Dear Lars,

We are very grateful for your in-depth review of our manuscript and your comments are extremely helpful and no doubt will improve our manuscript.

General comments

It is not clear from the methods description how many times each collar was visited for CO2 measurements during 8 am to 6 pm on the individual measurement days. This information on data coverage is important and must be included.

-Daily data coverage is indeed very important and in all studies there is a trade-off between spatial (many plots and sites) and temporal coverage (many measurements and visits). We have included additional information in this respect by editing the M&M and including the number of measurements per visit which varied between 2 and 9 (see comment below).

It seems somewhat unfortunate to pool all data from each site to derive models for GPP and Reco. This reduces the information on spatial variability, e.g., among the deep and shallow sites. The Reco models applied at site A and B are very different and apparently without any exponential term for temperature response? I think a model like eq.2 usually comes with an exponential term? Note also that parameter b appears twice in eq.2?

-Agreed. In pooling the data, we do reduce some information, particularly in regard to the response sensitivities at the various sites. However, we found that by pooling the data, we were able to considerably improve the goodness of fit of the models, as we were now able to incorporate a much wider range in environmental variables (e.g. LAI and WT). We feel that approach is vindicated by good r2 values and a satisfactory 1:1 fit between observed and modelled data, and with the independent test data. We feel that this then provides for a much more robust estimation of annual CO2-C balances at these sites.

Equation 2 was incorrectly written in the manuscript. The correct equation which was used for the model is now included:
Reco = (a+(b x WT)) x (exp (c x ((1/T_{ref}-To) – 1/(T_{5cm}-To)))

A treatment of uncertainties in the annual C balances is missing. How was the effect of uncertainty in modelled response parameters addressed? It seems only spatial uncertainty was addressed (by running the models for individual collars and presenting data as mean and standard deviation). Generally, details on reported variability are missing or unclear; I have provided suggestions below.

-Uncertainty in the annual estimates of GPP and Reco has now been addressed by summing up the maximum and minimum standard errors associated with each of the model parameters (following the methodology used in Drösler 2005 and Elsgaard et al. 2012). The largest deviation from the mean was used as an approximate SE estimate of GPP and Reco as applied in Elsgaard et al. 2012. As NEE is not directly modelled, uncertainty in the annual NEE estimate was calculated following the law of error propagation as the square root of the sum of the squared standard errors of *GPP* and R_{eco}. This approach to uncertainty in the annual estimates is now described in the M&M and the results shown in Table 4.

Detailed comments

Abstract comments:

Introduce abbreviation GHG at first mention (line 4, rather than line 9)
Line 13 and 14: confusing to report C per m2 and N2O per ha; I suggest to change units and unify area
Line 13: spell out net ecosystem exchange, NEE, at this first mention in abstract
Line 13: in information like (NEE 233 g C m-2 yr-1) I suggest to introduce a comma for clarity, i.e., (NEE, 233 g C m-2 yr-1). This applies throughout the manuscript.
Line 16: suggest not to use '+' at occasional instances

Abstract responses:
Introduction of abbreviations are now following editorial rules (both abstract and main text should stand alone and therefore and therefore abbreviations are spelt out upon first mentioning in each.
L 13 units changed so that all results are now in g m-2
L13 advice from editorial team will be sought for preferred way to show information (NEE, 233 g C m-2 yr-1) or (NEE: 233 g C m-2 yr-1) or (NEE 233 g C m-2 yr-1)
L16: all occasional "+" have been removed in front of positive values.

Introduction comments:
Line 20 (p 5559): maybe use ' year' rather than 'annum'
Line 12-13 (p 5561): I guess biomass C export should also be included here in the elements mentioned for calculation of NECB?

Introduction responses:
L20 (5559) 'annum' was replaced by 'year'
L12 (5561) 'as well as biomass C export,' added to the sentence.

Material & Methods comments:
Line 8 (p 5562): (annual runoff, c. 586 mm) rather than just (c. 586 mm)
Line 22 (p 5562): IPCC citation should be 2014 rather than 2013
Line 9 (p 5563): Holcus lanatus
Line 16 (p 5563): '…each collar…' change to '…each collar for GHG measurements (see below)…'
Line 1 (p 5564): define PPFD at first mention here (rather than p 5564, line 23); is PPFD used here as equivalent to PAR then PAR can be replaced (p 5563, line 26)
Line 2-4 (p 5564): why are linear regressions used to develop site specific relationships, when data in Fig. 1 are shown with fitted cubic regressions? Indeed, it seems inappropriate to use linear regressions through the origin to fit to data in Fig. 1?
Line 5-9 (p 5564): it seems these 5 line rather belong to the next subheading on 'Greenhouse gas measurements'
Line 11-12 (p 5564): (n = 7 for site A_d, n = 5 for site A_s)
Line 21-24 (p 5564): so you had two series of soil temperature recordings at 5 and 10 cm?
Line 1 (p 5565): '…chamber method (Alm et al., 2007), between 8 a.m. and 6 p.m.' change to '…chamber method (Alm et al., 2007). Measurements were done between 8 a.m. and 6 p.m.'
Line 1-2 (p 5565): no need to define NEE again and no need to give units of PPFD again
Line 9 (p 5565): (PP Systems. UK). Change to (PP Systems, UK).

Line 22-28 (p 5565): I appreciate the consideration of low fluxes which are valid even though r2 is not high. Indeed r2 is not a good quality indicator of robust fluxes (see e.g. Görres et al. (2014) Agriculture, Ecosystems and Environment 186: 64–76.)
Line 4 (p 5566): stick to 'fortnightly' rather than 'biweekly' (which can also mean twice a week)
Line 7-9 (p 5566): 'Four 50mL samples were withdrawn into 60mL polypropylene syringes from the chamber headspace at 10 min intervals over a 40 min period and then injected…' This is a little ambiguous as samples taken over a 40 min period at intervals of 10 min would amount to 5, rather than 4, samples. Does this imply that no t=0 measurements were taken? Please clarify.
Line 9 (p 5566): I guess you used Exetainers®; this could be specified as these come with stoppers known to withstand storage of sampled gases
Line 22-25 (p 5566): what is the reason for being more strict on r>0.90 when it comes to CH4 and N2O fluxes? The situation is analogous to the CO2 fluxes, and you risks to discard a number of valid, but low fluxes.
Line 22 (p 5567): do you mean (see Sect. 2.2) ?
Line 1: delete 'in the soil'
Line 2: WT already defined; just use WT
Eq. 2 (p 5568): parameter occurs twice? Change T to T_{5cm}
Line 6 (p 5568): delete degree sign in front of K
Line 12 (p 5569): NEE already defined

Material & Methods responses:
L8 (5562): text inserted: (annual runoff, c.586 mm)
L22 (5562) IPCC citations updated throughout to 2014
L9 (5563) small cap for 'lanatus'
L16 (5563) 'for GHG measurements' added
L1 (5564) PFFD is defined and used instead of PAR.
L2 (5564) the LAI data was analysed using a linear model with a quadratic term (because the height is the only term squared, this polynomial regression still qualified as a linear model). Polynomial replaces 'linear' for clarity in the text.
L5-9 (5564) Lines shifted to next paragraph
L11 (5564) text edited (*n* = 7 for site A_d, *n* = 5 for site A_s)
L21-24 (5564) yes, the weather station recorded soil temp at both 5 and 10 at each site.
L1 (5565) text edited '…chamber method (Alm et al., 2007). Measurements were done between 8 a.m. and 6 p.m. (2 to 9 measurements per collar per day)'
L1 (5565) abbreviation used and units removed
L9 (5565) (PP Systems, UK)
L22-28 (5565). reference to support this point is acknowledged and included.
L4 (5566) text edited 'fortnightly' instead of 'biweekly'
L7-9 (5566). There was no measurement at t=0 and this is now clarified in the text.
L22-25 (5566). The same criteria as for CO_2 applied here. Discarded fluxes were not necessarily low fluxes but were obviously non-linear (due to leakages or other errors during measurement). Text edited.
L22 (5567) changed to Section 2.2
L1 (5568) deleted 'in the soil'
L2 (5568) WT used as already defined
Eq. 2 corrected
L6 (5568) degree sign removed
L12 (5568) NEE used

Results comments:
Line 23-44 (p 5570): 'In Year 2, Site A received similar to Year 1, above long-term average precipitation (1193 mm) driven by high values during June and July, September and October, and December and January. Both the mean…' suggest to rephrase to 'In Year 2, Site A again received higher precipitation that the long-term average (1193 mm) driven by high values during June and July, September and October, and December and January. However, both the mean…'
Line 5-6 (p 5571). It seems from Fig. 2 that cumulative Y1 and Y2 PPDF are very similar, so is your statement correct?
Line 18-20 (p 5571). The model coefficients for R-eco cannot be compared as they are part of two very different models, so rephrase this sentence.
Line 26-27 (p 5571): this qualitative statement could easily be backed up by statistics
Line 4-7 (p 5572): Again, statistics of bias is recommended
Line 27 (p 5572): change 'a-1' to 'yr-1'
Line 5-8 (p 5574): Clarify. '…the highest amount at 265 (27) kg N ha−1 yr−1 compared to 107 (45) and 80 (12) kg N ha−1 yr−1 in Site Ad and Site As. N biomass exports were 17 to 19% lower in Year 2 in Site Ad and As respectively.'
E.g.,:
'…the highest amount at 265 (± 27) kg N ha−1 yr−1 compared to 107 (± 45) and 80 (± 12) kg N ha−1 yr−1 in Site Ad and Site As (data are mean ± standard deviation with n = 7 for Site Ad and n = 5 for site As). Compared to Year 1, N biomass exports were 17 to 19% lower in Year 2 in Site Ad and As respectively.'
Line 12-13 (p 5574): I see no high emission in November ?
Line 19-21 (p 5574): reading this sentence is rather confusing: 'Annual CH4-C emissions differed between years with higher values (1.3±1.09 g CH4-C m−2 yr−1) observed in Year 1 and lower values (1.4±1.1 g CH4-C m−2 yr−1) in Year 2 (Table 4).' How is it 1.3 ends up as a higher value than 1.4? I would rather say the values are identical. Also, again, the basis of the used statistics on variability should be clearly stated.
Line 24-25 (p 5574): keep consistency and cite as Year 1 and 2
Line 7 (p 5575): delete '.' in parentheses
Line 8 (p 5575): I suggest to use same area unit as for other fluxes, i.e., 0.16 g N2O-N ha-1 yr-1
Line 18 and 19 (p 5575): specify that percentages refers to total fluvial C fluxes
Line 15-17 (p 5576): restructure this sentence to more clearly indicate what the basis for variation shown in parentheses represents. This is highly unclear as the numbers are very different from the uncertainties shown in Table 4 ? Also note, that 342 g C m-2 yr-1 probably should be 358 g C m-2 yr-1 ? Finally, 'g' is presently omitted in line 16.

Results responses:
L22 (5570) Text edited "In Year 2, Site A again received higher precipitation that the long-term average (1193 mm) driven by high values during June and July, September and October, and December and January. However, both the mean…"
L5 (5571). Data re-checked and statement corrected "There was no significant difference between Year 1 and Year 2 at Site A but PPFD values were consistently higher in Year 2 during the period June-December, except for July."
L18 (5571) Correct, Reco models differ between sites and therefore cannot be compared. Text edited: "The relationships between GPP and the environmental variables differed between the study sites as demonstrated by the different model coefficients derived for equations (1) given in Table 3."
L26-27 (5574) Text edited to include statistical analysis. Validation of the model showed a strong agreement between modelled versus measured GPP fluxes (Fig. 4) and independent test data (r^2 = 0.86 at both Sites A and B).

L4 (5572) text edited "As with GPP, the relationship between observed and modelled R_{eco} was generally strong (Fig. 4) (r^2 = 0.63 at Site A and r^2 =0.54 at Site B using independent test data)."
L27 (5572) 'a-1' changed to 'yr-1'
L5-8 (5574) edited text as suggested. '…the highest amount at 265 (± 27) kg N ha–1 yr–1 compared to 107 (± 45) and 80 (± 12) kg N ha–1 yr–1 in Site Ad and Site As (data are mean ± standard deviation with *n* = 7 for Site Ad and *n* = 5 for site As). Compared to Year 1, N biomass exports were 17 to 19% lower in Year 2 in Site Ad and As respectively.'
L12 (5574). The previous sentence states that high emissions are recorded during the summer months. We just wanted to add the information that both months of April and Nov show high fluxes (as high as say July) but not the 'highest'.
L19 (5574) (see Ref. 1) Data has been re-checked and is correct but indeed the sentence is not correct. We propose to change the sentence in the text to: "Annual CH_4-C emissions were not significantly different between years with 1.3±1.09 g CH_4-C m^{-2} yr^{-1} in Year 1 and 1.4±1.1 g CH_4-C m^{-2} yr^{-1} in Year 2 (Table 4)."
L24 (5574) text changed to Year 1 added for consistency.
L7 (5575) dot deleted in parentheses
L8 (5575) changed units and text: 0.16 g N_2O-N m^{-2} yr^{-1} and kept m-2 for consistency (not ha)
L18 (5575). As noted by Ref #1, 'total waterborne C fluxes' or 'total fluvial C fluxes' were added to the various percentage for clarity purpose.
L15 (5576) 342 changed to 358 (see Ref 1 same comment). Numbers in the sentence are however the same as in Table 4. Uncertainties based on NECB for each collar as not all same uncertainty associated with all the terms.
L16 (5576) 'g' added

Discussion comments
Line 11 (p 5579): for clarity: (30% more in Year 1 and 35% more in Year 2 compared to site As)
Line 14 (p 5582): 'maybe' change to 'may be'
Line 28-29 (p 5582): '…temperate climate, emitted only in very small amounts when the mean annual water table was around – 23 cm.' suggest to write '…temperate climate, and that CH4 was emitted only in very small amounts when the mean annual water table was around – 23 cm.'
These data fits well with a number of recent compilations, e.g., Audet et al. (2013) Ecological Indicators 34, 548-559.
Line 4 (p 5583): N_2O
Line 8 (p 5583): '(closer to the IPCC default values for nutrient rich shallow drained).' This is a little insider style: Include mentioning the soil type .
Line 3-4 (p 5586): subheading: '4.5 Implications for reporting and climate change mitigation strategies CO2, CH4, N2O and DOC emission factors' ; please rephrase for clarity
Line 16-17 (p 5586): suggest not to use '+' at occasional instances

Discussion responses:
L11 (5579) text changed as suggested to (30% more in Year 1 and 35% more in Year 2 compared to site As).
L14 (5582) 'maybe' changed to 'may be'
L 28 (5582) changed text as suggested: '…temperate climate and that CH4 was emitted only in very small amounts when the mean annual water table was around – 23 cm.'
L4 (5583) changed N_2O
L8 (5583) rephrased to '(closer to the IPCC Tier 1 default values for shallow drained nutrient organic soils)'
L3 (5586). During the transfer to pdf, the first heading was 'merged' with the second sub-heading. It should read:

"4.5 Implications for reporting and climate change mitigation strategies"
Then next line:
CO_2, CH_4, N_2O and DOC emission factors

This could be in effect 4.5.1 but since there is no 4.5.2 we felt this sub-heading is sufficient. The editorial team could advise here perhaps.

Tables comments:
Table 1: Table heading: change '…from both locations…' to '…from the two research locations…'
The data entries are not aligned with the years; move all data down one row
Table 2: State the data source of Table 2, so methods can be evaluated. E.g., related to LOI-to-OM conversion and OM–to-C conversion if C was not measured directly
$NH4_+N$ rather than $NH4_-N$ or is typing intended to be NH4-N?
Table 3: heading: insert 'and' between (GPP) and ecosystem respiration. Put Reco in parentheses. What about T0 from eq. 2? Was that modelled or assumed to a fixed value; please specify. Further, in the text r2= 0.85 is mentioned for GPP model at Site B (p 5571, line 26) – should this be the correct R2 to include in the Table (rather than 0.72) ?
Table 4: As done for NECB, the other parameters should be spelled out. State the basis for the SD and SE values, i.e., specify n and that they represent the spatial uncertainty. NECB Site Ad sums to 358 rather than 342.

Table responses:
Table 1: changed text to "from two research locations". Figures were aligned and will be checked upon typesetting.

Table 2: C and N were directly measured and the instrument stated as a footnote. Ammonium was measured and is usually abbreviated NH_4-N. Typesetting gives a wrong impression of 'negative'. This will be double checked after new typesetting of paper.

Table 3: inserted 'and' between (GPP) and ecosystem respiration and put Reco in parentheses.
T0 is set at 227.13K. This has been added to the text.
r2= 0.85 is the correct value for GPP model at Site B (p 5571, line 26) and was corrected in Table 3.

Table 4: 242 changed to 358.
Parameters have been spelt out. The uncertainty has now changed and is fully explained in the text and we feel that it is not necessary to include it again here for each parameter as the caption is already lengthy.

Figures comments:
Figure 2: panel identifier (a) is missing from upper panel. Note umol rather than µmol on y-axis, panel A. Caption: delete '(a)' in front of PPFD. Maybe write PPFD in full text in the figure caption
Figure 3: specify what error bars in panel a represent. At what depth were VMC measured and are the low contents in May believed to be real? Maybe you could have included precipitation data to substantiate the dynamics.

Figure 4: I suggest to spell out Reco and GPP in the figure caption. Has the accordance between data and 1:1 lines been tested? And just a detail, h-1 is used in caption whereas hr-1 is used on axes.
Figure 5: amend y-axes to 'CO2 flux (g CO2-C m-2). Figure caption: state the basis for standard deviations (e.g., standard deviation of means for CO2 fluxes for all collars at a site, i.e., with n = 7 for Site Ad etc…)
Figure 6: I suggest to delete this figure
Figure 7: Again, indicated source of variation.
'…grazing events…' change to '…simulated grazing events (cuttings)…'
y-axis: put cm day-1 in parentheses
Figure 8: amend y-axis to read: 'CH4 flux (mg CH4 m-2 h-1)'
Figure 9: cf comment to Fig. 8 and note 'µ' should be used on y-axis. Further the source and amount of fertilizer could be mentioned in the caption
Figure 10: spell out LAI in caption and indicate units on x-axis. Please provide some info on the nature of the points in Fig. 10. I guess each collar is represented 2 times (Year 1 and Year 2) to give 10 and 14 points. But what about LAI; when are the measurement done, or are the data a seasonal average? Please specify.

Figures responses:
Figure 2: The panel identifier (a) was not missing from the upper panel in the original and must have been lost during the upload. µmol on y-axis, panel (a) was added. Caption: deleted '(a)' in front of PPFD. PPFD has been written in full text in the figure caption.

Figure 3: error bars in panel (a) were specified in addition to the depth that VMC was measured at. We believe that the low VMC contents measured in May are the response to a very dry spell from mid April/early May (caused by a combination of low RF and increased temperatures). This is perhaps more clearly seen in the figures below, which show hourly RF (mm), and where T5CM and VMC data (recorded at 10 min intervals) are overlain. The low VMC values observed for this period are supported by WT data from the site (not shown), which showed that WT levels in early May were deeper than we were able to detect with our experimental set up. We agree that precipitation data would add useful information to this graph; unfortunately the sensor logging rainfall at our site malfunctioned in late December. We do not feel that it would be useful to provide rainfall data in such an incomplete form for the manuscript but are happy to incorporate it here in our response to your comment.

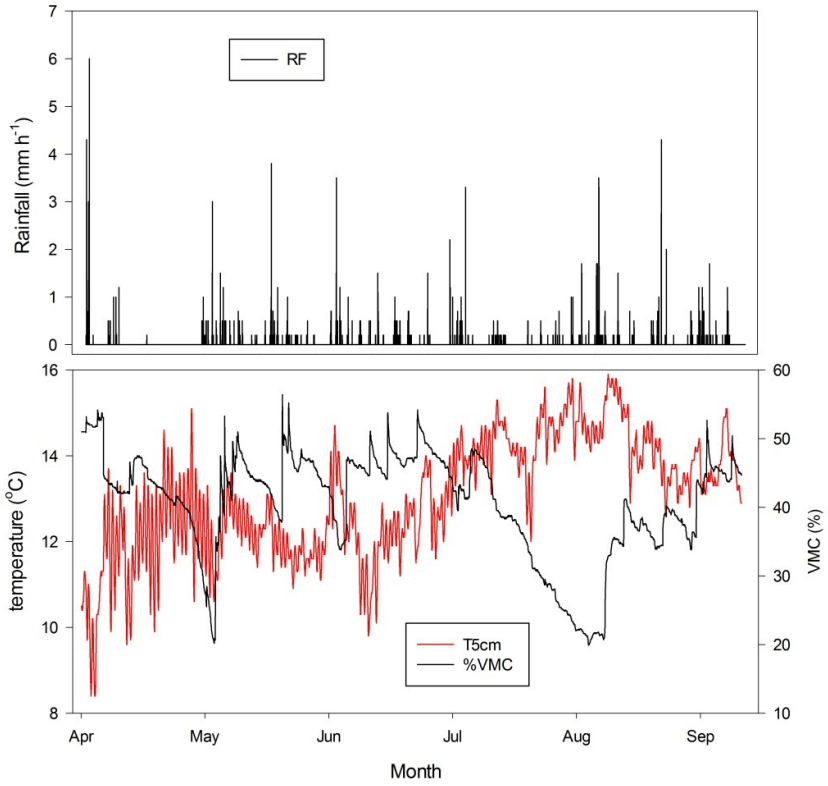

Figure 4: Reco and GPP has been spelt out in the figure caption. Accordance between data and 1:1 lines is explained in the text. h-1 is now used instead of hr-1.

Figure 5: amended y-axes to 'CO2 flux (g CO2-C m-2). Edited what standard deviations are.

Figure 6: Fig 6 has been removed.

New Figure 7: added source of variation and changed text to "…simulated grazing events (cuttings)…' and put cm day-1 in parentheses

Figure 8: amended y-axis to read: 'CH4 flux (mg CH4 m-2 h-1)'

Figure 9: edited µ on y-axis.
Information related to fertilisation event was added in the caption

Figure 10: LAI spelt in caption and units on x-axis. New information in caption pertaining to the nature of the points. 'Mean annual' added to LAI and Year 1 and 2 combined. Information regarding when and how LAI was measured is in the text and repetition in this caption unnecessary. This is now Fig. 6 in the revised manuscript.

Thank you.
Dr Florence Renou-Wilson