We would like to thank the reviewer for her/his evaluation of our manuscript and for the suggestions that were given. We have carefully gone through the points that were raised. In the following, we provide a response in blue colour.

**Anonymous Referee #2**

**Received and published: 11 January 202**

Thank you very much for letting me read this very interesting manuscript. CO2 flux sites are too few at the African continent, so your work providing two new sites in Africa is of very high relevance for all of us working within both the Eddy covariance community, and within general climate and Earth system sciences. This manuscript set out to investigate the impact of different grazing regimes on land atmosphere of CO2. And having this very nice data for two adjacent semi-arid savannah site in South Africa is a fantastic opportunity to make some very interesting analysis. This manuscript thereby has great potential in increasing our understanding in carbon cycle dynamics for semi-arid savannah landscapes. However, I have some major concerns regarding the presentation of the manuscript as outlined below.

**General comments:**

1) As mentioned, the main aim is to study the impact of grazing by comparing two adjacent sites with similar meteorological and hydrological conditions but with different grazing regimes. However, no real analysis comparing the two sites is provided.

The manuscript will be reorganised with a few new figures, but also some of the key figures will be kept. These are Figs. 2, 3, 5, and 9. They already provide a direct comparison of either environmental conditions or fluxes. The new figure will further stress a direct comparison NDVI between sites (Year I & Year II) and seasons (Dry & Wet).

It is claimed that there is a significant difference. But not uncertainty estimates around either the budgets, or the environmental conditions are provided, and it is hence impossible to see if the differences are significant. I doubt that there is a significant difference between the two sites, given the high variability of the fluxes, and that the fluxes of the two sites still are relatively close to each other.

Uncertainty estimates were made for cumulative monthly, seasonal and annual fluxes. We used an honest and transparent way of reporting uncertainties, it is in our opinion mostly a discussion whether the two sites can be called "different" or not and then of course whether the conclusions we draw hold. Nevertheless, we will add further analysis investigating the differences between two sites. See our detailed response to major point 2.

You hypothesize that the HG regime reduces the ecosystem carbon sink potential by altering vegetation cover, decreasing above-ground biomass (AGB) and gross primary production. But no test of this hypothesis is presented. Such an analysis must be provided in order to draw the conclusions presented in the manuscript.

We understand the reviewer's point. We will review the interpretation of NDVI by doing a statistical difference test for the NDVI data between sites, Years I-II & III-IV and the growing & dry seasons. Also, we will add bar plots to compare NDVI between sites (Year I & Year II) and seasons (Dry & Wet).

2) Instead of focusing the results on the main aim, that is to study the impact of the grazing on the budgets comparing the sites, the results are basically just one long report of various CO2 flux budgets for different temporal averaging periods. No results of an actual analysis comparing the sites is provided. It is fine to have a section in the beginning of the results describing the hydrological and meteorological conditions as well as a first presentation of the fluxes. However, while reading the results I was continuously waiting for the actual results to start. The main focus of a results section should be to fulfill the aims set out in the introduction. As it is now it is too much focus on reporting flux budgets, and to o little on comparing the difference between the sites. I would recommend to streamline the results substantially, and move quite a lot of the currently presented results/figures to an appendix/supplementary information.

We would like to point the reviewer to Figs. 2, 3, 5, and 9, which already provide a direct comparison of either environmental conditions or fluxes. In the revised version of the manuscript, we plan to include statistical tests to demonstrate whether or not datasets are significantly different. Analyses will cover differences between sites, differences between growing seasons of sites, differences between years, probably binning year 1 and 2 vs. year 3 and 4 to take different grazing into account. Our conclusions will be aligned accordingly. We will also develop an aspect of the relationship between the water availability and fluxes by introducing additional graphs (SWC/P & fluxes).

3) The Introduction and methods section reads very well, but both the Results and the Discussions must be streamlined with the aim of the paper. The conclusions drawn in the discussions are not firmly based on the presented results (see further comments below).

We will perform further statistical tests, introduce two new figures (see our comments above), insert an additional table (with annual cumulative NEE, GPP and Reco) and move Figs. 4, 7, 8 and Tables 2, 3, 4 to the appendix. We will further revise and streamline the results, discussions and conclusions in relevant parts.

**Minor comments:**

Include standard deviation of the quantified sink and sources (L21).

We will include the standard deviation.

If “The two sites differed in soil heterogeneity and characteristics particularly in stone content (soil skeleton >2 mm for the HG site)” (L131). Should this not have a substantial influence on the difference in the CO2 flux budgets between the sites?

There was indeed a slight difference in the soil between the sites with unfortunately no detailed soil surveys available. We assume, however, that the small difference in soil characteristics may not play a significant role in driving differences in $CO_2$ fluxes between the two sites. Inter-site differences were only observed in year I and II. In these years, grazing intensity at each site remained the same, i.e. low at LG site, resting at HG site, and amount of rainfall was more or less the same. Thus, we attribute the inter-site differences to grazing and the inter-annual variability between year I and II to different rainfall distribution. We will further clarify this in the manuscript. Soil texture may surely have an influence on physical components like diffusion of $CO_2$, but we don't think that in these soil types it plays a major role as it can be seen in our results.

(L186-189) Why was it decided to use the night time partitioning method? Is there a strong relationship between CO2 fluxes and temperature? I think in general the respiration-temperature relationship is pretty weak for semi-arid ecosystems. I think the daytime partitioning method is better under these circumstances.

We used nighttime partitioning because it has the advantage that GPP+Reco = NEE, which is not the case for daytime partitioning, where GPP may become negative in single cases due to uncertainty in the fit. However, we see the reviewer's point and logical reason behind it. We will follow his/her advice and use the daytime method instead. Table 3 and Figure 5 will be updated accordingly.

I do not quite understand how the systematic errors were included in the uncertainty estimates. In equation 3, only random and gap filling bias is included? Whereas at L194 it is stated that systematic errors associated with advection, flux divergence and tilt correction, were taken into account. Where in the results is the uncertainty estimates presented and used?

The approach by Finkelstein and Sims. (2001) of estimating random errors was applied in this study. Also, bias errors from gap-filling of EC data were considered. We will modify the L194 accordingly. Lucas-Moffat et al. (2018) and Moffat et al. (2007) describe the equations to calculate bias and random errors, which are then summed up to give a measure of uncertainty of annual sums. The uncertainty estimates were used after the ± sign for cumulative NEE.

Why use both MOD13Q1 and MYD13Q1? (L215) Would it not be enough with one of the products. What extra info is gained by using both Aqua and Terra time series? How were they combined, given that only one time-series is presented in Fig 2.

We thank the reviewer for this comment. The use of both Terra and Aqua data would have enabled us to increase the temporal resolution of our NDVI time series from 16 to 8 days. However, even though it was originally planned to incorporate both the MOD13Q1 (Terra) and MYD13Q1 (Aqua) products, we ended up analyzing the NDVI from Terra only. Hence, our data set description from L208 to L219 was not entirely correct and we will revise it accordingly. It now only mentions the Terra NDVI product (MOD13Q1) which we used for our investigation.

The footprint is very short (L260). How was it calculated? There is no description in the methodology.

The tower height is just 3 m. The footprint estimation was performed according to the “simple footprint parameterization” described in Kljun et al. (2004). We will add this description to the methodology section.

I would recommend to move the separation of hydrological years to the method section.

The division of hydrological years was written where it was used and discussed for the first time (Section 3.1), for the purpose of facilitating the reading and interpretation of the results.

A lot of figures and tables present the same results. In the interest of streamlining the manuscript I would recommend to move a lot of presented results to an appendix, and instead focus the results on an analysis comparing the two sites to see if a significant difference between the sites can be seen.

We thank the reviewer for this suggestion and will move Figs. 4, 7,8 and Tables 2, 3, 4 to the appendix.

Table 4, What is behind the ±? One standard deviation based on inter-annual variability? Or is it the uncertainty from the uncertainty estimates?

The sign ± represents uncertainty based on the description in Section 2.3.4.

Please include uncertainty around the cumulative fluxes of Figure 9. I would also recommend to skip the final figure covering the full study period, it is not really of importance how they differ over a 4-year period. One extra with the average year for both sites could be interesting, to see if the two sites on average differ from each other.

We thank the reviewer for making this point. We will include the uncertainties in Fig. 9 and add an extra one with four-year average values. However, we think it is useful to keep four years of cumulative NEE to more clearly describe the overall picture.

What was the following conclusion based on: "the two investigated grazing regimes under similar climate, soil conditions and topography have highly influenced plant species composition and vegetation cover leading to implications for their role as potential grazing areas and/or efficient CO2 sinks". I doubt that the vegetation cover between the sites is significantly different (NDVI Fig 2). The fluxes also seem to be very similar at diurnal (Fig 3); seasonal (Fig 5 and 6), and if including the uncertainty, most likely also the inter-annual scale.

We will add estimation of the plant species coverage (%) in order to emphasize differences in species distribution between sites. The plant cover varies a lot over time as is normal with semi-arid systems. Variation in cover is mainly due to germination and growth of annuals, and secondarily to growth of perennials, especially grasses. The species assemblages are known to be affected by long-term grazing regime.

(L381) Please include standard deviations around the budgets, to make sure that the sites are significantly different from zero (really being sinks and sources) and from each other.

We will add uncertainty estimations here .

(L384) How can we tell that there is a difference in Aboveground biomass. The difference in NDVI seems to be minimal, is there any way to test if the difference is significantly higher? How can we tell that it was caused by overgrazing in the past? Could it not be the current grazing as well?

Quantitative species inventories are conducted by the agricultural research institute GADI and available (however mainly unpublished; as referenced), and it was not in the focus of this study to conduct further inventories. We will add a table based on the inventories, listing the most common species in both sites, to help the reader get an idea of the impacts of overgrazing on vegetation. We will modify L 384 to emphasize differences in plant species between sites.

(L385) During most of the resting periods there is no difference between the sites, and the site difference does not seem to be dependent on if it is resting or grazing periods.

I do not understand how a conclusion regarding the impact of the long resting period can be drawn in the discussions. First, there is still grazing going on, so there is no way the impact of the long resting period from the current grazing regime can be separated. Secondly, previously in the manuscript it was stated that the effect of the heavy grazing was still evident and that the grazing that was continued after 2017 warranted that the HG site could be used a heavy grazed site. In this case the long resting period should not have an impact. The heavy grazing is continued from 2017 and onwards. Could it not rather be so that the heavy grazing increases the CO2 uptake? (Tagesson et al 2016 in reference list). If this now really is the case.

The Heavy Grazing (HG) site was grazed by Dorper sheep using a 2-paddock rotational grazing system (120 days grazing followed by 120 days rest) at stocking rates approximately double that of the recommended rate as part of an experimental trial from 1988 to 2007. The site was ungrazed 2008–2017 but did not recover (palatable species did not come back after resting period). The Dorpers were reintroduced at a similar stocking rate in July 2017 (we have 4 years of measurement (Nov 2015 - Nov 2019)). In conclusion, we said that a long resting period, along with the transition of species from palatable shrubs and grasses to unpalatable grasses, affects carbon fluxes. We cannot conclude that heavy grazing increases $CO_2$ consumption. It can be said that species composition at the HG sites has altered and has been unfavorable for Dorper grazing due to overgrazing in the past. Thus the HG site is considered agriculturally degraded. In the same time, there has been a shift to an increased abundance of unpalatable drought-resistant grass species, favorable for carbon sequestration in such water-limited ecosystems.

L410 Please explain. I cannot see a statistically significantly higher NEE for HG than for LG in Fig 3. Quite the opposite, I see no significant difference?

We meant that the HG site had higher carbon sequestration rates compared to the LG site. We rephrase this sentence to make it clear.

(L432) Where is it shown that the inter-annual variability is caused by rainfall/SWC? A start of the growing season with start of the rainfall is no real surprise, that is the general case for semi-arid ecosystems (without dense tree cover). But it is not shown it in any figure; there is no place where a start of the rainy season is linked with the start of the growing season. It is also claimed that the inter-annual budgets are caused by the rainfall variability, but no actual analyze of such a relationship is presented.

We will develop this aspect further and add an additional graph that will present the NEE, Reco and GPP responses to the SWC.

L448-L463 This is not a discussion, it is just a long repetition of periods of rainfall and CO2 fluxes. Please do some analysis instead, present the results in the results section and then discuss these results.

We would disagree on the positioning of this paragraph in the discussion section; it more descriptively explains the difference between the first and second year (when the annual precipitation was almost the same for both years) and why our fluxes (especially annual cumulative NEE) acted differently in the Year I and II with same water availability. Thus, this explains the importance of the distribution of precipitation during the year in such water-limited ecosystems.

- The conclusion that “the high ratio of unpalatable vs palatable species made this site less suitable for its current use as sheep pasture” is not a conclusion of the presented results. Please present results that allows to draw such a conclusion.If current conclusions that HG has significantly higher CO2 uptake than LG holds for a statistical test, why is the n the conclusion not that the grazing regime of HG is better than LG?

The first sentence is derived from prior knowledge on studies conducted at the sites, as presented under the material and methods section. The reason we present it here is that it puts the $CO_2$ uptake finding in a more interesting light, and thus, is necessary background. We concluded that, unexpectedly, the site with continuous heavy grazing after a long resting period was more efficient as a $CO_2$ sink (due to transition to unpalatable grasses, which may be better able to compete for water). However, at the same time the value of the HG site for sheep grazing is reduced. Previous studies indicate relatively slow recovery from grazing: Seymour et al. (2010) reported that 20 years of recovery period in the Karoo degraded ecosystems restored grazing potential, while not returning all palatable species. As our study is conducted on sites where livestock grazing is conducted and studied, we believe that this provides an interesting additional angle to the interpretation of our results.