Responses to RC2

General Comments:

1. The manuscript presents a method to reconstruct surface ocean pCO_2 in the northern Gulf of Mexico (nGOM) using Machine learning. A 22 year, spatially resolved, time series of nGOM surface pCO_2 is created. The method itself seems robust, filling the gaps in the observational dataset using satellite observations. The dataset is then analyzed for spatial and seasonal patterns. Most of these findings have been previously described using direct observations or models and it is therefore unclear how this new method improves our understanding of surface pCO_2 in the region. Without new results the manuscript becomes a methodological paper, which is useful, but would be best published elsewhere.

Response: We thank the reviewer for acknowledging the robustness of our machine learning reconstruction. We fully agree with the constructive suggestion that we need to more clearly articulate how our new long-term, high-spatiotemporal-resolution dataset (22-year, 4-km, 8-day continuous pCO_2 product) resolves seasonal, interannual, and decadal variability in the nGOM, and how this advances our understanding beyond previous studies. In response, we have carried out additional analyses focusing on the decadal trends and have restructured the manuscript by separating the combined "Results and Discussion" section into distinct "Results" and "Discussion" sections (see our responses to General Comments 2 and 3 below). We believe the revised manuscript is no longer merely a methodological paper, but substantially improves our understanding of surface pCO_2 and air-sea CO_2 flux dynamics in the region.

2. The decadal trends are probably the most interesting result from this study. The manuscript should focus more on these trends. Thermal and non-thermal trends are interesting and provide additional knowledge for nGOM. However, the mechanisms to explain these trends are currently lacking.

Response 2: We agree with the reviewer that the decadal trends are the most valuable contribution of our study. In the original manuscript, we only presented the regionally averaged annual trends. In the revised version, we have significantly expanded the analysis by examining seasonal and spatial differences in these trends. The newly added Discussion Sections 4.2 and 4.3 present previously unreported findings on the contrasting seasonal trends and the long-term amplification of the seasonal amplitude of pCO_2 and CO_2 flux (see our response to General Comment 3 below). These additions provide a more complete mechanistic explanation of the decadal changes and directly address the reviewer's concern.

3. Finally, the format chosen by the authors, mixing Results and Discussion, does not help. This format often results in a lack of discussion, which is the case in the manuscript. The discussion is largely missing. Splitting Results and Discussion section would force the authors to clarify the results and to put them in perspective, which would likely greatly improve the manuscript.

Response 3: We thank the reviewer for this constructive suggestion. We fully agree that separating Results and Discussion could substantially improve the clarity and depth of the manuscript. In the revised version, we have restructured the manuscript accordingly, with the new Discussion section organized into four subsections, each addressing a distinct aspect of our findings. **Notably, subsections 4.2 and 4.3 are entirely new, and they present previously unreported results that significantly extend our analysis of the decadal trends.**

The revised structure of the Discussion section, along with a summary of its major contents, is outlined as follows:

4.1 Controls on seasonal and inter-annual variability

This subsection synthesizes the mechanistic drivers of the observed pCO_2 seasonality. We quantitatively separate the thermal and non-thermal components and discuss the contrasting roles of biological uptake in spring–summer versus vertical mixing in autumn–winter. We also interpret the interannual relationship between river discharge, nutrient loading, and pCO_2 variability.

4.2 Decadal changing rates of pCO_2 and CO_2 flux in different seasons

This subsection is entirely new. We decompose the decadal trends of the regionally and yearly averaged pCO_2 and CO_2 flux by season (see the newly added Figure 10) and reveal pronounced seasonal asymmetries: summer and winter show long-term decreasing trends in pCO_2, whereas spring and autumn exhibit significant increasing trends. For CO_2 flux, when accounting for long-term changes in wind speed, the winter CO_2 sink shows a clear strengthening trend. By combining these seasonal trends with the dominant controlling factors identified in subsection 4.1, we quantify the contributions of rising SST versus changes in biological uptake and vertical mixing, providing new insights into the drivers of decadal changes across different seasons.

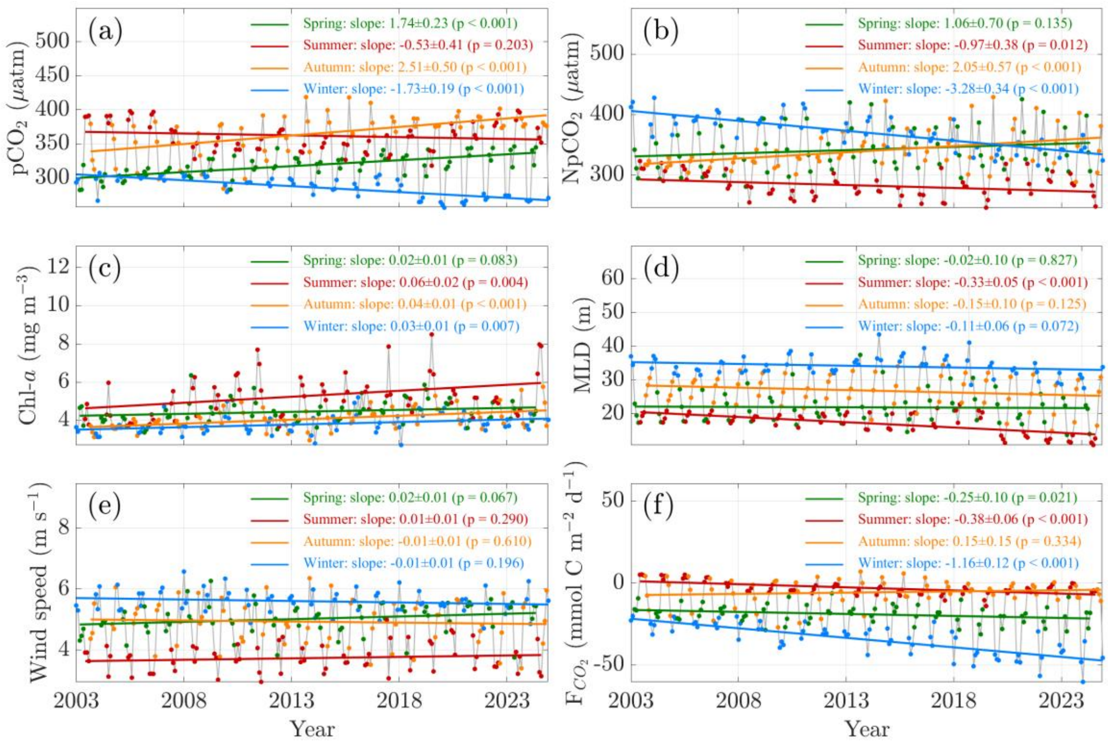

Figure 10. Seasonal long-term trends of (a) sea surface pCO_2, (b) $NpCO_2$, (c)Chl-a, (d)MLD, (e)wind speed, and (f) air-sea CO_2 flux (F_{CO2}) at different seasons. The solid lines represent the linear regressions for each season's monthly averages from 2003 to 2024: spring (green). Summer (red), autumn (yellow) and winter (blue).

4.3 Increasing seasonal amplitudes in pCO_2 and CO_2 flux

This subsection is also newly added. Following the discussion in subsection 4.2, we examine the long-term trends in the yearly maxima and minima of pCO_2 and FCO_2. As shown in the newly added Figure 11, we find that the yearly maximum of pCO_2 shows a slight increasing trend (0.36 μatm yr^{-1}), while the yearly minimum exhibits a significant decreasing trend (−1.39 μatm yr^{-1}), leading to a marked long-term increase in the seasonal amplitude of surface pCO_2. Correspondingly, the seasonal amplitudes of FCO_2 have also expanded over the 22-year period. This amplification of the seasonal cycle has important implications for the seasonal dominance of the winter carbon sink and the overall enhancement of the annual CO_2 uptake.

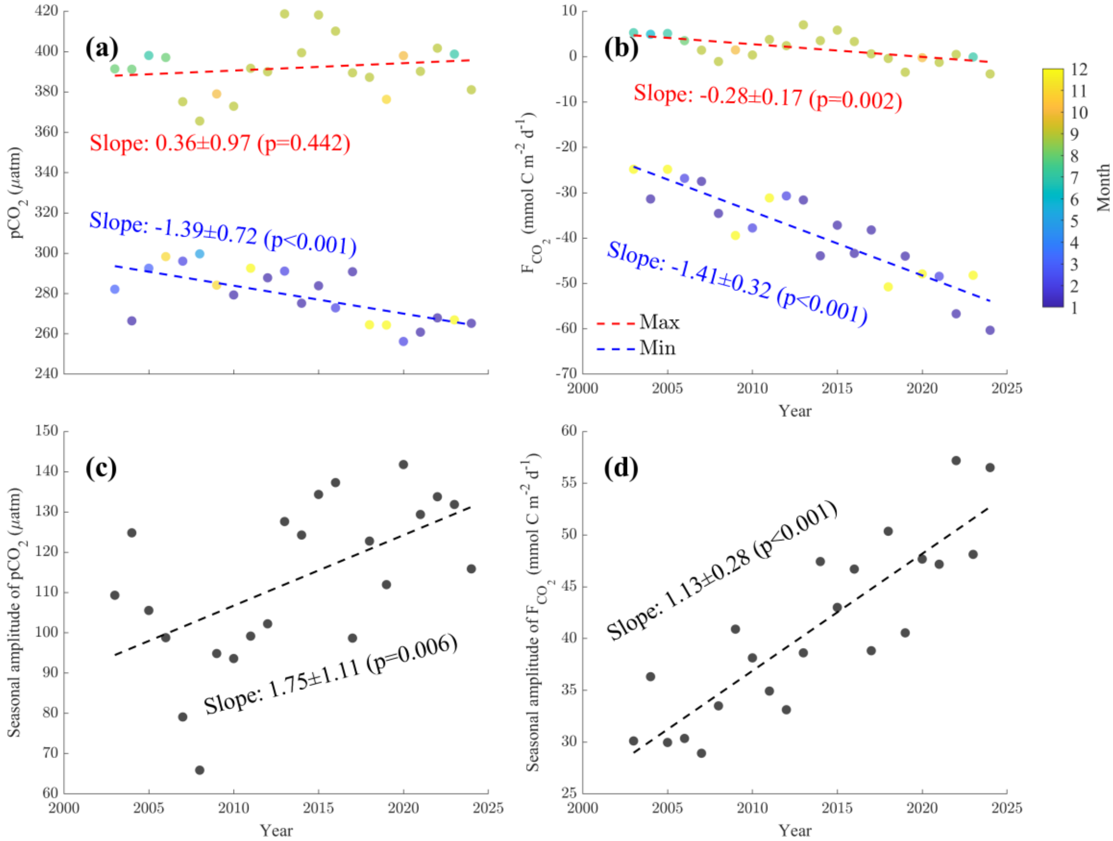

Figure 11. Long-term trends of the annual maxima and minima of (a) sea surface $p\mathrm{CO_2}$ and (b) air-sea CO_2 flux (F_{CO2}), and the resulting increasing trends in the seasonal amplitudes of (c)$p\mathrm{CO_2}$ and (d) F_{CO2}.

4.4 The effect of weakening westward Mississippi plume transport

This subsection discusses the spatial heterogeneity of the decadal trends. We demonstrate that weakening easterly winds have reduced the westward transport of the Mississippi River plume, leading to higher pCO_2 increase rates on the western shelf during the productive season. To strengthen the mechanistic link between wind-driven plume contraction and the elevated pCO_2 trend on the western shelf, we have added additional supporting evidences (see our responses to Specific Comments L436–L467 below).

We believe this restructuring has substantially improved the manuscript by clearly separating the presentation of results from their interpretation, and by providing a more thorough and integrated discussion of the underlying mechanisms and broader implications of our findings. We are grateful to the reviewer for prompting this important improvement.

Specific Comments:

L101: What do you mean by true values?

Response**:** We agree this phrasing is ambiguous. By "true values," we meant the actual decadal changing rate that remain uncertain due to limited observations and high spatiotemporal variability. We will revise the sentence to: *"In addition, the decadal trends remain highly uncertain in the dynamic nGOM due to limited observations and large spatiotemporal variations [Wu et al., 2024]."*

L157: Can you provide some information on the use of the internal validation dataset for hyperparameter tuning?

Response: We thank the reviewer for this question. In our study, the internal validation set (17.5% of the matched data) was used specifically for hyperparameter tuning. We trained multiple random forest models on the training subset and systematically varied key parameters, including the number of trees, maximum tree depth, and the number of candidate features considered at each split. The optimal parameter combination was then chosen based on the prediction error on the internal validation set. Because this subset directly influenced the selection of hyperparameters, its reported accuracy carries a degree of optimistic bias and cannot serve as a reliable indicator of how the model would perform on entirely independent observations. In contrast, the external validation set (30% of the data) was kept completely isolated throughout the entire model development process. It was not used in any fitting, tuning, or model revision, and was evaluated only once after the final model had been fully trained. This provides a more honest and unbiased measure of the model's generalization capability.

Section 2.2.2 will be revised to include additional details on the use of the internal validation dataset for hyperparameter tuning, and to more explicitly define the distinct functions of the internal and external validation subsets.

L196: Is this $pCO_{2,mean}$ as well?

Response: No. We confirm that the formula in L196 is intended for temperature normalization, i.e., to adjust the pCO_2 observed under varying sea surface temperature (SST) to a reference temperature (SST_{mean}), so that the resulting $NpCO_2$ reflects non-thermal effects. The variable in this formula should therefore be the actual observed pCO_2 at the time of measurement, rather than the long-term mean $pCO_{2,mean}$. We will clarify this point in the revised manuscript by explicitly stating the meaning of each term in the equation description.

L208: The model results deserve a bit more than 1 paragraph discussion.

Response: We agree. In the revised manuscript, we will expand Section 3.1 to provide additional validation on our model results (see our response to Specific Comment L222 below), and add a comparison with previous model performances in the region.

L222: Additional validation: Please provide more discussion of Figure S1 and include individual statistics.

Response: We agree with the reviewer that additional validation metrics would strengthen the manuscript. In the revised supplement, we will include a new table summarizing the statistical results for each individual cruise shown in Figure S1, including the root mean square error (RMSE), coefficient of determination (R^2), and mean bias. This will allow readers to assess the model's performance across different seasons and spatial regions more transparently. We will also add a brief discussion of these statistics in the main text to highlight the model's overall robustness and any cruise-specific deviations.

L236-253: This is already known, how does this new high resolution dataset improve our understanding of surface pCO_2 in the nGOM region?

Response: While the general spatial pattern (coastal low vs. offshore high) has been qualitatively described, our dataset provides quantitative, long-term climatological means and spatial heterogeneity at a resolution (4 km) that was previously unattainable. Specifically, we now quantify: the coastal-offshore gradient in absolute terms (310.9 µatm vs. 374.3 µatm); the annual amplitude (133.1 µatm in coastal vs. 88.4 µatm offshore); and the seasonal evolution of this gradient across 22 years (also shown in Figure S2). We will revise the text to emphasize that these quantitative constraints which provide a benchmark for future model-data integration.

Figure 3: It would be interesting to have a map of the spatial errors associated with these maps.

Response: We thank the reviewer for this suggestion. In the revised supplement, we will include an additional figure showing the spatial distribution of model errors (RMSE) corresponding to the climatological maps presented in Figure 3.

L264-276: It would be better to focus on pCO_2 and relate it to the other variables, instead of describing what is already known first.

Response: We thank the reviewer for this suggestion. In the revised manuscript, we will present the seasonal variations of key environmental parameters alongside pCO_2 in the order of the Figure 4 panels. The mechanistic discussion of pCO_2 seasonality will be moved to the new Discussion section, where it can be more appropriately contextualized (see also our response to General Comment 3 above). We believe these changes will improve the clarity and focus of the presentation.

L270-276: This would fit better in a proper Discussion section.
Response: Accepted. We will move the mechanistic interpretation of the seasonal cycle to the new Discussion section (see also our response to General Comment 3 above).

L281: Indeed it was already shown by Huang et al 2015, so what is the contribution from the new model?
Response: Huang et al. (2015) described the general seasonal pattern based on 13 cruises (2003–2012). Our continuous, gap-free, 22-year time series data reveal: 1) spatial patterns of seasonality with coastal vs. offshore differences, and 2) increasing seasonal amplitude over 22 years (see Response 1 to General Comments above). We will revise the text to acknowledge the earlier finding while clearly stating our new contribution.

Figure 4:
- Can you tighten the space between panels and increase the font size?
Response: We will adjust the panel spacing and increase font sizes in all figures to improve readability.

- Panel b: Practical Scale Salinity does not have units, see https://www.jstor.org/stable/43924646 Please update all figures/text accordingly.
Response: We have removed the salinity unit.

- Panel e: Supposedly there is a significant temporal variability that is not shown here.
Response: As the reviewer correctly notes, the seasonal climatology shown in Figure 4e represents the multi-year average seasonal cycle and therefore does not capture the long-term increasing trend in pCO_2, nor the associated changes in seasonal amplitude over the 22-year period. These decadal variations will be the focus of the newly added Discussion subsection 4.2 and 4.3 (see our response to General Comment 3).

- Panel g: Is this calculated from the mean of spatially/temporally varying $\Delta pCO2$ (as it should) or as the difference of the 2 lines in panel 4e? Visually it seems to be the latter, which I don't think is the right way to calculate ΔpCO2.

Response: We confirm that ΔpCO_2 shown in panel (g) is calculated as the mean of spatially/temporally varying air-sea pCO_2 differences. This is the correct bulk approach. The visual similarity to the difference of the two lines in panel (e) is coincidental because $pCO_{2,air}$ is spatially uniform.

L376-380: This is not shown, you should stick to the results or discuss based on the literature.

Response: We thank the reviewer for this helpful suggestion. We agree that the discussion of enhanced accumulation of respired CO_2 in shallow shelf bottom waters should be supported by literature references. In the revised manuscript, we will add citations to previous studies that have documented the mechanisms of bottom-water CO_2 accumulation on the nGOM shelf.

L396-398: These trends are not statistically significant and may be due to the time window chosen for the study. If there are indeed trends, please reference the corresponding studies.

Response: We accept this criticism. We will revise the text to clearly state that these trends are weak and not significant, and we will avoid overinterpreting them.

L402-404: This is an overstatement and doesn't seem to match the citations. Boyce et al is a global study, whereas Rabalais et al mentions the accumulation of nutrients as a reason for Chl increase.

Response: This sentence is revised to: *"The driver of this Chl-a increase is not fully understood, but may be related to a combination of factors, including warming-induced extension of the growing season [Boyce et al., 2010], changes in nutrient stoichiometry [Rabalais et al., 1996], and physical processes such as wind-induced mixing and coastal circulation [Li et al., 2022]."*

L406: Li et al (2022): Is this for NGOM or for the whole Gulf? This statement is misleading because it mixes processes that control open water chl and the Mississippi plume chl.

Response: We thank the reviewer for this clarification. We confirm that Li et al. (2022) reported spatial distribution of long-term Chl-a trends for the entire Gulf of Mexico (see Fig. 8b in their paper), and their results highlight a particularly strong increasing trend in the region of our study.

L408-409: The change in chl is highly variable though and may not be in sync with the change in SST. Indeed, the trend in $NpCO_2$ is not significant. L409-411: If you want to show this effect you need a more convincing figure.

Response: Thermally driven $TpCO_2$ exhibits a long-term increasing trend (0.52 ± 0.17 μatm yr^{-1}), while the non-thermal $NpCO_2$ trend (-0.25 ± 0.18 μatm yr^{-1}) is not statistically significant. The $NpCO_2$ trend can be result from increasing Chl-a during productive seasons (spring and summer) and/or warming-induced weakening of vertical mixing during cold seasons (autumn and winter). In the revised manuscript, we will testify this hypothesis in examining season-specific long-term trends in these variables.

Figure 9:

- Panel g: This is much more informative than Figure 4g and questions the need for Figure 4, particularly that panel.

Response: Fig. 4 shows the seasonal variations while Fig. 9 shows the decadal trends. They serve different purposes and are necessary for a complete picture. We will keep both figures but clarify their distinct roles in the captions and text.

L436: This is almost the same rate as the atmospheric pCO_2, can you comment? How did Kealoha interpret this trend in the northwestern shelf?

Response: The reviewer is correct that 2.30 μatm yr^{-1} is close to the atmospheric rate. This suggests that on the western shelf, pCO_2 is essentially higher than other regions in the nGOM. Kealoha et al. (2020) attributed the higher trend on the northwestern shelf to reduced biological uptake or enhanced vertical mixing. We will add this interpretation to our discussion.

L448: How significant is this trend? Can the units be interpreted like this: the dark blue regions correspond to 0.4% per year, which over the ~20 years cover by the study would be a ~8% decrease in the frequency of easterly winds?

Response: We thank the reviewer for the insightful suggestion and will provide quantitative descriptions on these results.

L449: Note that for this to be true the change in easterly winds need to be calculated over the productive season only.

Response: We agree. The easterly wind trend shown in Fig. 10d is the annual mean. We will re-calculate the trend for the productive seasons (spring–summer) and report that instead.

We will also revise the text to clarify that wind effects on plume transport are most relevant during the productive seasons.

L454-456: You need to provide better supportive data/figure to demonstrate this. L457-467: There is weak support for these conclusions.

Response: We thank the reviewer for these critical suggestions. To better understand the causes of the higher decadal pCO_2 increase rates observed on the western shelf, we will investigate two candidate mechanisms in the revised manuscript. First, we will examine the long-term changes in wind-driven plume transport during the productive seasons, following the reviewer's Specific Comments L448 and L449. Second, we will evaluate changes in vertical mixing intensity during the cooling seasons, which may also contribute to the enhanced pCO_2 trend.

Figure 10: Be careful with overinterpretation: these panels don't provide information about the significance of these trends.

Response: See our responses to Specific Comment L448 and L449.

L510: Can you provide the model code on GitHub? I wonder if sharing reconstructed data only is enough to meet BG standards.

Response: We will upload the model code and include the link in the revised manuscript.

Figure S3d: There seem to be sharp changes in derived pCO_2, presumably associated with the switch between models. Can you comment on that? The other variables have much smoother transitions in the seasonal cycle.

Response: We thank the reviewer for raising this important point. In our model development, we actually tested two training strategies: a single full-season model applied to all months, and the season-specific random forest model with four sub-models (Table S1). Both approaches yielded consistent results for the subsequent seasonal and decadal trend analyses. However, we found that the season-specific model consistently outperformed the full-season model in terms of agreement with *in situ* observations, which is not surprising given that the spatiotemporal distribution and dominant controlling mechanisms of surface pCO_2 differ markedly among seasons in the nGOM. Regarding the transition between seasonal sub-models, we also tested using both adjacent seasonal sub-models for the transitional months between seasons (e.g., May–June, August–September, November–December), and found that the differences between the resulting predictions were smaller than the overall uncertainty of the model (RMSE ~27.6 μatm). Therefore, no significant abrupt transitions were introduced at the seasonal boundaries. We have added a brief explanation of this transition strategy in the revised Methods section to clarify the procedure for readers.

Technical Comments, typos:

L72: seasonality in air-sea CO_2 exchange

Response: Revised as suggested.

L87: ...these limitations and achieve a continuous, satellite remote sensing... Something is missing in this sentence.

Response: Correct to: *"To overcome these limitations and achieve continuous, high-resolution reconstruction, satellite remote sensing combined with machine learning has emerged as a powerful alternative."*

L92: made a significant

Response: We have corrected this as suggested.

L100: suggestion: In this area, the true values of the decadal trend had large uncertainties...

Response: Correct to: *"In addition, the decadal trends remain highly uncertain in the dynamic nGOM due to limited observations and large spatiotemporal variations [Wu et al., 2024]."*

L102: this study attempt to develop

Response: Revised as suggested.

L106: Our strategy is to train dedicated

Response: Revised as suggested.

L132: Replace "gravel and detritus" by "dissolved and detrital organic matter"

Response: We have corrected this as suggested.

Figure 1f: Training

Response: We have corrected this as suggested.

L277: remove "clear", this terminology is very subjective.

Response: We have corrected this as suggested.

Figure 5: See earlier comment about psu.

Response: We have removed the salinity unit.

Figure 9:

- Same comment as earlier about the size of the font
- Panel c: psu: see earlier comment

Response: We will adjust the panel spacing and increase font sizes to improve readability and remove the salinity unit.

Figure 10: Panel e: psu: see earlier comment

Response: We have removed the salinity unit.

- L496: Acknowledgements (according to BG template)

Response: We have revised the Acknowledgements section according to BG template.

Figure S2: Can you provide the animation?

Response： We thank the reviewer for pointing this out. Figure S2 was originally included as an animated GIF in the submitted Word file. However, during the conversion to PDF, the animation was lost and only the first frame was retained. In the revised submission, we will provide the animation as a separate supplementary file.