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. 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. 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.

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 too 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.

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

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?

(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.

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?

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?

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

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

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.

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

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.

(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.

(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?

(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.

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?

(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.

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.

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 then the conclusion not that the grazing regime of HG is better than LG?