Comment on os-2021-66

Anonymous Referee #2

Referee comment on "A clustering approach to determine biophysical provinces and physical drivers of productivity dynamics in a complex coastal sea" by Tereza Jarníková et al., Ocean Sci. Discuss., https://doi.org/10.5194/os-2021-66-RC2, 2022

This manuscript is well written and outlines a clustering technique applied to a coupled physical-BGC model of the Salish Sea. The results of the clustering analysis are then used to identify and interpret the emergent properties and drivers behind the BGC dynamics. The clustering technique will be of high interest to a broad audience who are grappling with the interpretation of large datasets that more traditional methods will struggle with. On the other hand, much of the interpretation of the results are of a regional nature (e.g. applicable to the Salish Sea). Not being particularly familiar with this region, my comments relate to the methods used, rather than the interpretation of the results.

General Comments:

Lines 158 - 164: Noting that many of the clustering signals relate to variables that influence or are influenced by stratification, why is it that you choose to use depth integrated phytoplankton biomass? Would it not be better to look at the upper 20m (or similar e.g. above the Halocline as defined in lines 154-156)? Thus detecting the effect of stratification on nutrient supply to to the upper ocean.

Lines 174-177: This paragraph needs some clarification. Is the clustering done in univariate (e.g. independently for each signal), or in a multi-variate (e.g. for all signals at once) manner?

Section 2.3: How are the signals standardised such that the relative magnitude doesn't favour one signal over the other?

Section 3: Are you able to determine which of the signals are contributing the greatest

information content in dispersing the clusters? E.g. Is it the freshwater influx? Or halocline? Or do all signal contribute equally? I suppose that Figure 6 goes someway to answering this question, as the Freshwater Index and VED appear to have the most pronounced difference between then clusters.

Section 3: Is there substantial correlation between the signals, what is the impact of this correlation on the clustering algorithm?

Section 3 and Appendix A: This clustering approach may be a powerful method to diagnose the source of model error. Are there any indications that the model performs better/worse against observations in each of the clusters? For example, In the area of dark blue (Figure 5), how does the BGC model compare with obs in this region?