Referee comment on "A framework to evaluate and elucidate the driving mechanisms of coastal sea surface $pCO_2$ seasonality using an ocean general circulation model (MOM6-COBALT)" by Alizée Roobaert et al., Ocean Sci. Discuss., https://doi.org/10.5194/os-2021-70-RC2, 2021

Roobaert et al. assess the skill of the MOM6-COBALT model for representing the seasonal cycle of $pCO_2$ in coastal regions and develop a methodology for interrogating the processes driving seasonal and regional differences. They use the model output to interrogate the drivers of the seasonal signal in three regions where model skill is high. This study makes good use of data products for assessing the skill of models in reproducing seasonal coastal dynamics and the unique information that models can bring to coastal carbon research, however, a few issues should be addressed before publication.

1) In the method to assess the different processes controlling seasonal $pCO_2$ variability, the assumption that the coefficients (explained in lines 195-206) are constant in time needs to be explained. Are the coefficients truly constant in time if the goal is to understand how processes like freshwater discharge (a spring event in many regions) impact $pCO_2$? Doesn’t a spring-time river runoff event, for example, change the relationship between DIC, ALK, SSS, etc, which would not be reflected in coefficients derived from average conditions over 1998-2015?

2) In the methodological limitations section (3.1.4) it is mentioned the coastal-SOM-FFN climatology does have limitations in reproducing $pCO_2$ variability in some regions. Since this is the case, for the regions where there are SOCATv6 data (lines 325-326 state there are 45 grids with sufficient data), the paper should include model-SOCATv6 comparisons, especially for seasonal amplitude. Right now, Figure 4 does show SOCATv6 annual mean but Figure 5 does not show seasonal amplitude from SOCATv6, and seasonal amplitude is, as the authors state, underrepresented by coastal-SOM-FFN. Figures 4 and 5 should both include SOCATv6 as well as the residuals between model and SOCATv6 (for the regions where there is observational data). This is also an issue with the supplemental tables, where Table S1 presenting annual mean does include SOCATv6 but S2 presenting seasonal amplitude does not. The paper should include a more robust assessment of model-SOCATv6 seasonal amplitude comparisons, given seasonality, not annual mean, is the central focus of the study.
3) The ESRL atmospheric data are not properly cited. First, ESRL does not provide pCO2 as stated in line 115. The atmospheric community measures and provides xCO2. The authors need to properly cite the ESRL data source (not the current citation of Joos and Spahni) and explain how atmospheric pCO2 was calculated.

Minor issues:

Line 45: This statement seems to be Northern Hemisphere biased. Given this study used a SOCAT-based data product, Southern Hemisphere coastal regions are extremely underrepresented and many areas are likely not well characterized.

Line 290 / Figure 4: As stated earlier, it would be easier to see the model-data difference if a residual plot was included rather than ask the reader to compare Fig 4a and 4b.

Line 326: Many of these 45 grid cells with continuous pCO2 time series are likely buoy locations. Added to SOCAT in 2015, these continuous time series are an essential feature of SOCAT for seasonal assessments like this study, and make a strong case for a more thorough model-data comparison as mentioned previously.

Line 352-353: In some places like here the regions are only stated by their associated numbers, however, it is easier for the reader to understand the results if stated by their name and number as in lines 356-357.

Lines 388-390: This seems to be an important result of the study that should be included in the Conclusion section.


Lines 575-583: Description of the xCO2 data source is missing from this section.

Figure 3: This is another figure that could benefit from showing a MOM6-COBALT vs SOCATv6 comparison for pCO2.
Figure 6: If any of these regions have continuous pCO2 time series in SOCATv6, SOCATv6 should also be included.