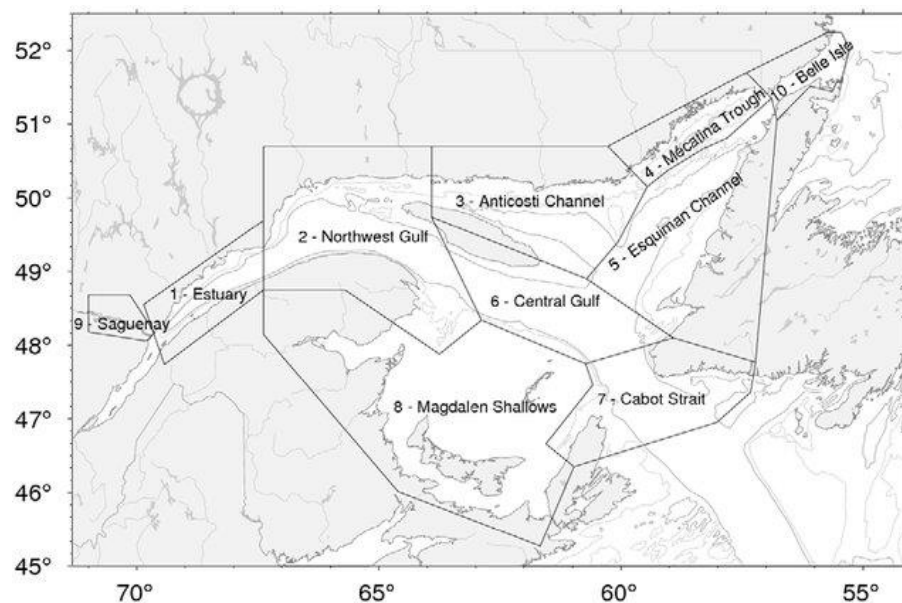


Figure 8. Oceanographic sub-regions of the Gulf of St. Lawrence (from Galbraith et al., 2015)

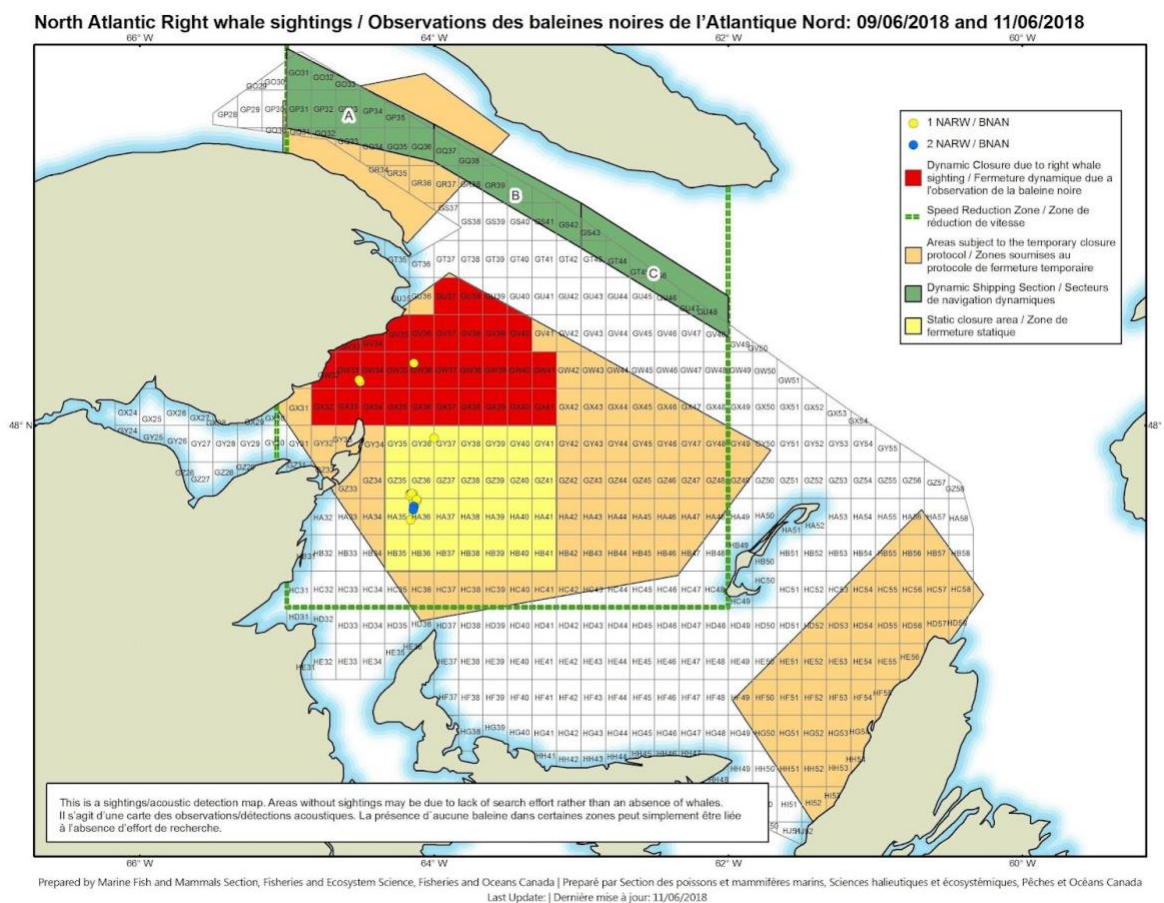


Figure 9. Example of monitoring grid for the management of the NAWR established by DFO in the Gulf of St Lawrence.

- Following a recommendation of IOCCG (2009), we could divide the entire GoSLM into a limited number of subregions represented by their similarity in terms of bloom phenology (so called phenoregion; e.g. Krug et al., 2017; Ferreira et al., 2021). This will enable an objective identification of areas with similar bloom phenology in a bid to investigate the phenology pattern across the entire area. Thus, we classify the entire region into different sub-groups using K-means clustering by iteratively classifying every pixel with similar phenology into clusters. The number of sub-regions (or clusters) will need to be

determined to fit the purpose of the NARW SDM modelling effort. Figure 10 shows an example of different phenoregions that the GoSLM can be partitioned into. We used K-means clustering to group the blooms with similar statistical behaviour into 6 classes. The four metrics selected for this clustering include B_{init} , B_{Dur} , B_{Peak} and B_{Freq} for the period 1998 - 2002. Region 1 (dark blue cluster) is observed mainly in GoSL and GoM, extending further south to the US east coast. Region 2 (light blue cluster) occupies the southern edge of the GoSLM, and covers most of the areas from latitude 42.50N to latitude 350N. Region 3 (cyan cluster) is found in the coastal strip along US east coast and GoM, including the southern strip of GoSL. It also covers the Lower St Lawrence Estuary. Region 4 (yellow cluster) is directly above region 2, with some parts of it within region 2. It also extends in a north western direction from longitude 450W to longitude 550W. This region, together with region 2, covers more than half of the GoSLM. Region 5 (orange cluster), sitting above region 4 and extending into it in the central area, partly extends to the GoSL and covers Newfoundland. Region 6 (red cluster), which is the smallest group, is mainly seen in small patches in a few areas along the US east coast.

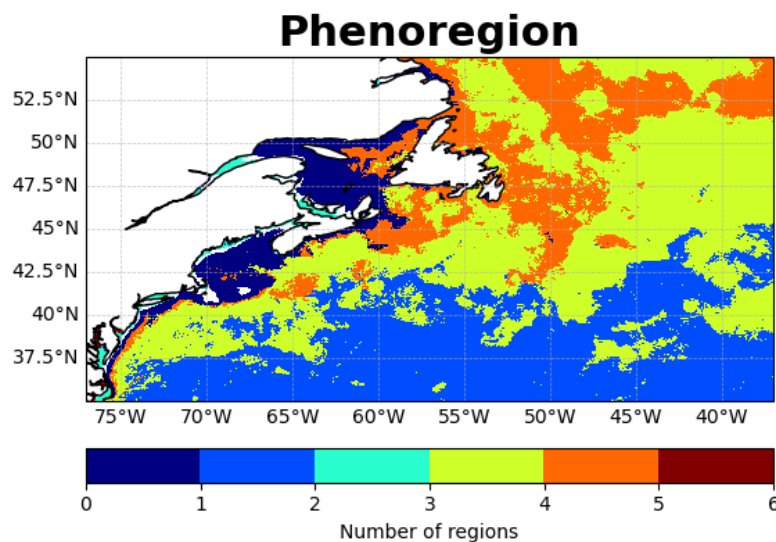


Figure 10. Example of phenoregions



Primary production metrics

Phenology metrics could also be computed with primary production (PP) rates as input instead of Chla biomass. Additionally, for the NRT computation, we propose to calculate the 45 days and 90 days running cumulative sum of primary production (PP_45D_accum, PP_90D_accum). As an example, Figure 11 shows the DOY of the annual maximum of PP for MODIS data time series (2002-2020). PP reaches its maximum value earlier in the warm water of the Gulf Stream in the southern portion of the domain. In the GoMSL, PP reaches its maximum in summer, but large interannual variability is observed. Interestingly, the PP maximum was reached earlier in 2017 in the GoSL compared to other years and to GoM.

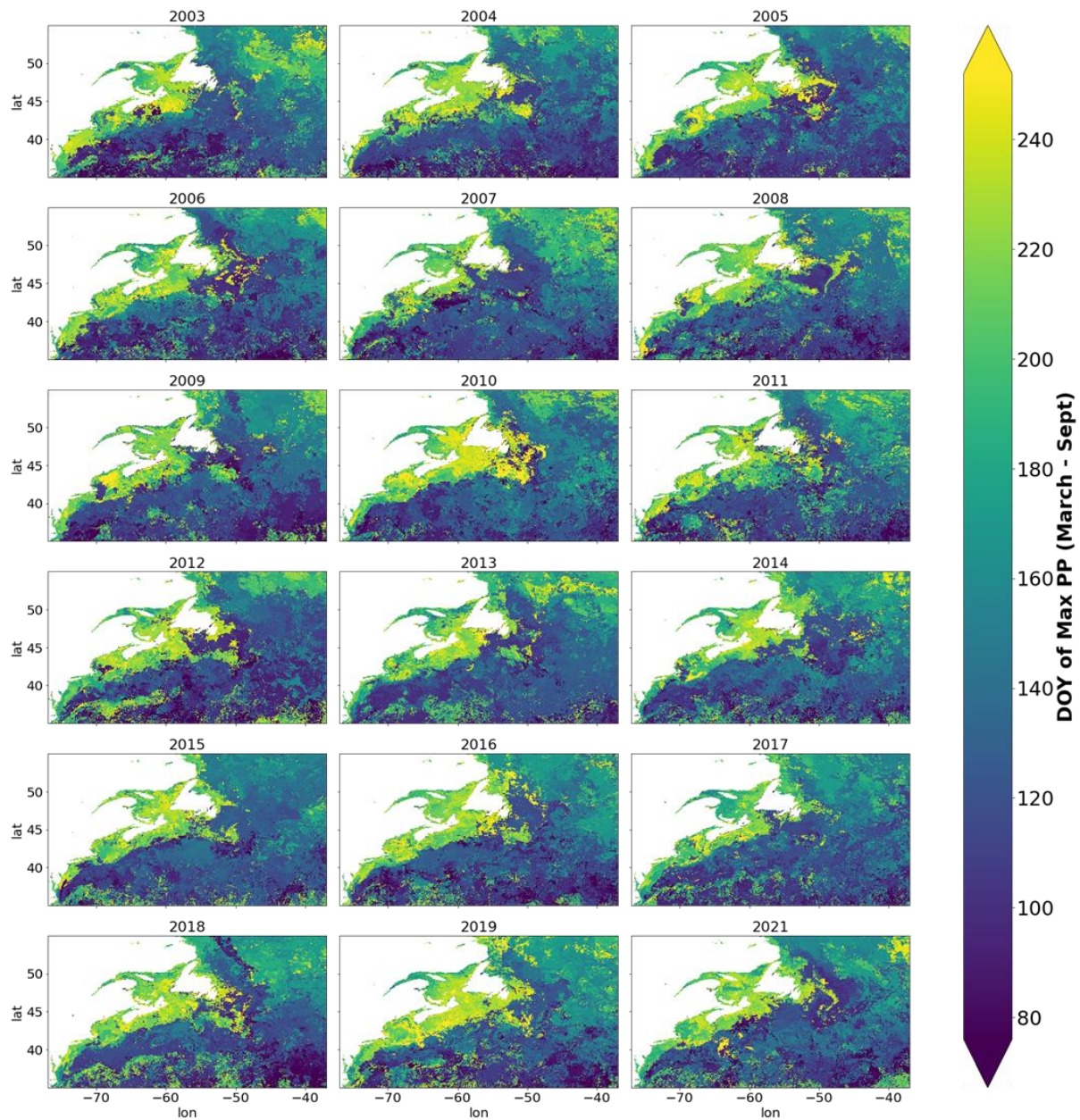


Figure 11. DOY of maximum primary production between March and September for MODIS time-series (2002-2020).

5. SUMMARY

The work presented here is part of the SIMBA work package 2 that supports the provision of critical data that will help to model NARW habitats and/or its preferred prey in specific areas at a given time, to support decision-makers and mariners make more accurate recommendations when determining restriction zones. We have reviewed the current literature to make our choice on the method of bloom metric extraction from satellite OCR data. We implemented a threshold-based method and applied it on L4 interpolated Chla product distributed by the CMEMS to test the approaches. Preliminary results are promising, but more efforts are underway and needed to improve the method, which will be applied to improved regional PCA-based Chla products. We can already start to work with the modellers for WP3 to test the potential of these new OCR products in the context of the NARW habitat prediction. While our work is ongoing, these preliminary efforts will allow us to get direct feedbacks about specific bloom metrics most appropriate for SDM in WP3.

6. REFERENCES

- Biggs, T.E.G., Brussaard, C.P.D., Evans, C., Venables, H.J. and Pond, D.W. (2020). Plasticity in dormancy behaviour of *Calanoides acutus* in Antarctic coastal waters. *ICES Journal of Marine Science*, 77(5), 1738–1751. doi:10.1093/icesjms/fsaa042.
- Blais, M., Galbraith, P.S., Plourde, S., Devred, E., Clay, S., Lehoux, C. and Devine, L. (2021). Chemical and Biological Oceanographic Conditions in the Estuary and Gulf of St. Lawrence during 2020. Research Document 2021/060. https://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2021/2021_060-eng.html
- Brody, S. R., M. S. Lozier, and J. P. Dunne (2013), A comparison of methods to determine phytoplankton bloom initiation, *J. Geophys. Res. Oceans*, 118, 2345–2357, doi:10.1002/jgrc.20167.
- Cole, H., S. Henson, A. Martin, and A. Yool (2012), Mind the gap: The impact of missing data on the calculation of phytoplankton phenology metrics, *J. Geophys. Res.*, 117, C08030, doi:10.1029/2012JC008249.
- Cushing, D. H. (1990), Plankton production and year-class strength in fish populations: An update of the match/mismatch hypothesis, *Adv. Mar. Biol.*, 26, 249–293.
- Ferreira, A., Brotas, V., Palma, C., Borges, C., & Brito, A. C. (2021). Assessing phytoplankton bloom phenology in upwelling-influenced regions using ocean color remote sensing. *Remote Sensing*, 13(4), 675.
- Fleming, V., and S. Kaitala (2006), Phytoplankton spring bloom intensity index for the Baltic Sea estimated for the years 1992 to 2004, *Hydrobiologia*, 554, doi:10.1007/s10750-005-1006-7.
- Galbraith, P.S., Chassé, J., Nicot, P., Caverhill, C., Gilbert, D., Pettigrew, B., Lefaivre, D., Brickman, D., Devine, L., and Lafleur, C. (2015). Physical Oceanographic Conditions in the Gulf of St. Lawrence in 2014. DFO Can. Sci. Advis. Sec. Res. Doc. 2015/032. v + 82 p
- Gavrilchuk, K., Lesage, V., Fortune, S., Trites, A., & Plourde, S. (2020). A mechanistic approach to predicting suitable foraging habitat for reproductively mature North Atlantic Right Whales in the Gulf of St. Lawrence. DFO Can. Sci. Advis. Sec., Res. Doc.(August), iv + 47 p.
- Gittings, J., Raitos, A., Krokos, D.G. and Hoteit, I. (2018). Impacts of warming on phytoplankton abundance and phenology in typical tropical marine ecosystem *Sci. Rep.*, 8, pp. 1-12, 10.1038/s41598-018-20560-5
- Greve, W., Pringle, S., Zidowitz, H., Nast, J. and Reiners, F. (2005). On the phenology of north sea ichthyoplankton. *ICES Journal of Marine Science*. 62, 1216–1223.
- Gohin, F. (2011). Annual cycles of chlorophyll-a, non-algal suspended particulate matter, and turbidity observed from space and in-situ in coastal waters. *Ocean Science*, 7(5), 705–732. <https://doi.org/10.5194/os-7-705-2011>
- IOCCG (2009). Partition of the Ocean into Ecological Provinces: Role of Ocean-Colour Radiometry. Dowell, M. and Platt, T. (eds.), Reports of the International Ocean-Colour Coordinating Group, No. 9, IOCCG, Dartmouth, Canada
- Khan C, Duley P, Henry AG, Gatzke J, Cole TVN. 2016. North Atlantic Right Whale Sighting Survey (NARWSS) and Right Whale Sighting Advisory System (RWSAS) 2014 results summary. Northeast Fish Sci Cent Ref Doc. 16-01; 12 p. Online at: <https://doi.org/10.7289/V59P2ZNN>

Krug, L. A., Platt, T., Sathyendranath, S. and Barbosa, A.B. (2018). Patterns and drivers of phytoplankton phenology off SW Iberia: a phenoregion based perspective, *Progress in Oceanography* (2018), doi: <https://doi.org/10.1016/j.pocean.2018.06.010>

Maritorena, S. and Siegel, D.A. 2005. Consistent Merging of Satellite Ocean Colour Data Sets Using a Bio-Optical Model. *Remote Sensing of Environment*, 94, 4, 429-440.

O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M. and McClain, C.: Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research: Oceans*, 103(C11), 24937-24953, 1998.

Pendleton, D. E., Holmes, E. E., Redfern, J., & Zhang, J. (2020). Using modelled prey to predict the distribution of a highly mobile marine mammal. *Diversity and Distributions*, 26(11), 1612-1626.

Platt, T., White III, G.N., Zhai, L., Sathyendranath, S., and Roy, S. (2009). The phenology of phytoplankton blooms: ecosystem indicators from remote sensing *Ecol. Model.*, 220, pp. 3057-3069, 10.1016/j.ecolmodel.2008.11.022

Platt, T., and S. Sathyendranath (2008), Ecological indicators for the pelagic zone of the ocean from remote sensing, *Remote Sens. Environ.*, 112(8), 3426–3436.

Record, N. R., Balch, W. M., & Stamieszkin, K. (2019). Century-scale changes in phytoplankton phenology in the Gulf of Maine. *PeerJ*, 7, e6735. <https://doi.org/10.7717/peerj.6735>

Racault, M. F. et al. (2015) Phytoplankton phenology indices in coral reef ecosystems: Application to ocean-color observations in the Red Sea. *Remote Sens. Environ.* 160(222), 234, <https://doi.org/10.1016/j.rse.2015.01.019>

Racault, M., Quéré, C. Le, Buitenhuis, E., Sathyendranath, S., & Platt, T. (2012). Phytoplankton phenology in the global ocean. *Ecological Indicators*, 14(1), 152–163. <https://doi.org/10.1016/j.ecolind.2011.07.010>

Salgado-Hernanz, P., Racault, M.-F., Font-Munõz, J. and Basterretxea, G. (2019). Trends in phytoplankton phenology in the Mediterranean Sea based on ocean-colour remote sensing. *Remote Sens. Environ.*, 221, 50–64.

Siegel, D., S. Doney, and J. Yoder (2002), The North Atlantic spring phytoplankton bloom and Sverdrup's Critical Depth Hypothesis, *Science*, 296 (5568), 730–733, doi:10.1126/science.1069174.

Staudinger, M. D., Mills, K. E., Stamieszkin, K., Record, N. R., Hudak, C. A., Allyn, A., Diamond, A., Friedland, K. D., Golet, W., Henderson, M. E., Hernandez, C. M., Huntington, T. G., Ji, R., Johnson, C. L., Johnson, D. S., Jordaan, A., Kocik, J., Li, Y., Liebman, M., Nichols, O. C., ... Yakola, K. (2019). It's about time: A synthesis of changing phenology in the Gulf of Maine ecosystem. *Fisheries oceanography*, 28(5), 532–566. <https://doi.org/10.1111/fog.12429>

Visser, F., Hartman, K.L, Pierce, G.J., Valavanis, V.D. and Huisman, J. (2011). Timing of migratory baleen whales at the Azores in relation to the North Atlantic spring bloom. *Marine Ecology Progress Series*. 440: 267–279. doi: 10.3354/meps09349

Yoder, J. A., and M. A. Kennelly (2003), Seasonal and ENSO variability in global ocean phytoplankton chlorophyll derived from 4 years of SeaWiFS measurements, *Global Biogeochem. Cycles*, 17(4), 1112, doi:10.1029/2002GB001942.

Xi, H., Losa, S. N., Mangin, A., Soppa, M. A., Garnesson, P., Demaria, J., Liu, Y., d'Andon, O. H. F., & Bracher, A. (2020). Global retrieval of phytoplankton functional types based on empirical orthogonal functions using CMEMS GlobColour merged products and further extension to OLCI data. *Remote Sensing of Environment*, 240(February), 111704. <https://doi.org/10.1016/j.rse.2020.111704>



SIMBA

Phenology Algorithm Theoretical Based Document

smartWhales initiative

Ref.: 9F040-190633/004

ATBD-Pheno

End of the document