Response to Reviewer #1

We thank the reviewer for their comments and input on the manuscript. We have addressed these comments as described below. All reviewer comments are presented in italic font while the author responses are displayed in standard font. Specific text that was added to the updated manuscript is provided in blue text.

The study by Johnson et al. applies a novel inversion technique, i.e., the Gaussian Process machine learning method to infer state-wide, source-specific CO_2 fluxes using OCO-2/3 XCO_2 data, illustrating that OCO-2/3 XCO_2 can be assimilated into inverse models to estimate sub-regional and source-specific CO_2 fluxes on a seasonal- and annual-scale. In general, this is a nice study that extended our knowledge on employing new artificial intelligence-based inverse modeling methods to infer CO_2 fluxes from atmospheric observations. I have a feeling, the current content looks a bit "thin", some advantages of the new method have not been clearly illustrated. Also, some detailed analyses should be added to make the presented results more sound. I would like to recommend it for publication after addressing the following issues.

We are glad the reviewer appreciates the novelty of this study and the importance of the inversion method applied. We have addressed the comments below in order to improve the robustness of the analysis.

Major comments:

1. The current results do not clearly illustrate the advantages of the GP/ML inverse method. Comparing the new method with old ones should help.

We have now improved the discussion of the advantages and disadvantages of the Gaussian Process Machine Learning (GP/ML) inversion system compared to standard atmospheric inverse modeling systems using the following text in the Discussions and conclusion section of the revised manuscript. This specific topic is illustrated in Jeong et al. (2025), that was not yet fully published when this manuscript was submitted for review and thus was not citable, which focuses on comparing the classical Bayesian inversion (CBI) and the GP inversion methods. We have added the following to the text in the revised manuscript:

"In evaluating the GP inversion method used in this study compared to linear inverse classical Bayesian inversion (CBI) models [e.g., $\mathbf{y}=\mathbf{K}\boldsymbol{\lambda}+\epsilon$, commonly used in atmospheric inversions; (Rodgers, 2000)], advantages and disadvantages become apparent. The GP-based inversion method employed in this work offers several advantages over classical methods, as highlighted by Jeong et al. (2025). Jeong et al. (2025) presents the advantages of the GP method through inversion results from different approaches, and we briefly describe them here, focusing on the key points.

First, Jeong et al. (2025) demonstrated through simulations using multiple inverse modeling methods for constraining CO_2 fluxes in California when assimilating OCO-2+OCO-3 XCO_2 observations that the GP inversion yields superior results compared to the CBI method. Specifically, the CBI method failed to capture the FF scaling factor accurately at the 68% confidence level. They repeated the inversion multiple times, and this result was consistent. Their work also showed that without a proper prior distribution, a simple linear regression produced a

physically implausible scaling factor for fire emissions. In our full Bayesian approach, we specify prior distributions for all parameters, including the scaling factors and kernel hyperparameters. Overall, the GP inversion offers substantial flexibility in modeling intricate, non-linear dependencies without needing a pre-specified model framework (Ebden, 2015).

Second, the CBI method typically relies on analytical solutions and lacks robust techniques for estimating crucial parameters, such as hyperparameters (e.g., variance of the diagonal elements) for the covariance matrix (i.e., the kernel). Consequently, many previous inversion studies have used prescribed values for hyperparameters. For instance, these studies often utilized known values from other work (e.g., Roten et al., 2023) or estimations derived from sensitivity analyses to construct the model-data mismatch covariance (Gerbig et al., 2003; Jeong et al., 2013; Johnson et al., 2016). Such approaches do not guarantee that the estimates are consistent with the observed data. In contrast, Jeong et al. (2025) showed that the GP method can infer the noise variance, with its median value closely aligning with the true value. This represents a significant advancement over previous approaches, as it enables the direct estimation of true noise variance from the input data. Also, the GP inversion intrinsically provides quantification of uncertainty, which proves advantageous in scenarios with limited data as in atmospheric inversions.

Third, the GP method includes the spatiotemporal kernel as its essential component, as shown in this work. While some previous work (e.g., Turner et al., 2020) used a covariance with both spatial and temporal components, many previous inversion studies have not used a fully spatiotemporal covariance. This is because it is not straightforward to estimate the hyperparameters for the spatiotemporal covariance in the CBI method based on analytical solutions (Jeong et al., 2025). For example, the work by Turner et al. (2020) did not estimate the covariance parameters in a way consistent with the data. Incorporating the spatiotemporal covariance as a core component of the inversion is a significant advantage of the GP method over the CBI method.

While the GP-based inversion has many advantages over CBI methods, the computational demands of the GP method increase significantly with larger datasets, potentially restricting its application in certain contexts (Williams and Rasmussen, 2006; Murphy, 2022), although recent development for high-performance computing (e.g., GPU-enabled tools) can alleviate this issue. Conversely, the linear inverse model, while less computationally demanding, assumes linearity and typically requires explicit assumptions about the underlying distribution, which may not always be valid and can lead to underestimation of model uncertainty (Wang, 2023). Overall, this study demonstrates the clear advantages of using GP-based inversion techniques and this modeling framework should be considered for application in future studies for constraining GHG fluxes when assimilating satellite retrievals."

2. The presentation of results:

- *Figure 2, I expect to see a time-series plot (if it can not be made for pixel scale, for regional scale also work), which can better show the performance of model optimization for prior/posterior and observations. Histogram plots for prior/posterior residuals (simulations minus observations) would also help.*

We agree with the reviewer that this timeseries and histogram information would be helpful for the reader. The figures below were added to the supplemental information and the following text was added to the manuscript: "A timeseries of daily co-located prior and posterior model predicted XCO_2 compared to OCO-2/3 observations during the year 2020 is also displayed in Fig. S6 (histogram of annual prior and posterior residuals displayed in Fig. S7)".

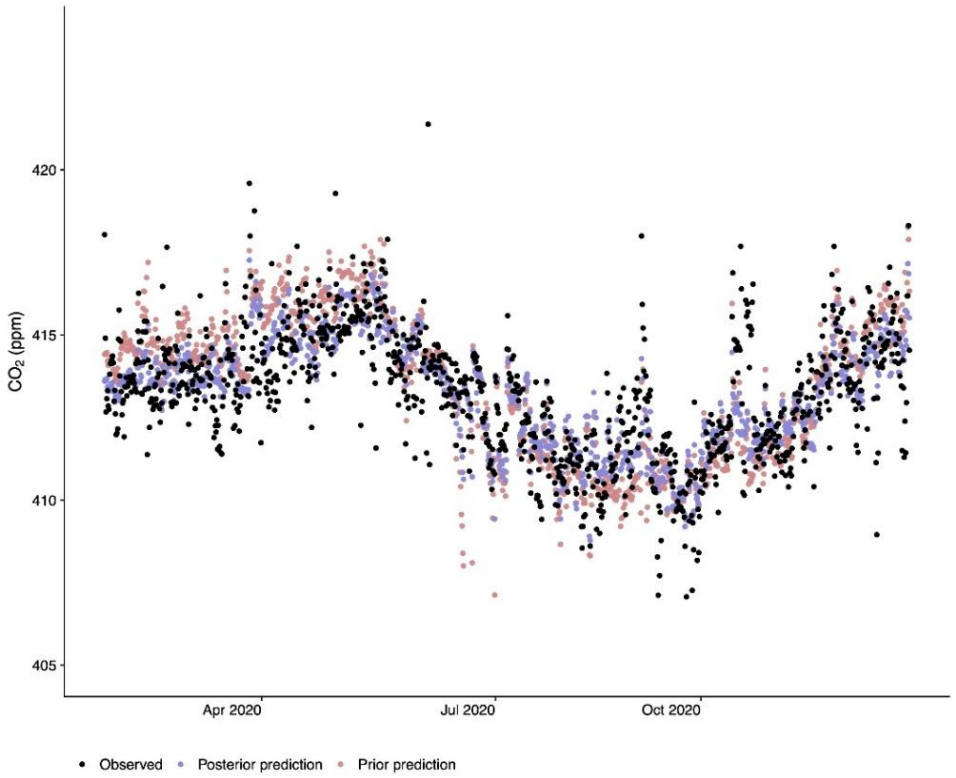

Figure S6. Timeseries of daily co-located prior (pink dots) and posterior (blue dots) model predicted XCO_2 concentrations (ppm) compared to OCO-2/3 observations (block dots) during the year 2020.

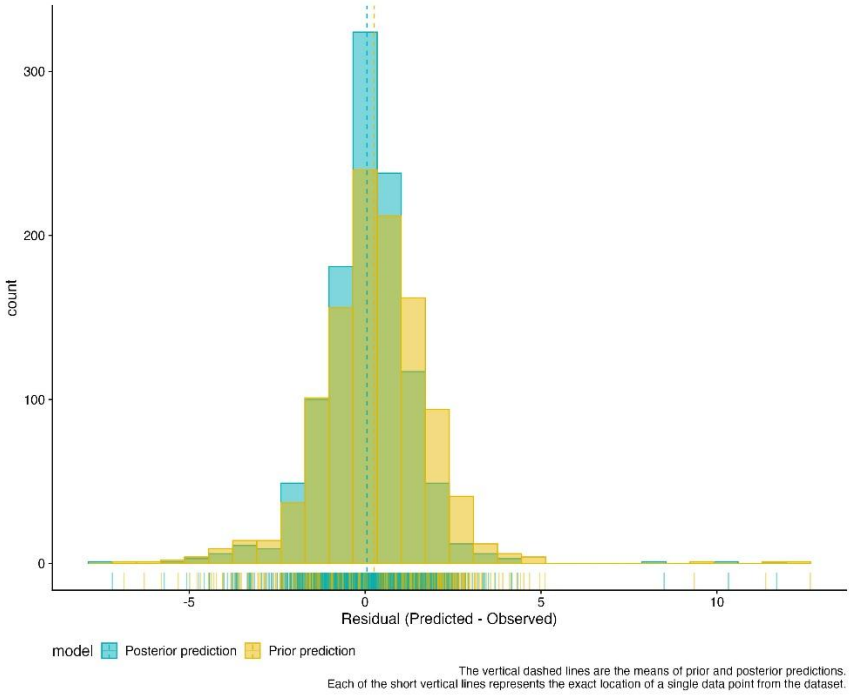

Figure S7. Annually-averaged error residuals for prior (orange) and posterior (blue) simulations (predicted – observed) of column-averaged XCO_2 (ppm) in prior and posterior model simulations for the year 2020.

- *Figure S1 showing the spatial distribution of the posterior fluxes in the supplemental files can be moved to the main text.*

Figure S1 shows the seasonally-averaged a priori CO_2 emissions used in GEOS-Chem simulations. The caption for this figure has been updated to state this. We have now added a map of the seasonally-averaged posterior emissions to the main body of the revised manuscript (new Fig. 3).

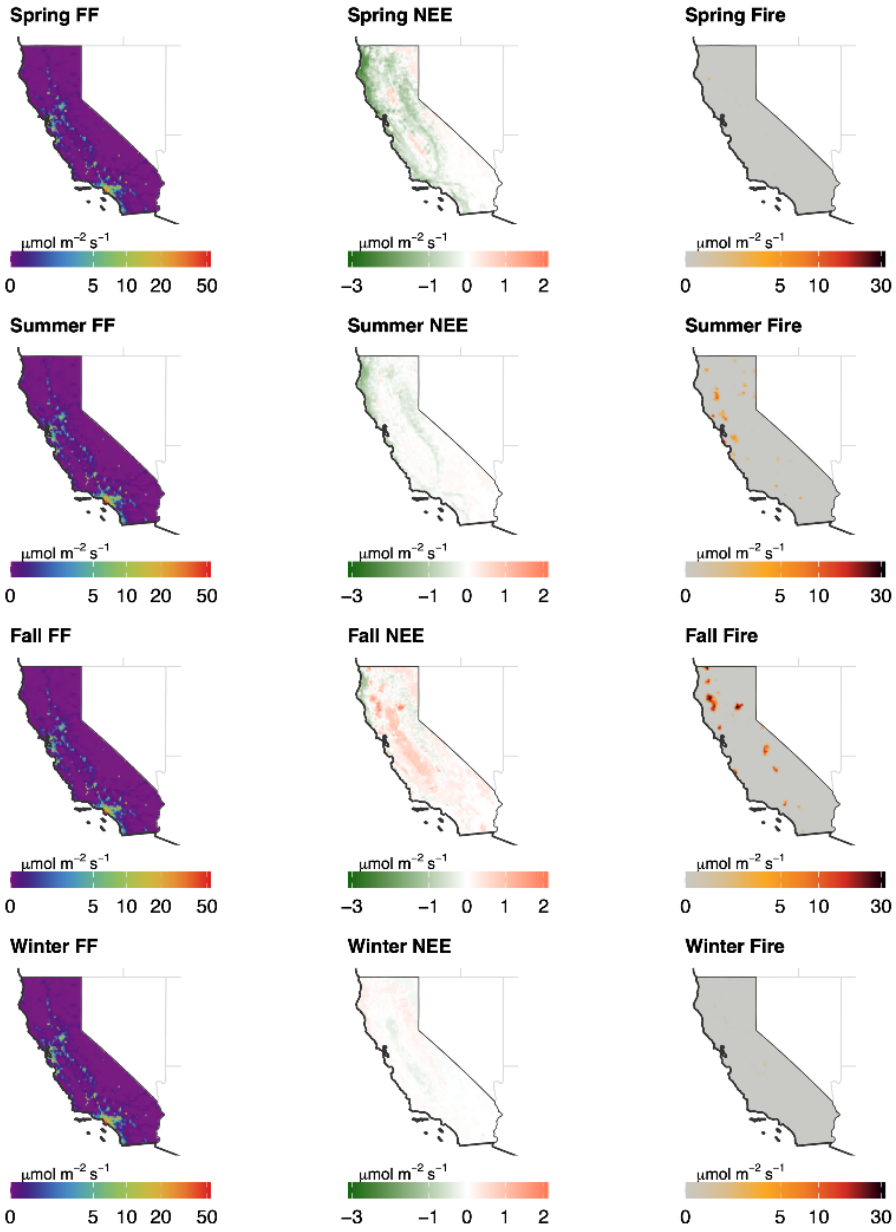

Figure 3. Seasonally-averaged 2020 posterior CO_2 emissions (μmol m^{-2} sec^{-1}) for the state of California. Emissions from the terrestrial portion of California are shown for FF (left column), Fire (middle column), and NEE (right column) for the spring (first row), summer (second row), fall (third row), and winter (fourth row) months.

- *I expect to see the seasonal cycle (with a monthly time-step) of the prior/posterior CO_2 fluxes, for fossil fuels, fires, and NEE. These would help to see if the constrained fluxes can indicate the impact of COVID-19, wildfires, and seasonal anthropogenic emissions. Currently, we only see the results by season.*

The figure below showing the monthly prior and posterior CO_2 emissions for California was added to the supplemental information. The addition of this figure resulted in the authors adding a couple

additional findings to the revised manuscript specific to the year 2020 for fossil fuel and NEE CO_2 fluxes such as: 1) "Interestingly, from the monthly-averaged state-wide emissions shown in Fig. S8, it can be seen that some months in the spring, summer, and fall of 2020 had posterior fluxes further reduced compared to the prior emissions emphasizing the strong reduction in GHG emissions due to COVID-19 lockdown restrictions" and 2) "From Fig. S8 it can be seen that during some of the summer months posterior NEE fluxes were near neutral compared to the large uptake suggested by prior fluxes. This is likely due to the strong drought and hot temperatures experienced in California during 2020 greatly reducing the CO_2 uptake during the growing season".

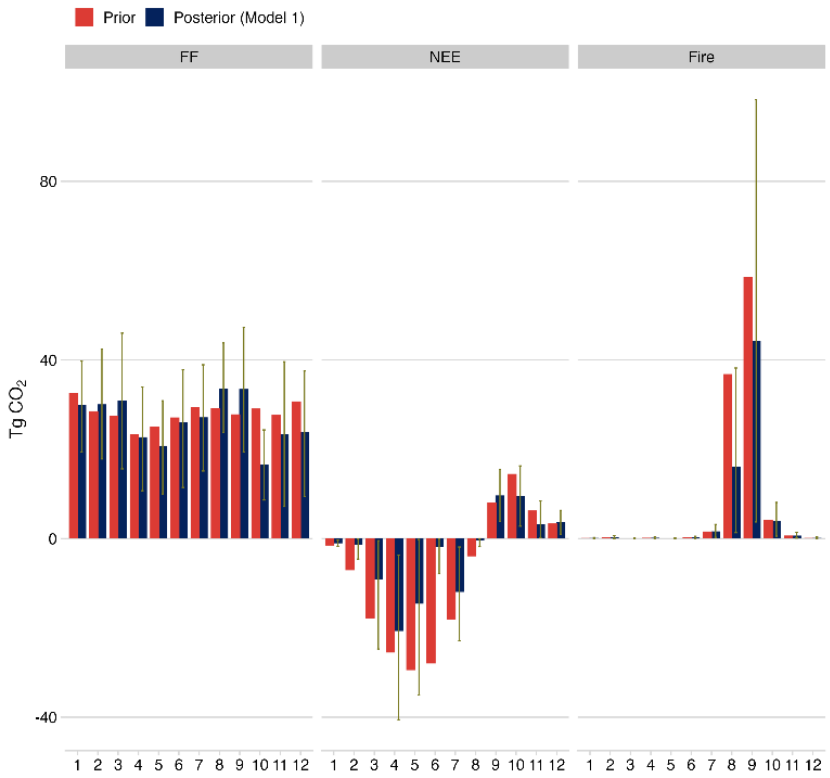

Figure S8. Monthly prior (red) and posterior (blue) CO_2 emissions ($TgCO_2\ yr^{-1}$) in California from fossil fuel, NEE, and fire fluxes during the year 2020.

3. Is it possible to perform one-year more inversion? So then we can better understand the performance of the inversion model in revealing the impact of disturbance from COVID-19, wildfires, and droughts.

We appreciate the reviewer's request to perform additional years of inverse model simulations. However, the authors feel this is outside the scope of this initial work. This specific study focuses on the demonstration of the GP/ML technique for constraining CO_2 emissions using satellite retrievals. Due to computational expense, we selected a single year with anomalous emission features in California (i.e., 2020) to demonstrate this technique. Multi-year studies using this inversion technique should however be conducted in future studies.

4. In Section 2.1, I see the boundary conditions were taken from GEOS-chem 4D-Var run at 4° ×5°, it is relatively coarse compared to the resolution of 0.5° × 0.625° for the inversion. I am not sure if this leads to some uncertainties for the inversion. Some higher-resolution BC, e.g., from CarbonTracker, might be better for the current inversion. Or some tests about the sensitivity of BC can be added.

Thank you for the comment. We agree that the boundary conditions from a higher resolution product could potentially reduce the uncertainty associated with the boundary conditions. To resolve this potential bias, including the background bias, we use a bias term (**D**) as shown in Eq. (5) of the manuscript. To provide more information about the bias, we added a new figure showing the distribution of the bias hyperparameter for each month in the supplemental information of the revised manuscript. We also added the following text in the revised manuscript [right after Eq. (5)].

"We show the probability density function of the estimated bias hyperparameter by month in Fig. S3. The median values range from –0.99 to 0.71 ppm, depending on the month. As noted, this value reflects the combined bias arising from atmospheric transport, boundary conditions, or other potential sources of error. This approach to addressing model bias has been applied in previous studies (e.g., Jeong et al., 2017). In this work, we included the bias term in the mean function [Eq. (5)]. As in Jeong et al. (2017), we model the bias term (**D**) as a single component in the GP mean function due to the lack of prior information needed to separate it into identifiable sources (e.g., transport or boundary condition errors). Introducing multiple terms without such constraints would risk overfitting and model instability.".

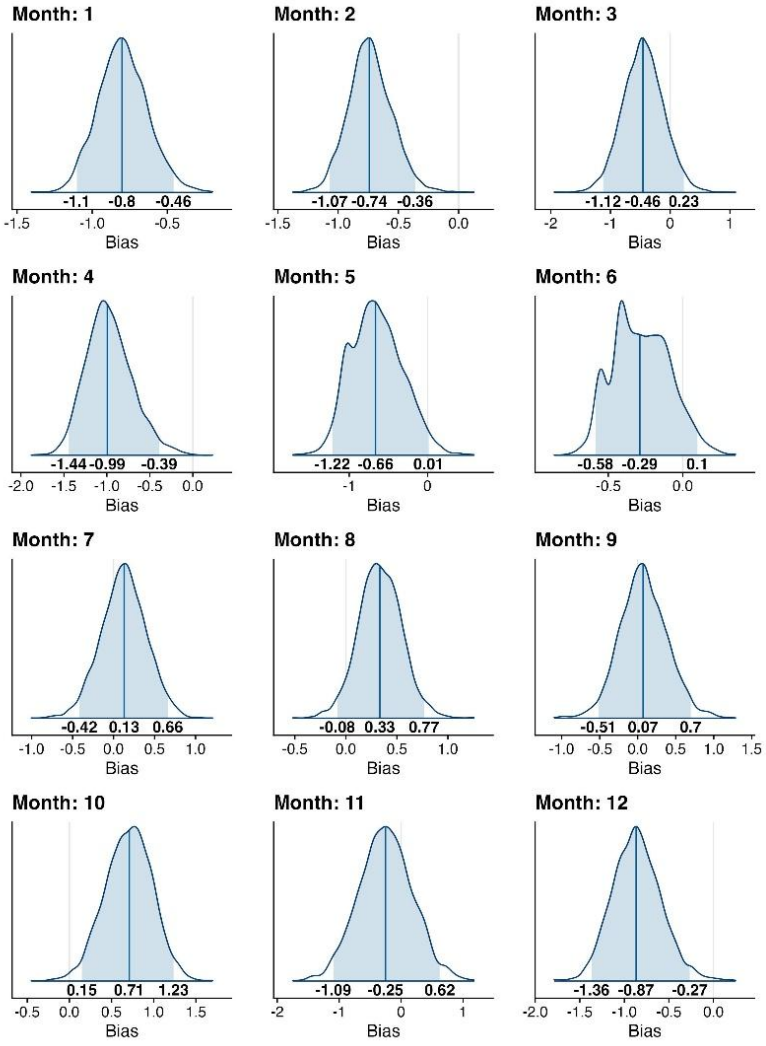

Figure S3. Probability density function for the estimated bias hyperparameter (**D**, ppm) in Eq. (5), shown by month. The shaded area represents the 95% confidence interval, while the vertical line indicates the median value. The bold numbers at the bottom represent the 2.5th, 50th, and 97.5th percentiles.

Minor comments:

1. Line 169, constraining-> constrain?

This was corrected.

2. Section 2.3, I expect to see a spatial map showing the data coverage of OCO-2 and OCO-3 XCO_2 observations over the study area.

We thank the reviewer for this comment. We agree that a figure of the observational coverage of OCO-2+OCO-3 would be useful for the reader. The figure below has been implemented into the supplemental information of the revised manuscript.

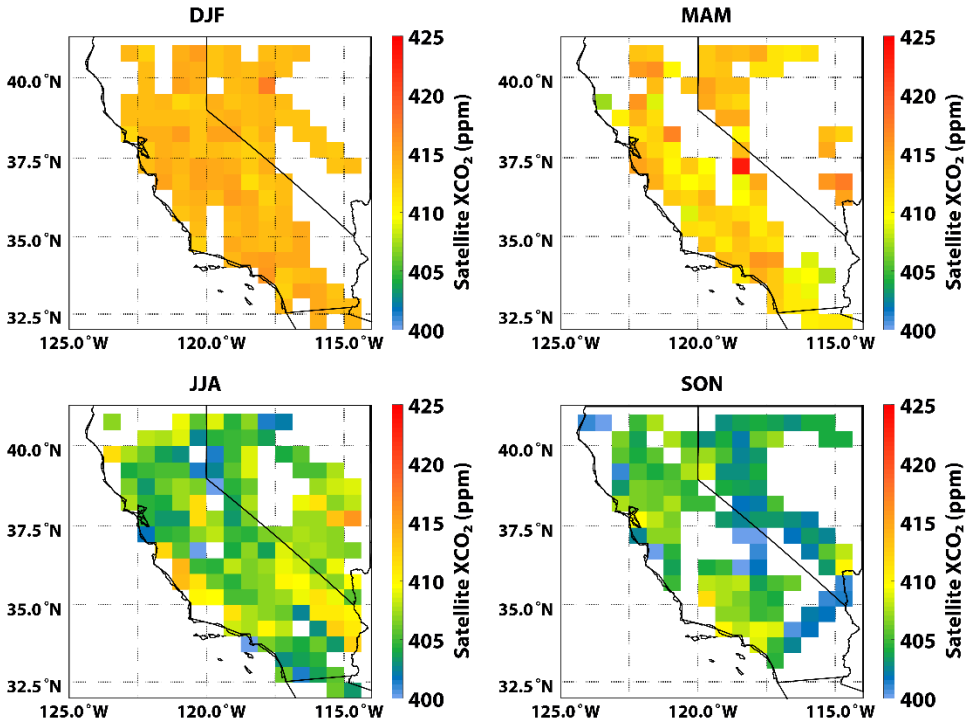

Figure S2. Seasonally-averaged XCO_2 concentrations (ppm) retrieved by OCO-2+OCO-3 during the year 2020 presented at the forward model spatial resolution of 0.5° × 0.625°.

3. How to optimally determine the hyperparameters of the GP/ML model? It is not clear.

We describe how the hyperparameters were determined in Sect. 2.4 of the original version of the manuscript and provide more details in Text S2 and Text S3 of the original supplemental information document. To improve the description of how this was done, we have now added the following text in Sect. 2.4 of the revised manuscript: "Using the prior distributions specified in Text S3, we inferred the posterior distributions of the GP hyperparameters using the NUTS algorithm, implemented via the PyMC framework. In our approach, all hyperparameters are estimated jointly. That is, the scaling factors are estimated simultaneously with other parameters, such as the kernel length scales (Jeong et al., 2025). This method enables full posterior inference of the hyperparameters, offering a more robust characterization of uncertainty than point estimation methods. More details on hyperparameter optimization for the GP model are provided in Jeong et al. (2025)."

References

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, J. Geophys. Res., 108, 4757, https://doi.org/10.1029/2003JD003770, 2003.

Jeong, S., Hsu, Y. K., Andrews, A. E., Bianco, L., Vaca, P., Wilczak, J. M., and Fischer, M. L.: A multitower measurement network estimate of California's methane emissions, J. Geophys. Res.-Atmos., 118, 11339–11351, 2013.

Jeong, S., Cui, X., Blake, D. R., Miller, B., Montzka, S. A., Andrews, A., Guha, A., Martien, P., Bambha, R. P., LaFranchi, B., Michelsen, H. A., Clements, C. B., Glaize, P., and Fischer, M. L.: Estimating methane emissions from biological and fossil-fuel sources in the San Francisco Bay Area, Geophys. Res. Lett., 44, 486–495, https://doi.org/10.1002/2016GL071794, 2017.

Jeong, S., Hamilton, S. D., Johnson, M. S., Wu, D., Turner, A. J., and Fischer, M. L.: Applying Gaussian Process Machine Learning and Modern Probabilistic Programming to Satellite Data to Infer CO_2 Emissions. Environmental Science & Technology, 59, 9, 4376–4387, https://doi.org/10.1021/acs.est.4c09395, 2025.

Johnson, M. S., Xi, X., Jeong, S., Yates, E. L., Iraci, L. T., Tanaka, T., Loewenstein, M., Tadic, J. M., and Fischer, M. L.: Investigating seasonal methane emissions in Northern California using airborne measurements and inverse modeling, J. Geophys. Res.-Atmos., 121, 13753–13767, https://doi.org/10.1002/2016JD025157, 2016.

Roten, D., Lin, J. C., Das, S., and Kort, E. A.: Constraining Sector-Specific CO2 Fluxes Using Space-Based XCO_2 Observations Over the Los Angeles Basin, Geophysical Research Letters, 50, e2023GL104376. https://doi.org/10.1029/2023GL104376, 2023.

Turner, A. J., Kim, J., Fitzmaurice, H., Newman, C., Worthington, K., Chan, K., Wooldridge, P. J., Köehler, P., Frankenberg, C., and Cohen, R. C.: Observed impacts of covid-19 on urban CO_2 emissions, Geophys. Res. Lett., 47, e2020GL090037, https://doi.org/10.1029/2020GL090037, 2020.