Reply to interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 13595, 2013.

## *Authors' reply to "Interactive comment on* "Improving the complementary methods to estimate evapotranspiration under diverse climatic and physical conditions" *by* F. M. Anayah and J. J. Kaluarachchi" by J. Szilagyi

szilagyi@vit.bme.edu

Received and published: 20 November 2013

My comment on this manuscript does not strive to be comprehensive. I just list a few problematic issues.

**Reply: The authors would like to thank Dr. Szilagyi for the time and effort made available to comment on the manuscript. The comments are constructive and helpful to better present the concepts and interpret the results.**

1/ In lines 27-29, pg. 13598 the authors mention some previous studies that used the complementary methods (CM) with "little success" and they list two of my recent works I was the principal author of: Szilagyi and Kovacs 2010, 2011. I am totally confused because in these studies the application of the CM was a clear success, as anyone can check. They also list in this context the recent study by McMahon et al. (2013) who concluded that the CM-based ET estimation methods are the best available practical ET estimation methods.

**Reply: First, the authors wanted to inform that the original complementary methods did not perform well specifically under different physical and climatic conditions. The authors did not attempt to undermine the work of others and instead, the authors were attempting to solicit more attention to the complementary methods and to show the need for improvements. Many other studies can be added here such as those of Hobbins et al. (2001) and Xu and Singh (2005) that showed the original complementary methods should be further studied.**

**As for the work of Szilagyi and Kovacs (2010, 2011), there was a systematic underestimation of 44% from measurements in one of the three EC sites used for validation. Although the difference was referred to the physical variation of that particular site (i.e., Hegyhatsal), the three sites still lie within a small area that shares similar climatic and geographic conditions. In the results of Szilagyi and Kovacs (2010), the authors corrected the ET estimates of the Hegyhatsal site due to the poor performance. The focus here is to say that the original complementary methods need further corrections and modifications to perform under variety of climatic and physical conditions.**

**As for the study of McMahon et al. (2013), Table 4 showed that CM-based models are not preferred or not recommended to use in most applications. It is only the Morton method that is preferred in the application of shallow and deep lakes. Additional studies failed to predict ET using the original complementary methods are shown in the supplementary material attached to the article in particular Sections S7 and S8.**

**We will edit the manuscript as needed for further clarity.**

2/ In the Penman equation the second, aerodynamic term accounts for local advection and not for "large scale advection effects" as the authors claim in line 21, pg. 13602. What accounts for large scale advection is the value of the Priestley-Taylor parameter, alpha, being larger than unity.

**Reply: As stated by Hobbins et al. (2001) when defining the terms of the Penman equation: "the second term of this combination approach represents the effects of large-scale advection in the mass transfer of water vapor and takes the form of a scaled factor of an aerodynamic vapor transfer term $E_a$."**

**As defined by McMahon et al. (2013), the aerodynamic component of the Penman equation accounts for regional drying power of atmosphere. However, alpha coefficient of the Priestley and Taylor (1972) is actually computed under advection-free conditions. It is known that the alpha value of 1.26 represents 26% of additional power of evaporation induced by advection (sensible heat transferred by wind).**

3/ In line 15, pg. 13604 the authors claim about the GG method that it does not need "surface parameters (temperature and vapor pressure)". I am asking them: which CM method asks for such values, because I am not aware of it, at least what concerns the CRAE or AA methods.

**Reply: The unique feature about the GG method is that it can estimate actual ET with no prior estimate of ETP. This is the first half of the sentence that is not true for either the CRAE method or the AA method. As for the other half which is describing the temperature and vapor pressure requirements, there is no indication that this is exclusively true for the GG method only. We will edit this sentence for clarity.**

4/ It would have been much more informative to use a mean BIAS value, not an abso- lute one, to see where the models overestimate and where underestimate EC-derived ET rates. From the published BIAS values this cannot be deduced, since they are all positive values, yet the authors discuss under and overestimation of the different mod- els under different climates before they do their analysis with the model-components.

**Reply: The authors decided to use the absolute value of bias to better assess and demonstrate the behavior of different models for accuracy. The problem of having the true value of the bias is that the results could be misleading especially when comparing a large number of model variations. This can be explained simply in the following example.**

**Consider two model variations A and B of bias range from -5 mm to 4 mm and from 2 mm and 3 mm, respectively. Which is better? Although the mean bias value of model A ((-5+4)/2=-0.5 mm) is smaller than that of model B ((2+3)/2=2.5 mm), yet model B is better as the discrepancy (bias) from the origin (true measurement where bias should approach zero) is lesser. The spread of bias values around the origin (zero value) is larger for model A which is an unfavorable for a better estimates of ET. It should be noticed that using the absolute value of bias simply avoids this confusion. Mu et al. (2011), for instance, had used the absolute bias value to assessing model performance since 46 AmeriFlux sites had been considered in this study. However, Mu et al. (2007) had applied their model on a fewer number of sites (19 AmeriFlux sites) and therefore used the bias value for comparison. Huntington et al. (2011) had used the percent bias which cannot have a negative value.**

5/ In line 28, pg. 13608 the authors say that the GG method has the lowest bias, but I do not think a value of 15.7 vs 15.5 marks a statistically significant difference, considering the errors in the EC measurements.

**Reply: Although the mean values of BIAS may not be statically different, but the AA model and GG methods both share the lowest mean BIAS value. Therefore, the average values of the different models when compared together, some characteristics of the distribution may be concealed. This is one of the major shortcomings of using simple arithmetic average. It is suggested to compare minimum and maximum values of the BIAS, for instance, and the difference will be clearer.**

6/ In lines 23-27 the authors discuss the study of Szilagyi and Kovacs (2010) and they say that at the third EC site the CM-based model gave a difference of 44% in ET rates in comparison with EC measurements. Unfortunately, they do not tell the reason why, which when explained turns out to yield the best ET results of the three sites. As is discussed by Szilagyi and Kovacs (2010, 2011), at that site the EC instruments were installed on a radio-transmitter tower at a height of 82 (as in eighty-two) m above the ground. Under certain wind directions the instruments were in the wind-break of the tower making the method unusable in such periods. Consequently, the derived ET and sensible heat rates added up to 44% less than the energy balance. Accounting for it, the CM-based monthly ET estimates explained 95% of the variance found in the EC measurements, with practically no bias. And this leads us to the question of footprints. The 82 m height of the EC instruments above ground translates into a footprint really comparable to the scale of the CM-based ET estimation methods: most likely the reason for the best, unbiased performance in comparison with EC data.

**Reply: In the original manuscript, there is no mention of such details (measurement heights of the other two towers) and whether there is a threshold to distinguish the validity of the complementary relationship. The authors used the results and information provided in the original manuscript. We will attempt to make these points clear in the revised manuscript.**

7/ The CarboEurope site (Bugac) from Hungary, listed in Table 1 has a measurement height of less than 2 m above the ground. I am not familiar with the other sites listed in Table 1, but I would risk to say that they may have comparable heights (i.e. a few meters). I ask the authors to list these values in Table 1. If I am correct then the footprints of the majority of these sites are just a tiny fraction of the scale the CM- based ET rates represent. Since surface properties, soil moisture status, vegetation may vary significantly at this fine scale (a few hundred meters) how representative are they then at the scale of the CM-based method? In my opinion a better validation would have been for the CM-variants to use water-balance data for the involved catchments.

**Reply: The height of measurement is not mentioned for each and every EC site. In addition, the canopy height matters and therefore, there is an argument whether to use the height to ground or to canopy. In few cases, it was found that sensors are set at different levels making it even more difficult to decide which one to consider. To avoid confusion to the methods, this parameter was not mentioned. The focus of this study is to explore the validity of the CM-based model variations globally under diverse conditions. The authors agree that some discrepancy in the results may be justified by this parameter while quality of data is of the greatest importance.**

**As for the water balance data, this is interesting; while this is beyond the scope of this work, the authors may consider this in future work.**

8/ I wish the climate of Hungary were Mediterranean as Table 1 claims. It is still conti- nental despite all climate change claims.

**Reply: While we agree with this comment, the aridity index is a metric used in this study comparing all sites based on specified criteria using long-term average values of annual precipitation and temperature. It is understood that some of the aridity index values may not represent the actual climatic conditions of some sites. The authors made a comparison using other aridity indices (not shown here, namely De Martonne, Thornthwaite and Mather, Budyko and UNEP indices) and found that the climatic class significantly varied among the different aridity indices for the same site. The aridity index used here provided the most stable results. The authors are happy to share this information if needed.**

9/ The winning GG18 variant is only slightly better than the original CRAE model. The R2 value is the same, the absolute BIAS value is 11 vs 15.7 mm/mo, and the RMSE value is about 20 vs. 27.8 mm/mo. Yet there is a big difference in input data require- ments between the two models: the former (GG18) needs wind velocity measurements, while the CRAE model does not (every other model inputs are the same). So the CRAE model performs almost the same as the GG18 model with fewer data input. Wind data is something not at all universally available historically. I still wonder if the GG18 model would outperform the CRAE model with the help of watershed water balance data.

**Reply: One may compare the original GG method and the GG18 model and come up with the same conclusion. Similarly, the CRAE2 model is also comparable with the original CRAE method (see Figure 3). In the same way, many similar comparisons among the models can be performed but probably will not lead to a focused conclusion.**

**The comparison is not between overall average values given by the CRAE method and the GG18 model. There are some other statistics (e.g., standard deviation) that show the accuracy (or distribution) of the ET estimates among the 34 sites. As discussed in the present study, one of the main problems of the CRAE method is that it fails to estimate ET under sub-humid and Mediterranean climatic classes (see Table 2). The discrepancy is clear when compared to the more extreme conditions (Table 2). The GG18 mode, however, shows a better behavior among the 34 sites and results are more consistent regardless of the climatic class as shown in Figure 6. The ET estimates of the GG18 model for the moderate-climate sites are comparable to those of either the wet or dry climatic classes.**