Comments from the reviewers are in black and responses from the authors are in blue. Responses from the authors are sometimes in the past tense because the authors have already been editing soon after receiving comments from the reviewers.

Authors' responses to the Reviewer #2

The topic of the submission is undoubtedly of interest to the readership of Biogeo-sciences Discussions as it presents some interesting data arising from an experimental manipulation and its impacts upon ecosystem functioning plus gas exchange in a floodplain. Of particular note is the attention to phenomena during the frozen season as well as the thaw season. However, my greatest criticism of the overall project is why investigate CO2 exchange in such temporal and spatial detail and yet completely ignore methane, and for that matter nitrous oxide exchange and changes both spatially and temporally. It is surely the balance of changes between these contrasting greenhouse gases, of differing radiative forcing strength and atmospheric concentrations that is key here? I miss this vital context entirely within the paper as currently submitted.

Please see also our response letter to the review #1 for a statement regarding this topic. The authors agree with the reviewer that methane and nitrous oxide are indeed important greenhouse gases. However, N_2O was not covered within this study, while CH_4 results were not included in this manuscript to avoid too much material for a single manuscript. Therefore, the objective of this study is to investigate drainage effects on exclusively the CO_2 fluxes, not all important greenhouse gas fluxes or budget/balance. This separation is necessary since we found substantial changes in ecosystem structure that can be linked to significantly dryer conditions in the manipulated areas, which in turn result in complex changes in CO_2 cycle processes. For example, we observed significant shifts in vegetation community structure, which substantially altered both GPP and ER, and a detailed analysis of these opposing effects is required to understand the subsequent shifts in NEE. Thus, changes in CO_2 fluxes itself is important to research.

The abstract is of a suitable length and is reasonably informative, but I do question the concluding remark. It is not sufficiently explanative in relation to the findings of the experiment.

It was corrected to clarify the intended message.

The remainder of the text is rather lengthy and could be substantially consolidated without losing any impact (in fact making it much more impactful).

Result and discussion sections are to be combined and some paragraphs/sentences will be corrected to make the manuscript more concise.

The number of figures is verging on excessive. They could be perhaps consolidated and some relegated to appendix/supplementary materials, providing the key focus the manuscript currently lacks?

The authors deleted 3 figures in the main text.

Page 2 line 25 onwards (and subsequent incidences) –the use and comparison of cumulative figures for 20 days in year 1 and 66 days in year 2 does not seem to me to be sufficiently clear to the reader.

Due to the limited period of field campaign, the measurements did not cover the whole growing season in both years. The flux interpolation was carried out for 20 and 66 days for each year because all necessary parameters were measured in parallel with fluxes during those time periods and the authors wanted to maximize the interpolation periods. To make the fluxes between two years comparable, the authors added results in the format g C m^{-2} d^{-1} as mean flux rates within the given observation periods. Of course even with this format, differences related to the length and/or parts of observation periods still have to be considered carefully when interpreting differences, but absolute values are not influenced anymore by the length of the time series.

The premise and importance of the study are generally both well explained within the introduction section of the manuscript.

Line 114: 'Reliable Prognosis"? What is this – it requires more careful explanation. Detailed information on this reference is described in reference section.

Line 118: what is the magnitude of such fluctuations? This information was added as suggested by the reviewer.

Line 120: the depth of the drainage ditch was? The total depth of the drainage ditch varied considerably in space, so a precise measurement of mean depth cannot be provided. However, we measured differences in terrain heights between ditch sections and the nearest plots along the transect and differences were minimum 50 cm.

Line 129 'affected' not 'effected'. How do you know that the drainage ditch had no effect at 600 m away – how did you ascertain this? The authors added some sentences to clarify this. Concerning the range of drainage ditch, there is no direct evidence that the undrained transect was not affected at all. However, ground-based vegetation community structure analysis before and after the drainage as well as larger scale analysis (WorldView, 2 x 2 m^2 resolution) indirectly infer that the drainage effects reached maximum 200 m outside the ditch. For more information, please see the responses to the reviewer #1.

Line 135: It is not clear to the reader what the rationale was for the 3 weeks sampling in 2013 and then 10 weeks in 2014. This should be made clear. This decision was mostly based on administrative and logistic constraints related to carrying out field work in this very remote part of Siberia. Please see also our comments within the response letter to review #1 regarding this topic.

Line 136: you introduce the term transects but have not done so before m- this is very confusing and should be addressed fully. Likewise, the labelling of transects is poorly defined. It was corrected throughout the manuscript as the reviewer suggested.

Line 141: The PVC collar was installed permanently, but how so and when, and to what depth was the soil isolated? For how long was the collar installed prior to sampling beginning? i.e. how long was there for recovery of the vegetation? Did cutting-in of the collar lead, as in many cases reported elsewhere, and in my own considerable experience, lead to vegetation damage/death in any case? The collars were installed approximately 15 cm into the ground, 3 weeks before the first flux measurement took place in 2013. During installation, we avoided damaging aboveground vegetation. Of course inserting the collars into the ground included cutting, and accordingly damages to belowground plant parts could not fully be avoided. However, in neither of the following field seasons where these plots were used, we found no noticeable plant damage or death around the collars, indicating that our field installations did not influence the vegetation substantially. Regarding the 2013 measurements, we believe that even if there was minor damage to the belowground plant parts, an equilibration period of three weeks should have provided enough time to avoid major effects on our flux data. Still, the authors agree with the reviewer that such potential implications should be discussed in the manuscript, so this information was added in the text.

Line 154: What evidence do you have that using ice packs effectively worked to keep the temperature constant as you claim? Air temperature was monitored with 1 Hz frequency while measuring fluxes, accordingly we were able to keep track of temperature gradients within the chamber. When temperatures increased more than 1 °C per minute, we started using ice packs by placing them inside the collar to keep temperatures stable. The total number of ice packs was adjusted until we found temperature conditions to remain at a stable level. Only then, the actual flux measurements were started. We added some more explanation in the text to clarify this.

Line 158: the units quoted need attention. The units are converted to g C basis.

Line 165 onwards: Your phrasing is not sufficiently clear here regarding the 'conflict' between choosing core sites based on WTD category This part was corrected to clarify this issue.

Line 170: define and quantify 'nearby' accurately please. It was corrected to make it clearer.

Line 174: It is not clear how the data from 2003 form a reference! The year 2003 represents conditions at our observation site before the drainage ditch was installed. All 2003 plots for vegetation sampling fall within the area that is now drained, and by co-locating sampling spots for vegetation community structure in 2003 and in 2013, we could directly assess the longer-term shifts in vegetation as a consequence of altered hydrologic conditions. Accordingly, 2003 data serve as a reference for pre-disturbance conditions. This issue was clarified in the text.

Lines 175-181: I find the explanation here insufficiently clear. You need to provide the reader with a much clearer explanation of what you did and why! It was clarified.

Lines 188-189: Aerobic incubation of soils – insufficient detail is provided here for this to be repeated. More explanation was added.

Lines 206-207: There is insufficient detail regarding why you chose August as your point of reference. August is the month within our datasets where thaw depths were deepest, and also differences in thaw depth dynamics related to the drainage were most pronounced. Please find a more detailed statement in our response letter to review #1.

Line 223: You do not describe the term -1. GPP. This should be corrected. It was clarified.

Line 269-70: how exactly did you add this error range in each case? Error ranges were calculated in each temperature and PAR bin for ER and GPP, respectively. When fluxes were interpolated over time, error ranges at each point were taken from the corresponding PAR and temperature bin that reflected current condition.

Section 3.3.1. I find this, as currently written, to be overly complicated in structure and terminology. The authors need to simplify their scheme substantially and really pull-out the key take-home points from the data. A key issue for me in this respect is that you are comparing data with different time periods of coverage between years. Surely a direct comparison between same time periods each year would be more helpful to the reader? In addition, your choice of use of cumulative data seems at odds with your choice of units. This requires correction. The authors agree with the reviewer's concerns. This part as well as the whole result and discussion sections will be re-written.

Line 396 onwards: low variability of what exactly? In the discussion section, I find the terminology again confusing – clearing this up substantially earlier-on in the paper and then following this through to the discussion would really help the reader comprehend the key messages much better than at present! Low variability was referring to relatively constant flux rates over time. To make the paragraphs easier to follow, this part was re-written and some sentences were added in the introduction.

Line 428: this is rather a weak initial statement – be more specific. It was corrected as the reviewer suggested.

Line 440: What direct evidence do you have to back-up this clear statement? References were added.

Figures:

(1) A useful conceptual figure, but much better reproduction is required. There seems to be a mis-match between continued drainage in the experiment and 'events' stated in the legend. This must be addressed. Figure legend terms do not match directly those used in the text of the manuscript. It was corrected as the reviewer suggested.

(2) You need to include the measurements taken at each observation site. It was corrected as the reviewer suggested.

(3) OK

(4) Grey points are not sufficiently clear on this figure. SD error bars – n = ? All figures that featured gray dots to represent individual measurements in the original submission were changed to box plot format in the revised manuscript version.

(5) Sub-seasons of 2014 are unclear – see my comments on this issue in the text of the manuscript also. There was no comment on this in the text, so the authors do not have the information that the reviewer is referring to. However, the definition of sub-seasons in 2014, which is used in several figures and tables, has been placed at a more prominent position in the revised manuscript, with references to this description added at all figures where the sub-seasons are still in use.

(6) OK – but why is this Figure 6 – surely this information warrants being at the start of the explanations!! It was corrected as the reviewer suggested.

(7) Clouds of grey points are hard to distinguish – rethink...such as boxplots for example. These would, I think be much clearer for the reader. It was corrected as the reviewer suggested.

(8) Ditto

(9) OK

(10) OK

(11) Comments as (7) It was corrected as the reviewer suggested.

Appendix figures OK.

There are numerous points in the text where small edits of the correct word are required for clarity of the narrative. There are too many to list here, but a native English speaker should be consulted to address these shortfalls. This is the worst such manuscript in this respect that I have read in some years. Please see also our statement regarding language edits in the response letter to review #1. The revised manuscript version will be checked by a professional language editor to improve overall readability.