Comment on essd-2021-329
Anonymous Referee #1

This well-written data paper collates and integrates several commonly used open access water quality data sets (time series), including those from Fisheries and Oceans Canada (now IISD) Experimental Lakes Area, the US EPA, and the US Northern Lakes LTER Program, as well as substantial datasets from Alberta (Canada), England (UK), and Sweden. The authors summarize some, but not all, of the main features of the data, and provide some illustrative relationships among in situ and external environmental parameters designed to stimulate further analysis. One key feature of the data is the calculation of daily Chl a “growth rates” (really, rate of change) that the authors use as a surrogate for phytoplankton population change. By identifying periods of growth > 0.5 ug Chl a/L/day, the authors examine how patterns of growth vary as a function of lake
trophic status, irradiance, and limnological parameters both before and during the windows of growth. This analysis is both the strength and weakness of the paper, as the protocol does not appear to be used a prior peer-reviewed methods or application paper and is insufficiently validated in the current submission to allow its reliability or utility to be determined. Basically, the paper appears half-way between a simple data paper and a methods paper, and so is not really a good fit to either category at the moment. Key comments are outlined below, arranged according to the line number of the paper.

Details

- Terminology, l 49 and elsewhere. I think the authors would be better served to use less jargon and be more precise in their description of regulatory processes. Phrases such as “Bottom-up” and “top down” appear derived from older food-web literature, and are used in a non-standard manner in the current paper. Here, ‘bottom-up’ appear to mean physical and chemical controls (physico-chemical), whereas ‘top-down’ seem to refer to grazing by zooplankton. I suggest they use these more precise description, as the current terms lack context when only one trophic level (phytoplankton) is actually described (as Chl a).

Similarly, this appears to be a paper about the rate of change of Chl a standing stock/crop, not production (an amount) or productivity (a rate) per se. Its ok to note that changes in Chl are used to infer these other processes, but try and focus this paper on Chl to avoid over-extrapolation of the analyses (see below). This gets to be an issue when discussing the ‘two phases’ of change in Chl, separated by what is commonly known as the clearwater phase (CWP). The latter occurs in productive (but rarely oligotrophic) lakes and arises when thermal stratification coincides with development of populations of large-bodies zooplankton (often Daphnia) to greatly reduce spring biomass maxima due to sedimentation (mainly diatoms) and herbivory (usually diatoms, flagellates). This decline in Chl does not necessarily correspond to declines in the rate of production (productivity), as available nutrients are usually elevated (because of recycling of biomass via homeostatic herbivores). Jim Elser and Bob Sterner have done a lot of work on this (see “ecological stoichiometry”). The main point is that rates of change in Chl do no necessarily correspond to changes in primary productivity, particularly through a typical annual phenology.
Introduction and elsewhere – literature cited. I suggest that the authors carefully review their citations, as I noted several instances when inappropriate references (e.g., marine, satellite) were used to describe or infer in-lake processes.

The presumption that bottom-up controls predominate in spring, whereas top down are most important in summer is not valid. As noted above (and in many many other papers), grazing control of phytoplankton is strongest during the CWP, and diminishes in summer when slow growing colonial cyanobacteria and algae escape grazer control. (see Carpenter and Kitcchell Trophic Cascade book etc). Without directly measures of grazing (or actual limitation by physico-chemical processes), the authors are setting up an unnecessary (and likely incorrect) framework. I think they might be better served by using the CWP research as a rationale for expecting 1 or 2 phases of phytoplankton biomass change ('growth windows') as they are well studied and explain the patterns seen in the analysis of Chl. The Plankton Ecology Group (PEG) model may be useful.

Data and methods. Perhaps I missed it, but I think the paper needs a table which explicitly lays out the data sources without requiring the readers to go to the actual database. As part of this, there should be a good array of summary data (dates samples, resolution, parameters, lakes, etc). While I recognized most of the sources of data, I did not find direct attribution of the data to individual investigators within the paper, which I think is inappropriate.

As part of this, I think the authors need to justify the re-publication of already open-source data. I suspect the rationale is that there is a new analysis of Chl a dynamics, but that in itself creates the problem that the method of Chl analysis has not be validated previously (mainly section 2.2). The simplest solution is to hold back publication of this data paper until the method is reviewed and evaluated. Alternately, the authors need to demonstrate more robustly that the method they use is reasonably artifact free (e.g., the statement that growth periods range 2 to 260 days is highly suspect – see below) and very likely dependent on time series resolution.
Data compilation – I think the authors need to provide much more information on how the data were harmonized. I know from research in the US LAGOs program (Soranno et al), this is a massive part of the compilation process, taking years in the case of large datasets. Just saying that chemistry was measured (l. 125-130) seems insufficient to me. How did techniques differ, does that affect the findings (or not), are old procedures replaced by new ones, etc.

Similarly, lake selection needs to be better justified. First, it makes absolutely no sense to include two low latitude lakes, famous or not, in a high latitude study (remove them; delete l. 130-135). Second, the authors do not appear to recognize that climate systems (Gulf Steam, NAO, etc.) affect latitudinal gradients, producing much different local conditions in the UK/Scandivia/EU than in North America (mainly continental Canada). I understand this is not a full analytical paper, however, the use of latitude is uncritical acceptance and can confuse patterns presented in this paper.

Fig. 2 – I like the workflow but is it also possible to see how the number of sites/data density changes through the process? (how much data is lost at each step). Or that could be in an appendix figure.

141-150. Please provide more information on data manipulation. This is too vague/unclear to allow replication.

153-155. I think the authors need to provide more summary data on the time series
themselves. A statistical overview of the resolution seems particularly important, as the vast majority of sites would be sampled at weekly or longer intervals, so would be expected to have substantial restrictions on the detection of growth window onset and duration. Monthly resolution could, arguably, make the growth window meaningless.

- SSR. The use of these data need much better justification. First, as shown from the map in Fig. 1 – the vast majority of lakes are not located very close to SSR stations. Second, cloud cover would be expected to greatly influence the receipt of solar irradiance, but would not be recorded well for individual lakes. I *like* the attempt to use SSR data to explain variation in growth windows, but find the approach was largely unsuccessful – likely because the data were inappropriate. I think the authors need to better justify the use of SSR or drop the analysis.

- 2.1.3 – the summary of lake characteristics is pretty limited. This would seem to be important if the data are to be used in other analyses (as inferred).

- Section 2.2. Detection growth windows. The paper hinges on readers accepting 0.05 ug Chl /L/d as a critical threshold for signifying 'growth'. Its use because it was the median value in lakes with a CWP (productive; 2 growth phases) is one possible justification; however, the use of the same threshold in oligotrophic lakes seems inappropriate as Chl values may be up to two orders of magnitude lower than in eutrophic systems. Basically, use of a single static metric over all lake types seems unjustified.

One possibility would be to do a sensitivity analysis. The selection of a single threshold immediately raises the question of how the data patterns would change in a different threshold were used. Normally (in a methods paper), the authors would conduct some form of sensitivity analysis to demonstrate that the findings were (not) robust to the precise value used in the study. (but again, the authors need to decide whether this Is a data paper or a full analytical report).
- Data set. As noted about, some summary characteristics of the lakes (morphometry), Chl time series (sampling resolution, etc) and chemistry would be useful to interpret the patterns seen in l 215-224.

- 220-221 and elsewhere. Try to be less declarative about the reason for unmeasured patterns. In North America, latitude is correlated inversely with human populations and activity. To an extent, this is true in the European continent.

- L 229-231 again cause of patterns two speculative (and see comments about CWP above)

- 253-263. This paragraph is overly speculative, as it is based on marine gradients, not those of lakes. See work of John Smol for patterns in Northern Canada.

- 282-285. Again, this is over-extrapolated (from a fossil study that does not prove mechanism). More importantly, the analysis of SSR is vague or unreproducible. “Sensitivity” to SSR is never defined, and there is no obvious relationship in panels of Fig. 9. This whole section in unconvincing because there is no clear relationship among variables, and because the SSR sites are not particularly close to the study lakes.
Key findings. In my opinion, too many of the findings infer mechanisms which were not proved rigorously in the paper. It’s fine to review the main patterns, but it’s inappropriate to infer causal mechanisms from the sorts of analyses presented. Again, I think the authors are trying to move this from a data paper to a methods or interpretive paper. While this may make the paper more scientifically interesting, I don’t find it appropriate for the format of a data paper.