

Major concerns

The manuscript led by Jarníková applies a clustering approach to ocean model output from the Salish Sea to evaluate regional differences including which physical factors are associated with phytoplankton growth. The manuscript is clear and well-written and has several interesting findings including a clear partitioning of physical processes and phytoplankton dynamics into different regions. I have a few major concerns and minor comments that should be addressed before I can recommend this manuscript for publication.

>>> R1.1 My first major concern is the lack of comparison phytoplankton model output with literature. Figure 9 shows that diatoms are the dominant phytoplankton group in the summertime. Yet Figures 5 and 7 in a recent observational paper by Del Bel Belluz (2021), whose study years overlap with this manuscript, shows that smaller cells (grouped as dinoflagellates and ciliates in this Jarníková et al.'s manuscript) are in fact usually dominant. In addition, from remote sensing data, Suchy et al. (2019) show that the spring chlorophyll concentrations were anomalously high in 2015 yet these high chlorophyll concentrations were not replicated by the model. I suggest that the authors add a section on model evaluation that compares model and observed phytoplankton functional groups and concentrations in the different regions. This would help solidify the model results.

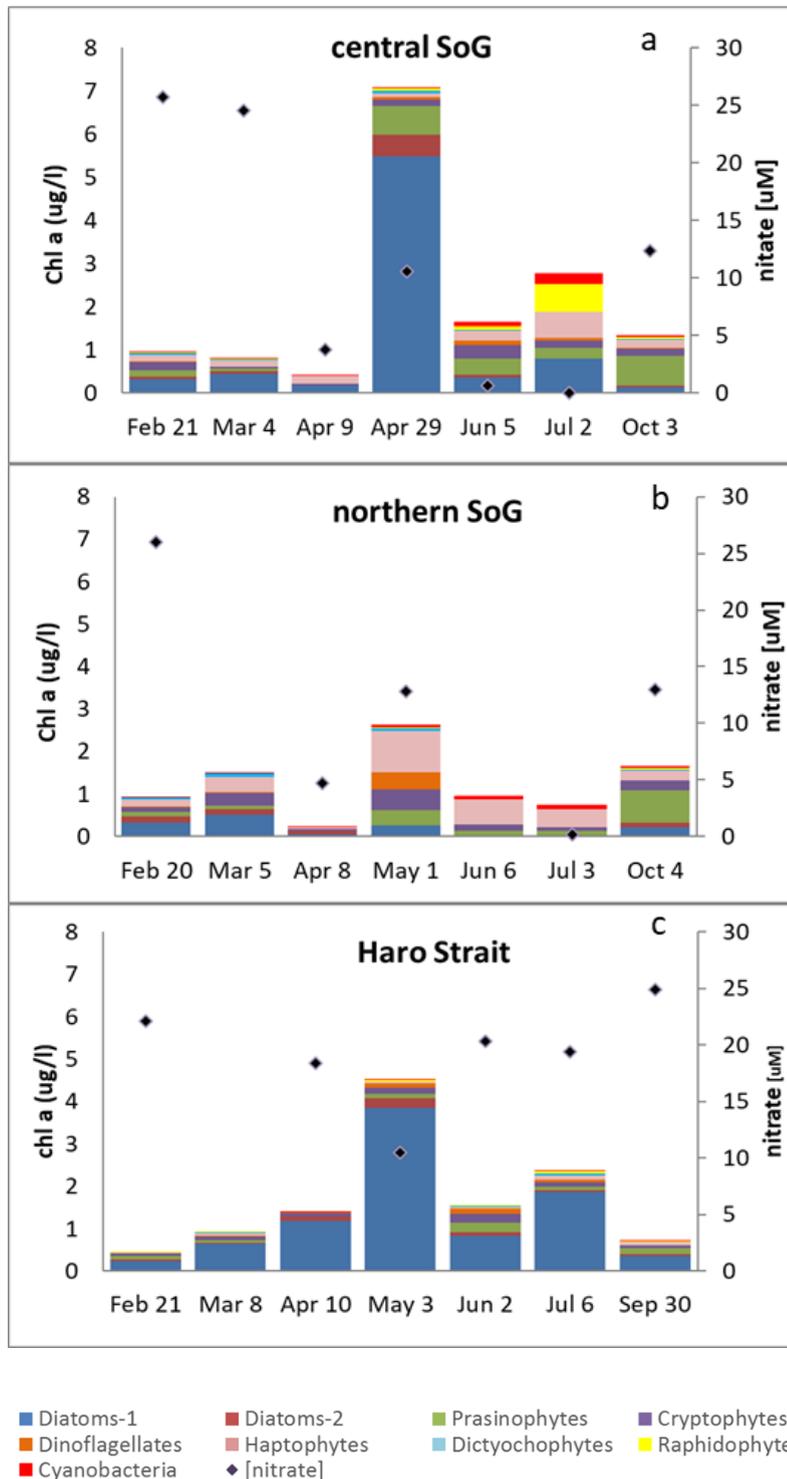
Thank you for this comment. We have clarified that when we discuss flagellates in our paper, we typically mean the small gleaner-type functional groups (for example, cryptophytes), not dinoflagellates, which would more closely align with the diatom class in behaviour. We now note this clarification in the text.

We agree that a functional group evaluation across the three main regions presented here would be an interesting and worthy study, and comparisons to HPLC-derived groups have been carried out but have not yet been published. Unfortunately, these data are not available in a published form at present.

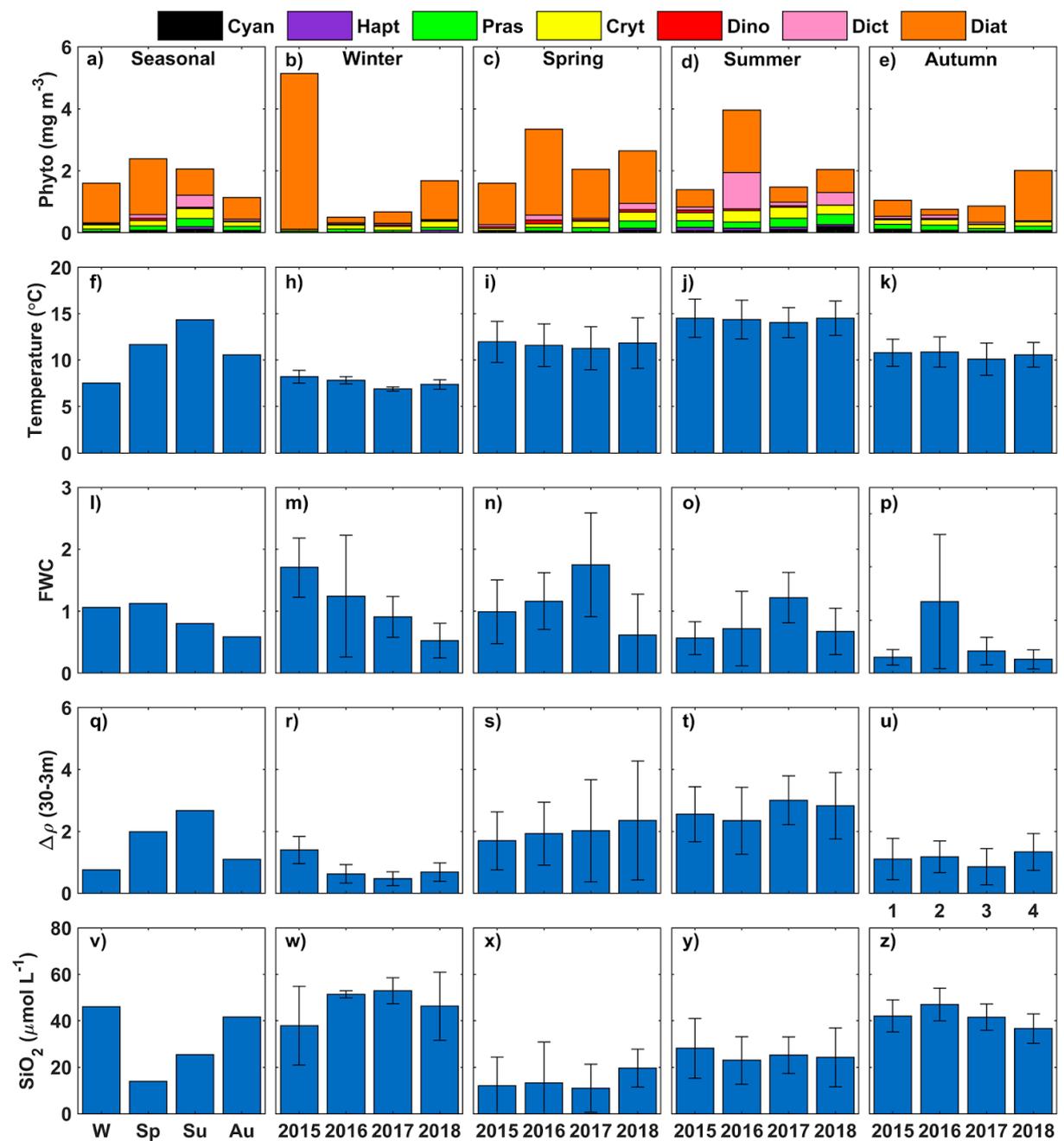
Our main focus is the overall regional patterns and function rather than an operational view. Comparing relatively few spot measurements of biological data, which tend to be patchy, at a single station and depth and finite amount of water, with model output, which in contrast presents the average of a larger volume (the model cell) is difficult at the best of times (e.g. Olson et al. (2020)). Unless there are many data, and well spatially resolved, the comparison is more likely to be anecdotal at best.

To illustrate the difficulty inherent in comparing model and HPLC results, it is worth noting that available HPLC data from the Northern Strait of Georgia do not necessarily agree between observational studies - while HPLC observations by Nemcek et al show almost no diatoms in the northern Strait of Georgia at any point in the year, Del bel Belluz Fig. 7d shows non-negligible presence of diatoms (though typically not diatom

dominance) during the same time of year. The HPLC data of Nemcek et al also show high diatom abundance throughout the observational period in the strongly mixed Haro Strait region, which agrees at first order with our model results.



Nemcek, Nina, Hennekes, M., and I. Perry (2020) Seasonal dynamics of the phytoplankton community in the Salish Sea from HPLC measurements 2015-2019 in State of the physical, biological, and selected fishery resources of Pacific Canadian marine ecosystems in 2019. (Eds.)



Del Bel Belluz, Justin, et al. "Phytoplankton Composition and Environmental Drivers in the Northern Strait of Georgia (Salish Sea), British Columbia, Canada." *Estuaries and Coasts* 44.5 (2021): 1419-1439. Fig. 7

Similarly, because Suchy et al use remote sensing data that senses the surface, i.e., top several millimeters, of the ocean, thus representing at best the top several meters of the water column, we cannot directly compare their chlorophyll observations. However, encouragingly, at dates, depths, and times where observations are available, our

biological model evaluation against nitrate, dissolved silica, and log-transformed chlorophyll performs quite well (Willmott Skill Scores ≥ 0.7 for all three metrics), giving us confidence that the model is reproducing large scale primary productivity patterns seen in in-situ observation.

In the new manuscript, we will add a brief, necessarily qualitative, discussion of the comparison of the models results to the observations from Del bel Belluz (2021), Suchy et al (2019), and the Nemcek et al HPLC data (published as a technical report in the Canadian Technical Report of Fisheries and Aquatic Sciences).

Amended text (line 103 in original manuscript):

Small flagellates (**representing functional groups such as cryptophytes**) have the lowest maximum growth rate while competing better at low nitrogen levels, low light, and higher temperature.

>>> R1.2 My second major concern is the interpretation presented in section 4.1, which suggests the difference in phytoplankton biomass between the central and northern Strait of Georgia is linked to strong summer wind events in the northern Strait of Georgia (lines 320 to 323). When I look at Figure 5, I see that summer winds are in fact stronger in the central Strait of Georgia so I don't think that the explanation holds. Could the authors please clarify how they came to this result?

We believe that the difference in phytoplankton functional groups between the Central and Northern Strait of Georgia is not because of the differences in wind (we note in the caption of Fig. 5 that summer winds are stronger in the CSoG, as you point out) but because in the NSoG, the comparatively weaker stratification allows for more upwelling and mixing of nutrients into the surface and the resulting summer opportunist blooms, even though the winds themselves are weaker. We have now added a sentence clarifying this interplay between wind and stratification - explaining that the difference is not due to the winds themselves.

Amended text (original manuscript, line 320, insertion in **bold**):

In the NSoG, nutrient drawdown also occurs, but episodic wind events lead to stronger upwelling and mixing due to the comparatively weaker stratification and inject sharp pulses of nutrients into the near-surface, leading to sharp, short-lived diatom blooms ([\cite{moore2022wind}](#)). **In contrast, in the CSoG, despite stronger summer winds, strong stratification limits nutrient injection into the upper layer and gleaner-type organisms continue to be favoured.**

>>> R1.3 My third major concern is that line 108 states that SalishSeaCast has been operational since 2014 and runs daily forecasts and nowcasts. Could the authors please explain then how this study analyzes data starting in 2013?

This confusion arises from my wording. The model first began to be run operationally (that is, running forecasts and nowcasts every day, similarly to a weather forecast) in 2014, but a hindcasted run is now available reaching back to 2013. Here we simply chose four sequential recent years. We've amended the text slightly to reflect this organization, and also updated the link to point to the ERDDAP server that contains this full hindcast.

Original text: (original manuscript, line 108):

SalishSeaCast has been operational since 2014 and is run daily with forecasts and nowcasts available at [\url{https://salishsea.eos.ubc.ca/nemo/results/}](https://salishsea.eos.ubc.ca/nemo/results/)

Original text: (original manuscript, line 117):

We analyzed four years of daily output from SalishSeaCast (2013-2016), using an unsupervised clustering algorithm (Ward's Euclidean Distance Method, see section 2.3).

Amended text:

SalishSeaCast has been run operationally since 2014, and results from 2013 to 2020 are available at [\url{https://salishsea.eos.ubc.ca/erddap/index.html}](https://salishsea.eos.ubc.ca/erddap/index.html)"

"We analyzed four years of daily output from a hindcast of SalishSeaCast (2013-2016), using an unsupervised clustering algorithm (Ward's Euclidean Distance Method, see section 2.3).

Minor concerns

>>> R1.4

- Line 2 – for clarity, I suggest that you list the four factors here

Thank you, that's a good idea. We've added them.

Original text: (original manuscript, line 2):

Here, we apply a hierarchical clustering algorithm (Ward's method) to four factors relating to stratification and depth-integrated phytoplankton biomass extracted from a biophysical regional ocean model of the Salish Sea to assess spatial co-occurrence.

Revised text:

Here, we apply a hierarchical clustering algorithm (Ward's method) to four factors relating to stratification (wind energy, freshwater index, watercolumn-averaged vertical eddy diffusivity, and halocline depth), as well as to depth-integrated

phytoplankton biomass, extracted from a biophysical ocean model of the Salish Sea. We then assess these factors for spatial co-occurrence.

>>> R1.5

- Lines 15 to 31 – I found this section difficult to read, with long sentences and no clear points. I suggest that this section is rewritten more concisely, with points related to this study made more clear
-

Thanks for pointing this out. We've rewritten this section, shortening the sentences and paragraphs restructuring for clarity.

Original text: (original manuscript, line 15)

Marine phytoplankton form the basis of the oceanographic food web and are responsible for approximately half of global carbon fixation (Field 1998). To predict changes in global climate it is necessary to understand the complexities in the underlying controls of marine productivity. The rate of primary productivity in the near-surface ocean is controlled by the availability of macro- and micro-nutrients and light, as well as temperature, which are in turn controlled by the interplay of stratifying processes and sources of mixing. The breakdown of the surface ocean stratified layer may permit mixing to reduce the availability of light for phytoplankton, inhibiting growth (e.g. Sverdrup 1953), or contrastingly bring nutrients from deeper waters to nutrient-depleted surface waters, thus stimulating growth. The breakdown of stratification may also inhibit the predatory action of zooplankton on phytoplankton by dilution, with net positive effects on phytoplankton growth (Behrenfeld 2010).

The interplay of different stratification regimes exerts control on the structure of ocean ecosystems (e.g. Legendre 1981), and changes in regime have been linked to shifts in phytoplankton community composition (e.g. Huisman 2004). The importance of phytoplankton in biogeochemical cycling, along with the position of phytoplankton at the base of the food web and impact on higher trophic levels, globally motivates the study of phytoplankton distribution and dynamics. Coastal regions are disproportionately more productive compared to the open ocean (Longhurst 1995) and typically have more complex mixing, circulation, and stratification dynamics than the open ocean, making resolution of phytoplankton biomass patterns difficult. Simultaneously, because both ocean stratification patterns and phytoplankton biomass dynamics may be expected to shift under anthropogenic climate change, there exists a need to characterize their dynamic structure and identify key drivers.

New text:

Marine phytoplankton form the basis of the oceanographic food web and are responsible for approximately half of global carbon fixation (Field 1998). To predict changes in global ecosystem functioning, it is necessary to understand the underlying controls on marine productivity. Primary productivity in the near-surface ocean is controlled by the availability of macro- and micro-nutrients and light, as well as temperature, which are in turn controlled by the interplay of stratifying processes and sources of mixing.

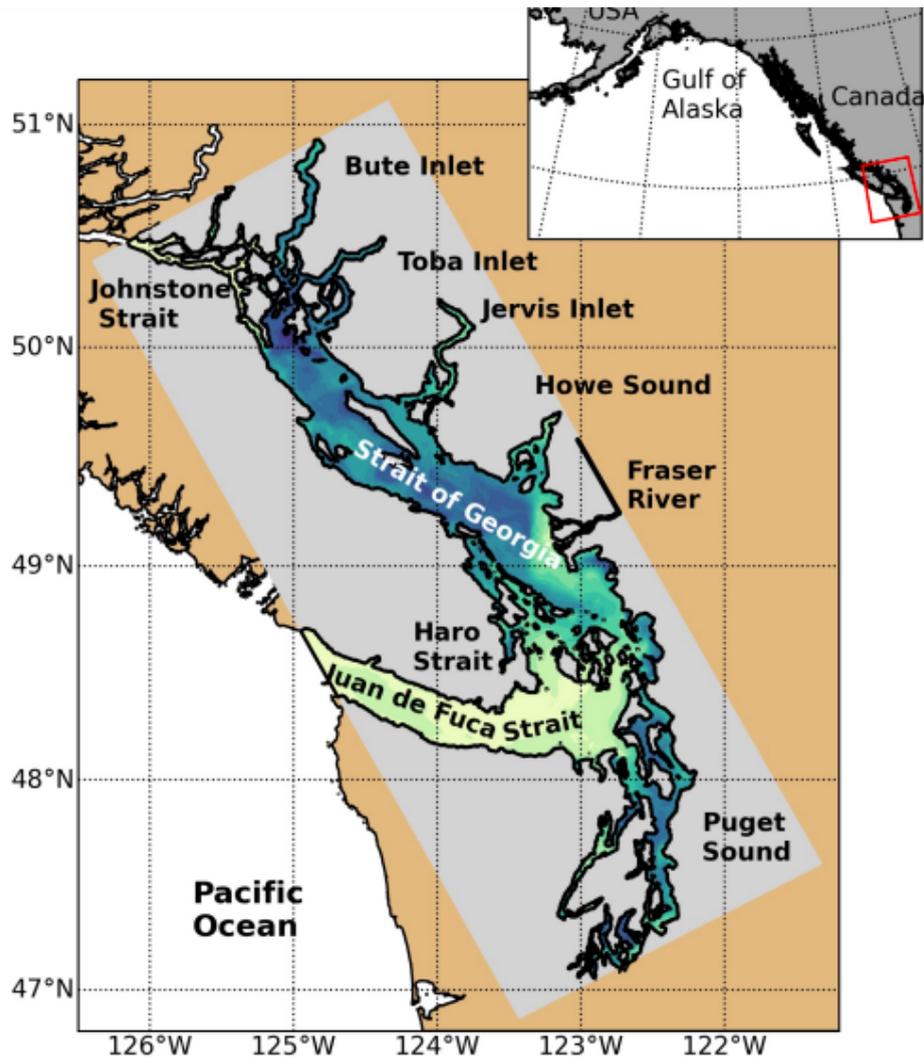
The breakdown of the surface ocean stratified layer may reduce the availability of light for phytoplankton, inhibiting growth (e.g. Sverdrup 1953), or contrastingly bring nutrients from deeper waters to nutrient-depleted surface waters, thus stimulating growth. The interplay of different stratification regimes exerts control on the structure of ocean ecosystems (e.g. Legendre 1981), and changes in regime have been linked to shifts in phytoplankton community composition (e.g. Huisman 2004).

The importance of phytoplankton in biogeochemical cycling, as well as their position at the base of the food web and impact on higher trophic levels, globally motivates the study of phytoplankton distribution and dynamics. Coastal regions are more productive than the open ocean (Longhurst 1995). Simultaneously, these regions typically have more complex mixing, circulation, and stratification dynamics than the open ocean, making resolution of phytoplankton biomass patterns difficult. Finally, because both ocean stratification patterns and phytoplankton biomass dynamics may be expected to shift under anthropogenic climate change, there exists a need to characterize their dynamic structure and identify key drivers.

>>> R1.6

- Figure 1 – I suggest that somehow the different regions (NSoG, CSoG and JdF) are labeled on the map. I know that the point of clustering was to help define these regions but it is difficult for people not familiar with this region to know what area the authors are referring to

After discussion, we have decided not to subdivide the Strait of Georgia into North and South on the map, because we feel these labels would clutter the map. The Juan de Fuca Strait is labeled, as is the Strait of Georgia as a whole, and we have latitude lines to orient the viewer. We also point out the North and Central subdivision in the caption: “The Strait of Georgia is often subdivided into the Central Strait of Georgia (CSoG) and Northern Strait of Georgia (NSoG).” We include the figure in question below, as a reference.



>>> R1.7

- Lines 42 to 47 – The link between phytoplankton and fish is still tenuous. I suggest that the authors add a clearer ecological motivation for this study or add some more information about why phytoplankton is important for higher trophic animals.

We respectfully disagree that the link between phytoplankton and fish is tenuous; this link has been studied extensively by ecologists and biological oceanographers. We list some examples below, and add a line about this context to the new manuscript.

For example, Platt et al. (2003) showed that larval fish survival is dependent on the timing of the spring phytoplankton bloom. Nearer to our ocean region, Malick et al. (2015) found that spring bloom timing correlated significantly with pink salmon productivity in British Columbia and Alaska populations. The link between spring bloom timing and fish population health has also been extensively studied. Chittenden et al. (2010) investigated the impact of release time on the survival and migration of both wild and hatchery-reared Coho salmon populations (*Oncorhynchus kisutch*) in relation to

both phytoplankton and zooplankton blooms. Finally, in recent years, Boldt et al. (2019) used the timing of the spring phytoplankton bloom and the peak availability of zooplankton prey to assess age-0 herring abundance and condition in the Strait of Georgia.

Boldt, Jennifer L., et al. "Bottom-up and top-down control of small pelagic forage fish: factors affecting age-0 herring in the Strait of Georgia, British Columbia." *Marine Ecology Progress Series* 617 (2019): 53-66.

Chittenden, Cedar M., et al. "Genetic versus rearing-environment effects on phenotype: hatchery and natural rearing effects on hatchery-and wild-born coho salmon." *PLoS One* 5.8 (2010): e12261.

Malick, Michael J., et al. "Linking phytoplankton phenology to salmon productivity along a north-south gradient in the Northeast Pacific Ocean." *Canadian Journal of Fisheries and Aquatic Sciences* 72.5 (2015): 697-708.

Platt, Trevor, Cesar Fuentes-Yaco, and Kenneth T. Frank. "Spring algal bloom and larval fish survival." *Nature* 423.6938 (2003): 398-399.

Amended text:

This coastal ocean is a region of significant ecological and cultural importance and provides habitat to many species of ecologically important megafauna, including the Southern Resident killer whales (*Orcinus orca*) and the local salmon populations. The ongoing significant decline of the local Coho and Chinook salmon (Preikshot et al. (2013)) has been implicated as a factor in the low reproductive success of the killer whale populations (Wasser et al. (2017)), which depend on these salmon as a food source. **The health of fish populations in the Pacific Northwest has been linked to spring bloom timing and phytoplankton abundance (e.g. Malick et al. (2015); Boldt et al. (2019)).** Thus, potential population declines in upper trophic levels further motivate the understanding of factors controlling the base of the food web.

>>> R1.8

- Lines 74 to 79 – This is where the question that the manuscript is trying to answer is defined yet I don't see a clear scientific question or objective here.

We see what you mean - the paragraph focuses on the procedural aspect of "what we did", as opposed to the "why we did it", and the clear statement of objective is lacking. We've added a general overview sentence to this last paragraph, as follows: **"Here our main goal is to investigate how physical dynamics in the Salish Sea objectively define regions of distinct of phytoplankton biomass and functional group composition."**

Original text: (original manuscript, line 74):

Here we extract model-available proxies for four separate factors related to watercolumn stratification: wind energy, freshwater index, watercolumn-averaged vertical eddy diffusivity, and halocline depth, and one indicator of primary

productivity (depth-integrated phytoplankton biomass separated by functional group). We then use a clustering analysis to discuss the three major regions of the Salish Sea in the context of the spatial patterns in the yearly signals of these factors, as well as to consider their interannual variability. We finally compare spatial patterns in stratification factors to spatial patterns in phytoplankton biomass and discuss possible linkages between the two.

New text:

Here our main goal is to investigate how physical dynamics in the Salish Sea objectively define regions of distinct of phytoplankton biomass and functional group composition. We extract model-available proxies for four separate factors related to watercolumn stratification: wind energy, freshwater index, watercolumn-averaged vertical eddy diffusivity, and halocline depth, and one indicator of primary productivity (depth-integrated phytoplankton biomass separated by functional group). We then **cluster each factor individually in order** to discuss the three major regions of the Salish Sea in the context of the spatial patterns in the yearly signals of these factors, as well as to consider their interannual variability. We finally compare spatial patterns in stratification factors to spatial patterns in phytoplankton biomass and discuss possible linkages between the two.

>>> R1.9

- Lines 93 to 94 – This wind data spatial resolution is too coarse to properly reproduce winds in most of the BC fjords. I suggest adding a sentence here that acknowledges that limitation yet explains why the model output from fjords is still accurate.

We agree that the wind product is too coarse to reproduce the local fjord wind patterns, and that this insufficient resolution is a limitation. I've added a sentence mentioning this limitation, and also mentioning that the HRDPS product performs well in the more open parts of the Salish Sea, as seen in Moore-Maley and Allen (2022). As the model evaluation doesn't focus on fjords (Fig. A1, A2), we don't discuss model accuracy in the fjords.

Original text: (original manuscript, line 92)

Atmospheric forcing, including winds and solar radiation, is derived from the High Resolution Deterministic Prediction System (HRDPS), a nested 2.5 km resolution operational atmospheric model (citet{milbrandt2016pan}).

New text:

Atmospheric forcing, including winds and solar radiation, is derived from the High Resolution Deterministic Prediction System (HRDPS), a nested 2.5 km

resolution operational atmospheric model (\cite{milbrandt2016pan}). The HRDPS model output is too coarse to accurately resolve atmospheric conditions in the northern inlets, but its wind fields have shown good agreement with observations throughout the Strait of Georgia (\cite{moore2021wind}).

>>> R1.10

- Line 126 – I don't know how a cluster is characteristic. Was there a way that clusters were quantified to decide which 4 to choose?

We see what you mean - we have rephrased this sentence, because the term 'characteristic' is very hand-wavy. We have also added a pointer to section 2.3, where we describe the cluster number selection. We include it in our response below. We also include the paper text about cluster number selection, for reference.

Original text: (original manuscript, line 125)

Several possible clusterings resulting from our analysis were visualized and compared for major differences. The most characteristic clusters are presented in the Results section, while an example visualization of all possible clusterings of one year of one of the variables is available in Fig. C1.

New text:

Several possible clusterings resulting from our analysis were visualized and compared for major differences (see Section 2.3). Results show the typical cluster structure for all four years for each individual factor (Section 3), while an example visualization of all possible clusterings of one year of one of the variables is available in Fig. C1. Here, we describe the signals.

Paper text on cluster number selection (un-numbered subsection of Section 2.3), for reference:

Cluster number selection

A common challenge in the application of clustering methods is the selection of cluster number. In hierarchical clustering, typical approaches include choosing a cutoff where the change in dendrogram distance between two cluster numbers is maximized. In our case, attempts to use objective metrics to determine cluster number, such as the Davies-Bouldin, Silhouette, or Calinski-Harabasz criteria, typically identified only two clusters in the a given dataset (not shown); though these may be the most prominent clusters, meaningful structure in the data persists at larger cluster numbers. Ultimately, our approach was to visualize several possible clustering outputs, with cluster number N varying from 2 to 15, and to visually compare emergent structure in multiple years of the clustering of a single variable (e.g. Fig. C1). In all variables, the same typical structure emerged at a relatively low cluster number (eg, N = 3-5) and persisted with increasing cluster number in all years. To facilitate comparison of clusters between years, we chose an N=5 for all years for all clusters, and are confident that the structures described are robust to a selection of a variety of cluster numbers.

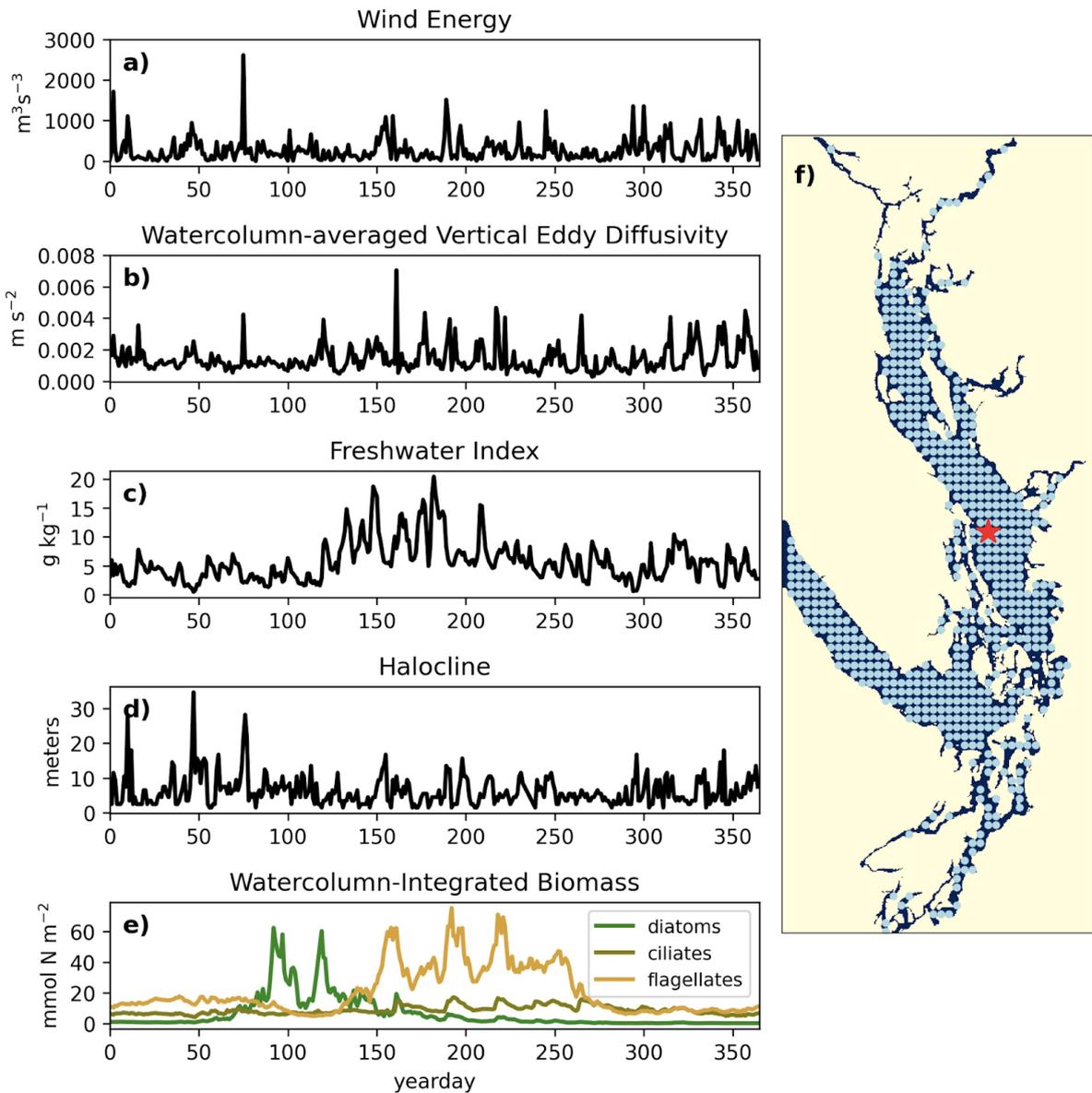
>>> R1.11

- Figure 2 – It took me a while to figure out what the vertical axes labels meant. I suggest that you make these clearer (i.e. add the description of the variables) to help the reader.

Though we understand this point, because we want to maximize the detail available in the signals themselves, we've chosen to leave the figure as is, with units on the y-axis and variables specified in the figure titles. For clarity, we have added information in the caption, as follows:

Figure 2 caption, amended:

Example yearly signals of clustered physical and biological factors from one station in the CSoG (red star). **The physical signals are as follows: a) wind energy (m^3s^{-3}), b) water column-averaged vertical eddy diffusivity (m s^{-2}), c) freshwater index (g kg^{-1}), d) halocline (m). The biological signal is water column-integrated phytoplankton biomass (mmol N m^{-2}), separated by functional group (diatoms, ciliates, and flagellates).** The remaining 570 stations used in the clustering are shown as blue points. Depth-integrated phytoplankton biomass signals are combined in series for clustering (see Figure 8).



>>> R1.12

- Lines 133 to 134 – I don't understand this sentence. Please clarify.

Because we are interested in the impact of wind on stratification, we extract the wind energy available for mixing from the model, which scales with the windspeed cubed. Because we have wind speed at hourly resolution, we cube it at hourly resolution and then take a daily average, in order to take advantage of this higher resolution product. We have rewritten this text to be more clear and also more detailed.

Original text: (original manuscript, line 133)

As wind energy available for mixing scales with the cube of wind speed, we cluster on the daily average of cubed hourly wind speed (Fig. 2a).

Amended text:

Here we are interested in the impact of wind on mixing of the water column. Therefore, because wind energy available for mixing scales with the cube of wind speed (Fischer 1979), we use the cube of wind speed as our signal here. To take advantage of the hourly resolution of the wind product, we use the daily average of cubed hourly wind speed (Fig. 2a).

>>> R1.13

- Lines 158 to 159 – I find this sentence confusing because I don't know what three yearly signals means.

We have rewritten this sentence to specify what we mean by these signals:

Original text: (original manuscript, line 158)

We extract daily-average depth-integrated phytoplankton biomass (mmol N m^{-2}) for each of the three phytoplankton functional groups to form three yearly signals (Fig. 2e).

Amended text:

We extract daily-average depth-integrated phytoplankton biomass (mmol N m^{-2}) for each of the three phytoplankton functional groups to form three signals (i.e. one year-long daily-resolution time series of depth-integrated phytoplankton biomass for each of the three phytoplankton functional groups, Fig. 2e).

>>> R1.14

- Lines 163 to 164 – How do the authors know that bottom-up effects dominate? Please explain.

Thank you for this question - because we haven't quantified or formalized the relative contribution of bottom-up and top-down effects on the standing stock of phytoplankton biomass, we have chosen to remove this sentence.

Original (removed) sentence: (original manuscript, line 163)

While factors other than primary productivity that influence the standing stock biomass (e.g. mortality, grazing, and sinking rate) are reflected in this clustering, bottom-up effects dominate the response seen here.

>>> R1.15

- Lines 176 to 177 – Why was the clustering applied to 4 years separately instead of applying clustering to the whole 4 year time series?

By applying the clustering to four years separately, we are able to see if the spatial patterns in the different variables persist in all four years (broadly, they do, which is interesting in its own right), and to comment on differences between years. We state this goal in the introduction to the paper (original manuscript, line 76).

Original manuscript, line 76:

We then use a clustering analysis to discuss the three major regions of the Salish Sea in the context of the spatial patterns in the yearly signals of these factors, as well as to consider their interannual variability.

>>> R1.16

- Lines 179 to 198 – As a non-clustering expert, I found these sections difficult to understand. I suggest that the authors simplify these sections so that they are geared towards non-experts.

We have changed the wording here to be more descriptive for a broader audience. We anticipate that the dendrogram difference is likely confusing to people who don't use clustering.

Original text (line 179):

A common challenge in the application of clustering methods is the selection of cluster number. In hierarchical clustering, typical approaches include choosing a cutoff where the change in dendrogram distance between two cluster numbers is maximized. In our case, attempts to use objective metrics to determine cluster number, such as the Davies-Bouldin, Silhouette, or Calinski-Harabasz criteria, typically identified only two clusters in the a given dataset (not shown); though these may be the most prominent clusters, meaningful structure in the data persists at larger cluster numbers. Ultimately, our approach was to visualize several possible clustering outputs, with cluster number N varying from 2 to 15, and to visually compare emergent structure in multiple years of the clustering of a

single variable (e.g. Fig. C1). In all variables, the same typical structure emerged at a relatively low cluster number (eg, $N = 3-5$) and persisted with increasing cluster number in all years. For easy intercomparability of clusters between years, we thus chose an $N=5$ for all years for all clusters, but are confident that the structures described are robust to a selection of a variety of cluster numbers.

New text:

A common challenge in the application of clustering methods is the selection of cluster number, as the clustering algorithm can produce anywhere between 2 and N clusters (where N is the number of signals being clustered). Typical approaches include choosing a cluster number where the difference in the mean signals of the found clusters, when going from cluster number N to cluster number $N+1$, is maximized. In our case, attempts to use objective metrics to determine cluster number, such as the Davies-Bouldin, Silhouette, or Calinski-Harabasz criteria (Maulik et al. (2002)), typically identified only two clusters in a given dataset (not shown). Though these may be the most prominent clusters, meaningful structure in the data persists at larger cluster numbers. Ultimately, our approach was to visualize several possible clustering outputs, with cluster number N varying from 2 to 15, and to visually compare how the spatial structure of the patterns changed with increasing cluster number (e.g. Fig. C1). In all variables, the same typical structure emerged at a relatively low cluster number (eg, $N = 3-5$) and persisted with increasing cluster number in all years. For easy intercomparability of clusters between years, we thus chose an $N=5$ for all years for all clusters, but are confident that the structures described are robust to a selection of a variety of cluster numbers.

>>> R1.17

- Figure 7 – It is hard to see the red dots

We appreciate that it is not ideal to have these dots as so small, but unfortunately, making them larger would send a 'highlighting' message that we want to avoid, and obscure other parts of the clustering. The exact location of the red dots is not a focus here: to characterize this specific cluster, all SalishSeaCast results, rather than every 10th point would need to be analyzed.

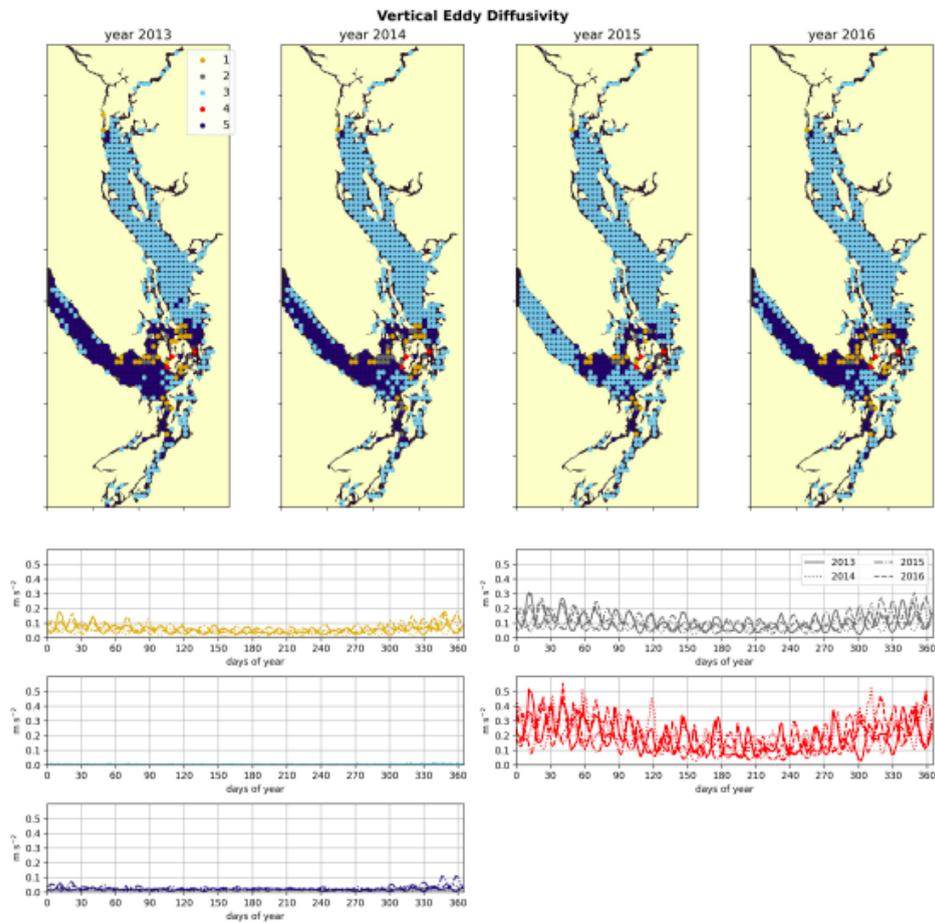


Figure 7. Clustering of the daily depth-averaged vertical eddy diffusivity signal. The domain is split into two major regions: the Strait

>>> R1.18

- Lines 233 to 234 – What is the source of the freshwater runoff data?

The source of the data is the measured Fraser river flow at Hope, British Columbia, from Environment and Climate Change Canada. We now include this reference in the caption of the accompanying Supplementary Figure C1:

Figure C1 caption:

Fraser river flow at Hope, British Columbia for the four modelled years, as implemented in Soontiens and Allen (2017). Data from Environment and Climate Change Canada (https://wateroffice.ec.gc.ca/report/real_time_e.html?stn=08MF005, accessed June 2021).

>>> R1.19

- Line 244 – I suggest that the authors define VED

'Thanks for pointing out this oversight - we now define the acronym VED to vertical eddy diffusivity in the section where we first introduce it:

Amended text: (original manuscript, line 136)

The vertical eddy diffusivity (**VED**) represents the strength of mixing in the system (Soontiens and Allen (2017)) and depends on the choice of vertical turbulence closure scheme. SalishSeaCast uses a k- ϵ configuration of a generic length scale turbulence model to estimate sub-gridscale turbulent processes (Umlauf and Burchard (2003)), with background vertical eddy viscosity and diffusivity both set to $10^{-6} \text{ m}^2 \text{ s}^{-1}$. We report a daily depth-averaged value here (Fig. 2b). Though average vertical eddy diffusivity reflects all sources of mixing and stratification present in the system, it is dominated by barotropic tidal activity, and we expect it to be highest at tidal mixing hotspots (Crean (1978)).

>>> R1.20

- Line 262 – I think it would be good to mention here that most of the Discovery Islands, where the highest tidal energy is, isn't resolved in this model

Thank you! We have added this mention.

Original text: (original manuscript, line 261)

In the highly-variable Haro Strait and Johnstone Strait regions, the spatial frequency of our sampling likely plays a role in our derived map of tidal mixing hotspots - as we sample only approximately every 100th horizontal model coordinate, we likely miss other high-VED model points in this subregion. Analysis of tidal mixing hotspots is not the focus of this work, but a full characterization of this tidally-mixed zone using a more refined clustering approach may be an interesting focus of future work.

Amended text:

In the highly-variable Haro Strait and Johnstone Strait regions, the spatial frequency of our sampling likely plays a role in our derived map of tidal mixing hotspots - as we sample only approximately every 100th horizontal model coordinate, we likely miss other high-VED model points in this subregion, especially channels that have width-scales comparable to our model resolution (0.5km), for example the intricate channel passages of the San Juan and Discovery Island groups in the Johnstone Strait region, which are known tidal

mixing hotspots (Fig. 1). Analysis of tidal mixing hotspots is not the focus of this work, but a full characterization of this tidally-mixed zone using a more refined clustering approach may be an interesting focus of future work.

>>> R1.21

- Line 269 – I'm not sure that you can classify the spring bloom in CSoG as early – based on Figure 8, it looks to me like the timing of the spring bloom is similar in all regions.

We have rephrased this text - we merely meant that diatoms bloom first compared to other functional groups, not that the bloom was especially early or late.

Old text (original manuscript, line 268):

The largest cluster (the CSoG - cluster 3/sky blue) is characterized by an early diatom bloom, followed by a transition to flagellate abundance in the summer months.

Amended text:

The largest cluster (the CSoG - cluster 3/sky blue) is characterized by diatoms blooming first, followed by a transition to flagellate abundance in the summer months.

>>> R1.22

- Figure 10 – With the exception of 2015, it is difficult to see that the spring bloom started earliest in 2015. I suggest that the authors change the color scheme to make this clearer.

We agree that it is difficult to see the distinction in the bloom timing, and we have tried to fix this with a different colour scheme, as suggested.

Original figure:

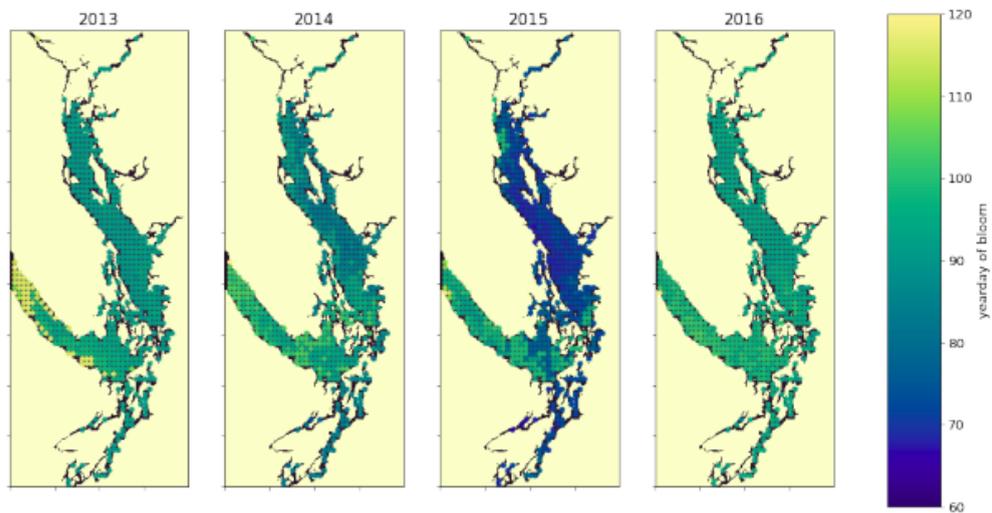
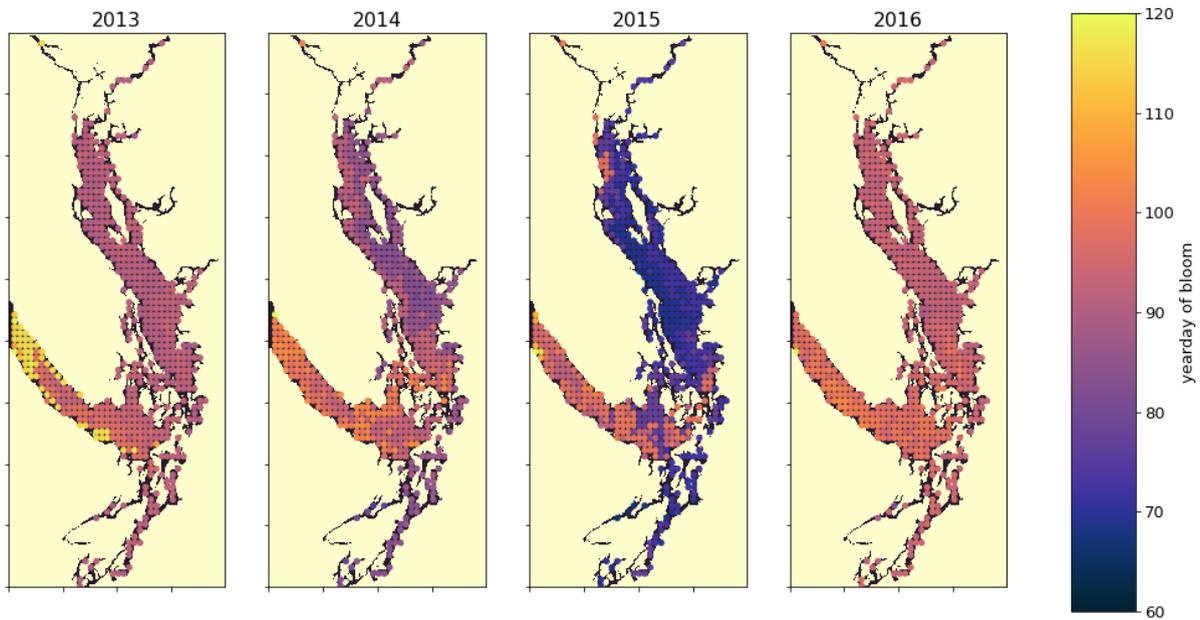


Figure 10. A spatial view of the onset of the spring bloom in the domain. Here the spring bloom is defined as the first peak in depth-integrated diatom biomass that is at least 30% of the maximum annual diatom biomass at that station. In all years, the spring bloom occurred earliest in the CSoG and subsequently in the NSoG before reaching JdF with a variable delay.

New figure:



>>> R1.23

- Lines 379 to 380 – I don't see this in Figures 8 or 10 that the spring bloom started in CSoG and radiated outwards

We agree that the difference in bloom timing between the CSoG and NSoG is quite subtle in the model, so we have amended the text to make clear that the main difference in bloom timing is between the SoG and JdF.

Text:

In all years, the spring bloom started in the CSoG and radiated outward toward to NSoG before eventually reaching JdF. This earlier bloom initiation in the CSoG than in the NSoG was also seen in multiple years of satellite observations of the region (Suchy et al. (2019)). The late bloom timing in JdF was likely driven by stronger mixing limiting light availability later into the year in JdF region (Fig. 10), consistent with the functional differences between JdF and the NSoG and CSoG discussed above.

Amended text:

Earlier spring bloom initiation in the CSoG with respect to the NSoG was seen in multiple years of satellite observations (Suchy et al. (2019)). In our results this progression within the SoG is almost indistinguishable and is followed by later blooming in the JdF. The late bloom timing in JdF was likely driven by stronger mixing limiting light availability later into the year in JdF region (Fig. 10), consistent with the functional differences between JdF and the NSoG and CSoG discussed above.

>>> R1.24

- Figure A1 – In cluster 5, which is labeled as JdF, why are there dots from the NSoG?

We did our evaluation based on the biological clusters, which broadly correspond to the geographic regions of the Northern Strait of Georgia, the Central Strait of Georgia, and the Juan de Fuca Strait. However, in several years, Bute Inlet clusters with the Juan de Fuca Strait, and temperature and salinity data are available in that region. To avoid 'picking and choosing' how we did our model evaluations, we kept these data in. We've added a note to that effect in the caption.

Amended caption for Figure A1, with additions in **bold**:

Model comparison with DFO CTD temperature and salinity data. The plots show modeled vs observed values for salinity and temperature for the entire model domain, as well as points matched only to the three major biological clusters - **the Northern Strait of Georgia, the Central Strait of Georgia, and the Juan de Fuca Strait** (cluster boundaries are specific to the year of observation). (**Note that in some years, Bute Inlet clusters with the Juan de Fuca Strait**). Because of the large amount of data available for comparison, a histogram view

is presented. The timeline and rightmost panel display observation times and locations. Summary statistics corresponding to these plots are shown in Table A1.

>>> R1.25

- Figure A2 – The caption says salinity but I think the authors mean silicate? Also, what are the three biological clusters mentioned in the caption?

Thank you for catching this typo! We have also included the names of the three biological clusters in the caption.

Amended text (caption, Figure A2):

Model comparison with DFO nitrate, **dissolved silica** and log-transformed chlorophyll data. The plots show modeled vs observed values for nitrate, salinity and log-transformed chlorophyll for the entire model domain, as well as points matched only to the three major biological clusters - **the Northern Strait of Georgia, the Central Strait of Georgia, and the Juan de Fuca Strait** (cluster boundaries are specific to the year of observation).