

Interactive comment on “Evapotranspiration over agroforestry sites in Germany” by Christian Markwitz et al.

Anonymous Referee #2

Received and published: 2 July 2020

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

C1

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 manuscript 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.

Overarching comments each concerning several points of the paper at once:

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. 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 energy balance closure gap. If then having to do the opposite, or a halfway correction,

C2

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. 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 introduced 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. 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.

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

C3

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. 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.

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. 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

C4

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

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

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

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

C5

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. 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. 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. 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.

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..."

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 languages). The major axis however is a statistical term

C6

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).

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 Reiffenhausen MC, but not so e.g. in table 3).

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

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.

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

C7

p26L05 (Sect. 3.5.3): "reduce" or "reduced"? The former simply repeats (and takes for granted, but this seems 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.

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.

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.

Interactive comment on Biogeosciences Discuss., <https://doi.org/10.5194/bg-2020-171>, 2020.

C8