

## ***Interactive comment on “Three perceptions of the evapotranspiration landscape: comparing spatial patterns from a distributed hydrological model, remotely sensed surface temperatures, and sub-basin water balances” by T. Conradt et al.***

**Anonymous Referee #1**

Received and published: 19 March 2013

### General Comments

This paper deals with the impact of different strategies for deriving evapotranspiration spatial distribution, in terms of the impact on the estimation of discharge in interior points. The available dataset on the investigated area, Elbe basin, seems to be detailed enough to endorse the study. The scientific approach adopted by the researchers is rigorous, but a general excessive verbosity of the presentation often make the analysis difficult to be followed. While the authors seem to have a good knowledge of the SWIM model, I have several strong concerns regarding the methodology adopted for

C402

the processing of remote sensing data. The adopted approach can not be clearly contextualized in the wide range of approaches already available in the literature, and it seems to introduce numerous approximations (not commonly made by other methods) that have to be justified (see details in Specific Comments). Additionally, since the authors' declared aim is to globally calibrate the remote sensing maps at year-scale, they are forced to introduce further assumptions to upscale the temporally sparse ET estimates to the year. This topic is still an open issue in the remote sensing scientific community, and the authors do not provide any evidence that the proposed approach does not introduce further biases not related to the thermal data itself. In summary, my opinion is that the reliability of the reported results are strongly affected by the consistency of the adopted approach, hence I suggest a major revision of the paper, with particular regards to the reconsideration of the assumptions made for the remote sensing-based model. I really believe that an improvement in the presentation of the results (which should also include some numerical indices rather than only qualitative comparisons) and a more careful processing of the remote sensing data would provide a significant contribute to the scientific community.

### Specific Comments

The authors refer to the same method with different “names” (e.g., SWIM, model, simulation all refer to the same approach). I understand the need of avoiding repetition, but in some circumstances it would be clearer for the readers if the authors label each method with one name, using it consistently through the text.

P1135-L10. It is not clear to me if the rain gauge stations are 501 or 853; please clarify this point. Moreover, since the authors say that heterogeneity in network density is one of the possible justifications for the discrepancies, it would be useful to see the spatial location of both fully instrumented and rain gauges, perhaps in Fig. 2.

P1135-Eq.1. To the best of my knowledge, Turc equation is based on solar radiation instead of net radiation. Additionally, I can not recall the dimensionless factor in the

C403

original formulation. A better reference for this approach should be provided, as well as a justification on the use of this alternative approach.

P1136-L1. The reported crop coefficients (land use specific factors) are quite high (> 0.9). Are these including also sparse vegetated areas?

P1136. The connection between EP and WU is not clearly highlighted. Also, how many soil layers are modeled? Is the temporal variability in root depth accounted?

P1136-L21. 30 cm is a rather thick layer for soil evaporation. Do you have any justification for this value?

P1138-L1. G is generally negligible only for full covered densely vegetated surface. Several formulations are proposed in the literature for accounting for daily G in case of sparsely vegetated areas. This can cause distortions in ET spatial distribution.

P1138-L24. The SEBAL model aims at reducing the inconsistency between remotely observed land-surface temperature and ground-observed air temperature. The use of a single Ta value or a Ta map is a different issue; this must be clarified. The use of conventional ground-measured temperature data does not overcome the problem of the inconsistency, which is the main issue in using thermal remote sensing data. The authors should rethink this assumption.

P1139-Eq.11. This equation is valid only for fully vegetated areas ( $T_e \approx T_a$ ). Since the SWIM uses this Eq. in the Turc model, this is a valid assumption; however, the same approximation is not valid in the framework of the surface energy budget ( $T_e \neq T_a$ ). Again, this can be a cause of distortions in ET spatial distribution.

P1140-L11. The temperature gradient in Eq. (15) is wrongly defined. It represents the gradient between aerodynamic temperature (see Norman and Becker, Agr. Forest Meteorol. 77:153-166, 1995) and air temperature above the surface. In your approach there are two strong assumptions: 1) the aerodynamic temperature corresponds to the surface temperature, which is not true for heterogeneous surface as hydrotopes

C404

likely are; 2) the ground-measured air temperature at 2-m corresponds to the value above the surface, which is not true for tall vegetation (i.e., forest but also crops taller than 2 m). The latter assumption is even stronger if we consider that Ta and land-surface temperature are not collected at the same time. Again, SEBAL model (as well as other commonly used thermal-based surface energy balance models) is specifically designed to circumvent both these problems rather than ignore them as your approach does. Given that one of the main goals of the paper is to quantify the value of the information provided by remotely observed land-surface temperature, this information must be used correctly within a physically based framework. In my opinion, considering also another limitation of the available dataset (i.e., absence of wind speed data), the author should consider to adopt a simpler (but widely tested) approach, as for instance the triangle method, rather than drastically simplify a physically based formulation by means of unreliable assumptions. I'm not at all convinced that the ET spatial distribution obtained under these assumptions (as well as the ones on Rn and G) reliably represents what a "standard" thermal-based method would provide.

P1142. Two different calibration methods are introduced here, however, only the first one seems to be used. This has to be clarified. Also, some discussion has to be made on the assumption of unbiased estimates for the two land-use classes. Does this explain differences in the ET maps (e.g., are the hydrotopes with high/low errors characterized by high/low forest fractions)? I'm wondering if this is another possible source of distortion in ET spatial distribution that it is not commonly present in remote sensing estimates.

P1143. This upscaling procedure is rather confused, especially because some terms are not well described: e.g., What ET<sub>tot</sub> is? Is it the annual average of ETSWIM (previously introduced)?

P1143. Is the linear relationship between Delta\_T and ET supported by any evidence? Additionally, the time invariance of aerodynamic difference is a further strong assumption, given that we well know its dependence from vegetation height and mass (as well

C405

as seasonality in wind speed that can not be accounted). This is another assumption not commonly made by thermal-based remote sensing approaches. I would suggest to consider at least the actual roughness parameter for each hydrotope parameters (in Rah), including its temporal variability. The optimization could be performed by calibrating the effect of wind speed at year-scale. This would probably partially reduce the discrepancies with commonly used approaches. In any case, effects of atmospheric stability are ignored (or lumped in the calibration procedure). This must be highlighted.

P1143. Another inconsistency in Eq. (24) is between the time-scale of  $R_n$  and  $\Delta_T$ . In fact  $R_n$  is a daily value while  $\Delta_T$  is measured at a specific time of the day. How do the authors deal with the upscaling of  $\Delta_T$  from one time-of-day to daily value? If another assumption here is made, this must be clearly highlighted.

P1144.sec.2.3. I agree that the water balance method is rather simpler than the other approaches; however, some more details must be provided. For instance, it is clear successively that it was applied separately for each gauged sub-basin, but this should be highlighted here. Was Eq. (25) applied at year scale (separately for the 3 years) and then the ET map averaged? Is  $P$  derived from the same interpolated fields used for SWIM?

P1144-L15. Given that the aim of the work is to evaluate the impact on interior gauges, the authors should demonstrate that the SWIM model performs accurately at least globally. If discrepancies between SWIM and water-balance are observed also globally, these should be minimized before to proceed in the analysis.

P1145-L17. Several methodologies are available in the literature (generally based on sinusoidal function), to reconstruct air temperature at a specific time-of-day from daily min, max and mean values. The correspondence between  $T_{a,max}$  and  $T_{a,overpass}$  is another assumption not really required.

P1145-L20. Is seven the number of images 99% clear in this study case? Please clarify that this number coming from the analysis of the data.

C406

P1145-L26 to P1147-L3. It is not clear to me what “no time-dependent weighting scheme... had been applied” means. Please detail more this consideration.

P1146-L7. This is another critical point. LST maps provide information on cloudy condition at the satellite overpass time, but nothing is said on the whole day. What about mixed conditions, when the sky is cloudy at the overpass time and clear the other daytime hours or vice versa? This point has to be clarified.

P1146-L9to15. In my understanding, the attenuation factor represents the relationship between  $\Delta T$  during “clear-sky” and “cloudy” conditions. Why do you assume that this is a constant value? Actually, it was observed in the literature (see e.g., Gallo et al., J. Appl. Meteorol. Clim. 50:767-775, 2011) that  $\Delta_T$  tends to be  $= 0$  under cloudy conditions, but it is a function of vegetation coverage under clear-sky conditions. This suggests that the attenuation factor is a function of vegetation coverage. It would be interesting to analyze the effect of attenuation factor  $= 0$  on the results (which is a rather simpler assumption than your approach, but likely more close to the reality).

P1146-L21. Is this map the 3-year average? Please clarify.

P1146-L21to22. Why mountainous area should have  $\Delta_T$  close to 0?  $\Delta_T$  is an indicator of water stress, which can occur both in mountainous and lowland areas. The small values in the upper regions can be related to an incorrect evaluation of the effects of elevation on  $T_a$ .

Sec.3.2. This should be the main section of the paper, and it is only 1 pag. (compared to 12 pages. of methodology). In my opinion the author should clearly state if they consider the water balance the “target” reference to evaluate the other two models and discuss the results accordingly.

P1147-L14to17. This is true in general, but when the outliers are removed from the analysis (see Fig. 12) the ranges of variability of SWIM and remote sensing are similar. Some more comments on this should be reported.

C407

P1147-L4to9. The correlation remote-sensing vs. water-balance and SWIM vs. water-balance are rather similar (on the German sub-basins). This seems to suggest no value in using remote sensing data. This point has to be deeply discussed, since it is the main topic of the paper.

P1148-L23to24. Is the density of ground measurements a real issue, especially over lowland areas where  $T_a$  varies smoothly? A map with the spatial distribution of the stations would be very helpful. Also, the SWIM model is affected by the same issue too, and network density is an even more relevant issue for precipitation. Some discussion on this should be added.

P1149-L5to10. This statement on the “remaining noise” must be supported by numerical evidences. What are the average bias and accuracy of the remote sensing estimates in this study case (assuming water balance as a reference)? Are these values comparable with the one obtained in the reported studies?

P1149-L12to16. This seems out of place in the discussion.

P1149-L21 to P1150-L4. These differences should be reported in terms of impacts on ET and compared to the differences with modeled (SWIM and remote-sensing) values. Are the differences in these sub-basins explained by the magnitude in water unbalance?

P1151-L23 to P1152-L2. I don't understand the sentence “the water balance approach does not seem to be more exact than the two other methods”. In my opinion, the fact that the two methods do not agree with the water balance is not a valid indicator of the accuracy of the water-balance. Only an external observation of ET (or storage) can quantify the accuracy of the water-balance approach.

If the authors consider the strong correlation between SWIM and remote sensing approaches an indicator of the good accuracy of these two estimates, what about the likely disagreement between discharges modeled in the interior points with the ob-

C408

served values?

P1152-L10. The reported reasons are in some cases true for all three approaches, and not only for the water-balance.

P1153-L13to21. This is not a conclusion in my opinion.

P1153-L22 to P1154-L6. This is not related for the result reported in this paper, and it is out of place (it may be part of the introduction).

P1154-L13to17. Again, the value of SEBAL (or other similar methods) is in minimizing the errors associated to the inaccuracy in thermal data calibration, as well as removing the need to separately estimate aerodynamic temperature and air temperature above the surface. The misunderstanding on the value of SEBAL (or other similar methods) must be clarified. The problem of spatially variable air temperature is also faced and addressed by SEBAL (or other similar methods) for applications over complex areas.

Technical corrections

P1131-L24. Spell-out the acronyms SWAT, SEBAL, as well as the other acronyms through the text (e.g., SWIM).

P1138-L15. The Stefan-Boltzmann equation reported here is redundant, since it is also reported in Eqs. (7) and (8).

P1138-L17. The correct value of the Stefan-Boltzmann constant is  $5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ .

P1139-L20. Is the effect of elevation accounted in  $R_{max}$ ?

P1140-L8. It is not necessary in my opinion to introduce the concept of Bowen-ratio (with 2 references), since it is barely used successively.

P1141-L6. This concept must be introduced early in the paper.

P1145-L2. In my opinion, it is not necessary to refer to other products not used in this

C409

study, as NDVI and LST nighttime maps.

P1145-L26. Blue-sky fraction is not defined.

P1146-L9. Please define “white pixels”.

References. The reference list is too long. Please review all the reported references and keep only the ones that are strictly necessary.

Table 1. Are the data reported here derived from SWIM? Please clarify.

Fig. 1. & 3. I’m not sure that these figures are really useful.

Fig. 5. Are these maps obtained as 3-year average? It is not clear the definition of “absolute” and “relative” cloud-freedom.

Fig. 6. Again, is this the 3-year average?

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 1127, 2013.