Author's response

Ms. Ref. No.: hess-2021-73

Revised title: The accuracy of temporal upscaling instantaneous evapotranspiration to daily values with seven upscaling methods

Author(s): *Zhaofei Liu*

It would be greatly appreciated for your kind reviewing to this paper. The manuscript is revised according to your valuable comments and suggestion. The revised manuscript should, after incorporating your valuable advices and suggestions, be improved greatly. For your convenience to re-review the paper, the corrections corresponding to your comments are described in detail as follows:

Responses to Editor' Comments

Comments to the Author:

Dear Author,

Thanks for your rebuttal, I think minor revisions are in order to clarify the reviewers' comment and clarify some points.

Looking forward to the revised manuscript,

Niko Wanders

****

Reply:

The manuscript has been revised according to reviewer's comments.

In addition to the comments of reviewers, the revisions to this paper are as following,

1. The scale bar in Figure 4 has been revised.

2. A funding number "XDA2006020202" has been added.

Responses to Referee #2 Comments

It would be greatly appreciated for your kind reviewing to this paper. Thanks very much for your valuable comments and suggestion. For your convenience to re-review the paper, the response corresponding to your comments are described in detail as follows:

This manuscript evaluates the performance of seven upscaling methods for predicting daily ET from instantaneous ET measurements and provides an analysis of the pros and cons of these methods. In general, the manuscript is well organized and well written. I only have minor suggestions as below:

1. The title seems not precise or a bit awkward, perhaps you can reformulate it a bit to something like below:

"The accuracy of temporal upscaling instantaneous evapotranspiration to daily values with seven upscaling methods"

Reply: Yes. In the remote sensing ET retrievals, a critical temporal upscaling step is upscaling from instantaneous to daily ET values. Therefore, this study is focus on evaluating the performance of seven upscaling methods for predicting daily ET from instantaneous ET measurements. The author has revised the title to "The accuracy of temporal upscaling instantaneous evapotranspiration to daily values with seven upscaling methods".

2. In this study, the author claim this is one of the first study to do such evaluation (seven methods) at a global scale, and based on such evaluation, some advantages and disadvantages of different methods were explained/discussed. On the other hand, it is not clear how all these methods will help us to understand better the physics/processes/mechanisms behind daily ET predication at a global scale across all different climate zones;

Reply: The physical process mechanisms behind daily ET predication at a global scale across all different climate zones are very complex. The data used in this study cannot guarantee the representativeness of each climatic region. Spatial distribution of the accuracy of upscaling methods is shown in section “3.3 Spatial distribution of the accuracy of the sine function and EF(Re) methods”. It shows that the performance is

poor in tropical rainforests (e.g., BR-Sa3, GH-Ank, ID-Pag) and tropical monsoon (PH-RiF) sites. As described in this section, "this may be due to irregular changes in the LE in these regions. For example, there is little seasonal variation in LE in tropical rainforest climate regions, and the fluctuation of daily LE data series is relatively small. This results in poor agreement between simulated daily LE and measured values (Fig. 5). However, the SD-Dem site, also located near the equator, was characterized by seasonal variation in LE due to the tropical grassland climate in this region. As such, the simulated daily LE at this site demonstrated greater consistency with measured values. Although the performance of upscaling methods was poor in agreement with the daily LE data, there was an apparent correlation between simulated daily LE and the measured data. For example, the R2 was higher than 0.30 and 0.40 at the GH-Ank and ID-Pag sites, respectively, while it was greater than 0.50 at the PH-RiF site." This is only one result that we found.

We believe the physical process mechanisms behind daily ET predication at a global scale across all different climate zones are very important, and need more further research at regional scales when more data are available.

3. The author suggested that the result of this study can help improve the accuracy of remote sensing ET products. However, there is no example/demonstration with the application of what the author claimed (e.g. Sine, EF(Re)), for producing remote sensing (RS)-based ET products. And there is also no intercomparison between such 'temporal upscaled' ET product with existing RS-based ET data.

Reply: As mentioned above, upscaling from instantaneous to daily ET values is a critical step in the remote sensing ET retrievals. This study is focus on evaluating the performance of seven upscaling methods for predicting daily ET from instantaneous ET measurements. The accuracy of remote sensing ET products might be improved by using the upscaling method with better performance. However, remote sensing ET data production is a huge project. That is not included in this study. But the results of this study can provide reference value for data producers.

4. Perhaps, the data script (e.g., download script / a readme file or so for how you access/download FLUXNET data [stations etc.]), and the processing script can be

opened and invite the community to engage better with this analysis, and see if further methodologies can be developed based upon your studies.

Reply: Yes. The FLUXNET data was downloaded from its website. The website is provided in the section "Data availability". However, it is the homepage of the website. The author has revised this website to the data download link. The link is https://fluxnet.org/data/download-data/. Details of readme file, station information, and other documents could be found in this link. All of the processing script is based on Visual Basic for Applications, which is embedded in Microsoft Excel. The author would like to share the processing script.

Some further minor suggestions as attached

"The intra-day distribution" from which method showed greater consistency with Gaussian function than the sine function?

Is this not expected?

Reply: In Figure 2-(b), the intra-day distribution of normalized LE is directly compared with the sine and Gaussian functions. It shows that the intra-day distribution of LE is more consistent with Gaussian function than the sine function.

As described in the section 3.1 Intra-day distribution of observed LE and its influencing variables, "The Gaussian function matched LE perfectly at any time during the day. The sine function slightly underestimated LE during the afternoon, and tended to overestimate LE from 6:00–10:00 and 15:00–17:00." Yes, this is not expected. This result is also presented in the Abstract as "The intra-day distribution of the LE showed greater consistency with the Gaussian function than the sine function."

Responses to Reviewer #2 Comments

The author studied 7 different methods to upscale instantanuous remotely sensed evaporation to the daily scale under different weather conditions. The manuscript is well structured, easy to read, and in good English. The paper is relevant and suitable for the HESS-audience and provides new insights. Most of my comments are minor (see below), but two points may need further explanation:

1) In the introduction it is stated that there already exist many (review) studies which focussed on this topic. Hence my question: how does this paper differ from those? What is the problem statement of this study? Is it that you investigate the topic also under non-clear sky conditions, plus the fact that you investigate the effect of 'time of satellite overpass'? If so, please elaborate on this and maybe refocus your study.

Reply: As mentioned in the fifth paragraph of the introduction (P3 L75-91), there are several studies focused on the evaluation of different upscaling methods. However, “In general, previous research has largely evaluated upscaling methods on a regional scale” (P3 L91-92). In addition, these studies have obtained different optimal methods at regional scales.

As described in P3 L83-85, “Based on 126 FLUXNET global sites, Wandera et al. (2017) evaluated three EF methods (Rs, Re, and (Rn-G)), finding that the EF(Rs) method yielded relatively better accuracy in daily ET simulations.”. Although this evaluation was at global scale, it only used three EF methods.

Therefore, this study is focus on evaluating the performance of seven upscaling methods at global scale. This is where our study is different from previous studies. The results are shown in the section 3.1 Intra-day distribution of observed LE and its influencing variables, 3.2 Accuracy of seven upscaling methods, and 3.3 Spatial distribution of the accuracy. This study also investigates the performance under non-clear sky conditions and 'time of satellite overpass', which are shown in the section of 3.4 and 3.5 respectively. In addition, Variability of simulation accuracy among different upscaling schemes and sites is also shown in the section 3.6.

According to your valuable comments, the fifth paragraph of the introduction has been revised, so that the characteristics of this study can be more clearly described. Please find revisions in the revised manuscript.

2) Did you check the energy balance closure of the FLUXNET data? And if it did not close, did you use the raw data or some kind of corrected data (e.g., assigning the gap in the SEB via the bowen ratio to H or LE?). Please explain and discuss how and if this affect your results.

Reply: Both of the raw latent heat flux data and the corrected data are included in the FLUXNET data. Corrected data by energy balance closure correction factor was used in this study. This is described in P4 L120.

This study tried to compare the evaluation results between the two data, and found that there was little difference between the two results, when the criteria used the root-mean-square error (RMSE), Nash–Sutcliffe efficiency (NSE), and determination coefficient (R^2). The difference of the evaluation results was relative greater when the criterion is the relative error. The difference magnitude was consistent with the energy balance closure correction factor. Overall, the difference could be found from the equation (8) (P6 L157). In this equation, the ratio of V_d to V_i is the same for simulations from two data. Therefore, the difference between two data simulations is expected to be the same as that between two data series themselves.

The new sentences "It should be noted that the FLUXNET includes both raw and corrected LE data. There was little difference between the evaluation results of the corrected data and those of the raw data." has been added in P16 L430-431.

Other comments:

- P2L51: I am not that happy with the acronym LE in case L is not equal to λ. In my opinion it is most clear if you define the latent heat flux as ρλE in W/m2, where λ= the latent heat of vaporization in kg/J, ρ density of water in kg/m3, and E the evaporation in m/s.

Reply: The acronym LE is referenced from the FLUXNET data document. They defined the latent heat flux as LE. This acronym is common in many literatures. The author would like to modify this sentence "the EF was defined as the ratio of the latent heat flux (LE=λET, where λ is the latent heat of vaporization) to the available energy flux (Rn-G) at the surface." to "the EF was defined as the ratio of the latent heat flux (LE=ρλE, where ρ and λ are the density of water and the latent heat of vaporization, respectively) to the available energy flux (Rn-G) at the surface." Do you think this modification is appropriate? Or could you please give some suggestions

about this modification?

- Eq2+3: What is the unit of LE_t? If this is a scaling factor (and thus dimensionless), I would recommend to change its name. Earlier you defined LE as the latent heat flux (in W/m2), so better to redefine it. Furthermore, 't' is not defined.

Reply: Yes. The name in these two equations needs to be changed. According to your valuable comment, it might better change the LE_t in Eq2 and Eq3 to $SINE_t$ and $GAUSSIAN_t$ respectively. $SINE_t$ and $GAUSSIAN_t$ represent instantaneous sine and Gaussian function values during daytime, respectively. "t" represents time. As you mentioned that, they are scaling factors, and dimensionless. The author would like to make revisions. The corresponding Eq4 has also been revised by the new names. In addition, the sum in Eq4 was replaced by the average.

*-Eq7: I would add ρ (density of water) to the LHS of eq7. Also I don't like the term ET for evaporation. It can be confused with E*T. Better call it E.*

Reply: The author has revised the "λET" to "λρE" in Eq(7).

- P10 L255-262: Other cause can be that remote sensing products only sense the top of the canopy and thus ignore the energy storage with in the canopy. Especially for forest this can be significant. See e.g. Coenders-Gerrits et al./ Jiménez-Rodríguez et al plus references herein.

References:

Miriam Coenders-Gerrits, Bart Schilperoort, César Jiménez-Rodríguez "Evaporative processes on vegetation: an inside look" (2020). pp 35-48. Book chapter in "Precipitation Partitioning by Vegetation: A Global Synthesis", editors John T. Van Stan, II; Ethan D. Gutmann; Jan Friesen; Springer.

Jiménez-Rodríguez, C. D., Coenders-Gerrits, M., Wenninger, J., Gonzalez-Angarita, A., and Savenije, H.: Contribution of understory evaporation in a tropical wet forest during the dry season, Hydrol. Earth Syst. Sci., 24, 2179–2206, https://doi.org/10.5194/hess-24-2179-2020, 2020

Reply: The author studied these two references carefully and found that it's true. Remote sensing products only sense the top of the canopy and thus ignore the energy

storage with in the canopy. Especially for forest this can be significant. According to your valuable comments, the author added this reason in the section “4 Discussion”. The two references are also added in the manuscript.

“In addition, remote sensing products only sense the top of the canopy and thus ignore the energy storage with in the canopy. Especially for forest this can be significant (Jiménez-Rodríguez et al., 2020; Coenders-Gerrits et al., 2020). This partially explains the poor performance in tropical rainforests regions.” has been added in P17 L463-466.

- P11 L293: Personally, I found an error of 36,7-25% not really 'satisfactory'.

Reply: This sentence is based on the previous sentence “When $0.4<\tau<0.5$, the simulated NSE had improved to exceed 0.70, and the corresponding R2 was greater than 0.75.” The relative error is approximately 10% (Figure 7). The words “low atmospheric transmissivity.” in this sentence are tend to confuse readers. The author has revised this sentence to “This indicates that remote sensing ET upscaling methods can achieve satisfactory simulation accuracy even when $0.4<\tau<0.5$.”

- Fig 3: Should the difference between S and M not be mentioned in the method section? Addtionally, it is logical that the error of M is lower than S if you follow the theory of error propagation. In M you have n=3 and thus the error reduces with a factor 1/ ✓n.

Reply: As mentioned in the caption of Figure 3, “S and M represent simulations by a single and multi-time values, respectively. For example, S10:30 is simulated by the ratio of a daytime value to a single time value at 10:30, while M10:30 is simulated by the ratio of a daytime value to the average of three-time values at 10:00, 10:30, and 11:00.” M means that the daily LE is simulated from the average of three-time values, but not from the sum of three values. According to your valuable comments, the author has revised the sentence in the caption “S and M represent simulations by a single and multi-time values, respectively.” to “S and M represent simulations by a single and average of multi-time values, respectively.” to avoid misunderstandings. In addition, as mentioned above, the sum in Eq4 has been replaced by the average.

- Fig 6: you interpolated the RE and RMSE over the entire world. But is this not

visually biassing your graph, since the global coverage of the FLUXNET data is not equally distributed over the world?

Reply: Yes. This may not be appropriate. The author revised the Fig 6 similar with Fig 4. The revised figure only shows the evaluation results of the observed sites, but not the interpolation results.