

Responses to reviewer comments

Partial energy balance closure of eddy covariance evaporation measurements using concurrent lysimeter observations over grassland

Peter Widmoser & Dominik Michel

We thank referee #1 for the comments. We respond to them below the referee's comment (in italic).

General Comment:

The manuscript is relevant for hydrological studies. The energy balance gap in eddy correlation (EC) measurement is an ongoing topic that deserves attention. The manuscript is well written. It is however rather brief and seems to be primarily readable for insiders. I am personally in favor of these short and concise manuscripts. It has the characteristics of a technical note. The editor could consider to publish it as such.

The positive feedback on the importance of the topic as well as on the writing is much appreciated. We leave it to the Editor to publish the paper as a technical note.

Major Comment

Throughout the manuscript d is considered to describe a systematic difference between LY and LE. Lysimeter measurements are systematically larger than EC measurements. However d can be negative or positive for different sites. Stating that d describes a systematic difference is confusing and not correct. d is simply the intercept of the linear regression model having the energy balance gap (epsilon) at the xaxis and LY – LE on the yaxis. If the energy balance gap (epsilon) is zero d will remain. We could even argue that we could drop d in the model since cLE does already quite a good job in correcting LE (table 3a in the manuscript). The description of d and the conclusions based on d should be corrected throughout the manuscript.

We changed in line 145 from “the y-intercept (d) can be interpreted” to “might be interpreted”.

We will, however, not correct the description of d and conclusions drawn from it. We have the following arguments:

- Our interpretation is correct from a statistical point of view
- Even if the interpretation of d may be irritating in the case of negative d as e.g. for Fendt 2013, we feel this irritation is helpful for a more detailed analysis of the data
- The new method proposed has been applied so far only to four stations. Further experience with data evaluations from more sites may decide about the d -interpretation

Specific Comments

- 1) *The formulation of line 30 to 33 is unclear to me. Was just the difference between LE measured with EC and LE measured with lysimeters smaller than the energy balance gap? I would suggest to reformulate this part and explain “reduced the differences”.*

This refers to the difference between EC and lysimeter data. The formulation was indeed unclear, as the numbers referred to the case without forced closure, which is confusing with the previous sentence.

We reformulated lines 30 to 32 accordingly: After forced closure of the energy balance, the difference between daytime LY and EC data on two fields could be reduced from -28.8% to 6.2%, respectively from -26% to -12.3%, with an accuracy....

- 2) *In line 37 partial evaporation closure is mentioned. Shouldn't this be partial energy balance closure? It is not clear what is meant.*

We corrected in line 56 (not line 37) to “partial energy balance closure”

- 3) *In the equations the dimensionless weights are in the form w_A , w_H , w_L . I find this confusing. I would suggest to use subscripts for A, H and L. Otherwise I could interpreted w_H as a weight times sensible heat, which is not the case.*

We feel it might be useful to apply the weight-notation in agreement with two previous publications in AGRFORMET. We leave it to the Editor to decide whether to change the weight-notations.

- 4) *LY is used for lysimeter LE in the equations. This is confusing. I think the notations should be reconsidered.*

We change the notation for lysimeter- and EC-values to LE_{LY} and LE_{EC}

- 5) *In line 312 the difference between LE and LY for humid climates is described as surprisingly little. I think this is not correct. The difference is large. 10 to 30 W/m² is similar to 0.35 to 1 mm/d which is equal to 128 to 365 mm/year. On a water balance in most regions of the world including Europe these differences are large.*

We changed the sentence on line 312 to: “They differ much less for humid climate with around 10 to 30 Wm⁻² (0.35 to 1.0 mm d⁻¹) than at Majadas-Station with around 30 to 60 Wm⁻² (1.0 to 2.1 mm d⁻¹)”.

- 6) *The formulation of line 320 to 325 is unclear. “The adjustments reached in this paper are higher”. Did the corrections/adjustments lead to better results? How come? If I am correct the literature citations in these lines have used full energy balance closure techniques with still large differences with lysimeter measurements right? (I haven't checked) This is something different.*

We replace line 320 by: “The effectiveness of our method is demonstrated by comparing our results given in Table 5 with the following results achieved by former authors:”

- 7) *Line 352. To my opinion better to reformulate this line. The presented manuscript is basically fitting a certain model, but that doesn't tell anything about what is best.*

We changed the sentence on line 352 to: "...gives, according to our knowledge, for the first time a fully rational method to partially close the energy gap and a more detailed description of the correlations between *LY* and *EC*-observations

- 8) *In line 359 the authors suggest to use 5 to 10 min resolution lysimeter data. I think this is unrealistic. There are no lysimeters other than dead weight compensated lysimeters that can measure accurately at such a fine resolution. Even presenting data on hourly intervals is to my opinion debatable. I would rather suggest to do the analysis on daily data or the sum of daytime data. The analysis would than be much less affected by lysimeter measurement errors and as proved in manuscript the correction weights for most situations are constant during the day.*

We can understand the reasoning of the referee. But we still feel it is justifiable to at least propose reduced time intervals. Higher time resolution of *LY*-data could can easily be retrieved. However, thorough filtering of the data is required for high resolution data, e.g. using the AWAT filter (Peters 2017, <https://doi.org/10.1016/j.jhydrol.2017.04.015>). As far as our method is concerned, Eq. 3 relies on statistics and its reliability is as such dependent on a high input number.

Technical corrections

- 9) *Typo: At the end of line 247 the word "und" should be "and".*

We corrected the error in line 247

- 10) *Figure 7a. Legend item "zero line" should be brown.*

We changed the color of the zero line in Fig. 7a

We thank for the careful, profound and challenging review of reviewer #2, which underlines our intentions (1) to draw attention to the subject and (2) demonstrate that the method proposed gives plausible results under various local conditions. We acknowledge their input on the discussion on the standard deviations, which made the manuscript more concise.

Some comments suggest adding additional figures and tables. We are, however, concerned about the length of the manuscript. Nevertheless, we will do our best to respond to reviewer #2 (italic). In case we cannot fulfill all expectations, we suggest to publish the paper as a technical note. We respond to the referee's comments (italic) below.

Note that the new nomenclature for the variables (e.g. LELY instead of LY, change upon input from Reviewer #1) is not yet used here for simplicity.

Main critics:

1) A lot of results are presented in this paper, but often without much comments by the authors. In that respect the paper looks more like a technical report which may form the basis of peer-reviewed paper. I urge the authors to take the reader along the circa eight tables and describe in text what the main message of each table is.

We changed the text accordingly at several places, see responses below.

2) Although the authors focus on the relation between LY and EC measurements, they also use the other observations of the surface energy balance, net radiation and surface soil heat flux in essential parts of their analysis. These observations should be described aswell in section 2.

Section 2 underwent major changes, including a more detailed description of the sites individually, comprising a list of the used sensors.

3) A rationale for the used method is lacking, given that the authors state that they are mainly interested in the relation of LY and EC evaporation observations.

We want to demonstrate the applicability of approach described by Widmoser et al (2018) to several stations as mentioned in the abstract. In our opinion, the rationale for this study is made very clear.

4) Section 4 is more a summary of results than a discussion. For example the part on standard deviations needs a discussion on what these comparison of SDV means and what can be learned of it. Now there are so many nice statistical results and apparently so little conclusions can be drawn. The question is whether these statistical techniques alone are sufficient to grasp the mechanisms behind the differences observed. Perhaps these should be accompanied by detailed case to case studies.

It is beyond the scope of this article to grasp all mechanisms behind the differences observed. We explain the meaning of SD-differences. See comment 21.

5) The text is not always as precise as it could be, some examples are given below. But there are more of these occasions. Please copy edit the text carefully on this aspect.

We will consider all comments below.

Specific comments

6) L45: *How is the evaporation fraction used to correct? Is that different from Bowen ratio preservation?*

This comment refers to our citation of Gebler et al (2015). We refer to p. 2151, 3 lines above Eq. 10 in Gebler. Their considerations are rather complex and we prefer not to explain them in detail in our article.

7) *Table 1: It would be nice to have the other information (measurement time interval, vegetation type, period of the day used) also in the table. This may require to turn the table by 90 degrees.*

We incorporate period of day, time interval and vegetation into Table 1 – or we add one table (1b) below Table 1 listing the requested information for all stations.

8) *What would be the influence of the oak trees at station Majadas on the flux observations.*

It is not the scope of this manuscript to investigate these influences and we refer to authors involved in the measurements at the site. However, since the oak trees are rather widely dispersed, we assume a mild influence on turbulence that would be significant to differences between lysimeter and EC measurements.

9) *S2.2: Here general error characteristics are given. Are the authors sure that these can be applied to the various sites used here. Are there any specific circumstances which may have an influence on the error characteristics. For example, how well are the conditions in the lysimeter kept comparable to the surrounding fields. Are there infrared surface temperature observations to judge possible inhomogeneities between lysimeter and surroundings?*

We are in contact with the PIs of the stations used in this study and trust that all instruments are maintained with the care needed. This includes keeping the lysimeter vegetation comparable to the surrounding area. Investigating differences between LY and EC originating from such conditions is not the scope of this article. We are citing articles dealing with this problem (Evetts et al. 2012) and other sources of uncertainty of lysimeters.

9) *S2.3: Here I have the same questions. I find these error estimates too general. It is always good to look at specifics of datasets/sites. EC measurements require all kinds of corrections. I miss a statement on the applied methods, and any differences in treatment per site.*

We are aware of the extensive correction procedure that is required to obtain accurate flux measurements when using the eddy-covariance method. However, we are in contact with the data providers of every single station presented and trust their capacity of correctly processing these eddy-covariance data. Describing the flux corrections in detail would take up an unnecessarily large part of the article.

10) *L113: Wohlfahrt and Widmoser (2013) apply the out-of-bound concept for corrected EC observations to judge whether these corrections lead to physically realistic values. Here you apply to the uncorrected EC observations, which may be physically unrealistic as this is the reason that you want to correct. Please clarify.*

Our screening according to the Out-of-Bound-concept avoids data combinations which

correspond to case 2 and case 3 in Fig. 1 of Wohlfahrt and Widmoser (2013).

We add at L113: According to this concept, the ratio $r1 = (ra+rc)/ra$, where ra and rc denote aerodynamic and canopy resistance, must numerically be within the range of 1 to infinity (see Fig. 1 in Wohlfahrt and Widmoser, 2013). Case 2 represents $r1 < 0$ and case 3 represents $r1$ is between 0 and 1. Data corresponding to case 2 and 3, are thus omitted.

11) S2.5: *To calculate the energy imbalance (epsilon) the authors also needs the available energy which is built from net radiation and the heat stored into the vegetation-soil system. I miss in section 2 a description of these observation for each site including error characteristics. It would greatly help when for each site a characteristic diurnal cycles are displayed of the components of the energy balance and the resulting imbalance (epsilon). This then should include a discussion if any peculiarities show up in these observations.*

We add figures of the diurnal cycles for observed A, H, LE, LY, and epsilon for all four stations in section 2.

12) S2.5: *The authors state that they are not interested in analyzing the full energy budget, but only the evaporative component. Alternatively the authors could have chosen to analyze the relation between the lysimeter and EC measurements. It would be nice if the authors could discuss the arguments for choosing not to follow that line.*

Certainly, one can achieve good correlations between LY and EC. But omitting epsilon-values would exclude any information about the contribution of LE to the energy closure. This is, however, the main objective of our paper.

13) Table 2 – 5: *Only very little comments are given by the authors to these eight tables. Some more wording to guide the reader towards important points to learn from each*

table would be very helpful.

We add to L173 (Tables 2): They highlight the substantial difference between the humid and dry stations in terms of the mean magnitude of evapotranspiration. Under moist conditions, in contrast, the dry station Majadas (M4) ranges around the same magnitude as the humid stations. We re-formulate L199 (Tables 3): Tables 3a and 3b show the absolute differences and their standard deviation between the EC-data presented in Tables 2a and 2b and LY-measurements. They indicate how the differences between EC- and LY-measurements mostly (except for F1) get smaller from observed (*DLoL*) to adjusted values of LY (*DLaL*).

For Tables 4, L228 is moved to L219 (beginning of the corresponding section).

In general, the findings from the tables are presented in the discussion.

14) Figure 1a and 2a: *The larger differences in the morning in fig 1 have disappeared in figure 2a. This must be related to the diurnal characteristics of epsilon. Addition of epsilon in fig 1 and discussion would be helpful.*

The change from the original data (oLE) to corrected (cLE) and adjusted (aLE) data involves the

interactions of σ_L , LY, epsilon, d and the regression of binned data. Especially the latter smoothes down the morning data.

15) S3.7 In section 2.5 the authors describe a binning procedure of the LE data for regressing and obtain wL . How does this relate to figure 5 where averages of wL per hour are given. Some extra wording would be helpful.

In line 266 we re-formulated the first sentence to: Figures 5a and 5b show the mean course of wL during daytime-hours using the average of all wL values at a specific hour.

16) L267: Bins ranging from 6 to 14. Is this the number of observations in each bin. Please be precise.

This is the numbers of bins. The bins were selected such that the number of data within all bins remained more or less the same. Depending on the sample size we had 90 to ca. 120 observations within a bin. This is described in section 2.5 (see pp. 148-150). For clarity, we re-formulated L267 to: The number of bins used in Fig. 5a per station varies from 6 (F1), 8 (F2, G2, RHB) to 14 (G1). The number of bins used for Majadas in Fig. 5b varies from 5 to 12, depending on the used period.

17) L273: Standard deviation in wL will among others depend on the statistical noise in the EC measurements. These can be large under convective low wind conditions during day time, and lower under the less convective conditions around sun rise and sunset.

We are grateful for this hint. However, it's rather during low turbulence conditions (evening through morning), that the signal to noise ratio of the EC system is small and EC measurements inaccurate, since the fluxes are very small. Eddy correlation in low turbulence conditions can then result in flux overestimation. We rather believe that the variability is lower in the morning/evening due to the mentioned smallness of the absolute values during stable/less convective conditions. We add in L275: ..., which relates to the fact that the absolute differences between LY and EC observations are comparably small during stable to weakly unstable conditions in the morning and evening. Additionally, in the discussion L343 is re-formulated accordingly, basically dropping the surprising character and deleting the vague statement about weather conditions (see also comments 22 and 23).

18) Figure 6b: what is the meaning of s_{11} in the labels?

s_{11} stands for data smoothed with a moving median filter with a window length of 11 timesteps. We add this information in the legend: A moving median filter (s_{11}) with a window length of 11 hours was used for smoothing.

19) Figure 6b: there is a remarkable drop in wL observed in the figure, but not mentioned in the text.

We add a comment on this at L347: The correlation of wL to the magnitude of evaporation is also indicated in Fig. 6b, where a drop in wL follows cLE_{EC} . This correlation is indicated in Fig. 6b, where a drop in wL follows cLE_{EC} .

20) S3.9: *Please explain what the value of these correlations are. One question that comes to my mind is: the authors use the result of the regression (wL and cLE) and look at their correlation. What can we learn from this?*

The authors rather use the regression of wL and LE and then derive cLE. The tables reveal, that wL and the magnitude of evaporation are positively correlated. But indeed, these tables present auto-correlated values, and their contribution to the study is small. We thus drop the whole section 3.9 and mention the correlation wL to magnitude evaporation in the discussion as the statement from comment 19).

21) L315-319: *What conclusions can be drawn from the summary of these results on standard deviations?*

We add the following statement on L319: The difference between $SD(LE_{LY})$ and $SD(oLE_{EC})$ is getting bigger since $SD(oLE_{EC})$ gets smaller after correction, whereas $SD(LE_{LY})$ remains the same.

22) L345: *See my comment #17*

see our comment under 17

23) L347: *“one might conclude that the high standard variations are rather related to weather conditions”. Where is this conclusion based on?*

we will drop this statement in view of our new statement on SD under comment 17.

24) S5, L352: *I would say that the best adjustment of EC to LY would be a direct regression of without the complications of epsilon and the full energy balance. And if this is the aim, why not use LY and refrain from EC?*

This is true, but without including epsilon one cannot get the contribution of EC to the energy gap, which is a strong part of the current study. And as the energy gap problem is characteristic to involving EC measurements, using lysimeters for evaporation instead of EC would obviously solve a part of the problem. We thus add in L364: If a high-precision lysimeter capable of resolving evapotranspiration as well as condensation is available complementary to an EC set-up, LE can directly be obtained from the lysimeter.

25) L358: *Note that also the statistics of EC observations will be come progressively worse when going to smaller time intervals. But combining scintillometry and EC- observations might be a way forward.*

The actual used time interval of EC measurements would not change, only the averaging window on the rawdata could be shifted, as mentioned in L360, so that an EC average would better correspond to a e.g. 5 min lysimeter value. The sentence is unclear and is re-formulated to: In a first step we recommend to perform the comparison of LY and EC based on 5 to 10 minutes lysimeter intervals, and center the one/half-hourly averaging window on EC raw data accordingly.

Textual comments:

26) L14: *“At the overall average” -> “Overall”*

Good suggestion, we use only 'overall'.

27) L15: *“which were partially closed with” -> “after applying “*

Good suggestion, as the term closed in regard of the measurements is misleading. Changed accordingly.

28) L16: *“remain high differences” -> “remain large differences”*

Will be changed.

29) L18: *“correction evaporation weights”. This looks like a defining term, but is never used in the main text, please be concise on terminology.*

True, we drop the 'correction'.

30) L19: *“correcting evaporation weights”. Yet another formulation never used in the main text.*

True, we drop the 'correcting'.

31) L29: *How is the energy balance gap defined? I would expect a value of 22-27 % for the magnitude of the gap.*

Good catch. This is true and we are aware of this. The numbers used refer to the closure, the gap thus is 22.0 to 27, as the R2 correctly suggests.

32) L30: *A comparison alone cannot lead to any reduced difference. I guess it is the adjustment of EC measurements with LY measurements that leads to this reduced difference.*

This has also been criticized by R1 and has been changed accordingly: After forced closure of the energy balance, the difference between daytime LY and EC data on two fields could be reduced from -28.8% to 6.2%, respectively from -26% to -12.3%, with an accuracy....

33) L31: *How do this percentages relate to the values of 73.2 and 78% on line L29.*

This is now answered with the re-formulation in comment 32).

34) L35: *“with” -> “of”*

Changed.

35) L36: *“an influence of the increasing plant height as against constant measurement height is suspected.” Unprecise wording, please correct.*

We re-formulate to: They reported substantially larger LY evapotranspiration rates compared to the EC measurements due to differences in plant growth in the LY and the EC footprint.

36) L38: *is -17% to -19% on a daily basis? Please be precise in formulation*

In L34 we add that Evett used daytime values. Additionally, we re-formulate the sentence to: ...mean differences from -17 to -19 % were found between the two measurements methods after...

37) L65: *Textually it would be nicer to start with some of the general information given below the table 1, and then introduce table 1.*

Good suggestion: We move the first paragraph (In addition to...) to the end of the section, and start the section with the second paragraph (Data were obtained from the following Institutions), followed by: Table 1 gives an overview on the location of the sites and time periods used.

38) *Table 2b: Some numbers are out of place in the last column, it seems.*

Those are line numbers being somehow jumbled into the table.

39) *L281: 212 weeks?*

Is changed to days.

Additional major changes to the original manuscript (not mentioned in response to reviewers)

Partial energy balance closure of eddy covariance evaporation measurements using concurrent lysimeter observations over grassland Peter Widmoser & Dominik Michel

- 1) Acknowledgements and description of data repositories adapted (according to our AC1 in public discussion)
- 2) Introduction newly written (as suggested by reviewer 2 and after discussion with the editor) to highlight the goal of the manuscript and make it understandable without knowledge of connected earlier publications
- 3) Material and Methods (Section 2) substantially changed according to suggestions by Reviewer 2 and the editor, such as detailed description of stations (including map) and instrumentation, providing units. Corrected instrumental description (0.01 mm accuracy instead of 0.02, resp. 7 W/m² instead of 15 W/m²) and thus mentions thereof throughout the manuscript.
- 4) All figures: re-arranged y-axis label and legend (changed legend items if necessary), color scheme made consistent
- 5) former Fig. 6b (new 8b): Wrong data version (different processing) were shown in original manuscript, is now corrected
- 6) line 323/324 in original manuscript: the word 'however' was removed, as it would convey a different (wrong) meaning of the sentence
- 7) At the end of Section 5 we added: "We recommend to test also high-quality flag 0 datasets (Mauder et al, 2013) for plausibility by the out-of-bound method, which may be derived from Wohlfahrt and Widmoser, 2013."
- 8) In Section 2.4 we added a comment on how the Majadas time series of wLE was extended (last paragraph, "In order to extend..."). Figure 8b was changed accordingly: We added the wLE-data based on an extended diurnal time period (wLE_long) to the wLE based on a shorter daytime period shown in the original manuscript (wLE_short).
- 9) At the very end of Section 5 (Summary and Conclusions) we added: "The method proposed here may also be applied if reliable sap flow measurements are available instead of lysimeter observations. We guess that an adoption of our method may apply to partial energy closure by heat fluxes if surface temperatures estimates are known from telemetry."
- 10) Using consistently 'evaporation' (as in the title) instead of 'evapotranspiration' throughout the manuscript
- 11) All hyphens between LY or EC and a noun (e.g. LY-measurements) have been deleted to make the manuscript cleaner and enhance readability.