

We would like to thank all five reviewers for their detailed, constructive and positive feedback on our original manuscript “Winter GHG emissions in a sub-alpine grassland”. We believe the comments improved the manuscript considerably. Here, we respond to all general and specific comments of each reviewer separately (regular font is the reviewer’s comments, italic font represents our answer).

Reviewer 1:

This manuscript presents an impressive dataset of gas flux measurements from a snow-covered site in Switzerland. The presentation of the data is generally appropriate, but in some cases there is a lack of transparency on how the data analysis was performed. While the introduction provides a good overview of the importance of measuring gas fluxes from snow covered sites, there is very little in the discussion or conclusions that puts the results of this study in a broader context. There are some comparisons of these data compared to other sites, but there is very little interpretation of what it means when the results either are similar or differ from other sites. Further, there is almost no discussion of the spatial component of this study. Finally, the writing would benefit from careful editing. In many cases, the authors have interesting ideas, but there are too many crammed into a single sentence, or paragraphs contain sentences that don’t provide a consistent narrative, and in some cases there are English usage errors. I have identified some examples of this in the specific comments below, but the authors would benefit from careful editing to make sure that individual sentences and the sentences within paragraphs tell a clear narrative throughout the manuscript.

Introduction There is useful information in this section, but the current structure needs some editing at both the paragraph and the sentence level. Why is the sentence starting on P403L8 a whole paragraph? Why is the general paragraph on N₂O and CH₄ emissions starting on P404L5 included? Why does the paragraph starting on P403L12 jump back and forth between what we know about GHG emissions generally and what we know about GHG in winter more specifically. One possible way to clarify the narrative of the introduction would be to first put GHG fluxes from ecosystems were put into a context of total GHG emissions (based on the data from the existing networks), followed by a discussion of GHG in snow covered ecosystems more specifically, including the methodological approaches and difficulties and what is know already about the magnitude of fluxes and the controls on the fluxes. Furthermore, the authors should be clear about the results of this work will contribute to our understanding the global significance of GHG fluxes in seasonally snow-covered systems or just subalpine/alpine grasslands. If it is the latter, they need to make a case for the broader significance of understanding GHG emissions there.

We would like to thank reviewer 1 for the kind words and revised the new version of the manuscript considerably in terms of both, paragraph and sentence level. We further strengthened the introduction as suggested: starting with a more general overview of GHG fluxes at the ecosystem level, then specifically focusing on winter emissions (current knowledge and limitations) and commonly used methodologies leading to the objectives of this study (see Page 2-5). Language editing was done by a native speaker which we hope further contributed to a better readability of the manuscript.

P402 L25- P403L2 This sentence is awkward

The sentence was rephrased to: “Measurement networks such as GHG-Europe or FLUXNET (Aubinet et al., 2000; Baldocchi et al., 2001) deliver fundamental data to investigate biogeochemical processes at the ecosystem scale.”

P403L12-P404L4 This paragraph is particularly confusing. Is the discussion specific to snow-covered ecosystems? What are the “complex processes” that we need to understand better? How come the first sentence suggests that “a profound understanding of such processes exists only for CO₂” when two sentences later it is suggested that “knowledge of fluxes for all three gases remains sparse?”

We restructured the introduction as follows: first, we are writing about the complex processes (directly following this statement) and then, we focus on the limited available knowledge for winter/snow-covered ecosystem in particular (Line 120 to 139). We further deleted “of fluxes for all three gases”.

P404 L12-P405L2 This paragraph switches back and forth between CO₂ and N₂O/CH₄. Why not complete the discussion of our limited knowledge of CO₂ and then discuss what we know about total GHG balance based on the even fewer measurements of CH₄/N₂O

This paragraph has been restructured with a clear distinction between each of the greenhouse gasses, highlighting the fair understanding of CO₂ exchange followed by a discussion of the limited knowledge on CH₄ first, followed by N₂O. We further included a short paragraph on the importance of all GHGs to the total GHG balance at the beginning of the introduction.

P405L5-7 It might be useful to have a table showing the studies that have compared methods including the means and some estimate of variability as well as the location of the study, snow depth, duration of study, frequency of measurements, etc. Or this could be included in the discussion.

We would like to thank reviewer 1 for the valuable comment on including a table giving an overview of the currently available literature on winter emission of GHGs, however we believe that such table would go beyond this paper and would rather fit into a review on winter emissions of GHGs and the different applied methodologies. In line 86, we refer to a recent review on winter CO₂ effluxes by Liptzin et al. (2009).

P405L7-9. This sentence doesn't make sense. Are you trying to say something about the hourly/daily/seasonal variability with a particular methodology compared to the difference among methodologies for a particular hour/day/season?

This sentence was changed to: “The comparability of the methods is controversial since each method covers different spatial and temporal scales. Therefore annual estimates of winter GHG fluxes are often variable as a result of the method used (McDowell et al., 2000; Björkman et al., 2010b).

P405L9-10 Another comparison between chambers and the gradient method is McDowell et al. (2000), which found no difference between techniques. McDowell N.G., Marshall J.D., Hooker T.D. and Musselman R. 2000. Estimating CO₂ flux from snowpacks at three sites in the Rocky Mountains. *Tree Physiol.* 20: 745–753.

We would like to thank reviewer 1 for pointing out this interesting publication, which we added to the methodological paragraph in the introduction.

P405L20 Where do the authors attempt to “identify the variables driving GHG emissions from different land-use type in a subalpine valley?” It seems like they looked at the variables related to the fluxes near the EC tower and they mapped the spatial pattern of fluxes, but there is nothing in the results suggesting what is driving differences in fluxes related to land use?

Our specific aims did not focus on the different land uses in this sub-alpine valley per se. However we included the transect measurements extending from the grassland towards other ecosystem types in order to better relate the magnitude of CO₂ fluxes found for the grassland to other typical vegetation types in the alps. Therefore we discussed the differences in CO₂ flux across ecosystem types. However we did not set up the measurement station etc. to compare different land use types.

P406L18 Delete “greenhouse gas”

Done

P406L20 Delete “majorly”

Done

P407L9-12 What is the frequency of the data used for the calculations (e.g. presented in figure 4 or compared to the physical data)?

Eddy covariance fluxes were calculated from the 20Hz data to 30min flux averages. After thoroughly checking the collected data and removal of low quality we used the remaining dataset for gapfilling. Gapfilling was performed using an online tool provided by the Max-Planck Institute for Biogeochemistry, Jena, Germany, which is based on the methods described by Reichstein et al. (2005). In order to compare the physical derived gradient measurements to the eddy covariance values we extrapolated our weekly profile measurements to daily fluxes, which were then summed up to derive a seasonal balance. In the revised manuscript, this procedure is explained in more detail.

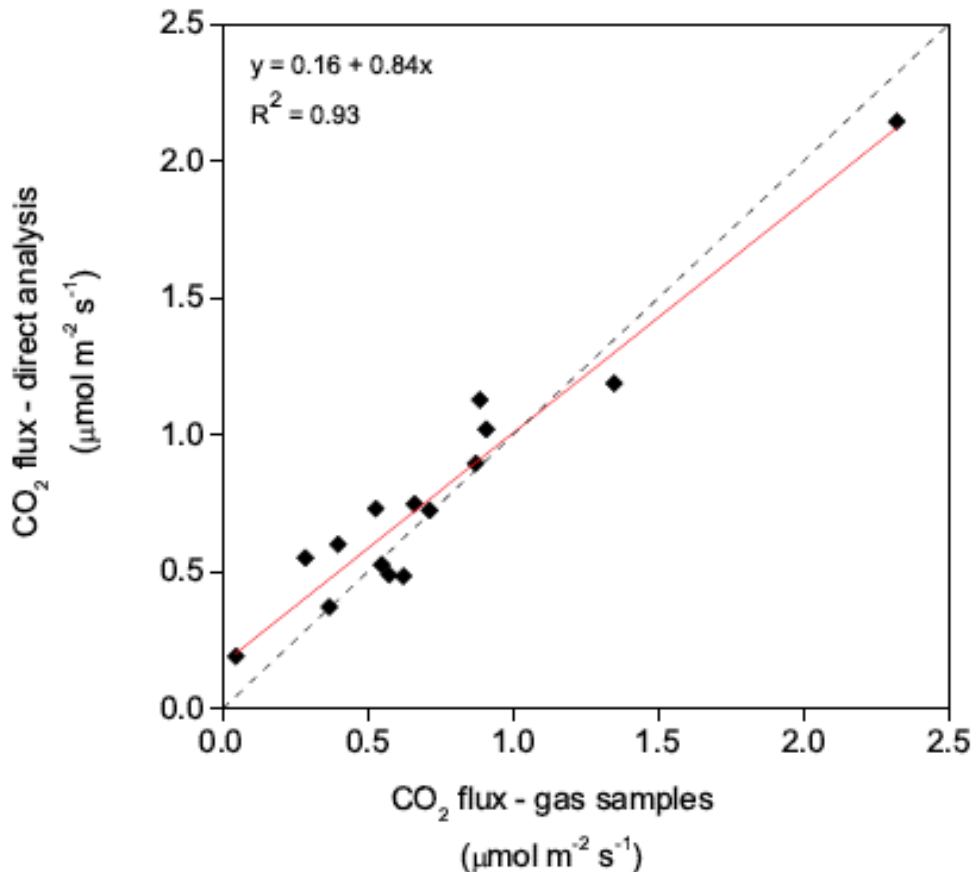
P408L14 fix the spelling of “tortuosity.”

Done, we fixed this through the entire revised manuscript.

P409L1-14 More details about the sampling and analysis are needed here. How is the air collected with the ski pole. Is there a pump? How much air is pulled from the snowpack and over what period of time? How many depths were sampled? Was CO₂ always measured with the IRGA in the field or was it also measured on the GC? In Lines 10-13, you should be specific about which detector was used to measure which gas.

Since three of the five reviewers asked for a more detailed description of the ski-pole

method we further extended this paragraph in the revised manuscript, explaining the two different approaches of deriving air samples using the ski pole ((1) with a pump to “feed” the LiCor with air and (2) syringes to collect gas samples for analysis with the GC. CO₂ was measured by both, infrared gas analyzer and GC (a brief comparison has been done in the master thesis by Steinlin (2011) indicating a good match of the two approaches ($r^2=0.93$, $y=0.84x+0.16$) see also the figure below). We further specified the detectors in order to clarify which detector measures which gas.



We specified the sampling and analysis of the ski-pole method as follows:

The ski pole had 10 cm interval depth markings along the pole to determine the insertion depth into the snow. The pole contained tubing inside and had a perforated tip allowing gas collection at any snow depth (Wetter 2009). Gas measurements were made at 10 cm increments either in the field via a portable gas analyzer directly connected to the ski pole or later in the laboratory following gas sampling. A different hole was used for each sampling date. Gas was collected from the ski pole using a micro diaphragm gas pump (NMP 015M, KNF Neuberger, Balterswil, Switzerland) to pull the air at a rate of approximately 0.4 l min⁻¹ through the infrared gas analyzer (LI-820, LI-COR Inc., Lincoln, Nebraska, USA) until CO₂ concentrations remained constant (usually after 30 – 60 s). In addition, gas samples were taken from the ski pole using a 60ml syringe which were immediately transferred into pre-evacuated 12 ml vials (Labco Limited, Buckinghamshire, UK) with a needle. In these samples, CO₂, CH₄ and N₂O concentrations were measured a few hours later by gas chromatography (Agilent 6890 gas chromatograph equipped with a flame ionization detector (FID) combined with a methanizer to measure CO₂ and CH₄ and an electron capture detector (ECD) to measure N₂O, Agilent Technologies Inc., Santa Clara, USA). For more gas chromatography details see Hartmann et al. (2011).

P409L15 The data from the automated system would provide an interesting comparison. However, since the method was not successful, there is not a compelling reason to include the description of the methodology or the short results section (3.5).

We think that the automated system would be a powerful approach and that other researcher may profit from our failure. Therefore, we left the automated system in the manuscript which is in accordance with the comments from four out of five reviewer. However, we shortened the result section and only briefly state possibilities for future enhancement of the system in the discussion paragraph of the revised manuscript.

P410L15-22 There needs to be more description of what was actually done here. Was the same gas sampling and analysis used here as in the weekly samples? What depths were sampled for the CO₂ measurements? It seems like there were two sampling strategies to look for spatial variability: first, two perpendicular transects were established in the grassland and CO₂, CH₄, and N₂O were measured and second, just CO₂ was sampled across a transect that spanned multiple vegetation types. It would help to have a map of the sampling design as part of Figure 1. Were there gradients in some environmental variables that would be expected to contribute to variability?

In general, the gas sampling was based on the ski pole method for both: weekly measurements as well as for the spatial measurements. The weekly measurement, samples within the snow profile were taken at 10cm increments resulting in commonly 4-5 concentrations per profile measurements, The intensive sampling campaign covering different ecosystem types were based on 3 measurements per profile only (10, 30 and 50cm depth respectively). This decision was made due to the large number of profiles (33 profiles for all GHGs and 150 profiles for CO₂), which needed to be sampled during a single day. We edited Figure 1 for better visualization of the transects determining spatial variability of all three GHGs in the grassland and for determining CO₂ flux variability across ecosystem types.

P411L24-25 What does “continuous snowfall” mean?

This was changed to “regular snowfall”.

P412L11 What does “followed a seasonal course” mean?

This was changed to “Fluxes of CO₂ calculated from the gradient measurements showed largest efflux rates at the beginning and end of the snow-covered period.”

P412L15-17 The mention of CO₂ concentrations doesn't fit here. More broadly, it would be helpful to make clear what time period is included in the flux values. That is, are they averages for all the samples during that month?

We removed the sentence about the CO₂ concentrations and added a statement that the presented results were based on weekly measured data only.

P412L17-24 It would be a useful comparison to calculate the EC values for the times that the gradient measurements were done to see how similar they are at the same

time and not just over the whole season.

We agree with reviewer 1 that such a comparison of EC values with gradient measurements would be useful however due to regularly occurring data gaps in the EC data during times when the gradient measurements were performed we were unable to perform such a comparison. Besides we remain critical about the applicability of such a comparison, since each method focuses on different temporal and spatial scales (gradient = plot scale, EC = ecosystem scale). Therefore, a comparison of direct measurements will most likely lead to different results. However, integrating measured data by either, EC or gradients, to derive annual balances seems more appropriate since gradient derived measurements of the 4 profiles account for the spatial heterogeneity on the grassland. Annual balances were than further derived gap-filling of both EC and gradient data. An additional paragraph on the calculation on annual balances was included in the revised manuscript.

P413L16-20 How was this model selected? Were the calculations done for every day of the snow-covered season? There needs to be a description in the methods of what this was done.

In the revised manuscript, we only used snow water equivalent ($r^2=0.80$) to extrapolate to seasonal CO₂ fluxes, because the combined model including more parameters did not improve the explanatory power significantly. In the Methods, we have described our approach to model seasonal fluxes as follows: “Gradient technique based seasonal flux estimates were derived by averaging the weekly profile measurements and relating the flux values to a set of environmental variables” and “The identified functional relations on a weekly basis were then used to extrapolate flux values for each day of the snow covered season and thereafter integrated.”

P413L21-25 Once again, there needs to be a description in the methods for how this analysis was done. For example, was a daily mean flux used to test the model? It is confusing that the seasonal values are so similar for the EC and gradient methods, but on shorter time scales the drivers of fluxes are not the same.

The methodological paragraph of the revised manuscript has been extended accordingly, further explaining modeled fluxes. Likewise we reduced the duration of the seasonal budget to time periods where the model could be verified. November and April were excluded from the seasonal cumulative – as also requested by two other reviewers.

P413L25-26 Can you compare month by month instead of using vague terms like “beginning and the end of the season?” Is there any statistical analysis here?

The comparison month by month is given in Table 2. We further extended the Table header in order to clarify our definition of winter as well as beginning and end of the season.

P414L12-14 What was the point of extrapolation for N₂O fluxes? Is it just for calculating the total seasonal emissions? I don't understand exactly how the running mean was calculated, but why not just do a linear interpolation between the days that were sampled?

The extrapolation of N₂O but also CH₄ was done in order to calculate total seasonal emissions. We decided to use a running mean, since we believe that simple linear interpolation may overestimate the contribution of spikes to the annual budget. Also from our experience on other grasslands (manuscript in preparation) the running mean approach lead to more reliable results than linear interpolation.

P415L1-10 Delete

In this case we disagree with reviewer 1. This paragraph was kept in the revised manuscript. The importance of this non-results in order to avoid such a setup by other researchers was also encouraged by 3 other reviewers. However we revised the discussion on this topic towards future setups in order to be able to measure Rn222 and CO₂ fluxes accordingly.

P415L13-15 How is the variability “stronger?”

We changed the sentence to “We observed a larger variation in CO₂ fluxes along the transversal cut of the grassland than along the longitudinal cut of the valley (Fig. 7a).”

P416 L6-11 Based on figure 8, it seems like a statistical test was actually used to compare among ecosystems. This should be included in the results section as well as in the methods section.

We believe that reviewer 1 refers to Figure 9 instead of Figure 8. Statistical comparison of CO₂ fluxes originating from different ecosystem were identified by an analysis of variance (ANOVA). We added this information in the M&M paragraph while we already stated the results in the current version of the manuscript P416L5-12).

P416L24 Do you know what the footprint is?

We estimated the footprint to not exceed the grassland, with the main wind direction occurring along the valley.

P416L25-P417L4 I don't understand what is being compared? Is it the value along both transects, is it the transects compared to the weekly sampling, is it the transects compared to EC?

We compared measurements collected on the two transects on the grassland with flux measurements from the permanent gradients surrounding the tower.

P417L8-9 Pressure pumping doesn't necessarily increase the CO₂ efflux, but it means that gradient measurements are underestimating the CO₂ efflux.

We agree with reviewer 1, pressure pumping does not necessarily lead to larger efflux. We discussed the potential impact of wind on the gradient method (“In their evaluation of CO₂ concentration profiles, measured at high frequencies in a subalpine meadow, Seok et al. (2009) concluded that potential errors are highest at high snow

densities (and hence late in the winter season). Moreover, their results revealed a strong impact of wind pumping, which decreased the CO₂ effluxes calculated by the diffusion method by on average 57% as compared to the actual flux.”

P417L21 “bares” is not the right word. Maybe contributes the most to the uncertainty? There is also some debate in the literature on how to calculate tortuosity (see discussion in Seok et al. (2009))

We considered this in the revised manuscript as follows: “The likely reason (for the discrepancy between gradient and EC-Based fluxes) is an incorrect estimation of gas diffusivity in wet and dense snow with snow porosity and tortuosity inferred from snow density measurements being the most uncertain variables (Seok et al., 2009).

P418L5 NO and NO₂ are not inert in the snowpack

“NO and NO₂” were removed in the revised manuscript.

P418L6-11 Is this discussion of liquid water apply all winter long or just when the snowpack is at zero during melt events?

Dissolving of CO₂ in water does not only apply during melt events. However, we expect the losses by dissolving of CO₂ in melt water to be larger during these periods than during peak winter season. Melting of snow during the peak winter season is very unlikely (see Figure 2, Tair and Tsoil-snow interface). Furthermore as shown by Sommerfeld et al. 1996 the magnitude of CO₂ flux in the liquid phase is two orders of magnitude smaller than the upward flux through diffusion of CO₂ in the gaseous phase.

P418L12-21 I can understand why liquid water might increase dissolved CO₂ and prevent emission to the atmosphere, but I don't understand why this would affect the methods differently. That is, I can believe that CO₂ fluxes would be lower during snowmelt but why would meltwater lead to the large discrepancy in the two methods?

As pointed out by reviewer 1 dissolving of CO₂ does not affect the two methods differently and we apologize for our misleading discussion. In the revised manuscript we moved the discussion of the different magnitudes in CO₂ flux to another more appropriate paragraph leading to a closed paragraph that discussed the losses of CO₂ in meltwater only.

P418L22-P19L13 This section should be condensed to make the point that tracers could be used. IS SF₆ injected into the snowpack to be used as a tracer? This is different than using radon. It seems like you have to know what the flux of radon is at the soil surface in order to validate the flux calculated with the gradient method?

We included the SF₆ traces as another possibility to derive reliable diffusion coefficients. In our study we intended to use radon in the automatic profiles while we further measured the radon flux at the soil surface during each sampling events. In the methods, we have added: “To quantify the actual ²²²Rn flux, we measured the increase in ²²²Rn with time in chambers placed on the soil surface in snow pits every week”

P419;L14-25 This paragraph doesn't belong in the discussion. It seems like the necessary portions are already included in the methods section.

We removed this paragraph in the revised manuscript.

P420L2-8. It is true that the pattern in this study doesn't match the data for the subalpine site in Liptzin et al. 2009 (Fillipa et al 2009 and Seok et al 2009 are not appropriate references here). However, it is similar to the conceptual zone II in Liptzin et al. (2009). There are many other studies in the Rocky Mountains report a mid winter (Jan or Feb) minimum compared to higher rates in late fall or spring. This is the pattern reported by Sommerfeld et al (1996) cited here as well as the Monson 2006a reference below as well as Brooks et al. 1997.

We would like to thank reviewer 1 for pointing out this mix in references and corrected the revised version accordingly.

P420L19-25. This sentence is awkward because it has too many ideas combined.

We have rewritten the discussion on temperature dependency.

P420L19-P421L7 This discussion is confusing. What would the mechanism be for SWE affecting CO₂ flux? Sommerfeld et al (2006) reports a similar correlation, but also does not provide a mechanism. The authors should provide some explanation for why this correlation is meaningful. The relationship for Monson et al. (2006b) does not seem appropriate here as it is a comparison of CO₂ flux and maximum snow depth among years and not in a given year.

In addition to a more thorough discussion on the temperature effects, we discussed the SWE effects as follows: "Snow water equivalent can be regarded as a measure for both (1) the progressing winter season with an increasing snowpack but a declining substrate availability as discussed above; and (2) the snowpack properties affecting the gas diffusivity and hence, the calculation of efflux rates using Fick's law." Thereafter, we discussed the discrepancy of the two methods. including a statement on the two fundamentally different transport mechanisms underlying the gradient and the eddy covariance method.

P421L26-29 How does the soil moisture at this site compare to the others cited in the text. That might help evaluate if the lack of variability

There was generally larger variation in soil moisture in the other studies cited in our current manuscript. This might be one of the reasons (mentioned P421L27) why we could not establish a correlation between CH₄ fluxes and soil moisture. Secondly soil moisture of 0.4 m³ m⁻³ as measured on the Dischma grassland does not imply full water saturation. Unfortunately we can not provide information on bulk density of the soil in order to estimated possible methane oxidation microsites in the soil.

P422L4-9 Are you suggesting that land management affects N₂O emissions? Or is the beginning of this paragraph is just a summary of N₂O emissions?

Land management does in general affect N₂O emissions from managed ecosystems.

However, in this specific case of measuring winter emissions we only summarized the N₂O emissions observed on this grassland. The statement on the effects of fertilization was given as it is common practice to fertilize these sub-alpine grassland with manure shortly before the first snowfall events.

P422L11 What does the reactivity of NO, N₂O and NO₂ have to do with the seasonal patterns of N₂O fluxes?

This was a misleading sentence, which we adjusted to N₂O only.

P422L18-20 Why not estimate porosity based on the bulk density?

As already mention before we did not collect bulk density samples from our site.

P423L10-17 It would be helpful to have more discussion of the spatial variation. Otherwise why include them? For example, there is no mention in the discussion of why the different land uses differ in CO₂ fluxes.

In the revised manuscript we extended the discussion concerning the spatial variation of CO₂ fluxes: “Our results indicated significantly higher winter CO₂ fluxes in the grassland than in the forest (Figure 8). This pattern is consistent with the observation of decreasing summer soil respiration rates during afforestation of a sub-alpine pasture (Hiltbrunner et al., 2013). In their study, the lower rates in the forest were explained with a smaller root turnover, a lower litter quality and a less favorable microclimate in the forest than in the adjacent grassland. In our study, the thinner snow cover in the forests (<30 cm) leading to colder soils (Groffman et al., 2006) might have contributed to the lower soil respiration rates in the forest.”, In addition, we adjustment the objectives.

P423L25-P424L4 Are the author’s suggesting that there would be lower gas emissions during a shorter snow-covered season because the snow-covered season itself is shorter? It would be more useful to have a discussion of whether a shorter snowcovered season would affect soil temperatures which seem to be highly related to at least CO₂ and CH₄ emissions. If snow cover started later or snow depths were lower, what would the consequences for gas fluxes be.

Prediction of changes in CO₂ fluxes with less snow cover and shorter snow covered periods remains challenging due to the complex interactions between environmental variables but also between processes driving either GHG flux. For instance later snow melt often leads to colder soils.

Now, we think that this discussion would be to speculative and removed it from the Conclusion.

Table 2. Why is April included when there are so few manual samples taken? Wasn’t snowcover assumed to be gone on April 4th? How were the modeled values calculated for the EC measurements?

Table 2 was adjusted and values were recalculated after implementing the revised model for CO₂ (gradients). EC data were gap-filled using a marginal distribution sampling approach – we therefore changed the wording from modeled to gap-filled.

We further included all GHGs in the new version of Table 2 and included values for November as well for April only where relevant and reliable.

Table 3. It is confusing that the EC technique is used in this table since the CH₄ and N₂O measurements are only for the gradient technique.

CO₂ budgets in Table 3 were based on both, gradient and EC technique as indicated by the letters a and b. We adjusted this to direct naming in the revised manuscript.

Figure 1. The intermediate scale is not needed, but it would be helpful to see where the various transects were.

We included another subfigure to visualize the location of the transects.

Figure 2. The changes in soil temperature during the snow covered season are difficult to see in panel b because the scale has to go up to 30 degrees. Can you just show the temperature during the snow covered season? The useful information is the variability between -0.5 and +0.5 degrees

The y-axis for the soil temperature panel was rescaled in order to better visualize the variation during the peak winter season.

Figure 3. What are the error bars on this graph?

Error bars denote standard deviation calculated from the 4-5 profile measurements of each sampling day.

Figure 4. I don't understand the gray polygon.

The gray polygon represents the diurnal variation in EC fluxes while zero represents the actual flux. We intend to show both, the uncertainty in the flux, derived from the Gap-filling tool (Reichstein et al. 2005) and the daily variation.

Figure 5. What do the plots look like for daily CO₂ flux from the EC method vs temperature on the days when the gradient method was done? Is the lack of relationship between temperature and CO₂ flux a function of the high variability in flux as seen in Figure 4 at some times scales whereas the longer term trends are clearer. Is this a signal to noise problem depending on the time scale of averaging?

We aimed at comparing exactly such time periods, however two major reasons occurred which made such a comparison impossible: (1) regular occurring data gaps either caused by instrument failure or by applying quality checks removed more than 90% of the comparable data points and (2) each method applies to a different temporal but also spatial scale leading to unreliable results.

Figure 6. Not needed

We disagree with reviewer 1 and agree with the other reviewers in keeping this picture in the revised manuscript.

Figure 7. It is confusing to have the two transects connected to each other in a different orientation than on the ground. Even if there is a point of overlap between the two transects, the graphs would be easier to interpret as separate panels for the two transects. How was the interpolation done? Was it done first along the vertical profile of each gradient and then the horizontal interpolation between sampling locations?

We separated the two transects in the revised manuscript to avoid confusion. Interpolation between concentration gradients was based on geostatistical kriging approaches, which we further explained in the M&M paragraph.