Respond to Reviewer #1

This study presents an analysis of CO2 mole fraction data collected at the ZOTTO measurement site in Siberia. These data are an incredible resource for the community. The authors use a variety of statistical tools and an inverse method to interpret these mole fraction data and come to some conclusions – some that agree and some that appear to be at odds with previous studies. I have noted some weaknesses in their analysis that need to be addressed before I can support this study being accepted for publication. There are a number of grammatical errors that should be addressed before publication.

Thank you for the constructive review and suggestions. We detail our responses below in blue. All references to the line number in our blue replies correspond to the line number in the track-change file.

Line 52: this list is not comprehensive. These authors appear to be leaning on CUP/CRP ideas presented elsewhere but not cited properly.

Thank you for pointing this out. We have added (Keeling et al. 1996; Pearman et al., 1981; Bacastow et al., 1985; Myneni et al., 1997) to the reference list of studies (line 54) that used long-term CO_2 records to monitor the dynamics of carbon exchange in northern ecosystems.

This reviewer notes that some of the primary references are used by Kariyathan et al, 2023. This is important because the primary references outline some of the issues raised by the present study. There is no theoretical way to remove a linear trend from any time series that describes a range of stationary and non-stationary processes – there will always be leakage between these variations irrespective of the method used to disentangle the individual signals. As a consequence of this, it is also difficult to directly split apart changes in maxima and minima associated with seasonal cycles. Although the authors have used error correlations to study some seasonal relationships. The CCGCRV method is a complicated but effective tool that involves fitting a series of harmonics but it cannot easily address non-stationary processes, which is most noticeable during rapidly varying environmental conditions that affect the amplitude and phase of the time series. This should at least be acknowledged somewhere, particularly because the authors study the influence of a heat wave and anomalous fire year.

We agree with the reviewer that the analysis of atmospheric time series is often a complex process because the data are usually highly autocorrelated and consist of periodic and irregular variations on both long and short timescales. However, as the reviewer mentions, while there is a need to apply filtering and curve-fitting techniques to obtain smooth and continuous data, there is no perfect way to remove a linear trend or diagnose changes in non-stationary periodic cycles such as seasonal cycles from any time series.

The first limitation of curve-fitting is that the application of a particular curve-fitting program in the analysis or decomposition of an atmospheric time series may introduce biases that could significantly influence the results and conclusions of an investigation (Nakazawa et al., 1997; Tans et al., 1989; Pickers and Manning, 2015; Barlow et al., 2015). Already in the previous version of the manuscript, we had created an ensemble of different parameter settings of CCGCRV as well as compared the influence of two different curve-fitting methods on our analyses. These results suggested that there were no significant differences in terms of trends and signals when we alternated different CCGCRV parameters as well as compared them with the HPspline program.

The second limitation mentioned by the reviewer is that key aspects of the seasonal cycle of a CO_2 time series are sensitive to different curve-fitting approaches since neither CCGCRV nor other existed harmonic-based curve-fitting methods can properly address non-stationary processes. In the revised version of the manuscript, we have acknowledged this limitation more clearly (Section 3.1 lines 334-337). However, we note that even considering this second limitation of the existing curve-fitting method, the smoothed time series derived from CCGCRV for ZOTTO data still represented the influence of 2012 summer wild-fire well (lower CO_2 concentration in summer 2012, Fig. 4b in the manuscript). This implies that our method can identify anomalous seasons, even if we have to acknowledge that the absolute magnitude of change may be somewhat different with a different curve-fitting method.

This reviewer is more than a little concerned with their method used by the authors to calculate their linear slopes, given the large year to year variations. Values will be disproportionately influenced by outliers. More robust methods include the Seigel or Theil-Sen estimators. These methods will provide a more rigorous assessment of any observed trends, particularly given the length and noisiness of the ZOTTO time series. As a consequence of using a simple linear regression, this reviewer is wondering whether the results presented will remain statistically inconsistent with previous studies. Certainly, eyeballing some of their figures it is hard to imagine assigning any non-zero trend. Depending on what they find, using a refined method to determine trends may influence results described later about correlations between seasonal temperature anomalies.

We agree with the reviewer's assessment and will change the results of the manuscript accordingly. We have applied the Theil-Sen methods to all our trend analyses. As in the Fig R1. below, the trends and their p-values after applying the Theil-Sen method do not change that much from the previous version of the manuscript. We have revised all our linear trend graphs (Fig. 5, 7, 9, 10, C1, C2, G1, G2, H1, H2, H5, H6) to Theil-Sen trend graphs.

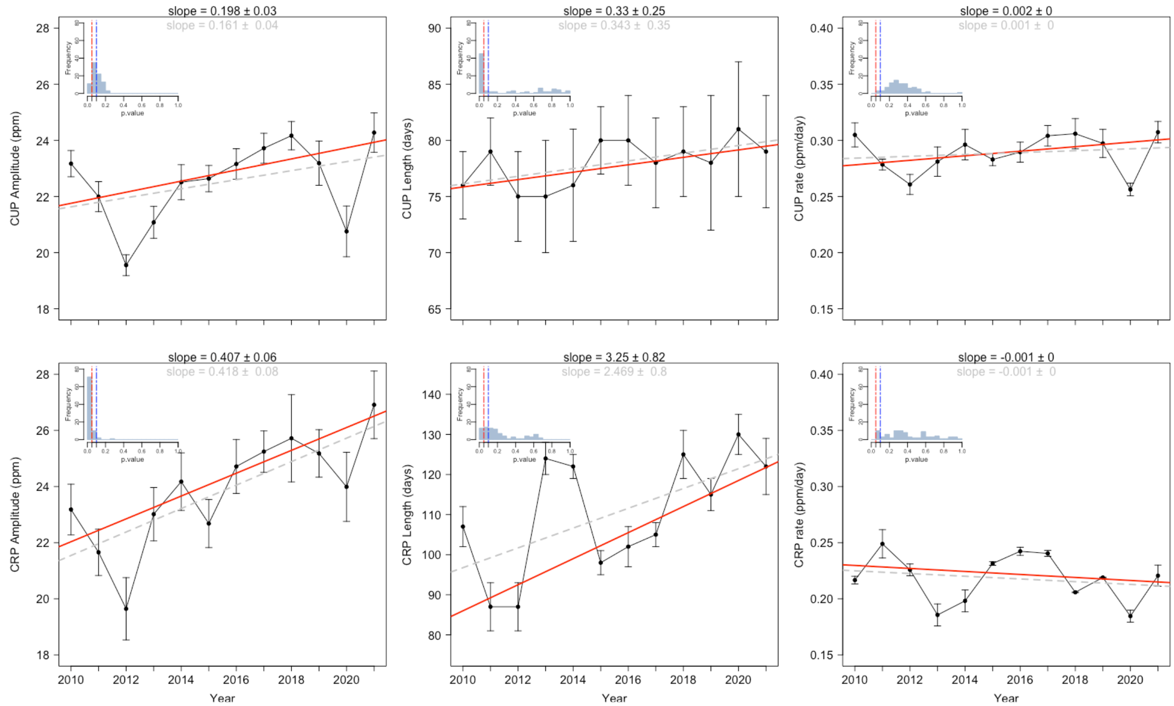

Figure R1. Trends and p-values of CUP and CRP amplitude, length, and rate after applying the Theil-Senn method, the grey colour results from the previous linear trend analysis.

Re inversions: these authors will be well aware that translating changes in atmospheric mole fractions to regional CO2 fluxes is complex, which in this case involves seasonal variations in atmospheric transport from lower latitudes. The imbalance between the size and distribution of observation networks at higher and lower latitudes may also render the posterior solution problematic. Is there a noticeable improvement in the model performance in describing the CUP and CRP metrics from the prior to the posterior fluxes?

The prior is the mean seasonal cycle of an inversion, such that the CUP and CRP calculated from it should not contain any year-to-year variations of the CO_2 flux, in particular no interannual variation similar to the observational data. If the CUP and CRP diagnosed from atmospheric concentrations simulated by forward transport of the prior contain any interannual variations, then these would come from either transport variability or from inhomogeneous sampling. We run the inversion from the prior fluxes and do the same comparison as in Figs H2-3 in the Supplement material. We did not find any inter-annual variations in the prior analysis and have added additional Figures (Fig. H5, H6) and discussion (lines 542- 545) in the supplement materials about this.

On a related note, this reviewer is not surprised by the influence of including this one site into a global inversion. Despite claims to the contrary, there are large gaps in our knowledge about the carbon cycle at high northern latitudes and across the tropics. It would be useful for this reader to understand what had to change elsewhere (via mass balance) due to this albeit small change in the regional carbon balance due to using ZOTTO data.

We understand that since the global carbon budget is closed at annual and longer timescales, adding ZOTTO data to the inversion affects regional uptake, with compensations in the rest of the world, dominantly at high northern latitudes and tropics as shown in Figure G4 in the manuscript. We pointed this out already in the previous version of the manuscript (lines 400-403): "This indicates that adding ZOTTO data shifted the estimated carbon uptake within the NH (Fig. G4)". To avoid misunderstanding, we will rephrase this sentence to "Since the global carbon budget is closed at interannual timescales, adding ZOTTO data to the inversion altered the estimated carbon uptake within the rest of the world accordingly to conserve mass, leading to higher carbon uptake spread widely across the NH tropical and mid-latitudes 20°-50° N."

Last point, re conclusions: Improved flux estimates will also come from satellite data, collected primarily during summer months when the observing geometry is favourable but also during other months via improved regional estimates of resolution transport model may improve regional estimates using sparse ground-based data, but whether the net impact is positive is debatable.

Thank you for pointing that out. We have expanded our conclusions to note that due to the sparseness and uneven distribution of the monitoring surface networks, it is unclear whether a higher-resolution regional transport model alone may better constrain regional fluxes (lines 429-434).

As the reviewer mentioned, satellite observations provide a more extensive and homogeneous coverage than in situ networks. Since they quantify a column-mixing ratio, and not only a local one like surface measurements, they provide a constraint on a larger fraction of the atmosphere than surface observations. This is considered an advantage since it should lead to a more complete representation of the atmosphere, but it also makes inversions using satellite retrievals more sensitive to model errors in the upper troposphere and stratosphere (Monteil et al., 2013). Important drawbacks of satellite retrievals are that they are available only for a limited range of atmospheric conditions (absence of clouds, low aerosol load) and that they are less accurate and more difficult to validate than in situ measurements. Most importantly, since the satellite data are constrained to mainly summer months at high latitudes, their data are of limited value to constrain the full seasonal cycle in inversion models given the lack of constraint on high-latitude on the full seasonal cycle and its phasing. Therefore, it is still unclear whether flux estimated from satellite data, where ground-based observation is sparse, will provide an improvement from flux estimated from ground-based observation. We will include this consideration in the conclusion section of the manuscript.

Minor points

- Line 162: which 78 sites? Suggest they are listed somewhere.

The 78 sites were represented in Fig. D1, an additional table to a list of these stations has been provided in the supplement material of the revised manuscript as Table D1 in the Appendix D section (line 477).

- Line 263: 1958 to 1961?

The sentence “Our finding is consistent with Graven et al. (2013), using aircraft-based observations of CO_2 from 1958 to 1961 to show that the seasonal amplitude at altitudes of 3 to 6 km increased by 50% for high latitudes.” has been re-written as: “Our finding is consistent with Graven et al. (2013), comparing 2009-2011 aircraft-based observations of CO_2 above the North Pacific and Arctic Oceans to earlier data from 1958 to 1961 and found that the seasonal amplitude at altitudes of 3 to 6 km increased by 50% for high latitudes.” (lines 270-273).

- Grammatical errors throughout. Worth a closer check by the authors when they revise their manuscript.

We agreed and will thoroughly address them in the revised manuscript.

Responses to Reviewer #2

This manuscript investigates the interannual variability of the carbon uptake and carbon release period in Central Siberia. The presented analysis is based on CO_2 observations carried out at the Zotino Tall Tower Observatory (ZOTTO) in central Siberia, during the period 2010 – 2021. This reviewer acknowledges the relevance of the ZOTTO dataset for the scientific community as it is collected in a region that is presently experiencing significant impacts of climate change.

Results based on CO_2 measurements revealed that amplitude and length of carbon release and carbon uptake period increased during the analysed period (2010 – 2021). However, data show that the growth of the amplitude of carbon release period is larger than the growth in carbon uptake period amplitude, suggesting that the enhanced carbon uptake during the growing season was offset by the autumn/winter carbon release.

The manuscript is generally well written, and presented results are an important contribution to the knowledge on the effect of climate change on terrestrial ecosystems.

Thank you for the constructive review and suggestions. We detail our responses below. However, some aspects of data analysis and related discussion should be clarified:

• Time series of the target tank is used to evaluate the quality of CO_2 mole fraction measured at ZOTTO. This reviewer noted a small jump in the time series of target tank between 2018 and 2019. The average value before the jump seems to be lower than the value for the following period. In the opinion of the reviewer, there could be a potential bias introduced in the ambient measurements. Could the authors comment on this?

Thank you for pointing this out. Unfortunately, we have not been able to find a satisfactory explanation for this jump. The values differ between the 2019-2022 (i.e., after the jump) average of the target tank is 404.67 ppm and the 2009-2018 average is 404.62 ppm. This bias of 0.05 ppm could have been the result of many smaller changes in the measurement set-up. We note that this bias is one order of magnitude smaller than the difference between flask measurements and the continuous in-situ data and three orders of magnitude compared to the averaged amplitude of the seasonal cycle of CO_2. We, therefore, believe that this bias is comparatively small and does not have a significant impact on our analysis.

• Lines 122-125. Authors are applying a despiking methodology to remove unreliable CO_2 observations but it is not clear to this reviewer the meaning of “unreliable”: is this the definition assigned to concentrations mainly affected by local sources? Moreover, this reviewer is wondering if the removal of unreliable data is impairing the capability of CO_2mole fraction dataset to detect the effect of extreme events. Finally, this reviewer advises authors to show the percentage of removed observations.

Thank you for pointing this out. This procedure removed “unreliable” outliers that we consider to be caused by local effects and which despite being relatively sparse, do not represent the large-scale seasonal variation that the fitting function should capture. Obviously, by removing those extreme local effects data points, we also removed the extreme events in terms of daily scale. However, our study focuses on seasonal and inter-annual scales. We have checked the removed data points, and they appeared to be random and are not particularly concentrated on

the time of the day or any particular seasons or years. This percentage of the removal data is 2.4% of the total data. After the noisy data removal process, we still could see the abnormal curve shape in 2012 (when an extreme event happened) in Fig. 4b (black line). We, therefore, would say that the removed unreliable data is insignificant in detecting the effect of extreme events in terms of seasonal and inter-annual scales. As suggested by the reviewer, the percentage and “quality” of the removed data points have been added to this section in the revised manuscript (line 130).

• Lines 135-136. Has the consistency of the assumption that there are not significant changes in the curve shape of the season over the years been tested?

We have not tested the significance of changes in the curve shape of the season. We mentioned the change in the curve shape of the season over the years is negligible (line 138) but have not addressed why. The ability to isolate changes in the phase and amplitude of the seasonal cycle with fidelity could be determined by using Monte Carlo numerical experiments as in Barlow et al. 2015. Their study shows that the errors associated with independently identifying changes in phase and amplitude that can result in the misinterpretation of seasonal signals are more pronounced when using the detrended CO_2 seasonal cycle as opposed to using the time derivative of a time series. This result informed our study. Our use of the time derivative of a time series can provide a more robust estimate of the key dates that define the CUP without taking into account the changes in the shape of the seasonal curve compared to the conventional use of zero crossing date derived from the detrended CO_2 seasonal cycle. We have added this discussion and reference to Barlow et al. 2015 study in the revised manuscript (lines 138-141 and lines 324 - 333).

• Section 2.5. The choice of the threshold value of the spatial root mean square (RMS), used to determine the region of influence, should be explained. Moreover, the impact of different RMS thresholds on the extension of the region of influence should be addressed. Finally, this reviewer suggests including a description of the land cover in the region of influence. This could help readers to get an idea of the type of ecosystems embedded in this area and affecting observations collected at ZOTTO.

Thank you for your feedback. We have tested different RMS thresholds on the extension of the region of influence, but it was not included in the manuscript. We tested varying the threshold between 10% and 50% and decided 40% would be the most presentable for the region of the influence for the ZOTTO station (Figure R2). The ecosystem cover in this region of influence comprises Pinus sylvestris forest stands (about 20 m in height) on lichen-covered sandy soils. As suggested by the reviewer, the description of the land cover in the region of influence has be added to the revised manuscript (lines 184 -185).

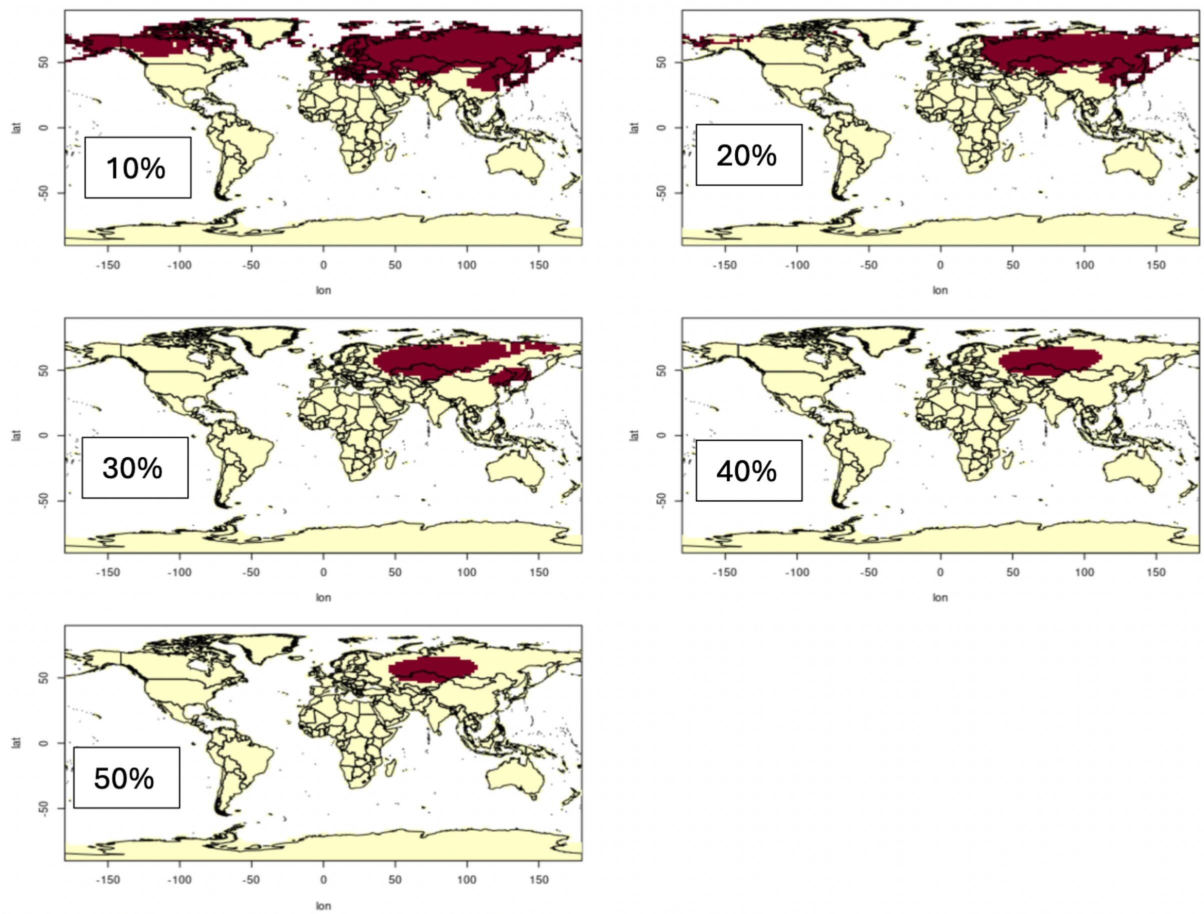

Figure R2. Dicerent RMS thresholds (10-50%) on the extension of the region of influence for the ZOTTO station.

• Section 3.1. Authors found that there is not a significant trend in the timings of CUP, while there is a significant increase of the CUP length. How is it possible to have an increase in the CUP length when the timings (onset and termination) are not changing? Moreover, authors are stating that the heat wave in 2020 induced an early onset of CUP, but the error bar associated to the estimated CUP onset in 2020 (Figure 5) is very large, casting doubt on the author's statement.

One possible explanation for the fact that we did not see a significant trend in the timings (both onset and termination) of CUP but the CUP trend is that: the changes in timings (slight decrease and increase trend for onset and termination respectively) are too small, the time series of 10 years is not long enough to see significant changes. CUP length is the result of both onset and termination. Therefore, slight insignificant changes in both onset and terminations could still result in more visible significant changes in CUP length over 10 years. We also did additional regression of onset and termination timings for both CUP and CRP (Fig. R3 below). There is no clear regression on both cases.

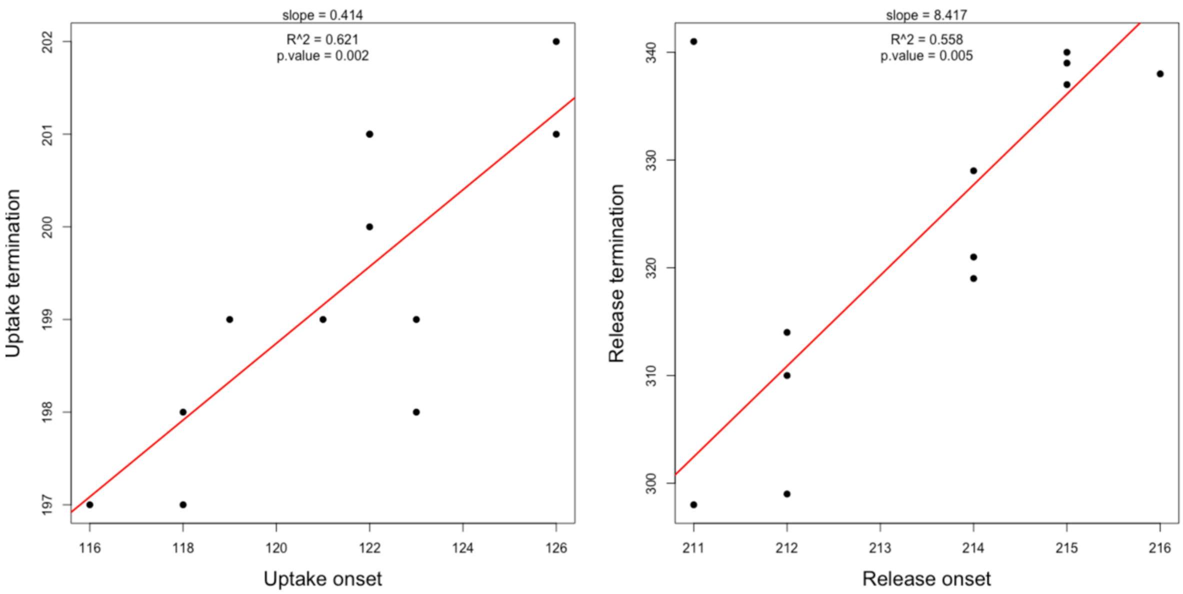

Figure R3. Regression of timing of onset and termination for CUP and CRP.

• Finally, authors are claiming for a significant jump in the CUP length in 2020 but looking at Figure 7 the jump is visible in the CUP amplitude (and CUP rate), not in the CUP length.

We understand that the use of "a significant jump in the CUP length" in line 305 is misleading. What we meant was that during 2020, when the Siberian winter-to-spring heatwave occurred, there was only an increase in the CUP length due to the early spring onset but not in the CUP amplitude. In line 305, "a significant jump" has be replaced with "an increase" in the revised manuscript.

Minor points to be addressed are listed below:

• This reviewer advises authors to cite the data repository where they retrieved the CO_2 mole fractions measured in atmospheric stations used for both inversions s10v2022 and s10v2022+Allstations.

A list of stations and their data repository have been added in the Appendix D (line 477).

• Lines 120-122. Ranges of short-term and long-term cut-off values tested are different from those reported in Table B1.

This has been checked, and correct values have been added.

• Line 205: change "later" to "layer".

"later" to "layer" has been changed in line 211.

• Caption of Figure 4: change “Thoning et al. (1996)” in “Thoning et al. (1989)”.

“Thoning et al. (1996)” has been changed to “Thoning et al. (1989)” in the caption of Figure 4 (line 223).

• Line 374. Add “are” after “s10v2021+ZOT”.

“are” after “s10v2021+ZOT” has been added in Line 395.