Final response to the reviewers comment from Anonymous Referee #2 on the manuscript bg-2020-171: "Evapotranspiration over agroforestry sites in Germany"

We thank you for your detailed feedback, suggestions and helpful comments on the manuscript. In the current document we give a point-by-point answer on above referee report. We show first the referee comments **(RC)** and second the authors answer **(AR)**. Specific changes in the revised manuscript are marked as green text as part of the authors response.

1. RC: *The authors measured evapotranspiration (ET) over pairs of adjacent agroforestry (AF, tree lines plus crop or grassland) and tree-free reference fields (MC for monoculture, only crop or grassland) at five sites in Germany over up to 2 years with 3 different methods. Plain eddy-covariance (EC) was used during campaigns, roving between sites due to limited gas analyzer availability. An energy balance method (ECEB) yielding ET as residual of EC measurements of the sensible heat flux and the non-turbulent energy balance terms, as well as a low-cost (EC_LC) method introduced elsewhere by the authors were operated continuously over the 2-year period and validated against EC. The paper presents - a comparison of the methods, in particular of the continuous methods versus EC - a detailed analysis of the energy balance closure (EBC) problem for the concerned methods (EC and EC-LC), and - a comparison of ET between AF and MC, between sites and years, and possible explanations on the result of this comparison, in particular on AF vs. MC The study presents the results of an impressive amount of work, applied according to best practice and including innovative elements, to questions relevant to BG and helpful for land use management decisions. The paper is well written and in terms of content and methods integrity could be published as is. However, the choice of the authors to treat so many dimensions (5 sites with 2 plots each, 3 methods, 3 above research questions) in one paper, makes the clear presentation of methodology and results a particular challenge. In this respect the readability of the mansucript could be improved in several ways. As far as I am concerned, all these improvements are optional; implementing some of them would qualify as minor revision and probably describes my recommendation best.*

1. AR: Thank for very much for your positive feedback and the detailed and valuable comments. We are confident that the quality of the manuscript will improve after considering your suggestions.

2. RC: *1. Whereas the abstract doesn't specify on the nature of the "monoculture" (MC) and the start of the introduction explicitly (and correctly) states that tree plantations can be monocultures as well, it becomes clear only later (maybe only in section 2.1 if I didn't overlook anything) that MC in this study refers exclusively to the crops/grass without trees (as opposed to a dense tree monoculture without a deliberately cropped understorey, which could be just an as relevant and logical comparison partner). Maybe it would be better to replace MC by something like e.g. NT for no-tree. If not I suggest to clarify earlier what MC in this paper means.*

2. AR: We prefer the current abbreviation of AF and MC. We clarified in the abstract and the main text what exactly the two abbreviations refer to.

In the Abstract:
"Therefore we hypothesize that short-rotation coppice agroforestry systems have higher water losses to the atmosphere via ET, compared to monoculture agriculture without trees (MC)."

In the introduction:
"The cultivation of fast growing trees with annual crops or perennial grass-lands on the same piece of land is an example of agroforestry (AF) \citep{Morhart2014,Smith2013} and has numerous environmental benefits relative to monoculture (MC) systems consisting only of crops or grasses without trees \citep{Quinkenstein2009}."

3. RC: *2. The fact that the authors seem to have tried (if I didn't misunderstand) both, down-correcting EBEC results to yield ET estimates with an EBC gap (sect 2.3.1, p8L17) and sometimes up-correcting EC and EC-LC results (Eq. 7-9, table 5), makes it hard to follow the interpretation of the results, particularly in places where the authors try to explain differences between methods / fields with their different EBCs (p14L32 / Sect. 3.3). Ideally it should be stated somewhere clearly that you present all results with (then down-correcting ECEB) or without (then up-correcting EC and EC-LC) an anergy balance closure gap. If then having to do the opposite, or a halfway correction, is still urgently needed for particular tasks in particular places, such as e.g. gap-filling between ECEB and EC-LC, it should be made clear at these points that this is the only purpose and usage of that "other" correction approach.*

3. AR: We treated the data in two different ways:
→ 1. we neither corrected the data up or down for the methodological comparison of the different methods based on the campaigns to explain potential differences between methods, as well as for the energy balance closure estimation (p15 Fig.4, p16 Table 3, p18 Fig. 5, p19 Table 4, p21 Fig. 6 of the initially submitted manuscript)
→ 2. for the comparison of annual sums of ET we did the 'half-way correction' of ET_{ECEB} (down-corrected, both for gap-filling of ET_{EC-LC} and ET_{ECEB}) and ET_{EC-LC} (up-corrected) to get closer to reality (as explained in Section 2.3 of the initially submitted manuscript) (p24 Fig. 8, p26 Fig. 9, p27 Tab. 5, p28 Fig. 10 of the initially submitted manuscript)
→ we will clarify the different up- and downscaling methods more detailed in the revised version of the manuscript, in the text supported by a table

In section 2.3 of the revised manuscript (“Gap-filling and energy balance closure adjustment”) we included a short explanation on the different gap-filling and energy balance closure adjustment procedures, as shown below:

2.3 Gap-filling and energy balance closure adjustment

For the comparison of ET_{EC}, ET_{ECEB} and ET_{EC-LC} and the estimation of the energy balance closure during the campaigns, we neither gap-filled the data, nor corrected the data for the energy balance non-closure. For the calculation of annual sums of ET_{ECEB} and ET_{EC-LC} data gaps were filled, as well as corrected for the energy balance non-closure by distributing the residual equally to H and LE. The residual was estimated by machine learning techniques.

In caption of Table 5 we included a statement that the uncorrected annual sums of ET are only used in this table and nowhere else, as shown below:

“Annual sums of energy balance closure corrected actual evapotranspiration, ET, potential evapotranspiration, ET$_0$, and precipitation, P, (\unit{mm\,a^{-1}}) for all sites, both set-ups (ECEB and EC-LC) and both years (2016 from April to December, and 2017 from January to December). We included annual sums of \unit{ET_{ECEB}} not corrected for the energy balance non-closure in brackets. The uncorrected annual sums of \unit{ET_{ECEB}} are only given in this table and were nowhere else used throughout the paper. The annual sums of ET and precipitation at Reiffenhausen for AF and MC in 2017 contain data from 01 January 2017 to 01 July 2017.”

4. RC: *3. In many figures an important correspondence between sub-panels (e.g. a-e) and cases (mostly sites, sometimes methods or periods) can only be established through the caption, which is even complicated by the letters changing their meaning with respect to site depending on whether one or two sub-panels are needed per site. I suggest to include the most important differences (e.g. site names) in the subpanels or next to rows or panels of subpanels, such that the figure can better stand alone. In Figure 9 quite suddenly abbreviations for the site names are introcudes which were nowhere used before (but might be useful for the above suggestion). It might also be worth thinking about re-naming the sites by characteristics relevant to the interpretation, e.g. crop vs. grass and/or the ranking of tree density.*

4. AR: we will include following abbreviations in following figures, referring to the initially submitted manuscript:
Figure 1: instead of (a), (c), (e), (g), (i) we write Dornburg, Forst, Mariensee, Reiffenhausen, Wendhausen; we remove the letters from the aerial photograph
Figure 2: we replace (a) by D-MC, (b) by D-AF, (c) by R-AF, (d) by W, (e) by F, and (f) by M
Figures 3 and A3: instead of (a)-(e) we write Dornburg, Mariensee, Forst, Reiffenhausen and Wendhausen
Figure 4: instead of the letters (a)-(i) we write D-AF, D-MC, F-AF, F-MC, M-AF, M-MC, R-AF, W-AF and W-MC
Figure 5: instead of the letters (a)-(i) we write D-AF, D-MC, F-AF, F-MC, M-AF, M-MC, R-AF, W-AF and W-MC
Figure 6: instead of (a)-(e) we write Dornburg, Forst, Mariensee, Reiffenhausen and Wendhausen
Figure 8: instead of (a)-(e) we write Dornburg, Forst, Mariensee, Reiffenhausen and Wendhausen
Figure 11: instead of (a)-(e) we write Dornburg, Forst, Mariensee, Reiffenhausen and Wendhausen
Figure A1: instead of the letters (a)-(i) we write D-AF, D-MC, F-AF, F-MC, M-AF, M-MC, R-AF, W-AF and W-MC
Figure A2: instead of the letters (a)-(h) we write D-AF, D-MC, F-AF, F-MC, R-AF, M-MC, W-AF and W-MC
Figure A4: instead of (a)-(e) we write Dornburg, Forst, Mariensee, Reiffenhausen and Wendhausen

5. RC: *4. Textbook knowledge that many others would present not at all or in an appendix is reported in the methods section. This is not necessarily a bad thing (although contributing to the overall length), but currently it is not done consistently. Equation 4 and 5 detail on quite straightforward conversion matters, and equations 12 and 13 on saturation vapour pressure and its slope, but on the other hand section 2.2.3 (p7L26) merely states that "mole fraction was calculated using measurements of relative humidity, air temperature and air pressure...", although this conversion involves at least as many reproduction-relevant decisions (and maybe partly same equation(s)) as the ones mentioned before. Ways out could be e.g. either to drop all these details, or insert an appendix section where such equations are gathered, some of which could then be referred to from multiple points of the paper if necessary.*

5. AR: To keep the main text concise, and given that some of the equations were already described in Markwitz and Siebicke (2019), presenting the EC-LC set-up, we moved Eqs. 4-6 and Eqs. 11-18 to the appendix and included conversion formulas from section 2.2.3. The Appendix is now structured as follows:

A Derivations
A1 Half-hourly ET rates and soil storage flux
A2 Water vapour mole fraction C_{H2Ov} from the thermohygrometer
A3 Canopy resistance

6. RC: *Further comments on the analysis:*
5. p08L19 (Sect 2.3.1): I may be overlooking something, and given how little we know about the source of the EBC problem your solution might be as good or bad as the more widespread Twine partitioning, but I do not understand why the latter cannot be applied to your data. Mathematically a Bowen-ratio conserving correction is equivalent to correcting both fluxes by the same factor 1/EBR, without any explicit need to know/compute/introduce the Bowen Ratio itself (and even if this was the case there would probably be an analytical or iterative solution to the problem). So if the available H (from EC) is already subject to the closure problem and does not need to be "down-corrected", the only thing left to do is to multiply the residually determined LE with EBR to get the desired estimate of a "non-closing / EC-like" LE.

6. AR: The suggested solution of multiplying LE_ECEB with the EBR is somehow limited by the fact that the EBR from EC was only available for the duration of the measuring campaigns and this would require the prediction of the EBR. Another solution would be to multiply the mean EBR (derived as the slope between H+LE and Rn-G from the campaigns) with the 30-min LE_ECEB. This would be even less accurate due to missing the temporal variability of the EBR throughout the day and the year. As already stated, the main question is not which method to use, it is rather the question how the residual gets partitioned to the different components. From our point of view, the current solution was the only viable option given the current data.

7. RC: *6. p11L05 (Sect. 2.6 / Equation 18): After an elaborate description of how the Penman-Monteith approach is used to infer conductances, the simpler (humidity-free) Priestly-Taylor approach is introduced for the Budyko analysis, although alternatives consistent with Penman-Monteith (e.g. FAO grass reference ET) exist. Was there a particular reason for this decision? Luckily it probably affects all sites similarly (more similar than in a global study mixing very humid and arid sites) and seems only to be needed in Fig. 10, even there only slightly changing X axis position but not the overall pattern.*

7. AR: Indeed, ET_pot could have been calculated as well with the PM approach, which required additional variables such as the aerodynamic conductance, derived from EC; the aerodynamic conductance was gappy, whereas meteorological data (Rn, G and Ta) were gap-filled which allowed us to compute annual sums of ET_pot; hence, more practical reasons led us to apply the more simple Priestly-Taylor approach; we compared the two approaches (PM and Priestley-Taylor) for a sample period and the differences were small, and yes, the effect is the same for all sites.
We gave an explanation as well in the revised version of the manuscript:

In order to investigate the relationship between the evapotranspiration index ($\sum \mathrm{ET}/\sum \mathrm{P}$) and the dryness index ($\sum \mathrm{ET_P}/\sum \mathrm{P}$) we used the simplified Priestley Taylor equation after (Priestley and Taylor, 1972) to calculate a potential evapotranspiration. For this analysis we used the simplified Priestley Taylor equation due to the lack of gap-filled data required in Eqs. (14) and (15). The Priestley Taylor approach depends only on available energy and air temperature, which both were available as gap-filled data sets. A comparison of the two approaches (Penman-Monteith and Priestley-Taylor) for a sample period yielded marginally differences between $\mathrm{ET_P}$. However, the Priestley Taylor potential ET was only used for the specific analysis and affected all sites the same.

The Priestley Taylor equation (Priestley and Taylor, 1972) is then given as

$$ET_{PT} = \alpha_{PT} \frac{s}{s+\gamma}(R_N - G) \quad (18)$$

with the Priestley-Taylor coefficient, $\alpha_{\mathrm{PT}} = 1.25$.

8. RC: Further comments on technical / presentation details:
7. p01L14 (abstract): Consider rewording "superior performance" to make clear that this indicates superior agreement with the widespread EC method. This is not necessarily identical to superior performance in capturing true ET.

8. AR: We changed this accordingly:

"Root mean square errors of LE_{EC-LC} vs. LE_{EC} were half as small as LE_{ECEB} vs. LE_{EC}, indicating a superior agreement of the EC-LC set-up with the EC set-up compared to the ECEB set-up."

9. RC: *8. p01L17 (abstract): There is an ongoing debate whether, how much and how we should continue to base conclusions about differences on significance (e.g. Amrhein et al., Nature 567:305, DOI: 10.1038/d41586-019-00857-9). While reporting p-values in a paper for the sake of completeness cannot do much harm (without wrong interpretation), care should be taken especially where wrong use in the past was particularly popular, and one of these cases is inferring that a difference is nonexistent or unimportant from a "failed" (insignificant) test. This sentence (and versions of it in the main text) comes somewhat close to suggesting something like this (although not explicitly claiming it). It might be more convincing to show (as done in the main text) how small the difference actually was (maybe compared e.g. to the mean ETs or to the inter-site, inter-period, or inter-method variability that was probably at the bottom of the significance test) and then it could still be mentiond if wanted (here or elsewhere) that the difference was also statistically insignificant (which depends strongly on sample size even if all the means and variances stay the same, and unlike conclusions from significant results, conclusions from insignificant results have the property to become the more likely the smaller the sample size). Also note that if keeping reporting the p-values somewhere, they should be rounded to a reasonable number of digits; especially for the second one at L23, $p = 0.0007$ or $p < 0.001$ would be sufficient.*

9. AR: We removed the p-values and statistics in the Abstract as shown below:

"[...] With respect to the annual sums of ET over AF and MC, we observed small differences between the two land-uses. We interpret this as an effect of compensating small-scale differences in ET next to and in between the tree strips for ET measurements on system-scale. Most likely, differences in ET rates next to and in between the tree strips are of the same order of magnitude but of opposite sign and compensate each other. Differences between annual sums of ET from the two methods were of the same order of magnitude as differences between the two land-uses. In contrast, we observed higher mean evapotranspiration indices ($\sum$ET/$\sum$precipitation) across sites for a drier than normal year (2016) compared to a wet year (2017) independent of the land-use or method. This indicates that we were able to detect differences in ET due to different ambient conditions with the applied methods.

We conclude that agroforestry has not resulted in an increased water loss to the atmosphere indicating that agroforestry in Germany can be a land-use alternative to conventional agriculture."

We will change the explanation in the same way in the main text.

10. RC: *9. p02L08 (Sect. 1): "comparable" reads strange in this context. Basically they are, arent't they? As far as I know the term monoculture does not distonguish between agriculture and forestry. Also see comment 1.*

10. AR: Yes, we removed the term "comparable" and rewrote the sentence:

"SRC plantations are monoculture systems with a single tree species grown."

11. RC: *10. p02L32 (Sect. 1): "such as" reads strange in this context. Maybe ", i.e." instead?*

11. AR: We changed this accordingly:

"For agroforestry systems we formulated the same hypothesis, i.e. system-scale evapotranspiration over agroforestry systems is higher compared to monoculture agriculture without trees."

12. RC: *11. p03L01 (Sect. 1): depend*s**

12. AR: We changed this accordingly.

13. RC: *12. p03L29 (Sect. 2.1): While reporting the access date of an URL is important if that URL is a source of data/information that couldn't be replaced by a better source, in this case the URL more has the role of an advertisement or a reference to further information for interested readers, and what exactly they will find at the project site if it still exists is not relevant to the paper. For this an access date seems inappropriate. If you weren't asked to add it during the access review, I'd suggest to remove it.*

13. AR: We were asked to include the access date to all the URL's in a previous publication, so we just added it here as well. It seems to be a journal requirement. But, we can change this later, if required.

14. RC: *13. p04L01 (Figure 1): Maybe add a scale bar to the aerial views (or one scale information for all if they are the same). I wonder how wide the elongated MC strip at Forst (b) was, how different the management west and east of it was, and how this is reflected in Sect. 3.2 (footprint analysis).*

14. AR: We will add a scale at each of the sites because they are not the same. The strip at Forst MC is 48 m wide and the management at the east and the west of this strip was always the same, but different from the MC strip. The crop type at the MC strip was always the same as in between the tree strips at the agroforestry system. As shown in Fig. 3 the footprint extended beyond the MC strip, hence, fluxes at the MC were also affected by the nearby crop fields.

15. RC: *14. p05L1 (Table 1): System size. Specify if it refers to AF, MC or the sum of both.*

15. AR: The system size referred to the AF system only and we changed it from “System size” to “Agroforestry system size”.

16. RC: *15. p05L18 (Sect. 2.2.1): Did I understand correctly that this required at least two available Li-7200? If yes clarify, if no reword sentence.*

16. AR: Yes, we deployed two LI7200 in parallel in 2017. We clarified this:

“During the field campaigns the standard set-up was extended by an enclosed-path infrared gas analyser (LI-7200 , LI-COR Inc., Lincoln, Nebraska, USA). In 2016, the campaigns were conducted separately at the AF and MC systems with one available gas analyser, whilst in 2017 both systems were sampled simultaneously with two available gas analyser.”

17. RC: *16. p06L01 (Table 2): ppp for pressure seems unusually long/complicated. Also, in the row that is solely about ppp it looks a bit lonely (and hard to understand) without the long explanation "Atmospheric pressure". I acknowledge that you aimed at consistently giving the long name only upon first occurrence in the table, but here would be space and reason enough for an exception. Or maybe the row could be switched with the BME280.*

17. AR: We named it P_A for atmospheric pressure.

18. RC: *17. p07L19: "unpublished data" and then no matching entry in the reference list is a bit vague. If there is not even an internal report to refer to (which could then be listed in the references), "pers. comm." would probably be more appropriate, and at any rate in this case the institutional affiliation of Schmidt et al. should be given e.g. in the acknowledgement, to ensure traceability.*

18. AR: We changed this:

“Marcus Schmidt (pers. comm., Georg August University of Goettingen, Buesgen Institute, Soil Science of Tropical and Subtropical Ecosystems)”

19. RC: *18. p07L28 (Sect. 2.2.3): Even though referring to a publication about the method where all this can probably be read in detail, not mentioning that there was a (probably large) spectral loss correction falls back behind the information given in the introduction (p03L17), and will make readers looking for this information in the methods section (the most logical place) wonder if and how this method could work at all.*

19. AR: We included more back-ground information on how the set-up worked in this section and for reproducibility also equations on how we transformed RH, TA and PA readings into a water vapour mole fraction into the Appendix.

Please find here the extended section 2.2.3:

2.2.3 Low-cost eddy covariance (EC-LC)

The EC-LC set-ups comprised of the same ultrasonic anemometer uSONIC3-omni as used for the EC and ECEB set-ups plus a compact low-cost relative humidity, air temperature and pressure sensor (BME280, BOSCH, Germany, Table 2). Water vapour mole fraction was calculated using measurements of relative humidity, air temperature and air pressure from the low-cost thermohygrometer. A derivation of the water vapour mole fraction from the low-cost thermohygrometer is given in Section A2. The turbulent water vapour fluxes were calculated as the covariance between the vertical wind velocity and the water vapour mole fraction from EC-LC, as per the principle of the eddy covariance method (Baldocchi, 2014). The cheaper but slower thermohygrometer had inferior spectral response characteristics compared to a gas analyser of fast response. The mean spectral correction factor of the thermohygrometer was 42% larger than for the LI-7200 fast response gas analyser for reference, with a 78% larger mean time constant of the thermohygrometer compared to the LI-7200. The mean time constant of the thermohygrometer and the LI-7200 was 2.8±1 s and 0.6±0.3 s, respectively (Markwitz and Siebicke, 2019). Spectral losses in the high-frequency range of the thermohygrometer were corrected by the fully analytical correction method of Moncrieff et al. (1997), which was explicitly recommended for either open-path sensors or closed-path sensors of heated and very short sampling lines. A detailed description and application of the EC-LC set-up for evapotranspiration measurements over AF and MC is given in Markwitz and Siebicke (2019). Evapotranspiration from EC-LC was neither gap-filled for the methodological comparison nor for the analysis of the energy balance closure due to the risk for new errors and artefacts from the respective gap-filling method.

And please find here the Appendix from the revised version of the manuscript:

A2 Water vapour mole fraction $C_{H_2O_v}$ from the thermohygrometer

In the following section we present the derivation of the water vapour mole fraction $C_{H_2O_v}$ from relative humidity, air temperature and air pressure from the low-cost thermohygrometer. The derivation was already given in more detail in Markwitz and Siebicke (2019).

The water vapour mole fraction, $C_{H_2O_v}$, was derived from the definition of the specific humidity, q, as the quantity of water vapour per quantity of moist air. The latter two quantities were expressed as the density of water vapour, $\rho_{H_2O_v}$, and moist air,

ρ_m, respectively. The density of moist air is defined as the sum of the density of dry air, ρ_d, and the density of water vapour.

$$q = \frac{\rho_{H_2O_v}}{\rho_m} = \frac{\rho_{H_2O_v}}{\rho_d + \rho_{H_2O_v}} \quad \text{(A4)}$$

We then replaced the density of water vapour and the density of dry air in Eq. (A4) as per Eqs. (A5) and (A6), respectively,

$$\rho_{H_2O_v} = \frac{C_{H_2O_v} \cdot M_{H_2O_v}}{V_m} \quad \text{(A5)}$$

$$\rho_d = \frac{p - e_a}{R_d \cdot T} \quad \text{(A6)}$$

with the molar mass of water vapour, $M_{H_2O_v} = 18.02\,\mathrm{g\,mol^{-1}}$, the molar volume of air

$$V_m = \frac{\Re \cdot T}{p}\,(\mathrm{m^3\,mol^{-1}}), \quad \text{(A7)}$$

the universal gas constant, $\Re = 8.314\,\mathrm{J\,mol^{-1}K^{-1}}$, and the specific gas constant of dry air, $R_d = 287.058\,\mathrm{J\,kg^{-1}\,K^{-1}}$.

Solving Eq. (A4) for $C_{H_2O_v}$ leads to the water vapour mole fraction

$$C_{H_2O_v} = \frac{q\,\Re\,(p - e_a)}{p\,M_{H_2O_v}\,R_d(1-q)}. \quad \text{(A8)}$$

The specific humidity in Eq. (A8) was calculated as a function of relative humidity, temperature and air pressure measurements from the thermohygrometer:

$$q = 0.622 \cdot \frac{e_a}{p} \quad \text{(A9)}$$

The actual vapour pressure, e_a (kPa), in Eq. (A9) was calculated from an approximation of the saturation vapour pressure, $e_*(T)$ (Stull, 1989) and from relative humidity, RH,

$$e = \frac{RH \cdot e_*(T)}{100} \quad \text{(A10)}$$

$$e_*(T) = 0.6112\,\exp((17.67T)/((T + 273.15) - 29.66)) \quad \text{(A11)}$$

(A12)

20. RC: *p08L27 (Sect. 2.3.2): This sentence at a first glance seems to contradict the sentence at the top of the same page. Maybe start like this: "Unlike for the methodological comparison and energy balance analysis, a gap-filling of EC-LC could not be avoided for [this and that, surely not for annual ET sums]. Therefore, for these analyses..."*

20. AR: We changed this in the revised version of the manuscript:

> Unlike for the methodological comparison and energy balance analysis, a gap-filling of ET_{EC-LC} could not be avoided for the calculation of annual sums of ET. Therefore, for these analyses we gap-filled the half-hourly ET_{EC-LC} with half-hourly ET_{ECEB} and corrected both ET_{EC-LC} and ET_{ECEB} for the surface energy balance closure as follows

21. RC: *p09L10-15 (Sect. 2.4): At the beginning consider replacing "As the" by "By". Citing software tools / packages can be useful when a) advertising that the own code can be made available to the reader or when b) Reproduction of results depends on using the same tool (mention package, e.g. because the method is so complicated it might give different results in other langauges). The major axis however is a statistical term independent of and introduced before R, and if correctly implemented in the package should yield the same result as any self-written implementation. Therefore it seems more important to refer to a statistical textbook or paper - e.g. Webster 1997, Eur. J. Soil Sci. 48:557, doi:10.1111/j.1365-2389.1997.tb00222.x (which by the way also provides in its "calibration" section support for your decision in other places to treat variables to be filled as "dependent" (Y) variables in a regression).*

21. AR: We considered the publication and changed the text:

> **2.4 Energy balance closure estimation**
>
> The energy balance closure (EBC) was quantified in two ways:
>
> 1. As the linear regression between the available energy (R_N - G - S), and the sum of the turbulent flux components (LE + H). We applied the major axis linear regression (Webster, 1997), which assumes equally distributed errors in both time series. We interpret the slope between the available energy and the sum of the turbulent fluxes as the closure of the surface energy balance. A slope of one and an intercept of zero corresponds to perfect energy balance closure. In the present study both the slope and the intercept were considered as variable.

22. RC: *p13L01 (Fig. 3): Cannot see MC footprint in subpanel d, is this somehow related to the inavailability of a campaign at Reiffenhausen mentioned at p08L24? But footprint modelling only relies on data measured anyway by the EC-H setup needed for ECEB and EC-LC? Maybe it would be good to state in a prominent place (or each time a particular result seems to be missing, e.g. in Fig. 3d, Fig. 4 between g and h, Table 3 row Mariensee EC-LC, Fig. 5, Table 4, Fig. 9) what was the reason (in most cases it seems to be the missing campaign at Reifenhausen MC, but not so e.g. in table 3).*

22. AR: The footprint at Reiffenhausen MC is missing due to the unavailability of a campaign there. Yes, the footprint estimation depends only on mentioned variables, but since the campaign did not take place, we decided to not include the data for the particular site and time period. It would distract from the interpretation of ET during the campaign over the AF in Reiffenhausen. In the footprint climatology for the whole year (Fig. A3 in the appendix) we did include Reiffenhausen MC, as this information is used to explain potential differences in annual sums of ET. We clarified where and why data were missing in the following figures and tables of the revised version of the manuscript: Tables 3-5; Figures 3-6, 8, 9, A1, A2, A5

23. RC: *p17L26 (Sect. 3.4.2 / describing Fig. 6): Is "square" a commonly recognized or self-explaining description of this kind of diel curve?*

23. AR: we rewrote the sentence to:

“The diel cycle of the EBR for the first group of sites (Dornburg and Wendhausen,...) show a strong increase from 6 am to 8 am, followed by a flat period between 8 am and 2 pm, and a strong increase thereafter until 6 pm.”

24. RC: *p20L29 (Sect. 3.4.3): It took (me) several readings to understand how and why you changed magnitudes, after talking about measured data all the time. Basically the idea of this whole section is simpler and more straightforward than it looks, and if needing to shorten the paper, this (writing it simpler or dropping it completely) would be my first suggestion. It can be reduced to the message "the importance of a relative uncertainty in a flux for the EBC scales with the magnitude of that flux". Even this effect probably vanishes when looking at absolute rather than relative errors / uncertainties, and even though it is not completely irrelevant for deciding how much to invest into improving which flux, it could probably also be demonstrated in a more general way with symbolic maths or a thought experiment.*

24. AR: We removed this section.

25. RC: *p25L04 (Sect. 3.5.2) "related" reads strange in this context, maybe "plots with an ET index".*

25. AR: We changed this accordingly:

"The figure indicates first that plots with an ET index larger than one were water limited, [...]"

26. RC: *p26L05 (Sect. 3.5.3): "reduce" or "reduced"? The former simply repeats (and takes for granted, but this semms save to me) what the cited references state, while the latter implies a claim that it can be seen well in your data, which should then however be confirmed by a clearer statement.*

26. AR: 'reduce' represents better what we wanted to say!

27. RC: *p26L01 (Sect. 3.5.3): The methodology section preferred aerodynamic conductance, here aerodynamic resistance (the inverse) is used. Consistently using only resistance or conductance could help to avoid confusion.*

27. AR: We changed it to 'Aerodynamic resistance' in the methodology section.

28. RC: *p31L06 (acknowledgements): Data from other sites than your own seem to have been used only in one place of the appendix, Fig. A6, if I didn't overlook something. If it is needed at all (there seems to be little connection to the main text), the small amount of sites used there seems to suggest that acknowledgements to the individual site PIs is at least as, or more, important than to the (for this number of sites quite lengthy) list of networks.*

28. AR: We shortened it accordingly:

"We are thankful for the provision of eddy covariance data acquired and shared by the FLUXNET community."