Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-257-AC1, 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access

HESSD

Interactive comment

Interactive comment on "Estimation of surface energy fluxes in the Arctic tundra using the remote sensing thermal-based Two-Source Energy Balance model" *by* Jordi Cristóbal et al.

Jordi Cristóbal et al.

j.cristobal@alaska.edu

Received and published: 6 October 2016

AC: We would like to thank the reviewer for their insightful comments and suggestions, which we believe have significantly improved the quality of this manuscript as well as our research.

RC_1: p9, l24–27 The method to estimate fG is not clear to me. How do you estimate the fraction of absorbed PAR by the green vegetation? Is it equal to PAR incoming - PARreflected in your model? This would also include PAR absorption by bare soil,dead plant material, mosses and other elements. Guzinski et al. (2013) actually suggests to use a different method, based on NDVI and EVI (as you mention on page 14). Do you have another reference that actually recommends the PAR ratio method?

Printer-friendly version

Discussion paper

AC_1: We would also like to note to the reviewer that we have accepted his/her suggestion to use the Guzinsky et al. (2013) method based on the EVI and NDVI to estimate the fG. According to Fisher et al. 2008, fG is defined by FAPAR / FIPAR where FAPAR is the fraction of PAR absorbed by green vegetation cover and FIPAR the fraction of PAR intercepted by total vegetation cover. Due to a lack of FAPAR observations, we estimated fG using only FIPAR as suggested by Anser (1998) and as the results show this might have caused an overestimation of fG at the beginning and at the end of the growing season contributing to model-measurement disagreement. Although this was not a major point according to the reviewer, we have re-run the model using Guzinski et al. (2013) approach to estimate fG. The new results yield a better model agreement, although it does not provide reliable fG values at the end of the season (mainly in September).

RC_2: 1. The authors stress the point that a remote-sensing based model can be applied at the larger scale (Title, Abstact p. 1, l. 17, 24, 26; Motivation p. 3, l. 1–12; Conclusions p. 15, l. 8–10). However, it seems that (except for the LAI, which is a minor point of the study) this was not done (p. 11, l. 23–24). This is a little bit disappointing after reading pages 1–3. Therefore I would suggest to force the model with satellite data only and compare the results. If this is beyond the scope of the paper, the authors should adjust the motivation statements.

This paper is focused on the local application with the tower micrometeorological and flux measurements representing local conditions in order to more reliably evaluate and refine the TSEB model for regional application to the Arctic tundra. As we stated in the conclusion section we will extend this research to regional scales using a TSEB-based model refined to be robust for the Arctic tundra using satellite inputs. To better clarify this objective, we have added text to the introduction motivating the need for localized testing in preparation for improvement of a regional satellite based energy balance model.

RC_2: 2. Section 2 is quite long given that the model description is published already.

CC BY

P4 l20 – p5 l12 could be omitted or moved to an appendix as the resistance terms and the sensible heat flux parameterisation are not discussed further in the manuscript. In this case, you could mention after Equation 11 that Hs is calculated as a function of the difference between canopy air temperature and soil temperature and of the soil resistance.

AC_2: Given questions raised by the second reviewer about the model formulation, we decided to retain the discussion of the TSEB formulations needed to understand the resulting refinements required to obtain good results (see discussions in sections 3 and 6).

RC_3: 3.1. You show two different approaches for estimating cG (Section 3.2). In both approaches you fit some parameters. However, if I understand it correctly, you use different data for fitting. On what data did you fit the parameters of the first method (p7, l23–24)? Why did you not use the same approach as for the second method, where you split the data set into a calibration and a validation subset? Are the data of all stations combined in a single data set? Do you take an equal amount of data points per station? Are the parameters fitted separately for month? Please describe the fitting approach in more detail in Section 3.2.

AC_3: We have improved section 3.2, 4.2 and Tables 3 and 4 to clarify these points. The Kustas et al. (1998) and Santanello and Friedl (2003) methods were evaluated against the same dataset used to evaluate all fluxes that had restrictions for balance closure, among others (see section 4.2). To fit and test the new cTG approach, data from the previous dataset with no restriction of balance closure and from 4 to 21 hours local solar time was used. Coefficients A, B and S were derived using 60% of all available data aggregated in 30 min timesteps for the whole summer period and the remaining 40% of the data were reserved for model testing. Table 4 shows the amount of data points per station (n) to derive model coefficients and test the cTG approach. To calibrate the cTG, for the Tussock and Heath flux towers n is similar (around 10 000) while for the Fen tower, less data were available (n ∼8 000).

CC BY

RC_4: 3.2. Would it be possible to use a proxy such as soil moisture to improve the fit?

AC_4: Soil heat flux plate measurements were corrected to account for soil heat storage using soil moisture from the water content reflectometers. This has been added to the text.

RC_5: 3.3. Although you mention that soil type and properties are important, none of your methods takes it into account. AC_5: We agree with the reviewer that soil type and properties are important to model G. In the original TSEB formulation, a simple approach based on the relationship between G and RNS was used (Eq. 13). This approach has less complexity and requires no soil texture and moisture information, which, unfortunately, is not routinely available over large areas. For continental-to-global applications of the TSEB, we are indeed finding that variations in the main parameters of the G formulation are required – for example over rock or desert sands. However, the modifications derived here help to better capture thermal characteristics of the tundra substrate.

RC_6: 4. In Section 3.3 you describe that you use two different Priestley-Taylor coefficients. Did you consider varying them with soil moisture or LAI? Are they valid for the whole Arctic, or only locally?

AC_6: The initial values of the Priestley-Taylor coefficients (PTC) we used in this paper were the originally proposed value of 1.26 for application of TSEB and a value of 0.92 averaged from the references found in the literature focused on Arctic tundra. As a starting point for the model we consider this range in PTC applicable for Arctic vegetation.

AC_6: In addition, to clarify how TSEB can adjust PTC for moisture conditions, the following paragraph has been added in section 2: “Under stress conditions, TSEB iteratively reduces ïĄąPTC from its initial value. The TSEB model requires both a solution to the radiative temperature partitioning (Eq. 2) and the energy balance (Eqs.

CC BY

6 and 7), with physically plausible model solutions for soil and vegetation temperatures and fluxes. Non-physical solutions, such as daytime condensation at the soil surface (i.e., LES < 0), can be obtained under conditions of moisture deficiency. This happens because LEC is overestimated in these cases by the Priestley–Taylor parameterization, which describes potential transpiration. The higher LEC leads to a cooler TC and TS must be accordingly larger to satisfy Eq. (7). This drives HS high, and the residual LES from Eq. (11) goes negative. If this condition is encountered by the TSEB scheme, ïĄąPTC is iteratively reduced until LES $\sim$ 0 (expected for a dry soil surface). However there are instances where the vegetation is not transpiring at the potential rate but is not stressed due to its adaption to water and climate conditions (Agam et al., 2010) or the fact that not all the vegetation is green or actively transpiring (Guzinski et al., 2013)."

RC_7: 5. Figure 2 does not demonstrate a relationship between TRAD and G, it merely shows that both variables exhibit a diel cycle (p11, l11–12 & p15, l1–2). Can you please provide more details on the expected relationship? I find that this is an important point as one of your main conclusions is that the approach using TRAD is better than using RN. If I understand your reasoning correctly, you assume that the relationship between TRAD and G holds for different vegetation types, times of the growing season and weather conditions. This point needs to be discussed in more detail. For example, a recent study by Juszak et al. (2016) showed that two different vegetation types with close to identical top soil temperatures differed in G by a factor of 2. It would be great if you showed evidence for this relationship under different conditions. I would at least expect to see scatterplots of TRAD and G as compared to RN and G and correlation coefficients. Of course you can use shifted time series to account for the time lack. RC_7: Figure 2 The temperature is not in Kelvin. I do not think it makes sense to take the mean of all available data as the station with most data will contribute more and biases can occur, for example if the coldest station on average starts measuring later during the year. I would prefer one plot per station, or a completely different graph (as explained above).

CC BY

AC_7: The axis title has been corrected. The relationship between TRAD and G and the definition of the new coefficient cGT has been explained in section 3.2 in which G is computed using Eq. 18. This method uses a phase shift proposed by Santanello and Friedl (2003) and is supported by the measurements illustrated in Figure 2. Figure 2 was only meant to show this phase shift and text in p11, l11–12 has been changed accordingly. Figure 3 and Table 4 show the behaviour of the new coefficient cGT derived from the TRAD-G relationship on a per station basis.

RC_7: 6. The results and discussion in Section 6 are for all stations combined. However, it would be interesting to read about the different (or similar) accuracies at the different vegetation types. This is particularly relevant if you want to conclude on vegetation dynamics and vegetation change (p14, l20–22). Figures 4 and 5 also reveal differences between the stations. For example LE is strongly overestimated at the tussock site. Why?

AC_7: Unfortunately, without detailed ground measurements to verify the assumed TSEB vegetation inputs (such as LAI), it is hard to identify any single factor that may have been a major cause for model-measurement disagreement, but overall the TSEB performance is considered satisfactory for all sites evaluated in this paper.

RC_8: 7. Why do you discuss the accuracy of RN (p12, l8–16; p14, l6–11, l24, l30–31, most figures and tables) and not of the incoming longwave radiation alone? If you use the shortwave radiation budget and outgoing longwave radiation from measurements and just compute the incoming longwave radiation in your model, it would be surprising if you found a substantial difference in RN. Did you use any of the remote sensing products (p12, l15–16) to justify your conclusion that 'this methodology scheme can be used to obtain reliable estimates of RN'?

AC_8: Downwelling longwave radiation results were discussed at the beginning of section 6.2 before discussing RN results. AC_8: We have not used remote sensing products to justify this conclusion. This sentence has been rewritten accordingly.

RC_9: 8. All Figure legends, scale bars and axis labels are far too small. Please increase the font size to about the same as the figure caption. Please also avoid to rotate the figures (in figures 6,8,9) and the axis labels.

AC_9: Figure legends, scale bars and axis labels have been increased. AC_9: Figures 6, 8 and 9 have been re-rotated.

RC_10: p1, l19 What is unique about tundra conditions?

AC_10: "Unique" was misplaced. It was supposed to be written before "parameterizations". In any case it has been removed from the text to avoid leading to misinterpretations.

RC_11: p1, l24–25, Section 2 How did you test the usefulness of the MODIS LAI? Maybe it would be helpful to compare the results of the three towers concerning the different LAI. Also, did you test if the model is sensitive to LAI variations? Which fluxes are influenced by LAI in the model?

AC_11: Unfortunately, we do not have LAI field measurement, thus, MODIS LAI usefulness was tested indirectly by means of the evaluation of the surface energy fluxes. The model is sensitive to LAI, since the radiation and temperature partitioning are affected by the LAI/fractional cover as well as the wind speed at the soil surface and LEc via the PT parameterization for the Rnc (Timmermans et al., 2007).

RC_12: p9, l21–22 Other comprehensive LAI data from close-by can be used as reference, e.g. Shaver and Chapin (1991); Shippert et al. (1995); Williams et al. (2001); Walker et al. (2003); Williams et al. (2006); Shaver et al. (2007); Sweet et al. (2015). In particular the study of Williams et al. (2006) has many details on different types. I am sure there are even more studies which measured LAI as the Imnavait Watershed and Toolik lake are very well studied. RC_12: p13, l30 An LAI of 1.7 seems to be quite high for the Imnavait Watershed. Did you compare with other data such as (Shaver and Chapin, 1991; Shippert et al., 1995; Williams et al., 2001; Walker et al., 2003; Williams

CC BY

et al., 2006; Shaver et al., 2007; Sweet et al., 2015)? Which vegetation type had this extreme value?

AC_12: References about reported LAI values in these previous works and alternative methods to estimate LAI in the Arctic tundra have been added.

RC_13: p1, l29 Omitting 'Near-surface or shelter level' would make the starting sentence more catchy.

AC_13: This has been deleted from the text.

RC_14: p2, l2–4 Less references would be enough.

AC_14: We have kept more recent and relevant publications.

RC_15: p2, l15 Do you really mean 'inconsistent', or rather 'sparse'?

AC_15: We meant spatially and temporally inconsistent, and we did not imply that the data is wrong in any sense. We have rewritten the sentence to avoid misinterpretations.

RC_16: p2, l18 What is an 'increase in peak vegetation'? Do you mean vegetation growth / activity / LAI?

AC_16: According to Jia et al., 2003 it is in the "peak vegetation greenness". The reference was misplaced. This has been changed in the text.

RC_17: p2, l19 Do fires contribute to the greening? Maybe it would make sense to exchange the first two sentences of this paragraph.

AC_17: Sentences have been exchanged in the text.

RC_18: p2, l24–25 As shown in the recent paper by Williamson et al. (2016), the albedo effects of shrubs may not be as clear. Also, wet surfaces and sparsely vegetated water may have an even lower albedo than shrubs (Gamon et al., 2012).

AC_18: We agree with the reviewer that wet surfaces and sparsely vegetated water may have an even lower albedo than shrubs. A reference to the Williamson et al.

CC BY

paper has been added to the paper.

RC_19: p4, l7 Does this mean that the model uses a spherical leaf angle distribution for all vegetation? How do the results change, if an erectophile distribution is used for the graminoid vegetation (fen, tussock tundra)?

AC_19: The assumed leaf angle distribution will affect the radiation divergence through the canopy layer and hence affect the net radiation partitioning between the canopy overstory and the soil/substrate. Without measurements to determine the leaf angle distribution, the default of a spherical leaf angle distribution is a reasonable one, particularly for heterogeneous surfaces having a mixture of vegetation species.

RC_20: eq. 1, 4–12 It is a bit confusing that R can be radiation or resistance, depending on the subscript. Maybe you could use 'r' for the resistance values?

AC_20: "r" has been adopted for resistance and changed in the text.

RC_21: p5, l25 The abbreviation TIR is not explained. Additionally, this paragraph suggests that the satellite data is used for the study. If this is not the case, delete the clause 'when daytime TIR satellite imagery is typically acquired'.

AC_21: we have expanded the TIR abbreviation. This paragraph refers to the original method development in which cG=0.3 was set. However, in order to avoid misinterpretations we have deleted "when daytime thermal satellite imagery is typically acquired" from the paragraph.

RC_22: Section 3.1 Why do you continue using the Brutsaert (1975) formula? Two comparison studies on empirical parametrisations of incoming longwave radiation found that other formulars described the data better, namely the Dilley and O'Brien (1998) clear sky formula and the Unsworth and Monteith (1975) cloud correction (Flerchinger et al., 2009; Juszak and Pellicciotti, 2013).

AC_22: Although there are other sky emissivity parameterizations which might give slightly better estimates of incoming longwave, the error in using Brustaert formulation

CC BY

in TSEB is minor compared to the errors in turbulent flux estimation. In fact from Table 5 in Flerchinger et al (2009) the RMSD from all sites measuring incoming longwave using Brutsaert (1975) is 27.2 W/m2 while for Dilley and O'Brien (1998) it is 23.3 W/m2. Regarding cloud correction, the Crawford and Duchon method is easier to apply since we do not have the data required for Unsworth and Monteith (1975) method.

RC_23: p7, l5 & p.7, l 25–29 Actually, in Eq. 12, not RN is used but RNS. Please make more clear which variable you use. And if you adjusted the model in case you use RN.

AC_23: This has been corrected in the text.

RC_24: p7, l8–14 Exchange this paragraph with the first paragraph.

AC_24: This has been exchanged in the text.

RC_25: p7, l15–17 Split the sentence in two parts as the 'while' does not follow easily on the first part of the sentence.

AC_25: This has been corrected in the text.

RC_26: p7, l23–24 Why does this sentence not appear in the results section?

AC_26: These are the values from the original model. We have rewritten the sentence to clarify the text.

RC_27: p8, l13 Remove '1.2.1'.

AC_27: This has been removed from the text.

RC_28: p9, l8 Are you sure you have several Dryas species (as indicated by spp)? Also, Dryas is a dwarf shrub species, so it would be more accurate to write '..., other dwarf shrubs, and lichen'.

AC_28: This has been modified in the text accordingly.

RC_29: p9, l11 What do you mean by 'vegetation-based measurements'? Maybe replace the term with 'canopy structure' or 'vegetation properties'.

CC BY

AC_29: We have changed the section title using "vegetation properties".

RC_30: p9, l29 Can you explain your choice of 1 for the clumping factor in more detail? What is a 'variable organic layer'?

AC_30: It is not an organic layer, it is a moss layer, and this has been changed in the text. As text says clumping factor was set to 1 based on the knowledge that Arctic tundra has a variable moss layer with little bare ground, thus, almost covering almost 100% of the ground. We used this approach for modelling purposes as we do not have actual data on the ground. However, a value of 1 seems a realistic approach for the study area.

RC_31: p9, l30 Vegetation height and the clumping factor are not variable. Can you estimate the uncertainty you introduce with this assumption?

AC_31: Over the growing season ground measurements indicated little change in vegetation height and density. Prior sensitivity studies (e.g., Zhan et al., 1996) indicate TSEB shows relatively small sensitivity to canopy height and fractional cover, which is related to the vegetation clumping factor.

RC_32: p10, l1–2 The sentence about future work should be moved to the discussion or conclusions.

AC_32: This sentence has been moved to the conclusions section.

RC_33: p10, l12–13 Why do you restrict the modelling to daytime conditions? It would be interesting to also test if the model is able to reproduce values at night. I am aware, that the incoming longwave radiation depends on cloud cover. However, you could interpolate the cloud cover during the night. How did you assess the presence of precipitation?

AC_33: Our testing is focused on daytime conditions for two reasons: First, EC flux observations used for validation are less reliable during night-time due to stable conditions and low wind speeds. Second, for transition to satellite applications, we are

CC BY

primarily interested in evaluating model performance during daytime satellite overpass times. Other techniques are typically used to upscale from the overpass time to daily total fluxes.

RC_34: Section 5 Using five different error estimates does not add additional information as compared to using only three. In your results, you rarely mention MAD and the information of MAPD and RMSE is largely the same. It is not very intuitive that in your notation the mean of ei is _X . You could use ei and _E or xi and _X (and the corresponding notation for oi and _ Y) instead.

AC_34: We have used five different error estimates to make the results section more comparable to other papers. Although we agree with the reviewer, in the literature you may find some studies in which MAE or MAD are only stated.

AC_34: ei and oi notations have been changed in the text.

RC_35: p11, l21 & Table 4 What is this flux subset? Please describe the choice of the subset in the methods.

AC_35: This was clarified in section 4.2 “Model inputs and evaluation dataset” and Tables 3 and 4.

RC_36: p11, l23–24 The first clause of the long sentence is out of place, it is an outlook and would fit better at the end of the conclusions.

AC_36: Sentence has been move to the conclusions section.

RC_37: p12, l2 To which method do the R2 and the RMSE value belong?

AC_37: Both methods yielded similar results. R2 was the same and RMSE for Brustaert (1975) and Jin et al. (2006) was 26 Wm-2 and 27 Wm-2, respectively. This has been clarified in the text.

RC_38: p12, l1–7 You found that the new method was not better than the original Brutsaert (1975) formula. However, this does not necessarily imply that the Brutsaert

CC BY

(1975) method is good. I would like to see a discussion of limitations and other potential approaches.

AC_38: Differences between methods for estimating clear sky incoming longwave radiation continue to be evaluated over different climate zones (e.g., Choi et al., 2008) and indicate that discrepancies tend to be relatively small compared to uncertainty in modelling the turbulent fluxes. Therefore, a detailed discussion is not warranted for this analysis (see also response above). RC_39: p12, l18 What is the 'evaluation subset'?

AC_39: This has been clarified in section 4.2 "Model inputs and evaluation datasets".

RC_40: p12, l30–32 The BR and RES methods need to be explained in the methods section. How does this description relate to the Priestley–Taylor approach you explain in the methods? Do the two methods refer to the canopy or the soil LE (eq. 10, 11)?

AC_40: BR (Bowen Ratio) and RES (Residual) methods have been referenced in the previous paragraph and they are intended to address the lack of closure of the flux station data used to evaluate the TSEB method. We compare TSEB to closed fluxes since the model requires energy balance closure while the measurements of H and LE using eddy covariance technique undermeasure these fluxes by 10-20% based on comparison with available energy (Rn-G). We used two methods: 1-a distribution of residual according to Bowen Ratio, with the acronym BR (Twine et al. 2000 and Foken 2008); 2- and LE was recalculated as the residual, with the acronym RES (Li et al., 2008). In order to clarify the text for these methods, we have introduces these acronyms in the previous sentence. These methods are well explained in these papers and, for the sake of brevity, we prefer to refer the reader to the original references.

RC_41: p13, l26 Is the fraction of vegetation cover not estimated from the PAR budget? Please explain this in the methods! How sensitive is the model to LAI?

AC_41: The fraction of vegetation cover (Eq. 3) is computed using LAI and not PAR. We have clarified this in the text.

CC BY

RC_42: p13, l30 Is fG a sensitive parameter?

AC_42: The value of fG modifies the estimated canopy transpiration (LEC) via the Priestley-Taylor parameterization (Eq. 10). It reduces LEC in direct proportion to its magnitude and has been used to adjust LEC based on crop phenology in other studies (e.g., Guzinski et al., 2015).

RC_43: p14, l26 As the interannual variability is not mentioned in the results, it should not be mentioned here.

AC_43: We have replaced "interannual" by "seasonal"

RC_44: p15, l3 'other models' is unclear. Do you mean 'G computation from RN'?

AC_44: Yes, this has been clarified in the text.

RC_45: p15, l3 As some readers start with reading the conclusions, it would be good to repeat that _PTC is used to estimate ET.

AC_45: This has been added in the text.

RC_46: p15, l6 Was the model sensitive to LAI? I would be surprised, as LAI (in the model) does not influence ET, albedo, or any of the other major fluxes. Otherwise this conclusion is not valid.

AC_46: LAI is used by TSEB (Eq. 3) to partition TRAD into soil and canopy temperature components, thus, it influences surface energy flux partitioning between the canopy and soil/substrate. The value of LAI also influences the radiation divergence and wind profile through the canopy layer and ultimately the soil and canopy aerodynamic resistances (Kustas and Norman, 1999;2000).

RC_47: p15, l8–10 On which result do you base this conclusion?

AC_47: We base this conclusion on the fact that the remote sensing-based TSEB model is able to capture the vegetation seasonal dynamics and contains the main

Printer-friendly version

Discussion paper

factors (LST, LAI, vegetation height/roughness) affecting H and LE partitioning. Thus with a multi-year time series of remote sensing observations from satellites are able to detect changes in vegetation cover conditions (LAI, canopy height and roughness) which in turn can affect LST and hence energy flux partitioning. This permits monitoring the impact of vegetation cover changes on the water and energy cycle at synoptic scales with satellite data.

RC_48: p15, l11–14 This seems very abstract. Maybe you could rather conclude on how to integrate more satellite data to apply the model to the regional scale.

AC_48: Methods described in this sentence are designed to estimate surface energy fluxes with satellite data. We have clarified this in the text.

RC_49: Figure 3 This graph is very important. However, it would be great if you could add uncertainties, or at least standard deviations.

AC_49: Standard deviations for the mean values have been added to this figure.

RC_50: Figures 4–6 In the caption, PTC should be a subscript. This way of plotting does not allow an evaluation of G, one of your main focusses. Also, it is impossible to tell the accuracy of LE. I suggest to use just one variable per panel and indicate the point density with colour (heat map). As this will result in four times more panels, I suggest to remove Figure 5 as the additional information is small.

AC_50: PTC has been subscripted. AC_50: Difference statistics between modelled and measured energy balance components are provided in tables 3 through 6. Having separate graphs comparing LE, H, RN and G would make it more difficult for the reader to have a sense of the relative magnitudes and scatter between the measured and modelled energy balance components. Showing the results in this manner gives the reader a better sense of the relative modelled-measured differences and which fluxes is the scatter the largest and most significant in the four components.

RC_51: Figure 7 The figure caption should be self explanatory. Please define fG.

CC BY

AC_51: This has been added in the caption.

RC_52: Figure 8 I would prefer to see a sample time series to 5-day averages of multiple stations.

AC_52: 5-day averaged fluxes displayed in the figures more readily indicates the seasonal behaviour of TSEB over the whole study period. A sample time series is too noisy and does not allow the seasonal dynamics of surface energy fluxes and energy partitioning to be easily determined or illustrated.

RC_53: Figure 9 Change the symbols to make the figure easier to read. With the tiny legend and the turned figure it is impossible. I would suggest to have the same symbol for the same variable, once filled (for observed) and once empty (for modelled).

AC_53: The figure has been turned, the legend has been increased in size and the symbols have been refilled.

RC_54: Table 1 Space missing between Longwave and incoming; the captions says 'Average and standard deviation for the input values were computed for each period and for each site.' However, there is just one value per site given. Which period is it for?

AC_54: This has been corrected in the text. AC_54: Average and standard deviations reported in this table were computed using all selected data from the full period of model evaluation (Period row) for each flux station. This has been clarified in the caption.

RC_55: Table 3 MAPD not MADP

AC_55: This has been corrected in the text.

RC_56: Table 5–6 One H misses the subscript.

AC_56: Sensible heat (H) is not missing the subscript. When using the residual method observed (from the flux tower) H is evaluated against modelled H.

CC BY

References Asner, G. P.: Biophysical and Biochemical Sources of Variability in Canopy Reflectance, Remote Sens Environ, 64, 234–253, 1998.

Choi, M. H., Jacobs, J. M., and Kustas, W. P.: Assessment of clear and cloudy sky parameterizations for daily downwelling longwave radiation over different land surfaces in Florida, USA, Geophys Res Lett, 35, Artn L20402 10.1029/2008gl035731, 2008.

Fisher, J. B., Tu, K. P., and Baldocchi, D. D.: Global estimates of the land-atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites, Remote Sens Environ, 112, 901-919, 10.1016/j.rse.2007.06.025, 2008. Guzinski, R., Nieto, H., Stisen, S., and Fensholt, R.: Inter-comparison of energy balance and hydrological models for land surface energy flux estimation over a whole river catchment, Hydrol Earth Syst Sc, 19, 2017-2036, 10.5194/hess-19-2017-2015, 2015.

Kustas, W. P. and Norman, J. M. Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. Agricultural and Forest Meteorology. 94:13-29. 1999.

Kustas, W. P. and Norman, J. M. A two-source energy balance approach using directional radiometric temperature observations for sparse canopy covered surfaces. Agronomy Journal. 92:847-854. 2000.

Timmermans, W. J., Kustas, W. P., Anderson, M. C., and French, A. N.: An intercomparison of the Surface Energy Balance Algorithm for Land (SEBAL) and the Two-Source Energy Balance (TSEB) modeling schemes, Remote Sens Environ, 108, 369-384, 10.1016/j.rse.2006.11.028, 2007.

Zhan, X., Kustas, W. P., and Humes, K. S.: An intercomparison study on models of sensible heat flux over partial canopy surfaces with remotely sensed surface temperature, Remote Sens Environ, 58, 242-256, Doi 10.1016/S0034-4257(96)00049-1, 1996.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-257, 2016.