Second Review of manuscript bg-2020-445 by Vogl et al.

The authors addressed the major comments of my first review in a more or less satisfying way (see detailed comments below). However, they overlooked or disregarded all my "Minor Comments", which is not acceptable. In addition, the author responses (and revisions) led to some additional issues that need to be addressed before the paper can be published.

Important: The page and line numbers I use in the following refer to the revised manuscript version with markups (ATC1).

COMMENTS

1) Please address the list of "Minor comments" in my first review. These comments have not been considered in this revised version (Some of the comments may have become obsolete due to other changes)

2) R1C1 response: I still miss the important discussion (in the Discussion section) about the assumption of scalar similarity, i.e. whether and why the results for H2O investigated here (with minor results also for CO2) can be applied to all other scalars, especially the ones for which REA is usually applied.
This is especially critical because the authors state on P24, Line 8: "Choosing the optimum proxy scalar is critical for the methods success". This sentence implies that a general similarity between all scalars is not expected.

3) P1, line9: In the context of major comment 2 above, this statement is over-ambitious especially concerning the part "...formulating universally applicable recommendations...", and I suggest to downgrade it to some extent.

4) R1C4 response: Such a strong effect of one single outlier makes the suitability of the used evaluation method questionable.

5) R1C6 response: The pooling of data from all three sites in the evaluation can be problematic. Does this imply that β_w and the w-statistics as well as β_T values are fully independent of the site conditions (canopy height, roughness, correlation between proxy scalar and scalar of interest, etc.) ? This issue needs some statements/discussion in the text.

6) R1C5 response: The newly added Table 2 is very important for the present study. There are two important questions arising from it that deserve some thoughts/discussion: i) can the average b_w be considered as a site independent constant?; ii) Would the use of an average constant b_w yield similarly good results like the use of half-hourly b_w-values?
The (successful) use of an overall constant b_w value would strongly simplify the REA measurements.

7) In various parts of the text and the abstract the terms "proxy-based approaches" or "proxy-dependent approaches" are used synonymously for the models 1 and 2. However, this is misleading and not correct, because the model 4 approach is proxy-based as well (and also model 3 uses w as proxy, though it is not a scalar). I suggest to use a more suitable and specific expression for models 1 and 2 like e.g. "dynamic scalar proxy approaches" or just "model 1 and 2 approaches"

8) Figure 10: correct the y-axis titles in the left and central panel to β_T.

9) Figure A1: For H2O in Fig. 9 an individually adjusted deadband width was used. Was the same deadband width also applied for CO2, or was it indivdually adjusted for CO2. In either case, does an individual adjustment yield the same optimum deadband widths for H2O and CO2?