Dear Chris, thank you very much for your very helpful and constructive comments, which we indeed appreciated. Below you will find our answers and replies to your points. From our point of view the most important ones are dealing with a) empirical model development, b) gap-filling procedures and c) exploration of additional controllers of GHG fluxes at the Höglwald site.

In short a few general comments on these three bullet points:

a) The manuscript was not thought to explore gap filling strategies for NO and N2O, but we picked up your idea (and also of Kim Pilegaard and reviewer #1). We therefore changed our strategy. We now use daily mean values (we stick with daily measurements since subdaily flux measurements are not as robust due to failures of single chamber measurements) for the development of empirical models (linear and non-linear approaches). Nevertheless, the idea behind the development of empirical approaches is not to gap-fill but to explore if easy to measure parameters can be used to simulate fluxes at different time scales.
b) The empirical relationships found were used for gap filling and in Table 3 we now provide estimates of annual fluxes with and without gap filling of data.
c) The point with the additional controllers of GHG fluxes is valid, though respective datasets are not complete. We now explore GPP values from tower measurements at the Höglwald site as well as N deposition (throughfall values) as drivers of soil trace gas exchange.

Reply to specific comments:

p12203, l12: could the authors be a little more specific with respect to the non-linear algorithms they tested? eg Hutchinson & Mosier (1981) model? Non-steady-state diffusive flux model by Livingston et al. (2006) ? Intercept method by Kroon et al. (2008) ?

At the time we started measurements we relied on the work of Hutchinson and Mosier (1981). This is now mentioned in the manuscript.

p12203, l18-23: please provide the dimensions of the chamber

Chamber dimensions are already mentioned on previous page 12202 line 22. All chambers have had the same dimensions (0.5 m x 0.5 m x 0.15 m)

p12204, l16-17: '...Gaps originating from instrumental failure were filled by linear interpolation between measured fluxes for calculation of cumulative annual emissions.' Was there a time limit for the data gap, eg a few hours up to a day, beyond which the linear interpolation method was not applied? A linear interpolation between two flux data points is fully justified if the two point are not too distant in time, but beyond several days this method can clearly bias the annual flux, especially for trace gases like N2O and NO, which are extremely dependent on instantaneous rainfall and short-term variations in soil water status. I would have found it much more logical to fill the flux data gaps using the regression algorithms developed in this paper (as is common practice in the CO2/NEE flux community and eg FLUXNET methodology). Why do all this work on environmental control of fluxes, and not apply this to gap-filling?

In the previous version we only calculated annual fluxes of trace gases if daily values were available for more than 80% of all days in a year (see Table 2). Measuring gaps were usually not longer than 10 days (in most cases <3-5 days), since equipment at the site was serviced in the first few years until 2002 2-3 times per week and thereafter at least once per week. This is now mentioned in the manuscript too. Nevertheless, we picked up the idea of gap filling data with the developed empirical relationships and compare gap-filled data with our simple empirical gap filling approach in Table 2.

p12204, l19-20: I doubt that air temperature measured at the weather station, presumably at 2-3m above grass (?), can be considered representative of air temperature in the understorey of the forest. We have air as well as soil temperature measurements within the forest. However, due to data recording failure, lighting events etc. the dataset on meteorological drivers is not complete. Therefore, we used the dataset from the German Weather Service. For gap filling soil temperature data – measured in-situ at the Höglwald Forest – we developed an empirical relationship of air temperature measurements at the weather station and continuously available soil moisture values as available for the Höglwald. We tested this relationship for those periods for which soil temperature measurements were available (see below). This is now also mentioned in the manuscript.

p12205, l17-21: '...To develop and assess empirical models, weekly aggregated data, monthly aggregated data and data aggregated within comparatively larger temporal resolutions...': Why not establish empirical regression models on the basis of hourly or daily fluxes? I understand that one might wish to use easily measured state variables like T and SWC as predictors of trace gas fluxes, but surely the hourly and day-today dynamics are essential to capture the overall flux variability? Besides, I doubt that the straightforward application of the algorithms developed here would have much predicting power at other sites than Hoeglwald; at the very least that would have to be scaled for N deposition (and how?), and to account for soil type, structure, pH, C/N ratio, dominant tree species, etc. For the purposes of the present paper, which was to explain the observed variance in fluxes at seasonal and inter-annual time scales, the finer resolution hourly data would provide a wider spread of flux values to stretch the range of the regression. I would suggest, if possible, to add to Table 3 the regression results using hourly and daily flux data; it may well be that the R2 values of the regressions are higher at weekly or monthly time scales than daily or hourly, but if that is the case then it should be demonstrated and used to justify the use of weekly/monthly fluxes.
We followed your advice and also established regressions for soil moisture/ soil temperature relationships and daily mean values of trace gas fluxes. This information is now given in a new, additional table (Table 4). We do agree that the relationships found for the Höglwald cannot be used straight forward for other sites too. But this was not the intention and is also mentioned in the discussion section. The idea is to unravel easy empirical relationships for a given site. This would e.g. allow to reduce the measurement effort for a given site without losing the chance to get an estimate for trace gas fluxes at an annual scale.
Therefore, we are still presenting empirical relationships for weekly and monthly time intervals, but comparing those now with gap filled daily values.

12205, l27-28: '...In addition, the under- or overestimation of the mean flux was compared to the measured mean flux...': this is a valid test only if the measured flux data capture during the comparison period is close to 100%; i.e. there are no flux data gaps during which environmental conditions and driver may lead to very different fluxes, which are represented in the model but absent in the measurements.

Point is taken, we now work with both gap-filled and measured data.

p12206, l1-3: How did the authors deal with bi-directional fluxes like N_2O and CH_4? Were the negative (uptake) N_2O fluxes removed prior to regression, since the log value of a negative number can't be calculated? For CH_4, presumably, the sign of (negative) uptake fluxes was changed to positive before log-transformation, but what happened to the few emission fluxes visible in Fig.2 ?

No, we did not remove negative values but worked with an offset (+20 µg N for N_2O and plus 20 µg C for CH_4). This information is now given.

p12206, l13-14: Why show the comparison for 2000 to 2003 only? Soil moisture was unavailable in 1996-99 and in 2004-06 only. Please show all years 1994-2010 in Fig. 1.

Done, see new Fig. 1

p12206, l17: what is the unit of the RMSE? %?

We clarified this. The unit is Volume %

p 12207, l4: Does the mean measured value of 932 mm include snow? If so, what is contribution of snow to total precipitation?

Data were measured with standard heated precipitation gauges by the German Weather Service. The data does not allow to judge if this was snow or not.

p12207, Section 3.3: the term 'aggregated' conjures up the notion of temporal integration. I would favour the use of 'averaged' instead of aggregated throughout this section and in Fig.2 as well as Table 3.

We followed your suggestion

p12207, l28 to p12208, l3: For CO2, Fig.3 does not show any significantly increased flux in February. The difference is really marginal. There may have been large differences during individual years with pronounced freeze-thaw cycles, but the difference is averaged out when taking the 15-yr mean (unlike N2O, which is clearly visible in red on the figure).

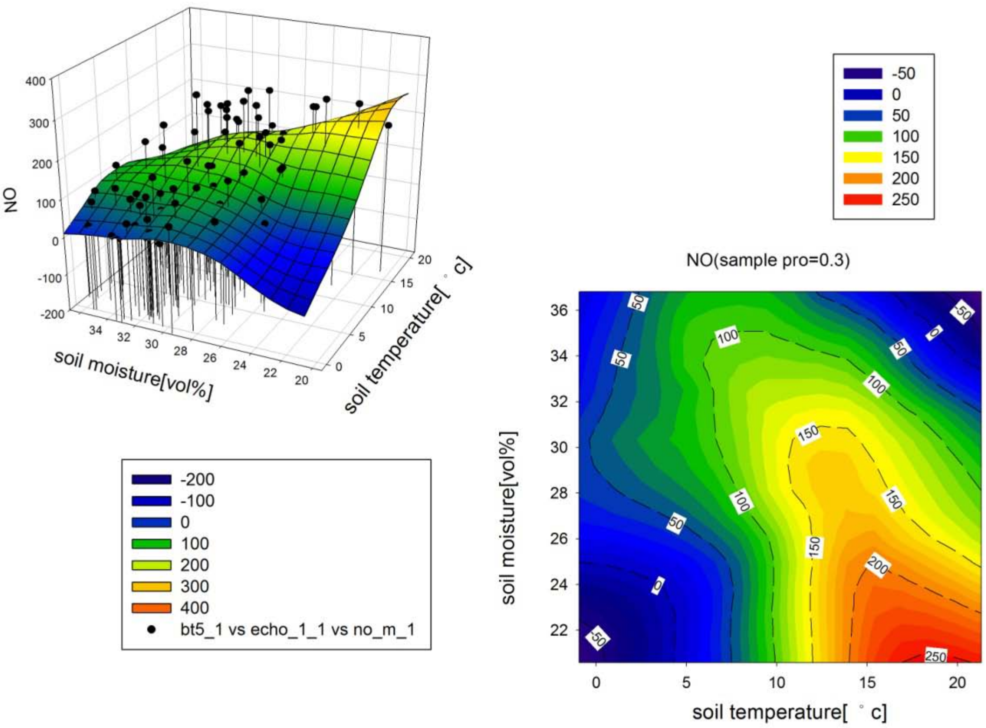
Point taken. We rephrased the sentence to "A notable finding was that monthly mean CO_2 emissions also slightly increased in February as compared to mean fluxes in January and March, due to an increased soil microbial activity during the period of soil thawing in some years (Fig. 3)."

p12209, l2-3: there is much too much scatter in Fig.6 for CH_4 to argue that CH_4 uptake is best described by a quadratic fit to soil temperature, and similarly for soil moisture on Fig. 5.

Agreed, but in the sentence before we were already pointing out that regressions are weak. We rephrased the sentence to strengthen the point that relationships are weak in general.

p12209, l6-12: I agree that the relationship of NO flux to temperature is, as one would expect, exponential, and that the impact of moisture is bell-shaped (Lorentzian ? in Fig. 5). However, I wonder why NO emission is clearly highest at SWC=22% (and much reduced at 30%) in Fig. 7 , while it is clearly highest at SWC=28-30% in Fig. 5. Could there be a problem with the calculation of the contour plot? Indeed I wonder that flux observations for NO and for the other gases only cover the soil temperature range -2 to + 15_C only, while the contour plots for CH4, NO and N2O extrapolate up to 22_C. Are there flux measurements at soil temperatures higher than 15_C, which are not shown in Fig.5, but which were used to calculate the contour plots of Figure 7 (except for CO2, which stops at 15_C) ?

This is obviously a misunderstanding while interpreting Fig. 7. Fig. 7 shows a planar fit to observed data, with temperature and moisture being the drivers of fluxes (see Fig. below). Which means, that the figure shows an interpolation. Thus the graph shows potential effects of moisture / temperature constellations for which no observations are available. We changed the captions of the respective figure (Fig. 7) to clarify this point. Also note that for regression analysis in Figs. 5 temperature measurements in different soil depths were used (we missed to give this information in the first draft) as compared to Fig. 7 were we only analyzed soil temperature in 5 cm soil depth.



p12209, l13-22 and Fig.6: the temperature effect on N2O for non-freezing conditions is likely exponential, and should be shown as such on Fig.6, in which the quadratic fit is non-sensical and misleading, being driven by freeze-thaw events on the left-hand side. I suggest redrawing an exponential fit, discarding soil temperatures below say 1_C.

Done, we agree that including freeze-thaw effects is not helpful at all.

p 12210, Section 3.5: I wonder why, having shown clearly on Figs. 5,6 and 7 that the controls of temperature, but especially moisture, on gas fluxes are highly nonlinear, the authors develop '...multiple linear regression analyses for predicting soil atmosphere fluxes...'. The log-transformation of fluxes prior to regression does justify the treatment of temperature as first-order when fitting the log values (being equivalent to an exponential fit overall), but for nitrification/denitrification products like NO and N2O, bell-shaped functions could have been built into the equation. This was done in Fig. 5, but then abandoned later on for the models; why?

Following also earlier comments by the referee we now also explore non-linear regressions. The performance is not significantly better, as can be seen from the newly introduced Table 4. Nevertheless, when using non-linear functions for calculating annual fluxes results are in tendency closer to measured values as compared to linear functions (Table 3).

p 12212, l13: suggest change 'well within' to 'at the lower end of'. This value of 7.91 tC/ha/yr was calculated from the means of available measured flux data in years when the number of days with data was >292 (Table 2). It would be good to compare these estimates with values derived from gap-filled time series, assuming that input data (soil T, SWC) needed to run the regression models were available during periods with no flux data.

Following the comments by the reviewer, we provide now gap filled annual emissions calculated by empirical relationships of emissions based on 5cm soil temperature and 10cm soil moisture. For details see above comments.

p 12214, l1-15: The hypothesis that long periods of winter freezing primes soil organic matter mineralisation and boosts annual soil respiration is an interesting and important one. However, the observation that the highest annual CO2 emission occurred during the year with the coldest mean temperature (1996 with 5.7_C) is not evidence enough to support this claim. For N2O the argument was much more convincing (although the processes involved are different), with cumulative winter N2O emission clearly linked to the duration of freezing, and many (14) contrasting measurement years, whereas CO2 measurements were available during 6 years only, and there may be confounding factors. Soil respiration has been shown to be strongly linked to gross primary productivity across sites (eg Migliavaca et al, 2010, GCB 17, 390-409), and interannual variations in GPP at one site may also result in different rates of soil heterotrophic respiration the following year. The high 1996 CO2 flux might have resulted from a high litter fall in autumn 1995, for example. The paper here would benefit from a discussion of other environmental controls of soil respiration, including interannual variations of fluxes of nutrients into the soil, such as in GPP (photosynthesis), and N deposition which can also vary by eg 30% from year to year, in relation to meteorology.
Further, what is the proposed uncertainty in the annual flux estimates? Can the 1996 CO2 flux be said to be significantly larger than the interannual mean?

We weakened our statement and now discuss also the potential contribution of C assimilation and N deposition as potential drivers of increased soil respiration in cold years. However, for a multi-year analysis and for the detection of carry-over effects of GPP on soil respiration in the following year, more data would be needed. Due to several reasons (mostly associated with funding long-term research), NEE measurements (and derived GPP) only started in 2006. Litter sampling was done in some years, but not in others. Same is true for N deposition. This incompleteness of data unluckily hampers a full assessment of drivers of interannual variability.

The uncertainty of soil CO_2 fluxes is surely significant. Besides questions of spatial representativeness of five chamber measurements for the entire stand, we also face the problem of data gaps (and gap filling). For that reason we do agree that our statement with regard to cold years soil CO_2 fluxes and priming effects need to be weakened. We did so, but still have it in the manuscript as a potential discussion point.

p12220, l1-3: '...Our failure to demonstrate such a relationship between soil moisture and CH_4 uptake rates is likely a result of the weekly and monthly aggregation of measurement data...'. I can only repeat my comment made above, about the opportunity of presenting regression results based on hourly and daily data, as opposed to and in complement of, weekly and monthly data as currently presented in the manuscript.

This point has been picked up (see above) and respective analyses are now presented.

p12223, l26-27: "...This rather low predictive power showed that simple regression models using measured soil environmental parameters hardly work to simulate soil N_2O fluxes." I certainly agree that predicting N_2O emissions takes more than a bivariate model; however, as suggested above, and considering the bell-shaped response (Fig. 5) to soil moisture at this site (and similar pattern for NO), a nonmonotonic approach (eg Lorentzian) would probably have yielded a better predictive power.

Also here we followed your comments and do provide now a much more detailed analysis.

Technical corrections

All technical corrections have been implemented