Reviewer 1 (R1):

We would like to thank R1 for the detailed comments.

1. **Energy budget closure problem at FLUXNET**. Energy budget imbalance has long been identified at FLUXNET sites. The imbalance is about -40% - +20%, indicating latent heat/sensible heat fluxes might be underestimated by up to 40%. Indeed, the energy imbalance is an existing fact we have to accept, I guess there is little can be done to overcome it in this particular study.

Response: Good point. We propose to include an intercomparison of ET_i upscaling results including both energy balance closure and non-closure in the revised version of the manuscript.

2. But my concerning is: if an ANN model is trained by FLUXNET data, how much confidence do we have when we apply it to satellite retrieval? The energy budget close problem affects the results in two ways: (1) the overall robustness of the proposed upscaling method (Rs method); (2) comparison of Rs method with the evaporative fraction based upscaling (EF method Eqn. 5). However, the exo-atmospheric irradiance method is not affected (Eqn. 6). I guess the authors must be aware of this issue; it would be better to literally discuss them in the results section.

Response: Regarding R1's concern on the impact of surface energy balance closure on the performance of ET_d evaluation, **it is important to mention that the implicit assumption in remote sensing based ET_i retrieval is the closure of surface energy balance**. Therefore, for the remote sensing retrievals, the energy balance closure problems will not affect the ET_d estimates in the current framework of ANN. **However, for the validation of remote sensing based ET_d retrievals, surface energy balance fluxes from eddy covariance measurements need to be closed.**

In the present study, the closure problem of surface energy balance will affect the evaluation statistics of all the three methods, and therefore, we propose to include an intercomparison of ET_i upscaling results including both energy balance closure and non-closure in the revised version. **As compared to the EF and R_STOA approach, the Rs method is more robust with regards to *ET* scaling on a daily time frame since the method carries maximum information on the cloudiness, which is a key limiting factor in upscaling of ET_i to ET_d.**

With reference to Eq. (1), the network developed is intended to develop an operational method to directly upscale ET_i (estimated from polar orbiting satellites) to ET_d based on the ratio of daily to instantaneous shortwave radiation (R_{Sd} and R_{Si}). Given there is no direct method to directly estimate R_{Sd} from remote sensing satellite, we trained an ANN with the FLUXNET observations of R_{Si} and R_{Sd}, and validated the model to predict R_{Sd} over independent sites, followed by using R_{Sd}/R_{Si} ratio to convert ET_i to ET_d. The datasets used for the ANN development covers a wide range of biome, climate, and variable sky conditions. Therefore, we assume the R_{Sd} prediction from ANN to capture a broad spectrum of radiative forcing, which is also reflected in the independent validation of RSd and ET_d (Fig. 5, Fig. 7, Table 2). The performance of this model for satellite retrieval of R_{Sd} (from R_{Si}) is dependent on the accuracy of R_{Si} retrieval(Loew, Peng, & Borsche, 2016). We shall make this point explicit in the conclusion section. Also, the distribution of sites over the tropics, Africa, and SE Asia are poor, and more sites over these regions are expected to make the ANN model more robust, which will also be mentioned in the revised manuscript.

3. **Cloudy-sky issue**. The biggest problem of the proposed upscaling method (Rs method) is that the ANN model does not include any information about "cloudiness". Therefore, model performance under cloudy-sky condition (or low atmospheric transmissivity) is much worse than clear-sky condition. One way to tackle it, is to use climatology precipitation data. Rainfall (highly related to cloudiness) has seasonal pattern, at least for some regions (e.g., tropical rainforest, savanna). Similarly, dry season-wet seasons could provide ANN model with additional information about "possibility" of the "cloudy-sky condition" during a certain time period. In Figure 7, the overestimation of ET under cloudy sky condition is "systematic", meaning there might be a simple way to "systematically" down regulate the ET as long as the ANN model knows it's a cloudy day.

Response: Including cloudiness as an input variable of the network during training process would significantly enhance the performance of the network. Use of daily precipitation as an indicator of cloudiness would have been the most appropriate approach in this circumstance. However the cloud information available from alternative sources e.g. from the Clouds and Earth's Radiant Energy System (CERES), the International Satellite Cloud Climatology Project–Flux Data (ISCCP-FD), and Global Energy and Water cycle Experiment Surface Radiation Budget (GEWEX-SRB) are available at coarse spatial resolution and there will be a scale mismatch. However, the precipitation data was not consistently available for most of the sites and the data gaps were significant to alter the sampling sizes. However for future studies, including cloudiness or daily precipitation as a variable in the training of the ANN to predict R_{Sd} is highly recommended. On the issue of systematic errors as a result of cloud conditions, we certainly expect overestimation or underestimation.

4. **FLUXNET site selection**. It was stated that the partition of data into training and validation was randomly selected. However, it's not clear whether the selected training sites are represent it cover a full range of (from dry to wet) rainfall regimes? For each vegetation type, how much percentage of data is selected to train the model? FLUXNET has more forest sites than grass/shrub sites. Are grass/shrub sites less represented in the training dataset? Following question: is the ANN model sensitive the FLUXNET site selection? This could be evaluated by doing e.g., 10 ensemble of random selection of FLUXNET sites. And check the difference among the resultant 10 ANN models?

   Response: Since this analysis was based on FLUXNET sites distributed across 0-90 degrees latitude north and south, the training datasets covers substantial climatic and vegetation variability. The percentage distribution of the training data according to vegetation type was; 23% crops, 31% deciduous broadleaf forest, 10% evergreen broadleaf forest, 20% evergreen need leaf forest, 8% grassland, 7% shrubs and 1% aquatic as indicated in table S1. The number of grassland and shrubs as indicated were relatively less as compared to the crops and forests sites. **However, biome specific error statistics (Fig. 9) indicted the absence of any systematic errors due to vegetation sampling with the exception of EBF. Availability of more EBF sites in the training datasets is expected to reduce the cloudy sky errors substantially. We shall elaborate this discussion in the revised manuscript.**

5. **Crop ET.** I think the proposed method might be only suitable for estimating natural terrestrial ecosystem *ET*. There is large bias of crop *ET* estimation (Figure 9). That could be due to irrigation? Land management? Those anthropogenic factors (largely alter land surface water budget) is not included in the ANN model and the *ET* estimation.

   Response: Yes, in the current framework the approach would be best suited for natural ecosystem. However, inclusion of daily soil moisture and rainfall in the ANN might improve the ET_d prediction in irrigated agro ecosystems. Given the rainfall and soil moisture measurements are not available in all the sites, we propose to use a subset of sites to test this hypothesis where rainfall and soil moisture information are available.

Further having many explanatory variables (e.g., land management, irrigation statistics, anthropogenic factors) to train the ANN, we risk overfiting the model and hence introducing bias.

6. **Vegetation control on ET.** The proposed upscaling method is based on the idea that higher available energy (Rs) lead to higher evapotranspiration (*ET*) (Eqn. 1). It basically assumes that the Bowen ratio does not change during the daytime, so that instantaneous ET/Rs is equal to daily ET/Rs. However, it ignores the important fact that *ET* is also mediated by vegetation via stomata control. For example, trees and grass have dramatically different stomata density, stomata size. Therefore, their stomata open/closure and its control on water vapor conductance are different. The question is: it is worthwhile to add biome type information in the ANN model? Is it possible to further improve the results (Figure 9) for forest sites by considering biome type information in the ANN model and *ET* estimates?

   Response: This is indeed a very good point and needs to be explicitly discussed in the manuscript. The stomatal and biophysical constraints are generally imposed in satellite based ETi retrieval schemes. However the carry over effects of the stomatal control on daily *ET* is indeed overlooked. We assume the inclusion of daily soil moisture and rainfall in the ANN framework will implicitly include the stomatal control at the daily time scale. The additional analysis proposed in the previous response would be helpful in this context. Therefore, instead of biome type information, we would rely on the daily soil moisture and rainfall for a subset of sites, and include a comparative analysis of the current ANN framework (without soil moisture and rainfall) with a modified ANN framework (including soil moisture and rainfall). The new results will also be explicitly discussed in the revised version of the manuscript.

Minor comments

Page2

7. L4. a key challenge in mapping regional *ET* using polar orbiting sensors

   Response: Necessary changes will be incorporated.

8. L6. On the terrestrial surface -> remove

   Response: Necessary changes will be incorporated.

9. L8. The approach relies on : : : -> remove

   Response: Necessary corrections will be made.

10. L16. derived from simple mathematical computation -> replace: e.g., solar zenith angle, day length

    Response: Changes will be made as suggested.

11. L20. Based on the measurements from 126 sites -> remove

    Response: Will be removed.

12. L20. Rs-based upscaling produced

    Response: Necessary changes will be incorporated

Page3

13. L7. ET variability is influenced by (1) available energy received, (2) soil moisture supply and (3) vegetation mediation. I think the third one is missing here. To be complete, the three key factors should all be fairly discussed in the introduction

    Response: Good point. We shall include the vegetation controls on ET in the introduction.

14. L9. "Therefore" is not appropriate here, there is no cause-effect relationship here. Better start a new paragraph and discuss the major challenges in Et upscaling

    Response: Agreed.

Page4

15. L19. Estimate R_{sd} form any specific time-of-day R_{SI} information. But isn't the value of this study is to predict R_{sd} based at satellite local crossing time (e.g., 10:30, 13:30)?

    Response: The aim of this study is to help develop an approach that would help in the upscaling of ET_i (retrieved at satellite overpass time) to ET_d. The value of this study consists of exploiting R_{Si} information at satellite local crossing time to predict R_{Sd} which is not directly retrievable from any polar orbiting satellites, so that the ratio of R_{Sd}/R_{Si} can be further used to upscale ET_i to obtain daily *ET* (ET_d) estimates (in the framework of eqn. 1). Currently we are limited to demonstrating with MODIS overpass times (Terra and Aqua), however in case there are new missions in the future with different local overpass time, the method would still be applicable. We shall make this description explicit in the revised manuscript.

16. L22. L22. In order -> remove

Response: Will do.

17. L24. ANN is a non-linear model. Multi-layer perceptron (MLP) is.. These sentences belong to method section.

Response: Necessary corrections will be made in the revised version.

Page5

18. L13. Cloudiness is a phenomenon. These sentences belong to discussion section.

Response: Necessary changes will be incorporated in the revised version.

Page6

19. L6. Two question: (1) Does Eqn. 1 assume the Bowen ratio is constant during daytime? (2) Does it ignore the night time *ET*, which could be large when surface wind speed is high?

Response: (1) There is no assumption of the conservation of Bowen ratio or evaporative fraction. According to eqn. 1,

$$ET_d/ET_i \approx R_{Sd}/R_{Si}$$

and

$$ET_d/ET_i = EF_d(R_N - G)_d/EF_i(R_N - G)_i$$

Where EF is the evaporative fraction, R_N is net radiation, and G is ground heat flux.

Therefore, eqn. 1 is based on the assumption that shortwave radiation is the principal driver of evaporative flux. Although *ET* can be limited due to both radiation and water, but in the water limited ecosystems the magnitude of ET_i will also be low due to low soil moisture availability and therefore and upscaling ET_i to ET_d in the framework of eqn. 1 may not introduce significant error. The evidence is already seen in Fig. 9 where shrublands showed relatively lower RMSE (despite being water limited) as compared to the forests. We shall extend this discussion in the revised manuscript.

(2) The analysis is based on 24-hour period, meaning night time *ET* contribution is implicitly considered. However, studies have ready shown that the nighttime ET in semi-arid regions contributes only 2 – 5% of the total season *ET* (Malek, 1992; Tolk, J, Howell, & Evett, 2006), and therefore does not appear to be significant.

Page8

20. L16. In a percentage ratio of 80:15:15. Is this right? Shouldn't be 80:15:5 or 70:15:15?

Response: The ratio should be 80:15:5, corrections will be made in the revised manuscript.

Page10

21. L9. We first evaluate the efficacy of the ANN method for predicting R_{sd}.

Response: Necessary changes will be incorporated.

22. L12. As obtained following the methodology described in the section 2.1 -> remove

Response: Necessary changes will be incorporated

23. L13. Showing -> including

Response: Necessary changes will be incorporated

24. L14. From the analysis it is apparent that -> remove

Response: Will be removed as suggested.

Page 11

25. L1. Figure 5 evaluates the R_{sd_pred} under different level of clear sky transmissivity

Response: Necessary changes will be incorporated in the revised manuscript.

26. L3. What if the ANN model includes "clear sky transmissivity, would model performance under cloudy sky condition be improved?

Response: We do not think so, because including clear sky transmissivity could make the modeling framework biased towards clear sky cases only.

27. L16. Using R_{sd_pred}/R_{si} as a scaling factor following eq. 1 -> remove

Response: Necessary changes will be incorporated

Page 12

28. L1. Figure 7 compares ET_{d_pred} against ET_{d_obs} for different level of daily. The overall RMSE, MAPE

Response: Necessary changes will be incorporated.

29. L4. Given that the overestimation is a systematic, is it possible to eliminate it or reduce it? The overestimation was due to the fact that during the specific time slot of interest (e.g., 11:30) the sky is clear while the sky is cloudy during other times. However, there could be another opposite case that sky is cloudy at e.g., 11:30 but clear at other times. It will probably lead to an underestimation of R_{Sd_pred}, and consequently underestimation of ET_{d_pred}. I am wondering why the latter is not the case at least in Figure 7.

Response: This is a very good argument. With the current framework of ANN, this systematic overestimation cannot be eliminated. However, with the inclusion of daily rainfall and soil moisture in the ANN model, such overestimation tendency could be reduced.

Regarding R1's argument on finding underestimation of ET_d from 1130 hr cloudy sky ET_i upscaling in a predominant clear day, such cases were also found in τ_3 category (Fig. 7) where clouds of data points clearly falling significantly below the 1:1 line, thus showing substantial underestimation of ET_d. We shall include this discussion in the revised manuscript.

30. L14.higher errors in ET_{d_pred} can be expected. Is there a way to overcome this problem?

Response: One of the probable ways to overcome the errors in cloudy sky is to incorporate daily rainfall and soil moisture in the ANN. This argument will be made explicit in the revised manuscript.

31. L24. Again, biome specific results are related to the clear-sky issue. Tropical evergreen broadleaf forests have high ET, water tends to re-cycle locally and generate rainfall. It's reasonable to see that cloudy sky condition is more frequent at tropical evergreen broadleaf forest than e.g., at grass land.

Response: Agreed. This point will be added in the discussion of the revised manuscript.

32. L27. ET estimations at cropland were much worse than grass. It that because e.g., irrigation? Land management? Or any other anthropogenic factors that are not considered in the ANN model? Page 13.

Response: Yes, the farm management practice especially irrigation might have impact on the output for example in a case where irrigation was carried out for three consecutive days yet the sky conditions were consistently cloudy would present a challenge. We shall explicitly mention this in the discussion section of manuscript.

33. L20. Based on Table 2, Figure 11, R_STOA method seems successful. Under clear sky condition, it was even better than the proposed Rs method. Further, over longer time scale (annually), there is no big difference between R_STOA and Rs.L20:

Response: Agreed and discussed also in the manuscript. As shown in Table 2, relatively lower RMSE of RsTOA for atmospheric transmissivity class above 0.75 reveals that under pristine clear sky conditions RsTOA can be successfully used to upscale ET_i. However, one of the main reasons for the differences in RMSE between Rs and RsTOA method for daily transmissivity above 0.75 could be due to the fact that if ET_i upscaling is performed from a cloudy instance for a predominantly clear sky day, then such RMSE difference between the two different upscaling methods is expected. These results also showed the probability of a hybrid ET_i upscaling method by combining Rs-method (for transmissivity between zero to 0.5) and RsTOA-method (for transmissivity greater than 0.5). However this hypothesis needs to be tested further. We shall discuss this explicitly in the revised version of the manuscript.

Page 16

34. L1. Briefly define what RsTOA-based method is, what is Rs method.

Response: Rs-TOA-based method is the upscaling method based on R_STOA and R_S method is the method based on Rs. The meaning R_STOA and Rs were earlier defined in the manuscript; please see Page 3 (L25 – L29). We shall further expound on it in the revised manuscript.

35. L4. ET_{d_pred} are defined early in the manuscript, consider the summary as an independent section. Better not to use these acronyms, or re-define it.

Response: Agreed, necessary changes will be incorporated

36. L21-25. This paragraph belongs to results & discussion section.

Response: Necessary changes will be incorporated

References

Loew, A., Peng, J., & Borsche, M. (2016). High-resolution land surface fluxes from satellite and reanalysis data (HOLAPS~v1.0): evaluation and uncertainty assessment. *Geoscientific Model Development*, *9*(7), 2499–2532. article. http://doi.org/10.5194/gmd-9-2499-2016

Malek, E. (1992). Night-time evapotranspiration vs. daytime and 24h evapotranspiration. *Journal of Hydrology*, *138*(1), 119–129. article. http://doi.org/http://dx.doi.org/10.1016/0022-1694(92)90159-S

Tolk, J, A., Howell, T. A., & Evett, S. R. (2006). Nighttime Evapotranspiration from Alfalfa and Cotton in a Semiarid Climate. *Journal of Agronomy*, *98*(3).