Response for Referee 2

23/11/2012

Major comments:

The two up-scaling methods used in this study were land cover classification/land cover types and LAI map derived from Quickbird satellite image. Although this part may be one of the most important information in this study, the description of how these two were derived was quite confusing. I read this part several times, but only able to understand this by the end of the results section. I am still a bit confused how and why the land cover classifications were separated as is. The authors need to answer the following questions regarding this description.
1) Why did the authors separate the landscape into different land cover classifications?
2) Why were the two methods chosen?
3) How were they separated?
The authors provided Hugelius et al. (2011) for detailed description, but they still need to provide some level of detail for this method in their study.
In addition, the description in section 2.1 needs to be revised accordingly.

The upscaling procedure is very essential content of the article, and we are grateful for the referee for pointing out that there might be some lack of necessary detail regarding this part. We revised the section 2.5 by adding the following information:

- ***a description of the segmentation process used Definiens Professional 5.0 software to group neighbouring pixels for land cover classification***
- ***more details of building the regression function for predicting LAI, and the formula of the regression function***
- ***a reference to section 3.3. and Fig. 6 concerning the dependence between cumulative CO_2 fluxes and LAI that was used for scaling up the CO_2 fluxes based on the LAI map***
- ***a few other minor corrections or additions***

Additionally, we refined the last sentence of the introduction, where the two data sets used for up-scaling are mentioned for the first time.

The idea was as simple as to use two alternative upscaling methods, and see how well they will perform. Our starting point was to produce a land cover classification similarly as in our earlier work in the same region: Heikkinen, J. E. P., Virtanen, T., Huttunen, J. T., Elsakov, V., and Martikainen, P. J.: Carbon balance in East European tundra, Global Biogeochem. Cycles, 18, GB1023, doi:10.1029/2003GB002054, 2004.
The possibility to use high-resolution satellite image and very detailed field data on LAI and plot scale CO_2 fluxes encouraged us to try another, alternative upscaling method based on LAI mapping. Here, LAI was predicted for each grid cell of the satellite image and the dependence between LAI and CO_2 flux was used for deriving the CO_2 fluxes from the LAI map.
We are not sure what the referee means by revising the section 2.1 accordingly, but we did add a reference to the section 2.5.1 on the land cover classification. We got the percent coverage of different land cover types from this classification.

We hope that the referee finds this explanation and the corrections satisfactory.

The part on ecosystem CO2 balance in different land cover classification was also not discussed in the discussion section. I suggest expanding this idea and answer the question why the patterns are how they are based on the land cover classification and LAI in the discussion. More description on this part would be very useful for the readers.

We were not completely sure if the referee means here the differences between the CO2 balances obtained by the two up-scaling approaches (land cover classification and LAI-map), or the differences in the CO2 balance between land cover types.

Concerning the differences in the results given by the two upscaling methods: Generally the agreement between the two up-scaling methods was very good considering the uncertainties in the CO_2 estimates, and both methods proved to work well in such upscaling studies. Unfortunately, we feel that the material from one year and from one study site does not provide very solid basis for discussing what caused the slight differences between the too methods, and such discussion would have been too speculative.

Another detail necessary for this manuscript was on modeling GP and ER. Modeling and up-scaling tundra GP and ER using different environmental variables have been done before. I am surprised how much the LAI explains variability in GP and ER in this study. But to support their conclusion, the authors need to provide more detail on each steps of modeling process. I'd like to direct the authors to similar studies conducted in tundra ecosystem: Lee et al. (2011) A spatially explicit analysis to extrapolate carbon fluxes in upland tundra where permafrost is thawing. Global Change Biology, 17, 1379–1393 and Williams et al. (2008) Upscaling Leaf area index in an Arctic landscape through multiscale observations. Global Change Biology, 14, 1517–1530. I recommend the authors first show correlations among the variables used in the models to show the readers how much auto-correlated each variable can be, and next show each step of model construction. The caveat of regression method in multiple variables is that the significance and confidence level increases with increased number of variables by nature. Therefore, each model needs to be compared by how much it is improved by adding more variables. I'm not sure after reading this manuscript whether the models were developed for each land cover classification and put together later for Figure 5, or if they were developed by pooling all the data from the beginning.

Certainly, there was a strikingly strong linear dependence between the CO_2 fluxes and LAI. LAI explained the very well the high spatial variability across all land cover type varying from water logged to dry and from peatlands to mineral soils with just a thin organic layer.

It seems that our explanation on how we actually used the LAI-CO_2 regression to up-scale was not fully clear, and we tried to make it more comprehensive by, e.g., adding cross-referencing between different sections and to the figures. Also the text was revised in many places, e.g., the second paragraph of the section 3.3. on the explanatory power of LAI for spatial variability in CO_2 fluxes. There, we explain how linear regressions between LAI and cumulative ER/NEE/GP were built for different seasons, and used for upscaling the plot scale CO_2 fluxes to landscape and regional level using the LAI map. Since the LAI-CO_2 dependence is linear, we could obtain the CO_2 balance for a certain area by inserting the mean LAI value to the regression function.

Modelling of CO_2 fluxes of chamber microsites based on their dependence on environmental parameters to produce hourly CO_2 dynamics is described in detail in section 2.3.2. Concerning the question if the data were pooled fro the model from the beginning: the ER was modelled separately for each chamber plot (three plots per microsite), while the three replicate collars belonging to one microsite were pooled for GP modelling. Citation from section 2.3.2.: "Response functions for ER were formulated individually for each chamber plot, while parameterization of the more complex GP functions was done at the microsite level (n = 3)." In this sense the modelling differs significantly from that described by Lee et al. (2011), were the data were pooled to three blocks (vs. our 30 chamber plots/10 chamber microsites). Very detailed description of the ER and GP models for each chamber plot/microsite would have required a lot of space, and we feel that it is not in the main scope of this paper.

Tied to this comment is that the auxiliary data collection was described only in supplementary information, but I think it would be more useful to be included in the main text to help the readers' understandings on how the data were collected. It seems that the environmental data were collected only from one weather station (actually two, but only one was used) and I'd like to know the authors' take on this, when the tundra landscape is very heterogeneous.

The basic meteorological data, such as wind speed and direction, air temperature, precipitation and PAR was collected just from one weather station, but this data is not affected by the surface where it is measured. The parameters showing high variability across the landscape, such as soil temperature and moisture, ground water level, active layer depth and leaf are index were monitored in several locations covering well the different land cover types.

As the referee claims, tundra is a very heterogenous environment and the soil and vegetation characteristics vary significantly across the landscape. We added some additional explanation about the auxiliary data collection to section 2.2., but would like to keep the more lengthy description on the instrumentation etc. in the supplementary part.

MINOR COMMENTS

- Throughout the manuscript, the word 'data' was treated singular. Please revise it to be plural.

The use of the word 'data' has been checked and revised.

- Throughout the manuscript, the term plot scale and microsites were used to describe the measurement scale, how did they differ? Does this need to be consistent? In addition, how do LCC, LCT, and microsites differ? Please clarify throughout the manuscript. This is also making Table 2 very difficult to understand.

We admit that the terminology related to different scales is quite complicated, but this is necessary separating the different levels used in this study. We hope that the use of the full terms 'land cover classification' and 'land cover type' instead of abbreviations makes the terminology more clear.

We were estimating CO_2 balance for three nested scales:

1) ***plot scale (chambers, snow gradient, lake measurements)***
2) ***landscape scale (EC + plot scale measurements area integrated to the EC footprint)***
3) ***regional scale (plot scale measurements area integrated to area of the QuickBird satellite image)***

These scales are explained in the last paragraph of the introduction section.

The terms 'land cover type' – 'microsite' – 'chamber plot' also refer to different levels of classifying the landscape. 'Land cover type' means a surface with uniform vegetation and soil type in general, while the word 'microsite' refers to a group of three replicate chamber plots or lake stations. Use of two different terms is necessary, since microsites do not always directly correspond to the land cover types observed by remote sensing methods. Table 2 and the last paragraph of section 2.5.1 describe the correspondence between of land cover types and microsites measured at the plot scale.

- P9946L15-18: I think there is too much logical gap from the previous sentence to this one. How was this achieved?

More description has been added to the abstract concerning the regional CO_2 balance. Additionally, a mistake in the lower limit of the regional CO_2 balance has been corrected (-67 -> -79)

- P9948L20: Describe what plot measurements were.

Description of the plots scale measurements ('the chamber technique in terrestrial surfaces and the gas gradient method and bubble traps in lakes') can be found just a couple of rows before this, probably it is not necessary to repeat it here. Maybe the term 'plot measurements' was confusing. We changed it to 'plot scale measurements' to be better in line with the rest of the manuscript.

- Section 2.4.4: Provide why this was conducted.

The referee is right that the section 2.4.4. about the footprint analysis was not very clearly tied to the context of the whole manuscript. The reason for performing the footprint analysis was to obtain the contribution of different land cover types to the EC signal, which allowed us to upscale the plot scale fluxes to the landscape level (= EC footprint) and further compare them with the EC flux. The section 2.4.4 has been now revised accordingly.

- P9954L24: Where were these 13 LCTs described?

The thirteen land cover types are shown in the Fig. 1A and Table 1. References have been added to the text.

- P9955L5-6: Where did this idea come from? Is there a reference or is this done conventionally?

The overall coverage of forest stands, sand and impacted tundra in the study region was only 2.7%, so we assumed that they have just a small importance to the regional CO_2 balance. Since we could not find any estimates on the CO_2 balance of equivalent forest stands in southern tundra, we decided to use a zero balance. This makes sense since the tree growth is very slow, and there is no significant peat accumulation. The coverage of sand and impacted tundra was even lower, and no published data was available there.

- P9956L13: 'hot' is this supposed to be 'high'?

The sentence has been revised.

- Section 3.4: I don't really understand the necessity of this section in this much detail with this many figures. The earlier graphs already explain much of seasonality. I recommend either downsizing this section or focus on growing season vs. non-growing season.
In addition, I think there is not enough information to show interannual variations unless the authors are going to go in depth on showing what caused the interannual variability.

We agree with the referee that seasonal variation of NEE is not very novel information and not in the core of this paper. The first paragraph of the section 3.4.1 has been deleted except for the max respiration and net uptake rates that were moved to the next paragraph (see our response for the referee 1). Also the figure 10 showing seasonal development of diurnal NEE has been dropped away.

Concerning the interannual variability, this section 3.4.2 may not be the most important considering the main message of this work and that is why we did not carry out very profound analysis on the topic. On the other had, we thought it is important to show even the limited information we have on the interannual variability when it is available. Moreover, there were not many differences in the cumulative CO_2 fluxes between the years, and we belief that the temperature was the main factor behind the difference see in upland tundra, as we claim in the paper.

- Section 4.2: I think it would be very helpful for the future research, if the authors can give some suggestions for addressing the problem of the difference between the two measuring techniques. In this section, the authors give a good overview of why this may be happening, but a suggestion for how to address this issue or how to fix this problem would greatly help the readers.

Referee is right that it would be very useful if we could give some suggestions for future research on how to deal with the differences between the two methods. However, based on the data that we have it is difficult to achieve this. In this section we go through several aspects that can be done in order to improve the performance of the methods used: e.g., keeping the chamber measurement time short and using non-linear calculation of the gas flux; avoid using open path sensor for EC measurements in cold climate, or if it has to be used, measuring the actual temperature of sensor hear; using high resolution remote sensing data for any vegetation mapping used in upscaling. A clear conclusion that can be done based on this study is that the two measuring techniques give significant added value to each other, and should be used in parallel when possible.

- Section 4.4: I don't really think that this study was much relevant to addressing this topic. The authors did not conduct this research under the assumption of changing climate. If they would like to include this section, the authors need to go over in depth description on how their model or observations change with climate and climate variability. Either route seems fine, but I believe this manuscript can do without the climate change effect.

The referee 1 had a similar concern on the relevance of this section to the paper, since we did not directly address the aspect of climate change with the set up of our study. We agree that the title of the section was probably promising too much. The section 4.4 had some discussion on the role of different land cover types on the regional CO_2 balance, which is very relevant for the paper, and was combined to the previous section about the regional CO_2 balance. Additionally, we wanted to keep the statement about the importance of landscape reorganization on the regional CO_2 balance due to the large heterogeneity in fluxes across land cover types. Although the heterogeneity of tundra GHG fluxes is not novel information, it was once again shown by our data.

- Table 4: What is WT?

WT is the abbreviation for water table level, i.e., the depth of the ground water from the peatland surface. The determination of WT has been added to the table caption.

- Fig3: What do these lines represent? What does Fen and B. nana tundra heath represent in the figure?

The dashed line is a 1:1 line, the solid line a linear regression fitted to the data. The two land cover types deviating from the 1:1 line, Fen and B. nana tundra heath, have been named.

- Fig4: What are the different colors in figure top-right represent? Indicate positive and negative values in the figure to make the figure stand alone from the text. How about labeling a, b, c, d for each panel to clarify the description of the figures? Were the data pooled from all the sites?

The missing details were added to the Fig. and the figure caption.

- Fig6: X-axis is missing

The x-axis has been added.

- Fig12: X-axis is missing

The EC data previously presented in Fig. 12 has been added to the figure on seasonal CO_2 dynamics of chamber measurements (previous Fig. 6).

- Tables S1 and S2 need more detail. The units and abbreviations are not described.

Explanations of the abbreviations have been added to the captions of Tables S1 and S2.