Referee 3

We are grateful for the detailed comments of this review that will help to improve our manuscript. Furthermore, we also would like to thank the referee for the supporting statement that the topic of our study is not only relevant for Biogeosciences and its audience; it is also relevant for science in general. Nevertheless, he/she had several criticisms that have to be considered in the revised version of our manuscript. General points are linguistic problems, the length of the chapters results and discussion and the misleading title. As stated in our reply to referee 1 and 2, we agree to that and a revised version of our manuscript will be professionally corrected by a native speaker. We will also change the title into "Regulation of CO_2 emission from temperate streams and reservoirs", and the manuscript will thoroughly be revised to reduce speculations in the result and discussion chapter before resubmission. Likewise to referee 2 the main criticism was the way k was calculated in the reservoirs. We followed the suggestion of referee 2 and recalculated k based on actual wind data. The obtained values are in the range of those calculated before. Thus, the message of our manuscript doesn't change.

There were several further comments that have to be considered in a revised version of our manuscript. We will answer all comments in the following:

It is very unclear which data (e.g. collection site, number of observations, sampling time) are pooled and used for the comparisons shown in the different figures. This should be made very clear in the methods part of the manuscript.

> We agree with this point and will include all mentioned details in the method chapter of the revised manuscript (see also our reply to referee 2).

Right now it is too confusing. In relation to this issue, it is surprising that no results from the statistical comparisons among systems are shown throughout the manuscript (e.g. Wilcoxon tests).

> All described differences are statistically significant. We have waived to include statistical details, to improve the readability of the manuscript. To avoid further confusions we will include all details in the revised version of our manuscript (see also referee 2).

I have great concerns about the way k is estimated in the reservoirs by assuming a constant wind velocity. This seems rather unacceptable. The authors should try to improve this by using additional data, or at least try to estimate the uncertainty associated with this assumption and incorporate it in their discussion.

> We also see this point which was already mentioned by referee 2. Thus, we decided to recalculate all k values from the reservoir based on wind data from a neighboring weather station. Comparisons with own wind measurements 1 m over the water surface had shown that these wind data are applicable to our system (see also referee 2). Furthermore, we also recalculated all fluxes for the reservoirs based on actual k values. Even if obtained data are in the range of those originally calculated and the message of our manuscript had thus not changed, we are confident that the recalculation of the flux data clearly improved the outcome of our study. All changes will be included in the revised manuscript (methods, results, discussion, figure 4 and 7, table 1 and 2).

The conclusions about metabolism being the main driver of CO2 evasion and the low importance of groundwater inputs seem a bit speculative based solely on the results from this manuscript.

> As stated in our reply to referee 1 we will consider the effect of lateral inflow on CO_2 variability in more detail in a revised version of our manuscript.

The upscaling of CO2 emissions to the whole catchment is very important; however, it should be improved and more details on the methods used should be provided in the methods part of the manuscript. In relation to this, it seems rather poor to assume a mean wetted with of 4m for all streams in the catchment. There are many approaches that could be used to improve this (see some of the references of the manuscript).

> We will add more details of the upscaling to the method chapter. As stated in our reply to referee 2, we estimated the rough error of the upscaling which improves the results clearly. We will include the details in a revised manuscript. However, we used only for the upscaling on the catchment level a mean width of 4 m for the streams. We used very detailed data for the upscaling on the stream level. We estimated the errors for this approach in our reply to referee 2.

Second, the authors should provide measures of uncertainty and discuss this issue in the context of their study and other studies.

We agree to this comment and will include statistical indicators like p values and standard deviation in the whole manuscript. We will also discuss possible uncertainties in the context of our study and other studies in more detail as we have already done it (cp. our reply to referee 2).

Finally, the authors should at least discuss the issue of not including measures of spatial variability in their emission estimates, especially in the reservoirs (there is growing literature on this issue).

As stated in our reply to referee 2, there were no strong lateral gradients of CO_2 in the reservoirs. For the purpose of this manuscript (regulation of CO_2 emission from lotic and lentic waters), the sampling point near the dam can be considered representative for the reservoirs (see also our reply to referee 2).

Specific comments:
P10023, L12-13: These references refer to measurements of reaeration in streams and not specifically to CO2 emissions.

We will replace the reference Wanninkhof et al. (1990) by the reference Billett and Harvey (2013, Biogeochemistry), which was a suggestion of referee 2.

P10024, L22: Delete "be" P10024, L24: "turbulence" rather than "turbulences".

We corrected both.

P10025, L11: I suggest avoiding "GHG" in this manuscript as much as possible since the only gas investigated is CO2.

We will delete GHG in this sentence and in the whole manuscript were it is appropriate.

P10025, L11: Tone down this sentence. There are previous studies that have measured CO2 emissions form streams and lakes in the same cathment (e.g. Buffam et al. 2011 GCB 17: 1193-1211).

We will rewrite the sentence:
"Only few studies exist where the factors influencing CO_2 emission in lakes and streams are directly compared in a temperate ecosystem."

P10027, L14: GF/F filters usually have an approximate pore size of 0.7um.

We apologize for this mistake. Because GF/F filters have an unspecific pore size, we will delete the stated pore size and write only GF/F.

P10027, L22: Why were different methods used for estimating CO2 concentrations on the streams and the reservoirs? Which is the best method? What are the problems associated with each method? How comparable are both methods?

The CO_2 concentrations were determined with different methods, because the samples are from different campaigns: The reservoir samples were taken during routine water quality monitoring while the stream data were obtained during a metabolism study (cp. Halbedel et al. 2013, Biogeosciences). Both techniques are widely used (Hope et al. 2004, Hydrological Processes; Striegl and Michmerhuizen 1998, L&O 1998) and the results are comparable. More details are described in our reply to referee 2.

P10029, L4: The mean wind speed of what? When and where these measurements made? Is there a meteorological station nearby from which wind measurements could be taken? I have great concerns about using this value for evasion calculations.

As stated in our reply to referee 2, we recalculated all k values based on actual wind data from a neighboring weather station.

P10030, L3: I could not find the results from the Wilcoxon tests anywhere in the manuscript.

We will include all necessary statistical details in the revised manuscript.

P100030, L24-25: These emission rates are just for the studied stream reaches and for the whole reservoirs? Please specify. Does it make sense to compare in this way these systems of different size?

This comment seems to be based on a misunderstanding. The emission rates are for whole streams. We used stream data from different organizations (Ute Enders: Unterhaltungsverband Holtemme, Detlef Cöster: Talsperrenbetrieb Sachsen-Anhalt and Otfried Wüster: Nationalparkverwaltung Harz) for an upscaling of the CO_2 emission from reach level to stream level (see also our comments to referee 2). We will add more details of the calculation in the method chapter.

P100031, L1-2: Can you provide a reference for this assumption?

This was for example described by Striegl et al. (2001, Limnol. Oceanogr.) and Karlsson et al. (2013, Geophysical Research Letters). We will include both references.

P10034, L7: I suggest using the median instead of the mean consistently throughout the manuscript, especially if you are using non-parametric statistics and your data are not normally-distributed.

The reviewer is right that not all of our data are probably normal distributed and non-parametric statistics are advised. However, if we have only few data and want to give a typical value and a measure for the data scatter, means and standard deviation are the usual way. The medians are given in the box plots.

P10034, L15-16: Confusing sentence. What do you mean?

We measured in a parallel study (Halbedel et al. 2013, Biogeosciences) the lateral inflow with chloride as conservative tracer. We found that besides Hassel, all other streams have less lateral inflow or an outflow in the adjacent soil or sediment. To avoid further confusions we will explain this in more detail.

P10034, L17-23: These arguments against the contribution of GW are not completely convincing. Please expand. Also, the results do apparently not show a significant relationship between metabolic rates (GPP, ER) and CO2 emission rates. That would be a good argument for the important contribution of metabolism to CO2 evasion. There was also no significant correlation between CO2 emissions and DOC. Thus, the conclusion about metabolism being the main driver of CO2 evasion seems a bit speculative based solely on your results (see general comment).

As already stated by referee 1 we will extend this part and include a discussion about the contribution of terrestrial lateral inflow (including ground water) to the CO_2 concentration in our streams. More details are given in our reply to referee 1. Nevertheless, the CO_2 concentration followed in most of our net heterotrophic streams a typical seasonal pattern, indicating a strong contribution of the metabolism on the variability of the CO_2 concentration. In addition, since major components of DOC can be recalcitrant (which is common in terrestrial forests or wetlands) and since besides dissolved organic material also particular organic material is known to fuel the metabolism, there must not necessarily exist a correlation between DOC and CO_2 concentration. We already discussed this point in our manuscript (P10037 L1-9).

P10035, L22: According to previous sentences your values were high and not low compared to those by Allin et al. (2011).

We checked this reference and our data (Table 1) again and found, that we published k values in Table 1 that were based on the calculation that was used by Allin et al. (2011, Journal of Geophysical Research; compare also P10035 L2-4). These k values (Rappbode k=9.5 cm h^{-1}, Hassel k=19.8 cm h^{-1}, Zillierbach k=18.7 cm h^{-1}, Ochsenbach k=14.8 cm h^{-1}) are in the range of those published in Allin et al. (2011) for small rivers and streams (Table 1: k=18.6 cm h^{-1} (mean)). But these are not the values we used for the flux calculation or correlation analysis (cp. P10035 L19-21, P10029 L17). The used k values are 9.05 cm h^{-1} (SD 2.0) for Zillierbach, 10.28 cm h^{-1} (SD 2.6) for Ochsenbach, 10.13 cm h^{-1} (SD 2.8) for Rappbode, and 12.08 cm h^{-1} (SD 2.4) for Hassel. We apologize for this inaccuracy. These values are low compared to those published by Allin et al. (2011). We will correct these values in Table 1. Furthermore, we will expand our discussion on the possible uncertainties when calculating k with different methods.

Please check. P10038, 15-25: Expand this part on upscaling and include parts in the methods section (see general comment).

We will include more detail about our estimations in the method chapter and discuss the need of more work regarding this topic in more detail in the discussion.

Table 1: Add standard deviations. Clarify if the measures are just for stream reaches and whole reservoirs?

We will follow this comment.

Figures 2, 3, 4 and 6: There are not just averages in this graph. Define the boxes (e.g. median, quartiles, outliers, etc.).

We will follow this comment.