Anonymous Referee #1
Report #2
Received and published: 21 January 2019
The comments and suggestions were provided in the annotated version of the document (attached as a supplement). The author's responses are in the sequence of the comments in the manuscript and generally contains (1) comments from Referees, (2) the author's response (in blue), and (3) author's changes to the manuscript (in red).

P1 L10-12 I think you need to state herein what this model is trying to achieve i.e. you are applying it to improve the accuracy of the ET estimate generated from coarse resolution data.
Response: Thanks for your suggestion. This sentence was rewritten as follows:
This study proposes a simple but efficient model (EFAF) for estimating the daily ET of remotely sensed mixed pixels using a model of the evaporative fraction (EF) and area fraction (AF) to increase the accuracy of ET estimates over heterogeneous land surfaces

P1 L20-23 Now that you have demonstrated that the technique does improve the ET estimates. I would recommend adding a sentence detailing how this research can prove to be beneficial to future hydrological applications just so that you adequately conclude on the relevance of this study.
Response: Thank you for this suggestion. We have added the following statement:
It is indicated that EFAF can reproduce daily ET with reasonable accuracy; can be used to produce the ET product; and can be applied to hydrology research, precision agricultural management and monitoring natural ecosystems in the future.

P2 L8 What do you mean by this? The manner in which it is phrased makes this description come across like it is a model issue rather than an issue associated with the spatial resolution at which the data is captured.
Response: Thanks for your suggestion. This sentence was rewritten as follows:
Therefore, models of ET estimates that perform well for fine-resolution remote sensing data (e.g., 30-m resolution Landsat data) may not be appropriate for coarser resolution data (e.g., 1-km resolution MODIS and AVHRR data)

P2 L10 derived
Response: Thank you for this suggestion. We have revised this word to "derived".

P2 L13-14 Chronological order? Check throughout and correct if necessary.
Response: We are sorry for that mistake. We have carefully reordered the references in chronological order throughout the manuscript.

P2 L15-17 From the manner in which this is described there appears to be no difference between the two types of calculations
Response: Thanks for your reminder. This sentence was rewritten as follows:
Distributed calculations are retrieved at fine resolutions and then aggregated to a coarser resolution, which is assumed to provide correct calculations in common scaling studies because

the fine-resolution calculation closely represents actual conditions, whereas lumped calculations aggregate fine-resolution parameters to a coarser resolution and then are retrieved at coarser resolutions.

P4 L12 used as

Response: Thank you for this suggestion. We have revised this part to "used as".

P4 L12 It is unclear what is meant here by simulate satellites.

Response: Thanks for your thoughtful suggestion. To avoid ambiguity, this statement was adjusted to "simulate the remote sensing images".

P7 L1-5 Is the EF of the mixed pixel and associated pure pixel manually determined?

If so, this can prove to be quite a laborious approach especially considering that there may be multiple land cover types which are present in a mixed pixel with the distance to the nearest pure pixel required to be known for each in order to successfully employ the EFAF approach. Please can this be commented upon in the text and acknowledge how this may limit the feasibility of this approach?

Once the EF and subsequently the LE of each sub-pixel is determined how is the daily ET for a particular region then estimated and displayed in a gridded format? The methodology described herein only provides a clear indication of how an individual mixed pixel is corrected but how is this process done for an entire image simultaneously correcting for all the land cover types that have been identified. This might be quite a simple task but it remains unclear how this was achieved.

Response: Thanks for your thoughtful comment. The EF of the mixed pixel and associated pure pixel is determined by computer program with nearest neighbour problem (NNP), which is a computer graphics problem. At present, there are many practical methods to carry out nearest neighbour search (NNS) quickly (Andrews, 2001; Zezula et al., 2006). With the development of cloud computing and parallel computing, we believe that NNS will become increasingly faster. So, we do not worry that this process may limit the feasibility of EFAF.

On the other hand, we used the matrix operation to estimate the entire image ET and displayed the ET results in GeoTIFF format in this study. We mainly focus on describing the process and the rationality of EFAF in this manuscript, and more technical issues regarding the application of EFAF will be considered in our following papers.

Therefore, we added the following sentence to Section 2.2

This process can be easily and rapidly implemented by computer program with matrix operation and nearest neighbor search (NNS) (Andrews, 2001; Zezula et al., 2006).

P10 L14-15 If images were only collected for 2012 can this be rephrased to state that 9 images were collected from June to September.

Response: Thanks for your suggestion. We have rephrased the statement as below:

Therefore, 9 images were selected for the study area under clear or partly cloudy conditions based on data quality metrics and artificial visual interpretation from June to September 2012, i.e., 30

June, 8 July, 27 July, 3 August, 15 August, 22 August, 29 August, 2 September, and 13 September.

Anonymous Referee #2
Report #1

Received and published: 10 January 2019

The comments and suggestions were provided in the annotated version of the document (attached as a supplement) . The author's responses are in the sequence of the comments in the manuscript and generally contains (1) comments from Referees, (2) author's response (in blue), and (3) author's changes to the manuscript (in red).

1) the introductionairy paragraph now also contains information regarding previous results of scaling methodologies. However these mostly talk about differences found on how aggregating input data to estimate lumped ET causes extra uncertainties in regard to distributed estimations. They do not however talk about efforts on actual disaggregation of ET. While DisAlexi is mentioned, the results of this model are not clearly shown. This (in my view) should be the benchmark to which all disaggregation models should strive. As such, it should be mentioned in the 'discussion section', however apart from the initial mentioning in the introduction it is not discussed.

Response: Thanks for your suggestion. I have added relevant sentences in introduction section and discussion section, as below:

Norman et al. (2003) proposed an approach called the DisALEXI model to estimate surface ET with the combination of low- and high-resolution remotely sensed data with little subjective endmember selection. Andersonet al. (2011) achieved upscale remotely sensed ET estimates by combining geostationary satellites and polar-orbiting satellite data and verified the consistency of ET estimate from high- to low-resolution based on the DisALEXI model.

(5) Mixed land cover types in a pixel are the major source of scaling errors in ET estimates (Chen, 1999). However, other surface variables may also be spatially discrete such as land surface temperature, surface albedo values, downward shortwave radiation and other factors, which cannot be ignored. For example, Norman et al. (2003) proposed the DisALEXI model to increase the accuracy of estimating surface ET considering the scale effect of atmospheric temperature. Anderson et al. (2011) have achieved in daily remotely sensed ET estimates by combining geostationary satellites and polar-orbiting satellite data using the DisALEXI model. EFAF have increased the accuracy of ET estimate by considering the scale effect of land cover types. The scale effect of other surface variables for ET estimate requires further investigation in EFAF. Addressing these issues forms the foundation of our ongoing work.

2) You specify the IPUS is a widely used EB model, capable of handlin multi-scale (VNIR,TIR) observations. The equations provided of this model however do not show this. In fact, the equations appear similar to all other EB models. Particularly because the TIR data used is of much lower resolution then the VNIR (300m vs 30m), the 'sharpening' of the TIR data should be explained more clearly, as the whole focus on this paper is on ET scaling issues.

Response: Thanks for your comment. In the case of input parameter upscaling (IPUS), the inputs of the energy balance model are first retrieved at finer resolution and then aggregated to coarser resolution. Subsequently, these coarser-resolution inputs are used in the one-source energy balance model named IPUS to obtain the four energy balance components at coarser resolution (Jiao et al., 2014; Peng et al., 2016). This model provides a feasible parameterization of the ET algorithm that uses only remotely sensed data and does not consider the surface heterogeneities at all. Therefore, this model matches well with our study as needed. On the other hand, this model was designed to simulate the remotely sensed images that have identical spatial resolutions in both the visible near-infrared (VNIR) and thermal infrared (TIR) bands and need not do the 'sharpening' of the TIR data.

Above all, we revised the sentences as below:

The input parameter upscaling (IPUS), a widely used one-source energy balance model that can handle the upscaling of all surface variables to a large scale before calculating the heat flux and does not consider the surface heterogeneities at all, is used as the lumped method in this study

3) In your response you have indeed clearified that advection is not included in the model. In fact this can be clearly seen in the transition zones in the transition zones between the oasis areas en the uncultivated ones. This however is not clearly highlighted upon on page 19 (where this is shown).

Response: Thanks for your reminder. We have revised the sentence as below:

In addition, larger dA values occurred mainly at the transition zones between oasis areas and uncultivated land because advection and its influences are not considered in EFAF. Addressing vertical and horizontal transport (such as oasis effects) at the same time would be excessively complex, and to our knowledge, such a process remains a huge challenge in the remote sensing of heat fluxes. However, as we can see, large positive and negative dA values existed in the same mixed pixel effectively (for example, the dA value on 2 September (Fig. 10)).

4) You specify that you exlude stations 07,08,10 and 15 from the results as they are located in the area's with pure pixels. However on page 21, you specify RMSE's for the pure pixels. In that regard, I am unsure what you mean. Please clearify and correct.

Response: Thanks for your good comment. EFAF is a model for calculating the ET over heterogeneous land surface (or in mixed pixels). To ensure a more reasonable validation of EFAF results, we should choose the validation sites located in the mixed pixels. Therefore we exclude the stations located in the pure pixels, i.e. EC07, EC08, EC10 and EC15.

On the other hand, in order to discuss the rationality of Hypothesis 2, we choose the EF for each pixel in the entire image, which is regarded as the correct value, to compare to the mean EF of its nearest pure pixel(s). This strategy can analyze the errors caused by Hypothesis 2 by analysing the deviation of the correct value.

The first paragraph of Section 4.3.2 was rewritten to avoid the vague meaning as below:

Hypothesis 2 states that the EF of each sub-pixel in a mixed pixel is approximately equal to the EF of the nearest pure pixel(s) of the same land cover type. It is indicated that the EF of a pure pixel can be regarded as the correct value. Therefore, we can choose EF for each pure pixel in the entire image to compare to the mean EF of its nearest pure pixel(s) to analyse the errors caused by Hypothesis 2. The RMSE, MBE and R^2 values were calculated for each maize, grass, bare soil and vegetable land cover type (Fig. 12).

5) In your study area you had the benefit of meteorological towers, so that you could have very good estimates on Ta, U, humidty.. etc. However in many case's such measurements are not present (and only singular values for these parameters are available for a whole study area, estimated by weather models). In my view, this is one of the most important uncertainties when scaling between coarse and fine resolution. Can you please comment on this in your discussion.

Response: Thank you for this suggestion. In fact, the effect of surface heterogeneity on remote sensing-based ET is caused mainly by heterogeneity of surface characteristics such as land cover types, land surface temperatures, surface albedo values, downward shortwave radiation and other factors. However, when scaling RS measurements over terrestrial surfaces, the scale effect caused by a density change is almost negligible; in general, mixed land cover types in a pixel are the major source of scaling errors (Chen, 1999). Therefore, EFAF focuses on the scale effect caused by land cover types.

We have added the following discussion in Section 5:

(5) Mixed land cover types in a pixel are the major source of scaling errors in ET estimates (Chen, 1999). However, other surface variables may also be spatially discrete such as land surface temperature, surface albedo values, downward shortwave radiation and other factors, which cannot be ignored. For example, Norman et al. (2003) proposed the DisALEXI model to increase the accuracy of estimating surface ET considering the scale effect of atmospheric temperature. Anderson et al. (2011) have achieved in daily remote sensed ET estimates by combining geostationary satellites and polar-orbiting satellite data using the DisALEXI model. EFAF have increased the accuracy of ET estimate by considering the scale effect of land cover types. The scale effect of other surface variables for ET estimate requires further investigation in EFAF. Addressing these issues forms the foundation of our ongoing work.

6) I find the modified 'error analyses' of the different hypothesis a good addition. I however miss how this can be used in terms of applicability. For instance, if the difference in dA is too large (for transition zones), one can specify that for those pixels, initial assumptions break down, and now estimations can be made. Similarly, one could specify that if R2 values fall below 90%, those data should not be included filtered out. is this possible?

Response: Thanks for your good suggestion. Regarding advection in the model calculation, to my knowledge, such a process remains a huge challenge in remotely sensed heat flux estimates. We have been aware of this risk and have removed the data contaminated by advection (a threshold "H+LE>Rn+G" is used to find advection effects) from EFAF results and the validation dataset when we undertook validation. To avoid non-values for images and better discuss the rationality of EFAF, we did not remove the data contaminated by advection in mapping ET results. Making EFAF more reasonable forms the foundation of our future work.

7) Finally, I find the most of the discussion section in fact not discusses the results, but is more a 'summary' of what the method does. Next to the obvious shortcomings that were highlighted in the main text, I would prefer if here additional information is given regarding other methods (such as Disalexi).

Response: Thanks for your suggestion. The results have been analysed in Section 4 and were not discussed in Section 5 to avoid repetition. I have added information about other methods in introduction section and discussion section, as below:

Norman et al. (2003) proposed an approach called the DisALEXI model to estimate surface ET with the combination of low- and high-resolution remotely sensed data with little subjective endmember selection. Andersonet al. (2011) achieved upscale remotely sensed ET estimates by combining geostationary satellites and polar-orbiting satellite data and verified the consistency of ET estimate from high- to low-resolution based on the DisALEXI model.

(5) Mixed land cover types in a pixel are the major source of scaling errors in ET estimates (Chen, 1999). However, other surface variables may also be spatially discrete such as land surface temperature, surface albedo values, downward shortwave radiation and other factors, which cannot be ignored. For example, Norman et al. (2003) proposed the DisALEXI model to increase the accuracy of estimating surface ET considering the scale effect of atmospheric temperature. Anderson et al. (2011) have achieved in daily remotely sensed ET estimates by combining geostationary satellites and polar-orbiting satellite data using the DisALEXI model. EFAF have increased the accuracy of ET estimate by considering the scale effect of land cover types. The scale effect of other surface variables for ET estimate requires further investigation in EFAF. Addressing these issues forms the foundation of our ongoing work.

I have noticed some citations to papers in the main text, which have not been included in the references. This should be addressed

Response: Thanks for your kindly reminder. We have added the missing references to the manuscript.