Review of
A regional hindcast model simulating ecosystem dynamics, inorganic carbon chemistry and ocean acidification in the Gulf of Alaska
by Claudine Hauri et al.

April 2020

**recommendation: accept with minor revisions**

1 General Comments

Hauri et al. introduces a new ROMS configuration of the Gulf of Alaska (GOA-COBALT) that improves upon previous efforts by introducing variable riverine and glacial freshwater fluxes to the modeled region. This is a solid overview paper with extensive evaluation of the simulation. However, I believe there are issues with vagueness/clarity on some key methods and statistics (*e.g.*, end-member analysis, linear Taylor decomposition, and correlations). I've also suggested some significant changes to the figures, which lack important information (bias between model and observations) and also use inappropriate colormaps which hinder interpretability. Lastly, I have pointed out a number of technical errors and suggestions that should be addressed prior to publication. That being said, I believe GOA-COABALT will make an immediate impact on the community. Congratulations on a great model release!

2 Specific Comments

1. I would recommend adding an additional sentence to the abstract to highlight the addition of variable freshwater forcing to this model. I know it's mentioned in L7-8, but it seems to be a substantial addition to regional modeling of the GOA and should be mentioned as such.

2. It would be nice to include the model mesh in Fig. 1 to make clear the horizontal resolution of GOA-COBALT relative to the features of the GOA, especially given this paper is associated with the public release of the model. Also to make clear it's not *e.g.,* telescopic. If 4.5 km is too fine to see in Fig. 1, you could have an inset where you zoom in on a sub-region and show the mesh.

3. L19-23, p4: Can you give some approximation of what this vertical resolution is? *E.g.,* N meters per grid cell in the nearshore and offshore.

4. L6-7, pg. 6: Does this mean that precipitation does not affect DIC or TA through dilution? This is typically standard in Earth System Models, while freshwater dilution from river inputs is less standard. Are there estimates for the impact of precipitation dilution vs. riverine/glacial dilution in the GOA? *I.e.,* is it a negligible term relative to rivers and runoff? This should be clarified in the text.

5. L16-17, pg. 6: Could you expand further on how DIC was normalized using the anthropogenic CO_2 estimates?

6. L22-23, pg. 6: Could you briefly describe this simulation that initialized the other variables in the text? *E.g.* what model it was, what time the output spanned. Was this a climatology initialization, a restart file?

7. L30-31, pg. 6: I assume this sentence through the end of Section 3.1 is describing the land hydrography model reviewed in Danielson et al. If this is the case, could you make it slightly more clear that the remainder of this section is summarizing Danielson et. al? I was confused by the correlation and p-value reported here, thinking it was referring to GOA-COBALT. If that is what's actually happening here, it should be more clear whether this is a pattern correlation, temporal correlation, over what domain, etc.

8. I would suggest changing the colormaps for the majority of figures in this text. Thyng et al. [2016] provides a good overview of colormap selection in "How to select an honest, effective, colormap." Every colormap used here is a red-to-blue *diverging* map, whereas most of this data is sequential and would be much more honestly portrayed on a sequential colormap. Some great colormaps available in python, Matlab, etc. are cmocean and Fabio Crameri's color maps. The main issue is that the red-to-blue colormaps diverge at an arbitrary value in this manuscript, causing a large visual distinction between red and blue regions that is not meaningful physically. The red-to-blue can be used for Ω and in the case of anomalies, but should be centered around 1 for Ω, since that is a critical threshold for that variable, and around zero for the anomalies. I imagine this sounds tedious, but I think it will drastically improve the visual presentation and interpretability of the output. More meaningful features will be apparent in the cross sections, *e.g.* in Figure 4, which will help the reader compare the model to observations. Below I compared a CESM hindcast run to ERSST observations as a demonstration. In the first example I use a red-to-blue diverging colormap. In the second, I use a perceptually uniform sequential colormap. I think the advantages will be more clear in the cross-section maps, but this still shows the differences and draws the eye away from the arbitrary divergent point at 15C.

ERSST (1980-2013)   CESM (1980-2013)
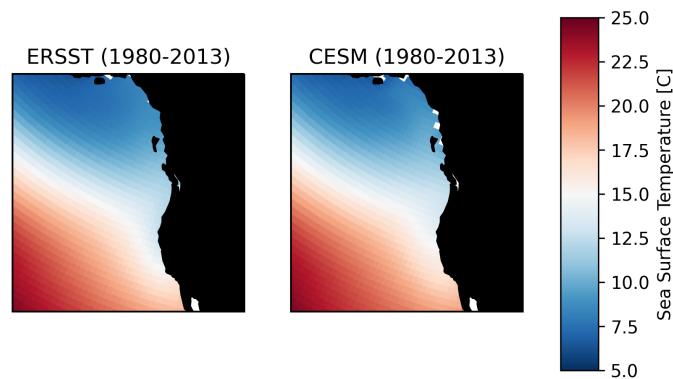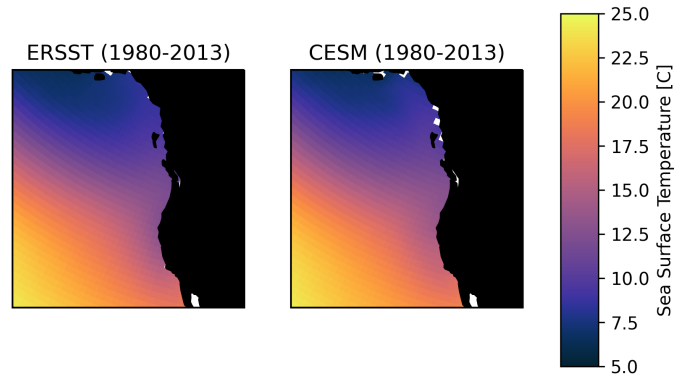
9. On a similar note, I am surprised that there is no third column in Figs. 2, 3, 4, 5, etc. showing the difference between the model and observations. Effort was made to interpolate the model and observations to the same grid, so it should be relatively straightforward to display the bias in the model by subtracting the two. I think showing this is crucial for the reader to see the regional expression and the magnitude of the bias. Currently, the reader relies on the author's highlighting of certain subregions of these biases in the text. For Figure 2 in particular, it's very hard to compare these by eye. On another note, in Figure 4, it looks like nearshore surface pH could be 0.2 units too acidic in the model, which would represent a 60% bias in the hydrogen ion concentration. Many of the quantitative arguments in the text about "overestimation" and "underestimation" will be made significantly more clear with the addition of a difference column (either raw or in percent bias).

10. Figs. 7 and 8: Is the white here due to missing data? If so, that should be made clear via something like gray or hatching since it could be misidentified as the value at the center of the colorbar. Although this would be alleviated in addressing (8).

11. L10, p14: Can you clarify in the text what these correlation coefficients represent? Is this the pattern correlation between the climatologies in the left and right columns of Figs. 4 and 5? Is this a temporal correlation of the transect average time series? What is the p-value for this correlation? I would suggest having a statistical analysis section of the methods that mentions your use of Pearson correlations, whether they are pattern or temporal correlations, and how you assess statistical significance. (See final specific comment as well regarding methods)

12. L3-4, p15: Can you quantify this through, e.g. RMSE or a correlation of interannual variability? This section investigates selected case studies of years and variables but doesn't do a bulk quantification to assess model skill.

13. L13-15, p15: Perhaps I am misreading this, but both photosynthesis and freshwater dilution should reduce DIC, and thus raise pH to more basic levels. So why is it surprising that pH is high, "despite the freshwater influence and its diluting character"? Or did you mean acidic by "high" here? Please clarify.

14. I found Section 5 very hard to interpret, and suggest that it is re-written and the methodologies here made more clear. Firstly, the end member analysis methodology should be made more clear. Admittedly I do not have a background in end member analysis, so perhaps it is clear to the informed reader what is happening here. As someone with a modeling background, I

first thought "end member" implied that multiple simulations were run and one at the edge of the distribution of riverine boundary conditions was selected for analysis. It's unclear what "non-zero" DIC means when all of the DIC range in Table 2 is non-zero. In general, it needs to be spelled out that this is an end member mixing analysis (I assume), and more care should be taken explaining the methodology here. Secondly, the linear Taylor decomposition should also be spelled out. I don't think every step of Rheuban et al. (2019) needs to be replicated here, but it would be helpful to the reader to have some of the key equations and assumptions. Particularly that the sensitivity terms aren't explicitly calculated, how anomalies are generated, etc. I would suggest an additional section to the methods summarizing the end member analysis and linear Taylor decomposition.

3 Technical Comments

1. I find the first sentence of the abstract awkward: "The coastal ecosystem of the Gulf of Alaska (GOA) is especially vulnerable to the effects of ocean acidification and climate change that can only be understood within the context of the natural variability of physical and chemical conditions." Is it the coastal ecosystem or the effects of OA/climate change that can only be understood within the context of natural variability? I wouldn't say that natural variability is a key topic addressed in this paper either. I would suggest revising this sentence to change its content or to make it more clear.

2. L3, pg1: "iron enriched" should be "iron-enriched" (and in other places in the manuscript)

3. L7, pg1: "high resolution" should be "high-resolution." I think this is used four times more in the text with changing usage of "high resolution" vs. "high-resolution."

4. L14, pg1: I would suggest dropping "As such,"

5. L16, pg1: "CO_2 sensitive" should be "CO_2-sensitive" and elsewhere in the manuscript.

6. Table 1: Is "alpha" supposed to be α? The formatting of this table is a bit difficult to interpret. E.g. the italicized sub-header, and it's not clear immediately that the a and b below explain the table values. Maybe this will be fixed on typesetting.

7. L7-8, pg2: I think this would be cleaner with something like "... high physical, biological, and chemical spatiotemporal variability across the GOA continental shelf."

8. L11, pg2: What is "its" referring to here? Grammatically it could relate to "this region" or "climate change and ocean acidification," among other interpretations. I would break L9-12 up into 2-3 sentences for clarity.

9. L14, pg2: It might be helpful to spell out why a seasonal increase in vertical mixing would lead to reduced carbonate concentrations here, since this is early in the introduction. Also what does seasonal "increase" refer to? Does it increase from winter to summer, summer to winter? Is the seasonality of mixing increasing with time?

10. L18, pg2: I find the use of "endowed" awkward here and elsewhere. I think it would be more simple to just say "contains low TA" or "is characterized by low TA" for example.

11. L20-21, pg2: I would split the last clause off to its own sentence. "exacerbating" should be "exacerbates."

12. L24-25, pg2: I would revise to something like the following for ease on the reader: "These two limiting nutrients lead to a phytoplankton community composition dominated by diatoms in the dFe-rich near-shore area and by small phytoplankton in the dFe-poor off-shelf area."

13. L4, p3: "impede" should be "impedes" in the current way its written. If you were to drop "coverage" it should be "impede."

14. L9, p3: I think "of the system" should be added to the end here. Or to expand more, mentioning that this is because the frequency of available observations causes aliasing issues or isn't sampled enough to cover the spatial and temporal decorrelation scales of the region.

15. L12, pg3: I would drop "successfully" here. How does one know that they are successful at representing the future?

16. L1, p4: Drop the comma following Ekman pumping.

17. L3-4, p4: I think this could be refined to "However, neither of these models simulate the influence of freshwater input along the coast, which exhibits high spatiotemporal variability." This puts the emphasis on the fact that they don't simulate freshwater input, rather than mentioning the variability first.

18. L9, p4: Should read "long-term anthropogenic trend."

19. L12-14, p4: This sentence is quite grammatically confusing as it stands. I suggest it is re-written entirely.

20. L14-17, p4: I would end this summary with a sentence explaining how this expands on past modeling efforts for the GOA.

21. L23, p4: Should be "eddy-resolving."

22. L23, p4: It might be worth explicitly mentioning that this resolves coastal upwelling scales in this region.

23. L27, p4: What does "the model" refer to here? Multiple experiments with the Coyle et al. model, or multiple experiments with multiple GOA models?

24. L33, p4: Is there supposed to be a comma following "energy-balance" in "energy-balance snow ice/melt,"? Also I don't think there's supposed to be a hyphen in energy balance.

25. L4, p5: Earth should be capitalized.

26. L17, p5: Water column shouldn't have a hyphen, as earlier in this line.

27. Section 2.2: Check past vs. present tense here. It varies in the first few lines.

28. L10-12, p6: This is an exact copy of the sentence in L33, p4. I'd suggest deleting the earlier case of it.

29. L13-14, p6: I might just spell out "DIC concentrations are higher than TA" instead of using "DIC > TA" which is difficult to read at first.

30. L3, p9: I would change "To sum up" to "In summary,"

31. L7, p12: "draw-down" should be "drawdown."

32. Section 3.3: Figure 5 is never cited here and should be in L7, pg13. Only Figure 4 is referenced once in this whole section–I would suggest adding more Figure references for clarity to the reader.

33. L10, p14: correlation coefficient should not be capitalized.

34. Figs. 7 and 8: Can the stations be translated to latitude for clarity as in the other Seward Line plots? Or at least designate which direction is offshore vs. nearshore?

35. L27-30, pg 14: Is this in reference to May or September? This would be made more clear by citing one or both figures here.

36. L7-8, p15: "Lowest surface temperatures of $< 3C$ are found nearshore in February and March" Should probably put (not shown) here since it directly follows the introduction of Fig. 9 and these months aren't included.

37. L8, p15: Should be "surface waters slowly warm"

38. L9, p15: Is "S^{max}" necessary here? I don't think this symbol is used elsewhere.

39. L27, p15: "In the following" should be dropped. Or turned into "In the following section,"

40. L9, p18: There should be no hyphen following "moderately" (an adverb).

41. L18-19, p18: Drop "interestingly" and "however" here.

42. If allowed by *Biogeosciences*, I would reference the relevant figures in the summary/conclusion as you work through the points.

43. L30, p18: "time-series" should be "time series" and/or standardized throughout. Both "time-series" and "time-series" is used in the manuscript.

44. L9, p21: "endmembers" should be "end members" or standardized to "end-members".

References

Kristen M Thyng, Chad A Greene, Robert D Hetland, Heather M Zimmerle, and Steven F Di-Marco. True colors of oceanography: Guidelines for effective and accurate colormap selection. *Oceanography*, 29(3):9–13, 2016.