Response Letter

Anonymous Referee #1

Received and published: 4 June 2018

The comments and suggestions are 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 contain (1) comments from Referees, (2) authors' response (in blue), and (3) authors' changes to the manuscript (in red).

P1 L24 More commonly referred to as the evaporative fraction

Response: Thanks for your suggestion. We have revised this phrasing throughout the manuscript.

P2L3-6 Reference? This statement can be argued otherwise. The data that is captured at the coarser resolution can be considered representative at that spatial scale (1 km). If validation/verification of the simulated ET were undertaken at this scale would there necessarily be a larger degree of bias? I think it is important to take cognizance of the spatial resolution associated with the data that is being assessed, as well as the data that it is being compared to. It would also be useful to define, at the outset what is considered fine and coarse resolution for this particular study.

Response: This argument is very important and helpful for understanding the nature of the spatial scale error in remote sensing. The spatial scale errors in remotely sensed ET (and other parameters inversed from remote sensing data) are mainly caused by the combination of nonlinear models and surface heterogeneity, which are more likely to occur in coarser resolution data (Hu and Islam, 1997; McCabe and Wood, 2006; Z. L. Li et al., 2013). Therefore, the data captured at coarser resolution cannot be considered representative at that spatial scale (1 km) without considering spatial scale errors. Accordingly, the validation/verification of the simulated ET undertaken at the same scale as coarse resolution data will not be able to mitigate the bias itself. Special techniques and more information are needed to eliminate these errors (Cammalleri et al., 2013; Ha et al., 2013; Kustas et al.,2003; Maayar and Chen, 2006). The authors have reported on an approach that is able to address spatial scale issues when estimating the daily ET, and it should be useful under most circumstances of coarse-resolution data (i.e., from 102~104 m). This statement was not sufficiently clear and has been revised as follows:

Studies have shown that different landscapes (Blyth and Harding, 1995; Bonan et al., 2002; McCabe and Wood, 2006; Moran et al., 1997) and subpixel variations of surface variables, such as stomatal conductance (Bin and Roni, 1994), or leaf area index (Bonan et al., 1993; Maayar and Chen, 2006), can cause errors in heat flux estimations. Models 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). The spatial scale errors in remotely sensed ET (and other parameters inversed from remote sensing data) are primarily arise from the combination of two factors, i.e., nonlinear models and surface heterogeneity, which is more likely to take place in coarser resolution data (Garrigues et al., 2006; Gottschalk et al., 1999;Hu and Islam, 1997; Jin et al., 2007;Z. L. Li et al., 2013; McCabe and Wood, 2006; Tian et al., 2002; Xin et al., 2012).

Bin, L., and Roni, A.: The Impact of Spatial Variability of Land-Surface Characteristics on Land-Surface Heat Fluxes, Journal of Climate, 7, 527-537, 10.1175/1520-0442(1994)007<0527:TIOSVO>2.0.CO;2, 1994.

Blyth, E. M., and Harding, R. J.: Application of aggregation models to surface heat flux from the Sahelian tiger bush, Agricultural & Forest Meteorology, 72, 213-235, 1995.

Bonan, G. B., Pollard, D., and Thompson, S. L.: Influence of Subgrid-Scale Heterogeneity in Leaf Area Index, Stomatal Resistance, and Soil Moisture on Grid-Scale Land–Atmosphere Interactions, Journal of Climate, 6, 1882-1897, 10.1175/1520-0442(1993)006<1882:IOSSHI>2.0.CO;2, 1993.

Bonan, G. B., Levis, S., Kergoat, L., and Oleson, K. W.: Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models, Global Biogeochemical Cycles, 16, 5-1-5-23, 10.1029/2000GB001360, 2002.

Cammalleri, C., Anderson, M. C., Gao, F., Hain, C. R., and Kustas, W. P.: A data fusion approach for mapping daily evapotranspiration at field scale, Water Resour. Res., 49, 4672-4686, 10.1002/wrcr.20349, 2013.

Ha, W., Gowda, P. H., and Howell, T. A.: A review of downscaling methods for remote sensing-based irrigation management: part I, Irrigation Science, 31, 831-850, 10.1007/s00271-012-0331-7, 2013.

Hu, Z. L., and Islam, S.: A framework for analyzing and designing scale invariant remote sensing algorithms, Geoscience and Remote Sensing, IEEE Transactions on, 35, 747-755, 10.1109/36.581996, 1997.

Kustas, W. P., Norman, J. M., Anderson, M. C., and French, A. N.: Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship, Remote Sens. Environ., 85, 429-440, http://dx.doi.org/10.1016/S0034-4257(03)00036-1, 2003.

Li, Z. L., Tang, B. H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A.: Satellite-derived land surface temperature: Current status and perspectives, Remote Sensing of Environment, 131, 14-37, http://dx.doi.org/10.1016/j.rse.2012.12.008, 2013.

Maayar, E. M., and Chen, J. M.: Spatial scaling of evapotranspiration as affected by heterogeneities in vegetation, topography, and soil texture, Remote Sens. Environ., 102, 33-51, http://dx.doi.org/10.1016/j.rse.2006.01.017, 2006.

McCabe, M. F., and Wood, E. F.: Scale influences on the remote estimation of evapotranspiration using multiple satellite sensors, Remote Sensing of Environment, 105, 271-285, 10.1016/j.rse.2006.07.006, 2006.

Moran, M. S., Humes, K. S., and Pinter Jr, P. J.: The scaling characteristics of remotely-sensed variables for sparsely-vegetated heterogeneous landscapes, Journal of Hydrology, 190, 337-362, http://dx.doi.org/10.1016/S0022-1694(96)03133-2, 1997.

P2L7-8 Should references be in chronological order or alphabetic order?Check throughout the manuscript.

Response: Thanks for your reminder. We have reordered the references in chronological order throughout the manuscript.

P2L10-12 What is the difference between distributed and lumped as discussed here?

Response: Thank you for this thoughtful comment. We have revised this sentence to briefly discuss the difference between distributed and lumped calculations:

To address the scale effect on energy fluxes, many studies have compared distributed calculations with lumped calculations. 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. Distributed calculations and lumped calculations may not be the same at different scales.

P4L3-4 state the resampling methodology that was used and why?

Response: Thanks for your thoughtful comment. We revised these sentences to briefly discuss the difference between distributed and lumped calculations:

Surface thermal dynamics controls energy partitioning and ET. However, the spatial resolution of thermal-infrared (TIR) images is usually not as high as the spatial resolution of visible near-infrared (VNIR) bands because the energy of VNIR photons is higher than the energy of thermal photons (Peng et al., 2016). The IPUS (input parameter upscaling), 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 as the lumped method in this study.

This model was designed to simulate satellites that have identical spatial resolutions in both the visible near-infrared (VNIR) and thermal infrared (TIR) bands and has been described in details in Peng et al., (2016). The energy flux components net radiation (Rn), soil heat flux (G), sensible heat flux (H) and LE are shown as below (Jiao et al., 2014; Peng et al., 2016).

P4L23 Provide the shortened energy balance equation and define all terms.

Response: Thanks for your careful reading of our manuscript. We have added Eq. (4) (P5 L1) to represent that the LE is estimated as the residual term of the surface energy balance equation.

Finally, LE is calculated as a residual item of the energy balance equation (Eq. (4)).

$$LE = R_n - G - H \quad , \tag{4}$$

Further details can be found in Peng et al. (2016).

P4L26 There are other methods which can/should be included especially if you are making reference to "various"

Response: This sentence was adjusted to avoid ambiguity as below:

EF is widely used to estimate the daily ET with RS data in different methods (e.g., the feature space of the Land Surface Temperature and Vegetation Index (LST-VI) (Carlson, 2007; Long and Singh, 2012) and SEBS (Su, 2002) models).

P5L3-4 The descriptions of the methodology presented herein are a critical aspect of the study and should therefore be described more clearly, so as to make the technique repeatable by other researchers. As it stands it is difficult to determine exactly how the values are being derived from just interpreting the equations below. See comments below, which detail critical aspects that require attention to improve the presentation of information.

Response: The description of the method was not sufficiently clear and was slightly misleading before. Section 2.2 (The EF of mixed pixels) was rewritten with an adjusted structure and detailed information. The derivation of the EF equation of mixed pixels and use of the equation are two main parts of this method, and each involves one key hypothesis. The first half of Section 2.2 presents Hypothesis 1 and the equation, and the second half presents how the equation is applied with the aid of Hypothesis 2. The first half is the theory and the second half is the technique.

2.2 EF of mixed pixels

(1) Equation for EF of mixed pixels

The EF is the ratio of LE and available energy (AE) (Rn-G), as follows:

$$EF = \frac{LE}{R_n - G} \ , \tag{5}$$

Studies have shown that the EF is quite stable over time and thus is well suited to denote the status of the surface energy balance for a certain period. For example, the EF is nearly constant during the daytime (Nichols and Cuenca, 2010; Sugita and Brutsaert, 1991) and thus, can be used for temporal scale extrapolation, i.e., from instantaneous LE at the satellite overpassing time to daily ET. EF is widely used to estimate daily ET with RS data in different methods (e.g., the feature space of the Land Surface Temperature and Vegetation Index (LST-VI) (Carlson, 2007; Long and Singh, 2012) and SEBS (Su, 2002) models).

In this section, the EF of mixed pixels is investigated and a novel approach is derived to estimate the daily ET of mixed pixels. In other words, EF is used for temporal scale extrapolation and spatial scale correction of the remotely sensed LE and ET at a coarse-resolution scale at the same time.

Because turbulence transferred by advection is always neglected in RS data, we only consider vertical turbulence. Therefore, the accurate LE (with scaling effects taken into consideration) of a mixed pixel can be weighted by the LE of its sub-pixels as follows:

$$LE = \sum s_i \, LE_i = \sum [s_i \cdot \frac{LE_i}{(Rn-G)_i} \cdot (Rn - G)_i] \ , \tag{6}$$

where LE denotes the accurate LE of mixed pixels, s_i the area fraction (AF) of sub-pixel *i*, and LE_i the LE of sub-pixel *i*. Eq. (5) and (6) can be combined as follows:

$$LE = \sum [s_i \cdot EF_i \cdot (Rn - G)_i] \ , \tag{7}$$

where EF_i and $(Rn\text{-}G)_i$ denote the EF and AE of sub-pixel *i* in a certain mixed pixel respectively.

At this step, a simplification is performed as described in Hypothesis 1:

Here, Hypothesis 1 is proposed as follows:

"The available energy (AE) of each sub-pixel is approximately equal to that of any other sub-pixels in the same mixed pixel within an acceptable margin of errors (e.g. 50 $W \cdot m^{-2}$ (Seguin B et al.,1999; Kustas and Norman, 2000; Sánchez et al., 2007)) and is equivalent to the AE of the mixed pixel."

Therefore, Eq. (7) can be transformed in to the following expression:

$$\widetilde{LE} = [\sum (s_i \cdot EF_i)] \cdot (Rn - G) \ , \tag{8}$$

where $\widetilde{LE}$ denotes the latent heat flux in mixed pixels based on Hypothesis 1. There is a minor difference between $\widetilde{LE}$ and LE that can be regarded as an error of Hypothesis 1, and it will be analysed below.

Rearranging Eq. (8) yields the following:

$$\frac{\widetilde{LE}}{(Rn-G)} = \sum(s_i \cdot EF_i), \quad (9)$$

Therefore, we have

$$\widetilde{EF} = \sum(s_i \cdot EF_i), \quad (10)$$

where $\widetilde{EF}$ denotes the EF of the mixed pixel, including the error of Hypothesis 1, which is quite small and can be neglected based on a data analysis (see Section 4.3.1). Hence Eq. (10) can be used as the solution to the EF of mixed pixels.

Using Eq. (10) makes calculating the EF of mixed pixels is straightforward since it only needs the AF of each land cover in the pixel, which can be easily obtained using a fine-resolution land cover map, as well as the EF_i of its sub-pixels, which requires a specific technique to get in operations.

(2) Calculating EF of mixed pixels

The EF_i of sub-pixels is required in Eq. (10); however, it is not available with coarse resolution data. In order to utilize Eq. (10), Hypothesis 2 is proposed:

"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."

The concept of "nearest" in this study is refers to the shortest distance between the centre point of the mixed pixel where the sub-pixel is located and the centre point(s) of the pure pixel(s) with the same land cover type as the sub-pixel in the study area. The concept is illustrated in Fig. 1.

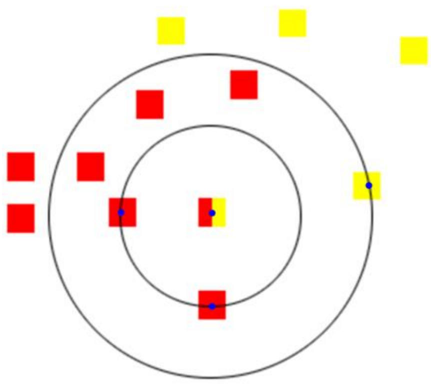


Figure 1: A sample graph of sub-pixel in a mixed pixel and its nearest pure pixel(s) of the same land cover type.

In Fig.1, there are two land cover types: Yellow and Red. The centre pixel, which is a mixed pixel, contains a red sub-pixel and a yellow sub-pixel. Hypothesis 2 examines which pixel the nearest neighbour. From Fig. 1, it is clear that the red sub-pixel has two nearest neighbours (red pixels with blue centres); thus, the EF of the red sub-pixel equals the mean EF of the two nearest pure pixels according to Hypothesis 2. The yellow sub-pixel has one nearest neighbour (yellow pixel with blue centre); thus, the EF of the yellow sub-pixel equals the EF of the yellow pure pixel.

This hypothesis is based on Tobler's First Law (TFL) (Miller, 2004; Li et al., 2007; Tobler, 2004), which states that "everything is related to everything else, but near things are more related than distant things". In other words, the most similar conditions, phenological patterns and physical characteristics exist between a sub-pixel surface and nearby (pure pixel) surfaces given the same land cover. Accordingly, the EF of sub-pixel *i* can be determined using EF of pure pixel(s) at coarse resolution scale based on Hypothesis 2.

Therefore, Eq. (8) may be reduced as above to the following:

$$\widetilde{LE} = (Rn - G) \cdot \widetilde{EF}, \qquad (11)$$

Eq. (10) and Hypothesis 2 together can be used to calculate the EF of mixed pixels and therefore the daily ET. Eq. (10) and (11) can be used together to correct the spatial scale errors of the instantaneous LE at the overpassing time.

In summary, by employing two key hypothesises, EFAF methodology is able to realize temporal scale extrapolation and spatial scale correction for remotely sensed LE and ET at a coarse resolution scale at the same time. The EF of a mixed pixel is expressed as the area-weighted EF_i of its sub-pixels with acceptable simplifications, which simplified the calculations, increased the accuracy, and facilitated its use for daily operations.

P5L5-6 Overall, I agree with this statement but I feel that this still needs to be framed within context i.e. capturing data at coarse resolutions and using this for localized applications may not be the most appropriate decision due to the effects that you have described. This needs to be stated in the introductory section and then reemphasized by stating what spatial scale the data is being verified at.

Response: This sentence was moved to the introduction section as general background, and the following information was given here as specific details on the method.

In this section, the EF of mixed pixels is investigated and a novel approach is derived to estimate the daily ET of mixed pixels. In other word, EF is used for temporal scale extrapolation and spatial scale correction of the remotely sensed LE and ET at a coarse resolution scale at the same time.

P5L13 The description herein of the methodology is a bit vague. Is the EF and AE of the sub-pixel calculated using finer resolution imagery? Essentially are you obtaining flux and EF estimates using finer resolution imagery and then determining the proportional contribution of these values to the flux and EF estimates obtained at the coarser resolution

Response: As mentioned in the response to P5L3-4, this is the first half of the methods and only includes the theory and method for deriving the equation. As explained in the second half, the use of the equation does not require the calculation of EF at a finer resolution. Rather, it needs the EF of pure pixels at a coarse resolution instead (see the revised Section 2.2 for details).

P5L16 Which is?

Response: The term “acceptable margin errors” refers to minor differences in available energy (AE) among the subpixels in the same mixed pixel. We assume that the available energy (AE) of each sub-pixel is approximately equal to that of any other sub-pixel in the same mixed pixel. In fact, there are minor differences among them; however, these differences are within an acceptable margin according to previous research (Seguin B et al.,1999; Kustas andJ.M. Sánchez et al., 2007). More discuss details are presented in Section 4.3.1.

P5L24 Similar to previous comment, as it stands this is too vague, is the EF of the sub-pixel for a finer resolution and then being compared to the EF of a pure pixel at the coarser resolution.

Response: The use of the equation does not require the calculation of EF at a finer resolution;

rather, it needs EF of pure pixels at a coarse resolution instead. The paragraph sequence here is incorrect and was rewritten. (see the revised Section 2.2 for details).

P6L4-6 Essentially the EFAF methodology is predicated on a combination of these hypotheses?

Response: This paragraph was deleted because its meaning was vague and not necessary. The following information was given as a summary of this section.

In summary, by employing two key hypothesises, EFAF methodology is able to realize temporal scale extrapolation and spatial scale correction for remotely sensed LE and ET at a coarse resolution scale at the same time. The EF of a mixed pixel is expressed as the area-weighted EF_i of its sub-pixels with acceptable simplifications, which simplified the calculations, increased the accuracy, and facilitated its use for daily operations.

P8L2 Define what are pure pixels and mixed pixels in the context of this study. Based on Figure 3, I would assume a pure pixel, is a pixel at the 300 m resolution which is entirely made up of 1 particular land cover class mapped at the 30 m resolution?

Response: Thank you for this suggestion. We have added the following statement:

The pure pixels at 300 m scale are entirely made up of one particular land cover type, and the mixed pixels are made up of two or more land cover types according to the land cover datasets with a spatial resolution of 30 m.

P8L6 This map is taken from Peng et al., 2016 and should therefore be referenced accordingly.

Response: Thanks for your kind reminder. The reference was added here as (Peng et al., 2016):

Peng, Z., Xin, X., Jiao, J. J., Zhou, T., and Liu, Q.: Remote sensing algorithm for surface evapotranspiration considering landscape and statistical effects on mixed pixels, Hydrology & Earth System Sciences, 20, 4409-4438, 2016.

P9L10 clear-sky or cloud free

Response: This has been revised as "clear-sky".

P9L11 what threshold was used here to decide this? for example less than 20 % cloud coverage over the study area or within the image?

Response: Thanks for your thoughtful suggestion. This sentence was rewritten as follows: the satellite data selected for this study were collected under clear or partly cloudy conditions based on data quality metrics and artificial visual interpretation. We combined data quality information with visual interpretation to select satellite images in this study, and quantity cloud detection was not performed.

P10L6-7 It may prove to be more beneficial to move this information further up within the section before presenting the land cover map for the study area.

Response: Thanks for your suggestion. This sentence was moved to the prior section presenting the land cover map for the study area immediately after the sentence beginning "The percentage of the numbers of land cover types".

The percentage of the number of land cover types (Yu et al., 2016) (Fig. 3) for the study area were extracted at a 300 m scale with 30 m land cover classifications, which were developed by Zhong et al.

(2014a) based on HJ-1/CCD time series.

P10L21-22 regarding system setup?

Response: Thanks for your reminder. This sentence was rewritten as follows:

The EC data were based on 30 min intervals; additional information regarding the system setup, data processing and quality control can be found in previous reports (Liu et al., 2011; Liu et al., 2016; Xu et al., 2013)

P10L22 Is there any particular reason for including a description of how these sensors were setup and excluding descriptions for the other sensors?

Response: Thanks for your reminder. This sentence was deleted since it is included in the above references and not necessary here.

P10L23-24 This also involves the use of soil temperature averaging probes and volumetric water content sensors

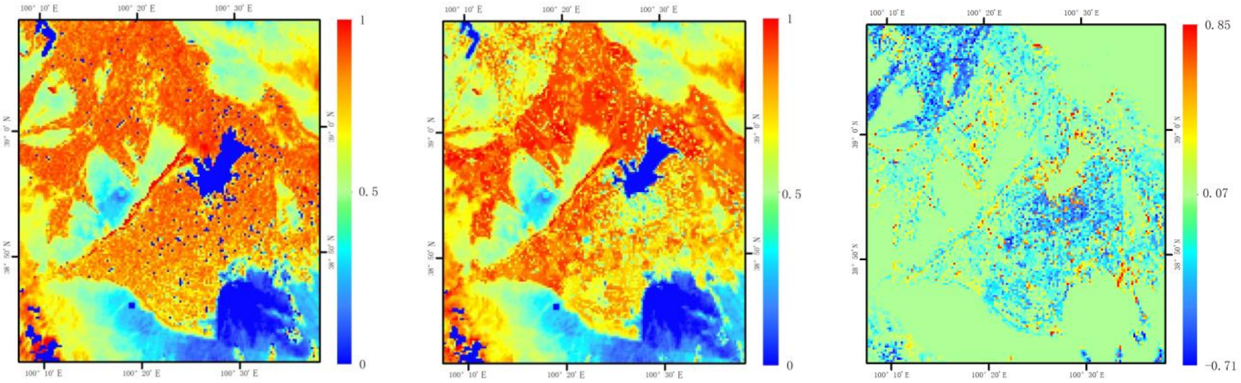
Response: Thanks for your reminder. This sentence was deleted since it is included in the above references and not necessary here.

P11L4 Is it possible to provide the footprint of the measurements?

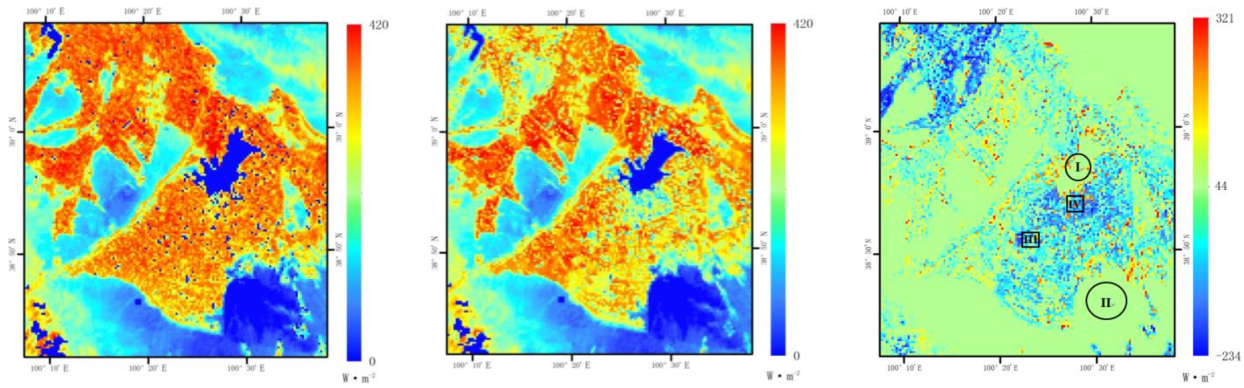
Response: Thanks for your suggestion. However, in our opinion, the footprint is a function of many variables, such as the tower height, wind speed and wind direction. Therefore, each site has different footprints on each day. If we want to provide the footprint of the measurements, a number of figures would be displayed in the manuscript. For the overall arrangement and the emphasis of the manuscript, we do not suggest providing the footprint of the measurements.

P13L4-7 It might prove to be useful to highlight these areas on the EF and LE maps as well.

Response: Thank you for this good suggestion. We have added the difference between lumped and EFAF (EF/LE) in Figures 5 and 6 and highlighted these areas in Figures 5 and 6 showing the difference between lumped and EFAF LE.

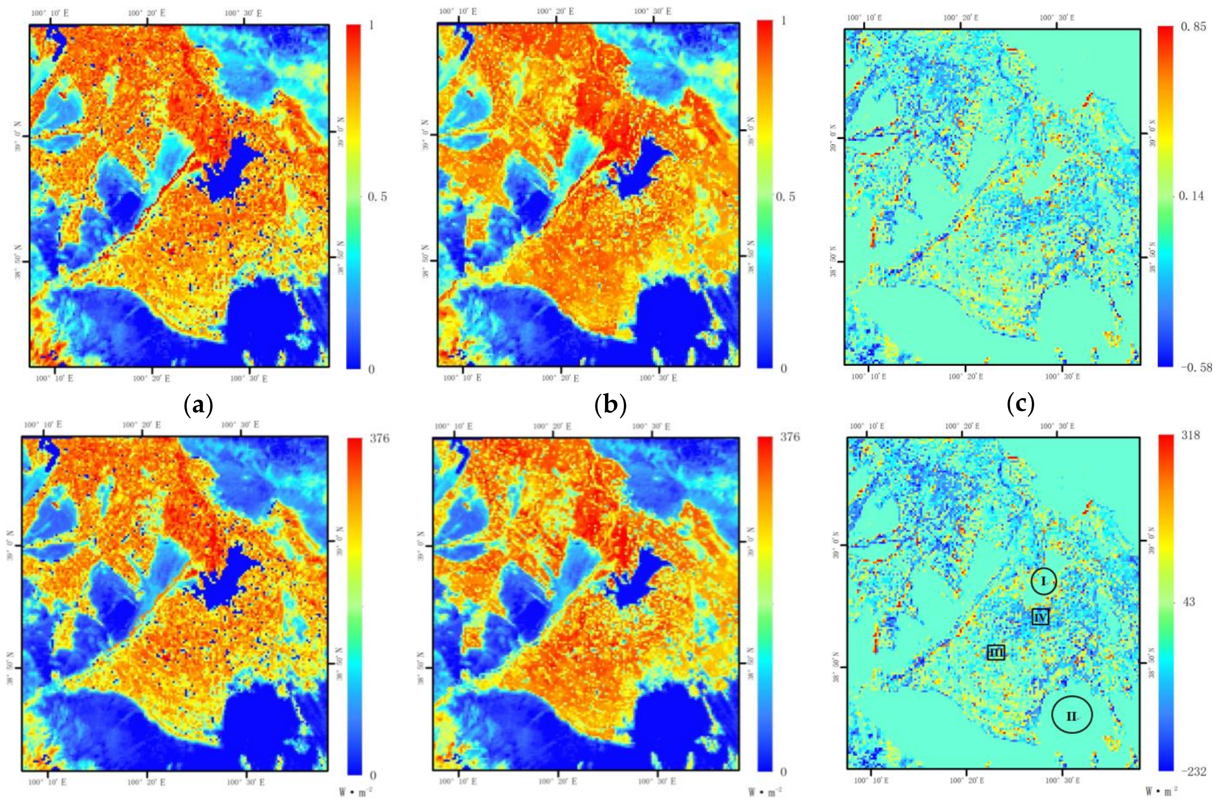


(a) (b) (c)



(d) (e) (f)

Figure 5. Maps of (a) lumped EF, (b) EFAF EF, (c) difference between EFAF and lumped EF (EFAF EF minus lumped EF), (d) lumped daily LE, (e) EFAF daily LE and (f) difference between EFAF and lumped LE (EFAF LE minus lumped LE) on July 8th, 2012



(d) (e) (f)

Figure 6. Maps of (a) lumped EF, (b) EFAF EF, (c) difference between EFAF and lumped EF (EFAF EF minus lumped EF), (d) lumped daily LE, (e) EFAF daily LE and (f) difference between EFAF and lumped LE (EFAF LE minus lumped LE) on August 22nd, 2012

P15L6 Just a suggestion to consider...EFAF was applied to improve LE and ET estimates through corrections to EF. However, this will also influence the sensible heat (H). It may prove to be beneficial to the study to also demonstrate how Lumped H and EFAF H compare against in-situ observations.

Response: Thanks for your good suggestion. The focus of this study is to discuss the scale effect of LE (or ET); therefore, we proposed the model named EFAF. Using this model to determine the influence on the sensible heat (H) will be discussed in future research.

P15L9 It may also be useful to perform some form of significance testing to demonstrate that EFAF method has improved the flux estimates to an extent that there is no longer any significant differences between the observed and simulated values.

Response: Thanks for your suggestion. A significance test is a good method of demonstrating that there are no longer any significant differences between the observed and retrieved values. However, in this study, fewer samples are provided due to the limitation on the number of observed values; therefore, identifying differences as statistically significant is difficult.

To better demonstrate that the EFAF method has improved the flux estimates, we added the decrease in the error percentage between the observed and retrieved values:

The correction effect of the EFAF method was most distinct at the EC04 site, and the RMSE at EC04 decreased from 5.36 to 2.72 MJ m^{-2} (about decreased by approximately 49.25%); this improvement stemmed from the fact that EC04 had the highest complexity of all sites. Maize-dominated pixels in EC04 included maize, vegetables, buildings and bare soil, at a ratio of 53:26:19:2, respectively. We conclude that maize and vegetables were land cover types with a high EF, while bare soil had a low EF. For buildings, the EF value was 0 in this study. Similarly, the difference of them againstbetween these estimates and the EC measurements had also declined from 4.12 MJ m^{-2} to 2.32 MJ m^{-2} (decreased by approximately 43.3%).

P16L1-2 This is the point that I was alluding to earlier. See my comment in the introductory section about bias at larger spatial scales.

Response: Thanks for your suggestion. As explained earlier, the spatial scale error of remote sensing estimation is not caused by scale mismatches between RS data and the EC footprint (see response to P2L3-6). However, validating the estimation could be affected by this scale mismatch. The point here is that even after spatial scale errors are corrected (partially) by the EFAF method, a validation still must to be performed at the same scale for estimation and field measurement. Otherwise, uncertainties will remain in the validation results, which explains why the footprints of the EC measurements were calculated and used in the validation. To clarify the footprint and scale match technique, the footprint results were provided before use (see response to P11L4).

P20L9-10 Any reference to support this observation?

Response: Thanks for your careful reading of our manuscript. We have added the relevant references here.

We consider these biases to be acceptable (Seguin B et al., 1999; Kustas and Norman, 2000; Sánchez et al., 2007).

P26L5-7 The method presented herein describes a novel approach to improve the accuracy of daily ET estimates when using coarser spatial resolution data. However I would like to see a discussion within the "discussion section" or a recommendation for future study in the "conclusion section"describing how this technique can be used to assist improving the accuracy of daily (everyday for a period of time) ET estimates when using coarse resolution spatial data. From my understanding the technique was only applied on days in which concurrent coarse and finer resolution data was available. Coarser resolution imagery such as MODIS images are often used for season and inter-annual assessments due to their high temporal resolution. However, finer resolution imagery is not available on a daily time step. Subsequently, the question I am putting forward is, can this approach be applied on a daily basis to improve the accuracy of ET estimates obtained using coarser resolution imagery and if so, how?

Response: Thanks for your kind suggestion. This information was added at the end of the conclusion section and can answer the question.

In brief, the estimated LE of pure pixels is considered accurate and used to calculate its EF. Based on this parameter, the equation for the EF of mixed pixels was established with two key hypothesises. A finer resolution land cover map is needed to search for “pure pixels” as well as to calculate area ratio of each land cover in mixed pixels. This process can derive the daily ET from coarse resolution remote sensing data with acceptable accuracy, and no other finer resolution data are needed in the EFAF method. Thus, this method may be applicable on a daily basis with daily coarse resolution imagery, such as MODIS, and only one finer resolution land cover map for a certain length of time, i.e., a week, month or season, as long as the land cover change is not extreme in that period. It is quite convenient for regional applications that need long-term running. This method can also be used as a correcting technique for LE estimations or remote sensing products since calculating the EF of mixed pixels is carried out after calculating heat fluxes that could be based on an energy balance equation or other methods at the very beginning. The application of the EFAF could be limited with very coarse resolution data since the probability of pure pixels becomes very low. In these circumstances, a compromise may have to be made between the “purity” of pure pixels and the searching distance for the pure pixels. Additional investigations are needed to evaluate the performance of this method with different remote sensing products.

Anonymous Referee #2

Received and published: 5 September 2018

This document contains the following: (1) comments from referees, (2) responses from the authors (in blue), and (3) authors' changes to the manuscript (in red).

General comments:

1) I find it strange that no reference is made to the only widely accepted disaggregation method currently producing high resolution ET: ALEXI/DISAlexi.

Response: Thanks for your kind reminder. This manuscript has been modified several times before submission, and this mistake must have been missed by the authors during this process. The references were added to the introduction section as below:

Classical satellite-based models such as the Surface Energy Balance Algorithm for Land (SEBAL) (Bastiaanssen et al., 1998), Surface Energy Balance System (SEBS) (Su, 2002), Atmosphere-Land Exchange Inverse (ALEXI) and an associated flux disaggregation technique (DisALEXI) (Anderson et al., 2011; Anderson et al., 2012),and the temperature-sharpening and flux aggregation scheme (TSFA) (Peng et al., 2016) have been developed to monitor land-atmosphere energy balance flux interactions.

2) The biggest concern I have with your approach is that your approach specifies two hypothesis that are used to upscale. However in this there is (in my opinion) two serious flaw: a. An underlying

assumption (that is not specified) is that the evaporative fraction at coarse resolution is correct. Considering that this evaporative fraction was determined over a coarse resolution (without considering subsurface heterogeneity) in the first place. As such oasis effects are not taken into account and can result in serious errors.

Response: Thanks you for this question. It was very helpful for the authors to interpret the method more clearly and improve its representation in the manuscript.

In this method, the LE and EF of pure pixels at coarse resolution were regarded as accurate and then used to calculate the EF of mixed pixels through the area fraction of each land cover in the mixed pixel and the corresponding EF of each land cover, which is represented by the EF of the nearest pure pixel with the same land cover. Therefore, the subsurface heterogeneity in the mixed pixels is considered while subsurface heterogeneity in pure pixels is ignored since the pixel is "pure", which is the underlying assumption and starting point of this proposed method.

The pure pixel is defined as the pixel with only one land cover type inside the pixel. This underlying assumption is acceptable and sufficient for the purpose of this study since the mixture of different land cover types is the most significant heterogeneity (Blyth and Harding, 1995; Bonan et al., 2002; McCabe and Wood, 2006; Moran et al., 1997; Peng et al., 2016) and should be considered initially. One may argue that a better method is available for defining pure pixels by using both land cover and other surface variables.

However, in our opinion, such a method may help to obtain purer pixels but will not help to obtain a better ET estimation since the probability of finding proper pure pixels for each land cover in mixed pixels becomes extremely low and reduces the applicability of the method. The following paragraph was added at the end of section "5 Discussion".

(4) The underlying assumption and starting point of this method is that the pure pixel is really the actual "purity" of the pure pixels; therefore, the EF of pure pixels is representative at least to surrounding the mixed pixels. Only land cover information was used to define pure pixels; therefore, subsurface heterogeneity in pure pixels caused by other aspects (such as variations in the surface variables) may have certain influences on the results. Including additional features in the definition of pure pixels may increase the complexity of the model and the difficulties of its application significantly.

b. Secondly, hypothesis 1 (having EFi = EF) only is valid for incoming radiation (optical and thermal). However considering that the outgoing radiation depends on LST, albedo and emissivity (each with greatly varying heterogeneity) this cannot be said for the net radiation consequently on the available energy. While for many agricultural site's the application might hold true, it cannot be stated as an overarching law. While this is kind of reflected in the text (as you change denotation from LE to LE_), the is not further touched upon at all.

Response: As noted, Hypothesis 1 (i.e., EFi = EF) is carefully used throughout the manuscript and its possible error was indicated in the first instance of its use ($\widetilde{LE}$ denotes the latent heat flux in mixed pixels based on Hypothesis 1). A section (4.3.**1 Error analysis of Hypothesis 1)** has been included to discuss this hypothesis and the errors. As shown in the manuscript, the errors are small (less than 7 W m^{-2}) and the hypothesis is acceptable.

c. While for hypothesis 2 at least some justification is provided (though one can argue what objectively is specified as ‘near’, no justification/argumentation for the 1st hypothesis is given.

Response: A section (4.3.1 **Error analysis of Hypothesis 1)** has been included to discuss Hypothesis 1 and its errors. As shown in the manuscript, the errors are small (less than 7 W m^{-2}) and the hypothesis is acceptable.

Hypothesis 2 was carried out based on Tobler's First Law (TFL): everything is related to everything else, but near things are more related than distant things. The term “near” refers to spatial distance in this hypothesis; it was added in the Section 2.2 to define the concept of “near” in the method. As below:

The concept of “nearest” in this study is refers to the shortest distance between the centre point of the mixed pixel where the sub-pixel is located and the centre point(s) of the pure pixel(s) with the same land cover type as the sub-pixel in the study area. The concept is illustrated in Fig. 1.

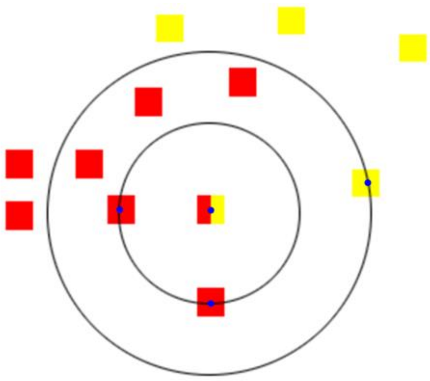


Figure 1**: A sample graph of sub-pixel in a mixed pixel and its nearest pure pixel(s) of the same land cover type.**

In Fig.1, there are two land cover types: Yellow and Red. The centre pixel, which is a mixed pixel, contains a red sub-pixel and a yellow sub-pixel. Hypothesis 2 examines which pixel the nearest neighbour. From Fig. 1, it is clear that the red sub-pixel has two nearest neighbours (red pixels with blue centres); thus, the EF of the red sub-pixel equals the mean EF of the two nearest pure pixels according to Hypothesis 2. The yellow sub-pixel has one nearest neighbour (yellow pixel with blue centre); thus, the EF of the yellow sub-pixel equals the EF of the yellow pure pixel.

d. Finally, at 30m resolution horizontal transport is becoming much more important (as you yourself indicate when considering EC footprints).

Response: Advection and its influences are not considered in this study because they are not the main concern of this work, and addressing two types of problems at the same time would be excessively complex. This concerns appears to focus on a mismatch in physics if we use field measurements contaminated by horizontal transport (such as oasis effects) to validate results without advection effects. The authors are aware of this risk and have removed the data contaminated by advection (a threshold “$H+LE>R_n+G$” is used to find advection effects) from the validation dataset. As for considering advection in the model calculation, to my knowledge, such a process remains a huge challenge in the remote sensing of heat fluxes.

Specific comments:

1) These shortcomings are reflected in that for EC4 (your most successful disaggregation site) still an error (2.7MJ) a factor 2 above any of your homogeneous sites (EC2,6,12 and 14) (each with errors below 1.2 MJ). This however is not touched upon in the text.

Response: Thank you for this kind reminder. The validation results shown in Table 3 and Figure 7 have confirmed the success of the proposed method in correcting the spatial scale error and estimating accurate daily ET from coarse resolution data. The outcome of EC4 is the most successful example when considering the relative RMSE change (RMSE decreased nearly 50%). However, additional errors are observed for the corrected ET compared with the homogeneous sites, and the possible reasons for these errors have been analyzed in lines 16-22 on page 16. The remaining larger errors in such pixels represent a reminder that this method has limitations in extreme conditions. More complex models should be built for such circumstances and more information other than land cover should be included when considering subsurface heterogeneity in order to obtain results that are as accurate as those for homogeneous sites. This ambitious goal will be the focus of future studies in spatial scale issues in the remote sensing of LE and ET.

These statements have been revised as follows:

The correction effect of the EFAF method was most distinct at the EC04 site, and the RMSE at EC04 decreased from 5.36 to 2.72 MJ m^{-2} (2.15 to 1.09 mm) (decreased by approximately 49.25%); this improvement stemmed from the fact that EC04 had the highest complexity of all sites.

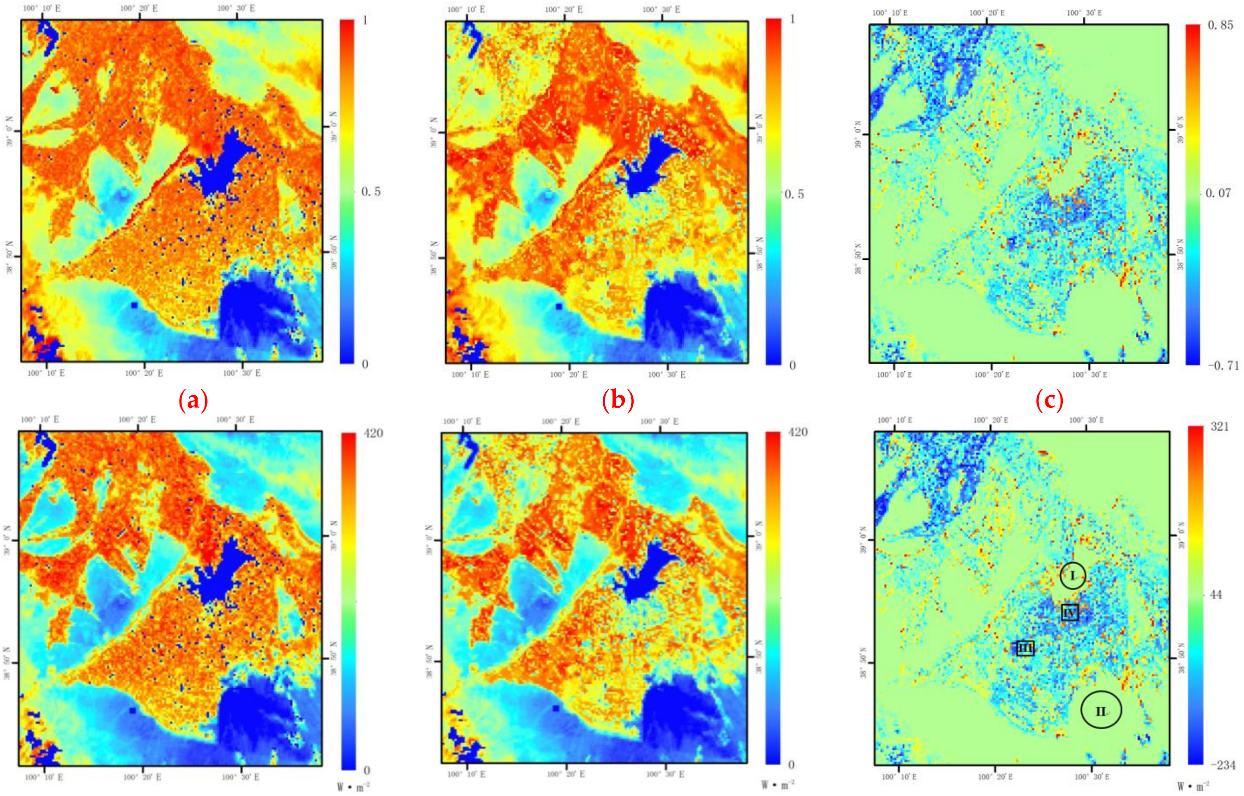
Maize-dominated pixels in EC04 included maize, vegetables, buildings and bare soil, at a ratio of 53:26:19:2, respectively. We conclude that maize and vegetables were land cover types with a high EF, while bare soil had a low EF. For buildings, the EF value was 0 in this study. For example, on 30 June, the EF of mixed pixels in EC04 was 0.81. However, the average EF values of the pure pixels positioned closest to maize and vegetables among the sub-pixels were 0.88 and 0.88, respectively and that of bare soil was 0.65. Therefore, when scale effects were taken into consideration, the EF of the mixed pixels was 0.70. Using the EFAF method, the daily LE of the mixed pixel where EC04 was located decreased from 13.57 to 11.78 MJ m^{-2} (5.45 to 4.73 mm). Similarly, the difference between these estimates and the EC measurements also declined from 4.12 MJ m^{-2} to 2.32 MJ m^{-2} (1.67 to 0.93 mm) (decreased by approximately 43.3%). Additionally, there were large discrepancies between the observed and retrieved LE values at EC04. Specifically, there are two points far from the 1:1 line in Fig. 8 (d), with values of 8.36 MJ m^{-2} (3.36 mm) on 27 July and 9.33 MJ m^{-2} (3.75 mm) on 3 August. Even after the EFAF method was applied, these values were 5.20 MJ m^{-2} (2.09 mm) and 4.59 MJ m^{-2} (1.84 mm), respectively, because EC04 was positioned in a maize-dominated pixel and the EC tower was located in a built-up area, thus generating errors associated with temperature retrieval that would create further errors in estimating Rn. For example, on 27 July and 3 August, the Rn observed by AWS for the EC station was 15.95 and 15.35 MJ m^{-2}, respectively, while the retrieved Rn of the pixels was 18.14 and 18.80 MJ m^{-2}, respectively. The remaining larger errors in such pixels are a reminder that this method has limitations under certain extreme conditions. More complex models should be built for such circumstances and more information other than land cover should be included when considering subsurface heterogeneity to obtain results that are as accurate as those obtained for homogeneous sites.

Technical comments:

1) Specifically figure 4 and 5. Here you want to show the difference between Lumped and EFAF (LE/EF) next to each other. In my view this could be better shown by 1 graph of Lumped LE/EF, and a 2nd showing the difference between Lumped and EFAF (LE/EF). At present the colouring of the maps

hide where specific improvements are made.

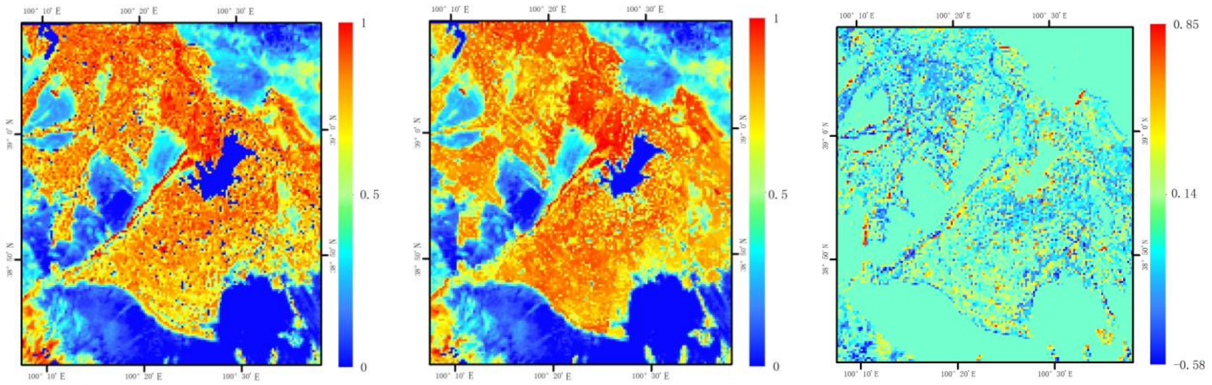
Response: Thanks for your good suggestion. We have shown the difference between lumped and EFAF (EF/LE) in Figures 5 and 6.



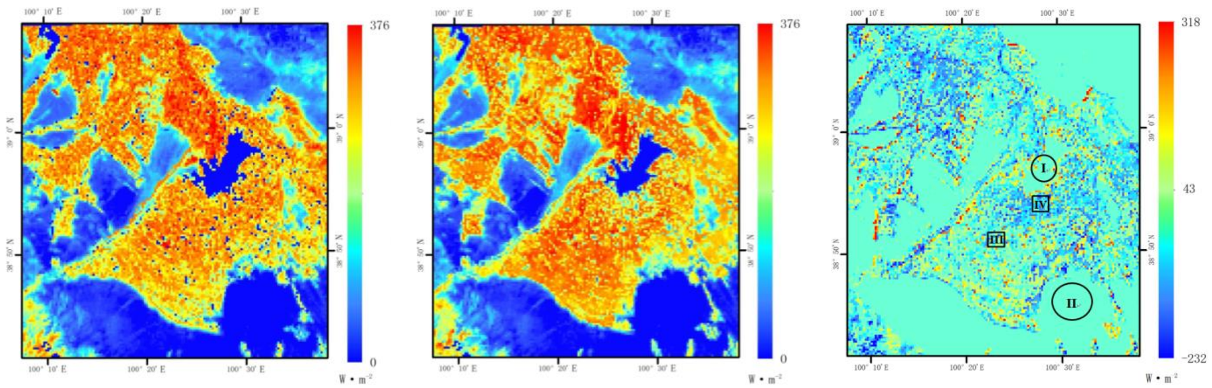
(a) (b) (c)

(d) (e) (f)

Figure 5. Maps of (a) lumped EF, (b) EFAF EF, (c) difference between EFAF and lumped EF (EFAF EF minus lumped EF), (d) lumped daily LE, (e) EFAF daily LE and (f) difference between EFAF and lumped LE (EFAF LE minus lumped LE) on July 8th, 2012



(a) (b) (c)



(d) (e) (f)

Figure 6. Maps of (a) lumped EF, (b) EFAF EF, (c) difference between EFAF and lumped EF (EFAF EF minus lumped EF), (d) lumped daily LE, (e) EFAF daily LE and (f) difference between EFAF and lumped LE (EFAF LE minus lumped LE) on August 22nd, 2012

2) You denote the validation-results in MJ/m2 instead of the customary mm/day. While this is simply a division by the latent heat of vaporization, denoting it in these units prohibits the comparison with other validation researches.

Response: Thanks for your careful reading of our manuscript. Accurate latent heat of vaporization is a function of temperature. If biases occur in the temperature measurement, the latent heat of vaporization and ET values will be affected. In this study, temperature values are external inputs; therefore, we cannot distinguish the errors from the EFAF retrieved or the temperature products. Hence, we use a constant value of latent heat of vaporization of approximately 2.49×10^6 $W\ m^{-2}\ mm^{-1}$ (Pan and Liu, 2003) to approximately calculate the ET as mm. Thus, in our opinion, the validation results in MJ/m^2 would be accurate compared with other validation results, and we have added results in mm as an auxiliary comparison.

Pan, Z., and Liu, G.: Evapotranspiration Research of Yellow River Delta Using Remote Sensing Method, Geo-information Science, 3, 91-96, 2003.

3) Also in the start of the manuscript you refer to results of intercomparison studies as 'biases' (while they should have been called errors/uncertainties), while you specify (in fiture 4.3) errors which cannot be qualified as such (as they do not refer to a comparison between ground measurements and retrieval), but instead are just variances of a single map.

Response: Thanks for your careful reading of our manuscript. For Section 4.3, we analyzed the approximate errors of two key hypotheses.

Section 4.3.1discusses the approximate errors of Hypothesis 1, which states that the available energy (AE) of each sub-pixel is approximately equal to that of any other sub-pixels in the same mixed pixel within an acceptable margin of bias and is equivalent to the AE of the mixed pixel. Therefore, the pixel values of a lumped 300 m resolution should be compared to the 10×10 set of 30 m pixels that they were drawn from in this study. However, obtaining simultaneous ground measurements of AE in 10×10 set of 30 m samples and a 300 m sample is difficult. Even if one or two datasets are obtained, generating the representativeness of the whole study area is difficult. Therefore, we consider the distributed retrieved values at a 30m resolution as accurate values and compare them with the 30 m resolution sub-pixel values, which have the same values as the lumped AE measured at a 300 m resolution from each mixed pixel. This method is relatively better for analysing the errors of Hypothesis 1 throughout the whole study area.

Section 4.3.2 discusses the approximate errors of Hypothesis 2, which 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. As for ground measurements of EF, the ground measurements of Rn, G and LE are required. The ground measurements of Rn and G are point-based observations, and those of LE are region-based observations because the footprints of EC measurements are considered. Therefore, using the ground measurements of AE (Rn-G) and LE to form a consistent spatial representation is difficult. For Hypothesis 2, the inherent significance is the use of the EF for each pure pixel as the correct value. Therefore, we can determine the approximate error caused by Hypothesis 2 by discussing the difference between the two nearest correct values, i.e., the EF of the two nearest pure pixels.

Above all, the errors of the two hypotheses are difficult to analyze by comparing retrievals with ground measurements. By analyzing the errors of the whole study area, we can better explain the rationality of the two hypotheses.

We have revised the word “bias” to "error" or "uncertainties" in the manuscript.

Estimating Daily Evapotranspiration Based on a Model of Evaporative Fraction (EF) for Mixed Pixels

Fugen Li[1,2], Xiaozhou Xin[1,*], Zhiqing Peng[1,2], Qinhuo Liu [1, 3]

[1] State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Beijing, 100101, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] Joint Center for Global Change Studies (JCGCS), Beijing 100875, China

Correspondence to: Xiaozhou Xin (xin_xzh@163.com)

Abstract. Currently, applications of remote sensing evapotranspiration (ET) products are limited by the coarse resolution of satellite remote sensing data caused by land surface heterogeneities and the temporal scale extrapolation of the instantaneous latent heat flux (LE) based on satellite overpass time. 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 accomplish this goal, we derive an equation for calculating the EF of mixed pixels based on two key hypotheses. Hypothesis 1 states that the available energy (AE) of each sub-pixel is approximately equal to that of any other sub-pixels in the same mixed pixel within an acceptable margin of error and is equivalent to the AE of the mixed pixel. This approach simplifies the equation, and uncertainties and errors related to the estimated ET values are minor. Hypothesis 2 states that the EF of each sub-pixel is equal to that of the nearest pure pixel(s) of the same land cover type. This equation is designed to correct spatial scale errors for the EF of mixed pixels; it can be used to calculate daily ET from daily AE data. The model was applied to an artificial oasis located in the midstream area of the Heihe River using HJ-1B satellite data with a 300 m resolution. Results generated before and after making corrections were compared and validated using site data from eddy covariance systems. The results show that the new model can significantly improve the accuracy of daily ET estimates relative to the lumped method; the coefficient of determination (R^2) increased to 0.82 from 0.62, the root mean square error (RMSE) decreased to 1.60 from 2.47 $MJ\ m^{-2}$ (decreased approximately to 0.64 mm from 0.99 mm), and the mean bias error (MBE) decreased from 1.92 to 1.18 $MJ\ m^{-2}$ (decreased from approximately 0.77 mm to 0.47 mm).

Index Terms: Evapotranspiration; Heterogeneous surface; Temporal scale extrapolation; Evaporative fraction; Area weighting

1 Introduction

Large-scale remotely sensed evapotranspiration (ET) estimates generally have a resolution that is too coarse for use in critical applications (e.g., drought assessment, water management or agricultural monitoring) (McCabe et al., 2017). Classical satellite-based models such as the Surface Energy Balance Algorithm for Land (SEBAL) (Bastiaanssen et al., 1998), Surface Energy Balance System (SEBS) (Su, 2002), Atmosphere-Land Exchange Inverse (ALEXI) and an associated flux disaggregation technique (DisALEXI) (Anderson et al., 2011; Anderson et al., 2012), and temperature-sharpening and flux aggregation scheme (TSFA) (Peng et al., 2016), have been developed to monitor land-atmosphere energy balance flux interactions; and in most cases, spatially variable inputs and parameters are based on assumptions of homogeneity of land and atmospheric surfaces (Sharma et al., 2015). However, surface characteristics such as land cover types, land surface temperatures, surface albedo values, downward shortwave radiation and other factors are spatially discrete. Studies have shown that different landscapes (Blyth and Harding, 1995; Bonan et al., 2002; McCabe and Wood, 2006; Moran et al., 1997) and subpixel variations of surface variables, such as stomatal conductance (Bin and Roni, 1994), or leaf area index (Bonan et al., 1993; Maayar and Chen, 2006), can cause errors in heat flux estimations. Models 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). The spatial scale errors in remotely sensed ET (and other parameters inversed from remote sensing data) are primarily arise from the combination of two factors, i.e., nonlinear models and surface heterogeneity, which is more likely to take place in coarser resolution data (Garrigues et al., 2006; Gottschalk et al., 1999;Hu and Islam, 1997; Jin et al., 2007;Z. L. Li et al., 2013; McCabe and Wood, 2006; Tian et al., 2002; Xin et al., 2012). To address the scale effect on energy fluxes, many studies have compared distributed calculations with lumped calculations. 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. Distributed calculations and lumped calculations may not be the same at different scales. Thus, their differences can be considered scale effects. Other studies have noted discrepancies between multi-sensor data aggregations. Moran et al. (1997) found a significant error of over 50% in sensible heat estimations of mixed pixels by comparing lumped and distributed surface fluxes for semi-arid rangeland in Arizona. Hong et al. (2009) found that peak values of ET at the pixel scale increased by 10%-25% following the up-scaling of surface fluxes retrieved by SEBAL from Landsat ETM+ at a 30 m resolution to MODIS at 250-, 500- and 1000-m resolutions. Ershadi et al. (2013) reported that input aggregation underestimated ET at the satellite image scale, with up to 15% fewer retrievals, and at the pixel scale by up to 50% relative to using an original fine resolution Landsat image. These results suggest that the spatial characteristics obtained from data of a specific resolution can only reflect characteristics observed at that resolution. For the heterogeneity of the geo-surface, RS data can synthetically reflect surface information. However, regardless of the spatial resolution, RS data inevitably neglect certain details due to the individual value of each pixel. Moreover, for fine resolution data, the process of up-scaling during smoothing inevitably results in the loss of geo-surface information, reducing the heterogeneity and leading to scale effects. Thus, at the pixel scale, determining whether the physical mechanism is suitable for application, identifying the applicable conditions and determining how to correct the scale effects are the three critical issues for remotely sensed ET estimates (Li et al., 2013).

Some studies have shown that the presence of different land cover types among sub-pixels can generate greater errors in surface flux (Moran et al., 1997; Kimball et al., 1999). Blyth and Harding (1995) proposed a patch model for estimating ET weighted by the area fraction (AF) of soil and vegetation at the pixel scale; the model hypothesizes that the heat transfer process involves significant levels of horizontal fluxes and that interactions among patches can be disregarded. This model structure and is relatively simple and has been widely used to map ET on a large scale (Norman et al., 1995) considering the contributions of surface fluxes from different components (vegetation and soil). However, such models only identify vegetation and soil when estimating ET and do not consider contributions from other land cover types (e.g., water bodies, buildings and snow) or vegetation types (e.g., trees, grasses and crops). 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). Maayar and Chen (2006) proposed an empirical algorithm that uses sub-pixel information on the spatial variability of leaf area index (LAI), land cover and surface topography to correct ET estimates at coarse spatial resolutions. However, an obvious weakness of this approach is that the coefficients must be adjusted for different models and study areas, which limits its applicability. Other studies that combine coarse resolution parameters with land cover maps have used different schemes for different land cover types to estimate ET at the regional scale (Hu and Jia, 2015; Mu et al., 2006; Mu et al., 2011; Peng et al., 2016). However, at the pixel scale, the low calculation efficiency of this method limits its application at a larger scale because the ET of each pixel must be estimated using sophisticated algorithms. Moreover, this method presents difficulties accurately describing surface information due to the coarse resolution of land cover maps.

Each of the above approaches reduces the error in ET estimates based on spatial disparities rather than both spatial and temporal disparities. Temporal scale extrapolation of instantaneous latent heat flux (LE) from satellite overpass time to daily ET is also crucial for applications of RS products. At present, the major temporal scale extrapolation methods include the method based on incoming solar radiation (Jackson et al., 1983; Zhang and Lemeur, 1995), the evaporative fraction (EF) method (Nichols and Cuenca, 2010; Sugita and Brutsaert, 1991) and the reference evaporative fraction method (Allen et al., 2007a; Allen et al., 2007b). The method based on incoming solar radiation uses a sine function to connect the instantaneous ET with the 24-hour trend in solar radiation, with the function expressing the relationship between instantaneous ET and daily ET. The EF method, which is the most widely used, extrapolates the instantaneous EF to the daily EF based on the characteristics of EF that remain constant over one day. The reference evaporative fraction method assumes that the instantaneous reference evaporative fraction which is calculated as the ratio of the computed instantaneous ET at the satellite overpass time from each pixel to the reference crop's (such as alfalfa's) ET, is the same as the average reference evaporative fraction over the 24 h average, and it then uses the reference crop's accumulated daily ET to obtain the daily ET. Chavez et al. (2008) compared different ET temporal scale extrapolation methods and found that the EF method generates values that are most consistent with the measured values.

Therefore, we propose a simple but efficient model (EFAF) to estimate the daily ET of mixed pixels. In this method, the daily ET of the heterogeneous land surface is estimated by calculating the EF of mixed pixels, and it only requires the area fraction (AF) of sub-pixels, which can be obtained from a high-resolution land-cover type map. The model was applied to an artificial oasis in the midstream of the Heihe River. HJ-1B satellite data were used to estimate the lumped fluxes at the scale of 300 m after resampling the 30 m resolution datasets to 300 m resolution, which was used to perform the key step of the model, i.e., correction of mixed-pixel EF and calculation of daily ET. Next, the EF of each pixel at a 300 m resolution was calculated using 300 m net radiation, soil heat flux, sensible heat flux and LE data at the satellite overpass time. The daily ET of the mixed pixels was retrieved from the EF of the mixed pixels and the available energy (AE) after temporal scale extrapolation.

2 Methodology

2.1 LE algorithm

Surface thermal dynamics control energy partitioning and ET. However, the spatial resolution of thermal-infrared (TIR) images is usually not as high as the spatial resolution of visible near-infrared (VNIR) bands because the energy of VNIR photons is higher than the energy of thermal photons (Peng et al., 2016). The IPUS (input parameter upscaling), 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 as the lumped method in this study. This model was designed to simulate satellites that have identical spatial resolutions in both the VNIR and TIR bands and has been described in details in Peng et al., (2016). The energy flux components net radiation (Rn), soil heat flux (G), sensible heat flux (H) and LE are shown as below (Jiao et al., 2014; Peng et al., 2016).

Rn is the difference between incoming and outgoing radiation, as follows:

$$R_n = S_d(1-\alpha) + \varepsilon_s L_d - \varepsilon_s \sigma T_{rad}^4 , \quad (1)$$

where S_d is downward shortwave radiation, α is the surface albedo, ε_s is the emissivity of land surface, L_d is the downward atmospheric longwave radiation, $\sigma = 5.67 \times 10^{-8} W \cdot m^{-2} \cdot K^{-4}$ is the Stefan-Boltzmann constant, and T_{rad} is the surface radiation temperature.

G is commonly estimated through the derivation of empirical equations that employ surface parameters such as R_n as follows (Su, 2002):

$$G = R_n \times [\Gamma_c + (1-f_c) \times (\Gamma_s - \Gamma_c)] , \quad (2)$$

where Γs is equal to 0.315 for a bare soil situation, Γc is equal to 0.05 for a full vegetation canopy, and fc is fractional canopy coverage.

The sensible heat flux (H) is calculated based on gradient diffusion theory:

$$H = \rho c_p \frac{T_{aero} - T_a}{r_a} , \quad (3)$$

where ρ is the density of air; c_p is the specific heat of air constant pressure; T_{aero} is the aerodynamic surface temperature obtained by extrapolating the logarithmic air temperature profile to the roughness length for heat transport; T_a is the air temperature at the reference height and r_a is the aerodynamic resistance that influence the heat transfer between the source of turbulent heat flux and the reference height. Finally, LE is calculated as a residual item of the energy balance equation (Eq. (4)).

$$LE = R_n - G - H \,, \tag{4}$$

Further details can be found in Peng et al. (2016).

2.2 EF of mixed pixels

(1) Equation for EF of mixed pixels

The EF is the ratio of LE and available energy (AE) (Rn-G), as follows:

$$EF = \frac{LE}{R_n - G} \,, \tag{5}$$

Studies have shown that the EF is quite stable over time and thus is well suited to denote the status of the surface energy balance for a certain period. For example, the EF is nearly constant during the daytime (Nichols and Cuenca, 2010; Sugita and Brutsaert, 1991) and thus, can be used for temporal scale extrapolation, i.e., from instantaneous LE at the satellite overpassing time to daily ET. EF is widely used to estimate daily ET with RS data in different methods (e.g., the feature space of the Land Surface Temperature and Vegetation Index (LST-VI) (Carlson, 2007; Long and Singh, 2012) and SEBS (Su, 2002) models).

In this section, the EF of mixed pixels is investigated and a novel approach is derived to estimate the daily ET of mixed pixels. In other words, EF is used for temporal scale extrapolation and spatial scale correction of the remotely sensed LE and ET at a coarse-resolution scale at the same time.

Because turbulence transferred by advection is always neglected in RS data, we only consider vertical turbulence. Therefore, the accurate LE (with scaling effects taken into consideration) of a mixed pixel can be weighted by the LE of its sub-pixels as follows:

$$LE = \sum s_i \, LE_i = \sum [s_i \cdot \frac{LE_i}{(Rn - G)_i} \cdot (Rn - G)_i] \,, \tag{6}$$

where LE denotes the accurate LE of mixed pixels, s_i the area fraction (AF) of sub-pixel *i*, and LE_i the LE of sub-pixel *i*. Eq. (5) and (6) can be combined as follows:

$$LE = \sum [s_i \cdot EF_i \cdot (Rn - G)_i] \,, \tag{7}$$

where EF_i and $(Rn\text{-}G)_i$ denote the EF and AE of sub-pixel *i* in a certain mixed pixel respectively.

At this step, a simplification is performed as described in Hypothesis 1:

Here, Hypothesis 1 is proposed as follows:

"The available energy (AE) of each sub-pixel is approximately equal to that of any other sub-pixels in the same mixed pixel within an acceptable margin of errors (e.g. 50 $W \cdot m^{-2}$ (Seguin B et al.,1999; Kustas and Norman, 2000; Sánchez et al., 2007)) and is equivalent to the AE of the mixed pixel."

Therefore, Eq. (7) can be transformed in to the following expression:

$$\widetilde{LE} = [\sum(s_i \cdot EF_i)] \cdot (Rn - G) , \quad (8)$$

where $\widetilde{LE}$ denotes the latent heat flux in mixed pixels based on Hypothesis 1. There is a minor difference between $\widetilde{LE}$ and LE that can be regarded as an error of Hypothesis 1, and it will be analysed below.

Rearranging Eq. (8) yields the following:

$$\frac{\widetilde{LE}}{(Rn-G)} = \sum(s_i \cdot EF_i), \quad (9)$$

Therefore, we have

$$\widetilde{EF} = \sum(s_i \cdot EF_i), \quad (10)$$

where $\widetilde{EF}$ denotes the EF of the mixed pixel, including the error of Hypothesis 1, which is quite small and can be neglected based on a data analysis (see Section 4.3.1). Hence Eq. (10) can be used as the solution to the EF of mixed pixels.

Using Eq. (10) makes calculating the EF of mixed pixels is straightforward since it only needs the AF of each land cover in the pixel, which can be easily obtained using a fine-resolution land cover map, as well as the EF_i of its sub-pixels, which requires a specific technique to get in operations.

(2) Calculating EF of mixed pixels

The EF_i of sub-pixels is required in Eq. (10); however, it is not available with coarse resolution data. In order to utilize Eq. (10), Hypothesis 2 is proposed:

"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."

The concept of "nearest" in this study is refers to the shortest distance between the centre point of the mixed pixel where the sub-pixel is located and the centre point(s) of the pure pixel(s) with the same land cover type as the sub-pixel in the study area. The concept is illustrated in Fig. 1.

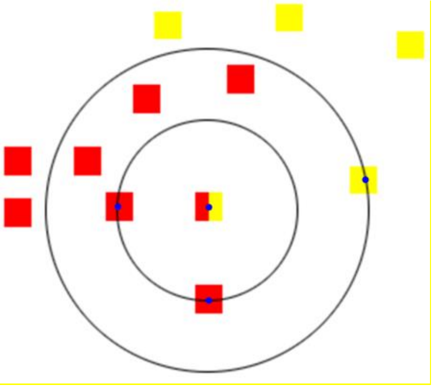


Figure 1: A sample graph of sub-pixel in a mixed pixel and its nearest pure pixel(s) of the same land cover type.

In Fig.1, there are two land cover types: Yellow and Red. The centre pixel, which is a mixed pixel, contains a red sub-pixel and a yellow sub-pixel. Hypothesis 2 examines which pixel the nearest neighbour. From Fig. 1, it is clear that the red sub-pixel has two nearest neighbours (red pixels with blue centres); thus, the EF of the red sub-pixel equals the mean EF of the two nearest pure pixels according to Hypothesis 2. The yellow sub-pixel has one nearest neighbour (yellow pixel with blue centre); thus, the EF of the yellow sub-pixel equals the EF of the yellow pure pixel.

This hypothesis is based on Tobler's First Law (TFL) (Miller, 2004; Li et al., 2007; Tobler, 2004), which states that “everything is related to everything else, but near things are more related than distant things”. In other words, the most similar conditions, phenological patterns and physical characteristics exist between a sub-pixel surface and nearby (pure pixel) surfaces given the same land cover. Accordingly, the EF of sub-pixel *i* can be determined using EF of pure pixel(s) at coarse resolution scale based on Hypothesis 2.

Therefore, Eq. (8) may be reduced as above to the following:

$$\widetilde{LE} = (Rn - G) \cdot \widetilde{EF}, \tag{11}$$

Eq. (10) and Hypothesis 2 together can be used to calculate the EF of mixed pixels and therefore the daily ET. Eq. (10) and (11) can be used together to correct the spatial scale errors of the instantaneous LE at the overpassing time.

In summary, by employing two key hypothesises, EFAF methodology is able to realize temporal scale extrapolation and spatial scale correction for remotely sensed LE and ET at a coarse resolution scale at the same time. The EF of a mixed pixel is expressed as the area-weighted EF_i of its sub-pixels with acceptable simplifications, which simplified the calculations, increased the accuracy, and facilitated its use for daily operations.

2.3 Estimation of daily LE

We use the EF method to extrapolate the temporal scaling of the LE. The EF method is based on the basic assumption that each component of the energy balance model remains relatively constant during the day and that the relative components of LE and AE (Rn-G) are constant (Nichols and Cuenca, 2010; Sugita and Brutsaert, 1991) Therefore, the daily LE can be expressed as follows:

$$\frac{\mathrm{LE_{daily}}}{(\mathrm{R_n - G})_{\mathrm{daily}}} = \frac{\mathrm{LE_{inst}}}{(\mathrm{R_n - G})_{\mathrm{inst}}} = \mathrm{EF_{inst}}, \tag{12}$$

$$\mathrm{LE_{daily}} = \mathrm{EF_{inst}} \cdot (\mathrm{R_n - G})_{\mathrm{daily}}, \tag{13}$$

where the subscripts “daily” and “inst” indicate daily cumulative and instantaneous values, respectively. To calculate the daily total ET from Eq. (13), it is necessary to determine the EF and the daytime total AE (Zhang and Lemeur, 1995). The daytime net radiation is obtained from the parameterization proposed by Bisht et al. (2005), in which the average daytime net radiation and then its integral are calculated as follows:

$$\mathrm{DANR} = 2 * \mathrm{Rn_{inst}}/\pi \sin[\left(\frac{t_{ovp}-t_{rise}}{t_{set}-t_{rise}}\right)\pi], \tag{14}$$

$$\mathrm{Rn_{daily}} = \int \mathrm{DANR}\, dt, \tag{15}$$

where DANR is the average daytime net radiation, $\mathrm{Rn_{daily}}$ is the daytime cumulative net radiation, t_{ovp} is the satellite imaging time, and t_{rise} and t_{set} are local sunrise and sunset times, respectively, representing times at which the net radiation shifts from positive to negative.

The daytime G is calculated from the by DANR and Eq. (2).

The flowchart of the EFAF shown below illustrates the (1) calculation of LE without a scale effect, (2) calculation of the EF of mixed pixels, and (3) extrapolation of the temporal scale (Fig. 2).

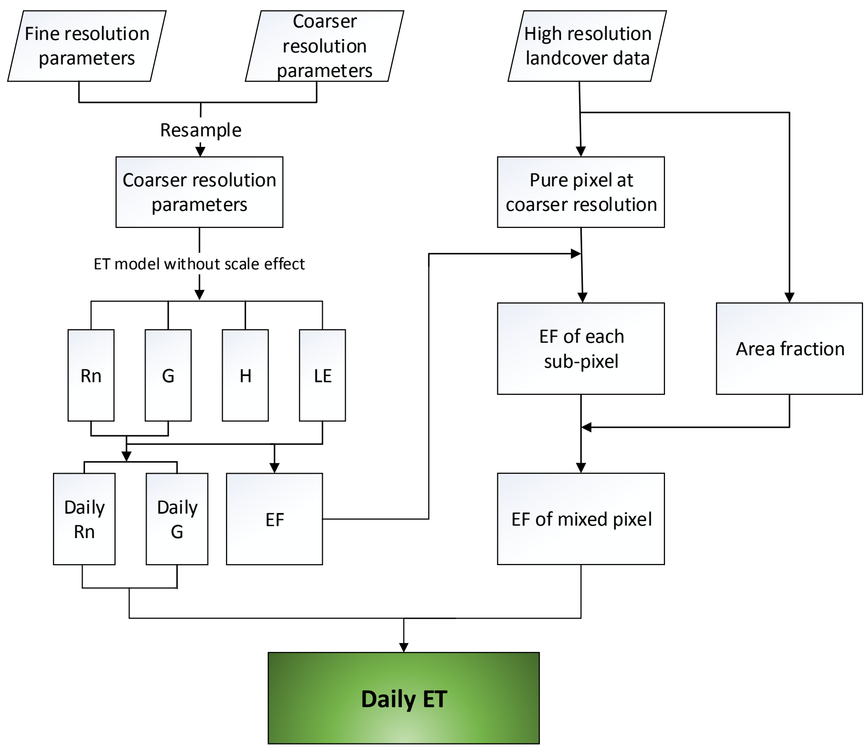


Figure 2: Flowchart of the EFAF, where trapezoids represent the input variables or parameters, and rectangles represent variables or parameters. The inputs of EFAF encompass the remotely sensed variables or parameters and meteorological forcing dataset. The abbreviations are defined as follows: Rn: net radiation; G: soil heat flux; H: sensible heat flux; LE: latent heat flux; EF: evaporative fraction; ET: evapotranspiration;

3 Study area and dataset

3.1 Study area

The study area is located in the Heihe River watershed in west-central Gansu Province, north-western China (Fig. 3). The Heihe River watershed has a land surface area of approximately 128,000 km ²and is the second largest inland watershed in north-western China (Gu et al., 2008). The Heihe River watershed includes the Zhangye sub-watershed, which covers a total land area of approximately 31,100 km ². The natural landscape of the study area is heterogeneous, including mountains, oasis areas, and desert (Ma and Veroustraete, 2006). The oasis is a typical farmland ecosystem located 8 km south of the city of Zhangye in which maize and wheat are the major crops. Large expanses of desert and mountains surround the central oasis. In this area, annual precipitation ranges from 100–250 mm, but potential ET levels reach approximately 1200–1800 mm yearly (Li et al., 2013)

Since 2012, an eco-hydrological experiment referred to as the Heihe Watershed Allied Telemetry Experimental Research (HiWATER) has been conducted in the area. An observation matrix composed of 17 eddy covariance (EC) systems and automatic meteorological stations (AMSs) was established across the landscape (Li et al., 2013).

The percentage of the numbers of land cover types (Yu et al., 2016) (Fig. 4) for the study area were extracted at a 300 m scale with 30 m land cover classifications developed by Zhong et al. (2014a) based on HJ-1/CCD time series. The pure pixels at 300 m scale are entirely made up of one particular land cover type, and the mixed pixels are made up of two or more land cover types according to the land cover datasets with a spatial resolution of 30 m. It has been shown that pure pixels account for 41.74% and mixed pixels account for 58.26% of the area. Such an area, with more mixed than pure pixels but with many of both, represents an optimal place to test the proposed method.

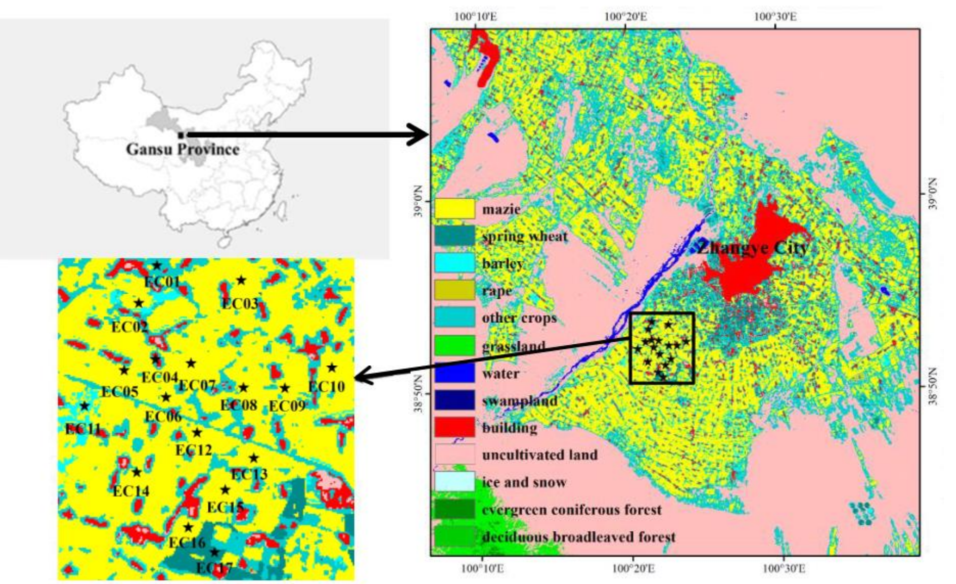


Figure 3: Distribution of in situ stations and land use classifications in our study area (revised based on Peng et al., 2016).

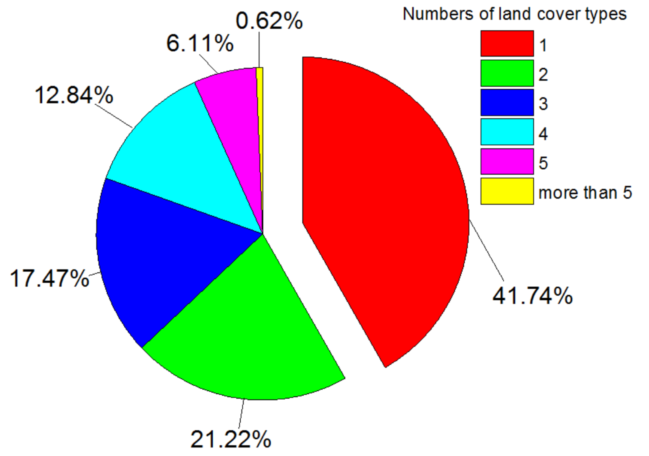


Figure 4: Percentage of the number of land cover types for the study at 300 m scale with 30 m land cover images.

3.2 Remote sensing data

The HJ-1B satellite (Table 1) was successfully launched on 6 September 2008 and follows a quasi-sun-synchronous orbit at an altitude of 650 km. After geometric correction, radiometric calibration, and atmosphere correction (Zhang et al., 2013; Zhong et al., 2014b), the image quality of the HJ-1B data is the same as that of Landsat-5 TM, and the data can be used for applications including environmental and disaster monitoring (Jiang et al., 2013). The calculation of ET levels represents one of the most important applications of the HJ-1B satellite data.

Table 1. Specifications of the HJ-1B main payloads

Sensor	Band	Spectral range (µm)	Spatial resolution (m)	Swath width (km)	Revisit time (days)
Charge-Couple Device (CCD)	1	0.43-0.52	30	360 (single) 700 (double)	4
	2	0.52-0.60			
	3	0.63-0.69			
	4	0.76-0.90			
Infrared Scanner (IRS)	5	0.75-1.10	150	720	4
	6	1.55-1.75			
	7	3.50-3.90			
	8	10.5-12.5	300		

The algorithms for most surface parameters used to estimate ET are applicable under clear-sky conditions. Therefore, the satellite data selected for the study area were collected under clear or partly cloudy conditions based on data quality metrics and artificial visual interpretation. The selected images were divided into nine groups (each group included the study area in the IRS and CCD data), from 30 June, 8 July, 27 July, 3 August, 15 August, 22 August, 29 August, 2 September, and 13 September 2012.

In this study, each component of the energy balance algorithm used to estimate the daily ET of mixed pixels was retrieved using the lumped method based on HJ-1B data (CCD/IRS). These components included surface albedo (Liang et al., 2005; Liu et al., 2013), downward shortwave radiation (Li et al., 2011), land surface emissivity (Valor and Caselles, 1996), land surface temperature (Li et al., 2010), the normalized difference vegetation index (NDVI), fractional vegetation coverage (FVC) (Peng et al., 2016), and LAI (He et al.,2012; Nilson, 1971).

Furthermore, 30 m resolution land cover classifications derived from HJ-1/CCD time series were used. Highly accurate 30 m land cover classifications for June to September 2012 based on HJ-1B data were developed by Zhong et al. (2014a). The major land use types included cropland for maize, wheat and vegetables (according to experiential knowledge, although it is considered as other crops in this classification), uncultivated land (including bare soils and Gobi Desert), water bodies, grassland, forests, and buildings.

3.3 HiWATER experiment in situ dataset

In situ data were provided by the HiWATER-Multi-Scale Observation Experiment on Evapotranspiration (MUSOEXE) over heterogeneous land surfaces of the HiWATER campaign, which was carried out at an artificial oasis in the Zhangye Heihe River watershed. During the HiWATER-MUSOEXE campaign, 17 EC towers and AMSs were arranged in two nested observation matrices (Li et al., 2013) to obtain ground measurements of radiation fluxes, meteorological parameters, and soil and turbulent heat flux. Details regarding the ground towers are shown in Table 2, and the tower distribution is shown in Fig. 3.

The in situ data are considered reliable based on various quality control measures. For example, prior to the main campaign, the performance of the instruments was compared in the Gobi Desert (Xu et al., 2013). After basic processing, including spike removal and corrections for density fluctuations (WPL-correction), a four-step quality control procedure was applied to the EC data. The EC data were based on 30 min intervals; additional information regarding system setup, data processing and quality control can be found in previous reports (Yang and Wang, 2008; Liu et al., 2011; Liu et al., 2016; Xu et al., 2013).

Energy imbalance is common in ground flux observations conducted over long periods. Common methods for forcing the energy balance include conservation of the Bowen ratio (H/LE) and the residual closure technique. Studies have suggested that computing the LE as a residual may be a better method for energy balance closure when the LE is large (with small or negative Bowen ratios due to strong advection) (Kustas et al., 2012). Therefore, the residual closure method was used in this study, because there was a distinct “oasis effect” on clear days (Liu et al., 2011).

Because this study focuses on mixed pixels of heterogeneous surfaces, we exclude some stations (EC 07, EC 08, EC 10, and EC 15) from our discussion, because they are located in areas with pure pixels. In addition, EC17 is in an area dominated by orchards. Orchards are considered other crops in our classification, and the complex vertical structure of orchard ecosystems can result in large gaps that are difficult to analyse. Therefore, EC17 is also excluded from our discussion.

Regarding the other observations, we conducted interpolation to fill null values in the observations. Linear interpolation (Liu et al., 2012) was used for missing values over intervals smaller than 2 hours, and the mean diurnal variation (MDV) method (Falge et al., 2001) was used for missing values over intervals greater than 2 hours. Next, energy residual methods were used to conduct the closure process. Finally, a Eulerian analytic footprint model (Kormann and Meixner, 2001) was used to calculate the source region and extract ground observation values, which can express the LE of the heterogeneous surface.

Table 2. Details of the Heihe River basin (HRB) in-situ stations

Station	Longitude (°E)	Latitude (°N)	Tower height (m)	Altitude (m)
EC01	100.36	38.89	3.8	1552.75
EC02	100.35	38.89	3.7	1559.09
EC03	100.38	38.89	3.8	1543.05
EC04	100.36	38.88	4.2	1561.87
EC05	100.35	38.88	3.0	1567.65
EC06	100.36	38.87	4.6	1562.97
EC07	100.37	38.88	3.8	1556.39
EC08	100.38	38.87	3.2	1550.06
EC09	100.39	38.87	3.9	1543.34
EC10	100.40	38.88	4.8	1534.73
EC11	100.34	38.87	3.5	1575.65
EC12	100.37	38.87	3.5	1559.25
EC13	100.38	38.86	5.0	1550.73
EC14	100.35	38.86	4.6	1570.23
EC16	100.36	38.85	4.9	1564.31
EC17	100.37	38.85	7.0	1559.63

4 Results and analysis

4.1 Results of the EFAF

The EFAF study was performed on crops that mainly grew during June, July, August, and September. We selected two days in different growing phases, 8 July (Fig. 5) and 22 August (Fig. 6), and compared the changes in lumped EF, EFAF EF, lumped LE, and EFAF LE on these days. The results showed similar changes in EF and LE.

Overall, there were no differences in EF and daily LE on either day between the city and the desert area that could be distinguished based on land cover data, because of the homogeneous surface of both land cover types (Fig. 5 and 6). For example, Area I in Fig. 7 represents the city of Zhangye, and Area II in Fig. 7 represents uncultivated land. The EFAF EF and EFAF LE values of both areas are the same as the lumped EF and lumped LE values because pure pixels were not corrected in this study.

However, the boundaries became blurred between buildings, which were given an LE of 0 in this study (Peng et al., 2016), and farmland; thus, the intersection of these land cover types resulted in "buffer pixels". For example, in Area III in Fig. 7, the EF and daily LE of pixels dominated by buildings (village areas with many villages) appears blue, denoting low EF and LE values without scaling correction; these areas appear orange after considering agriculture areas around the buildings. For the same reason, in the suburbs surrounding the city of Zhangye (Area IV in Fig. 7), an area of mixed pixels dominated by buildings appears blue, with low lumped LE values; the same area appears yellow or pale blue after considering the presence of vegetables.

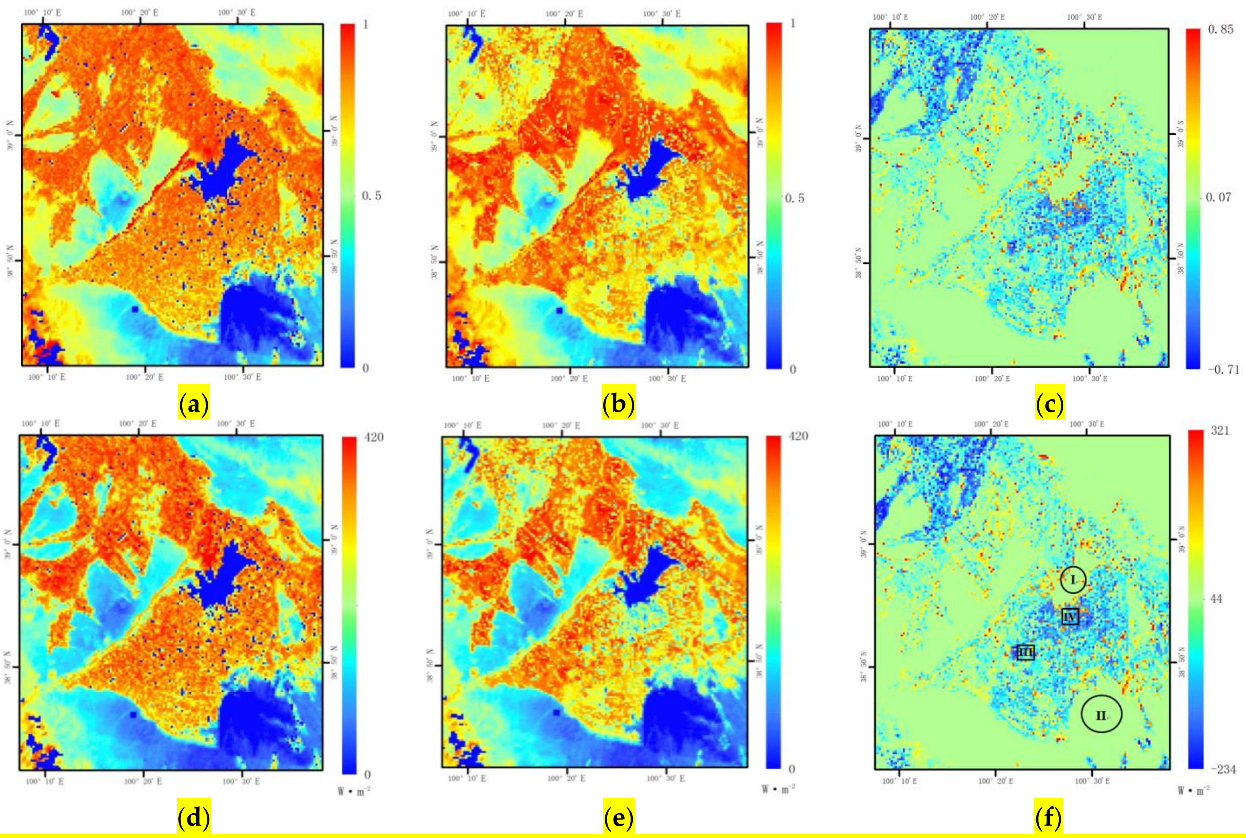


Figure 5. Maps of (a) lumped EF, (b) EFAF EF, (c) difference between EFAF and lumped EF (EFAF EF minus lumped EF), (d) lumped daily LE, (e) EFAF daily LE and (f) difference between EFAF and lumped LE (EFAF LE minus lumped LE) on July 8th, 2012

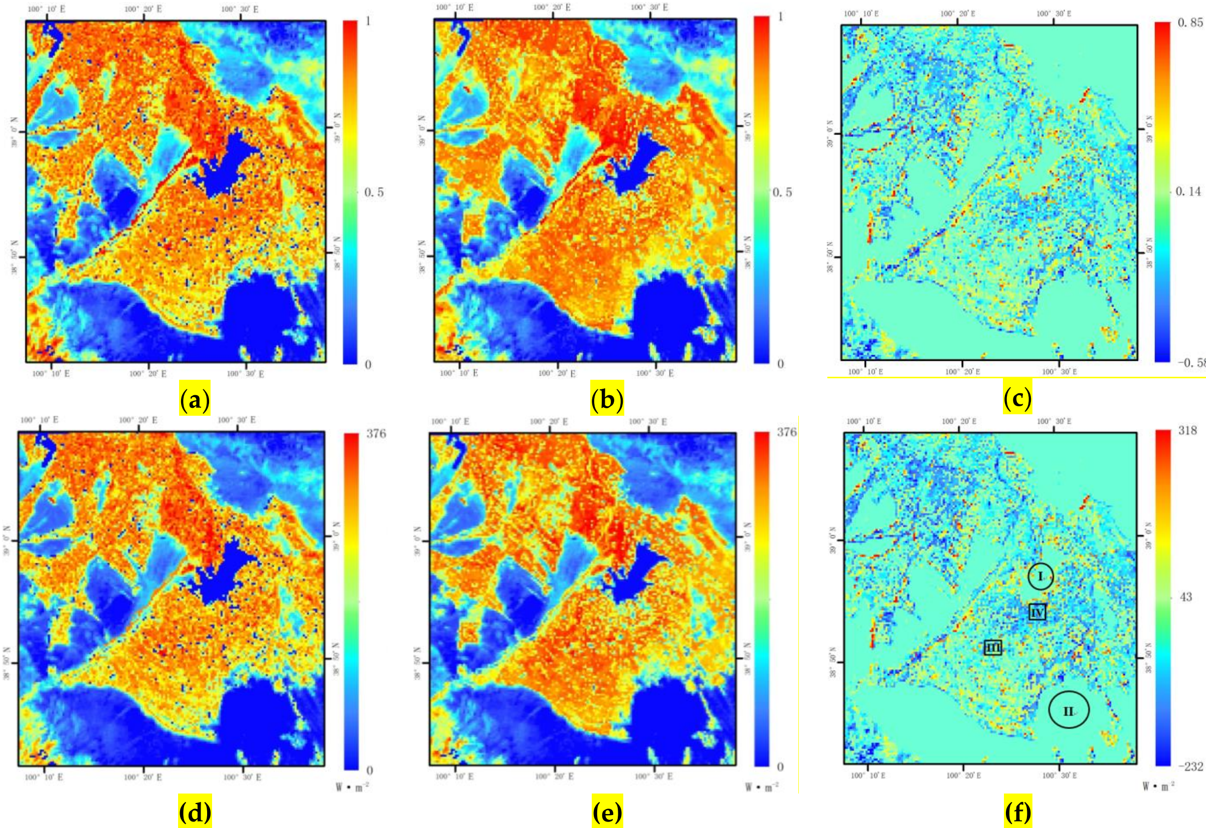


Figure 6. Maps of (a) lumped EF, (b) EFAF EF, (c) difference between EFAF and lumped EF (EFAF EF minus lumped EF), (d) lumped daily LE, (e) EFAF daily LE and (f) difference between EFAF and lumped LE (EFAF LE minus lumped LE) on August 22nd, 2012

The EF and LE values for pixels dominated by agriculture and including buildings decreased, likely because the area included villages whose EF was set to zero. For instance, in region IV (Fig. 7), pixels dominated by buildings and including cropland and pixels dominated by cropland and including buildings account for 20 % and 80 %, respectively, and the spatially averaged daily LE decreased from 8.98 to 7.39 MJ m^{-2} (decreased approximately from 3.57 mm to 2.97 mm, the latent heat of vaporization is approximately 2.49×10^{6} W m^{-2} mm^{-1} (Pan and Liu, 2003), the same below) on 8 July 2012. However, for pixels dominated by buildings, the spatially averaged daily LE increased from 0 to 4.70 MJ m^{-2} (increased from approximately 0 mm to 1.80 mm).

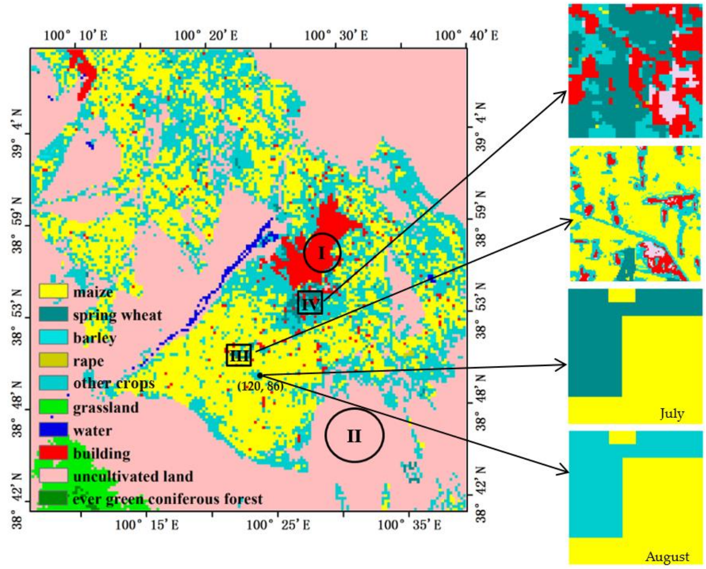


Figure 7. Land cover maps of the study area at 300 m resolution and certain regions at 30 m resolution. Area I and II represent the area of pure pixels, Area III and IV represent the area of mixed pixels, and (120, 86) represents an example point of different land cover types in different months.

In addition, the EF and daily LE decreased significantly on 8 July when the EFAF method was applied in the north-western and southern oasis areas of the study area. This change was less pronounced on 22 August. The EF and daily LE decreased slightly in the north-western parts of the study area and increased slightly in the south-central oasis area. The reason for this difference could be that the mixed pixels in this area mainly included maize, spring wheat, and barley. In July, spring wheat and barley were in a ripening stage, which is characterized by lower ET. However, by August, the spring wheat and barley had been harvested and replaced by vegetables, and the maize had entered its dough stage, which is characterized by reduced ET. The ET of vegetables was higher than that of the spring wheat and barley in July (Wu et al., 2006). These differences could have resulted in the increase in the EF and daily LE after the EFAF method was applied.

For example, the point located at coordinates (120, 86) (Fig. 7) included maize (58%) and spring wheat (42%). The mean EF of the pure pixels closest to the maize was 0.75, and the mean EF of pure pixels closest to the spring wheat was 0.65. Therefore, application of the EFAF method resulted in a decrease in the EF from 0.81 to 0.71 and in a decrease in the daily LE from 14.25 to 12.37 MJ m^{-2} (approximately 5.72 to 4.97 mm). In contrast, on 22 August, this pixel included maize (58%) and vegetables (42%). The mean EF of the pure pixels closest to the maize was 0.81, and the mean EF of the pure pixels closest to the vegetables was 0.86. Thus, application of the EFAF method resulted in an increase in the EF of 0.79 to 0.83 and an increase in the daily LE of 12.33 to 13.00 MJ m^{-2} (approximately 4.95 to 5.22 mm). Another reason for these minor changes could be related to irrigation, which occurred in the southern oasis area on 22 August (Peng et al., 2016). The EF of bare soil would likely increase because of greater soil moisture due to irrigation. As a result, the difference in EF values between agricultural land and bare soil decreased, as indicated in Fig. 6 (a) and (b).

4.2 Validation of daily LE

Daily EC measurements for LE were aggregated using a range of time series data based on the time at which net radiation shifted from positive to negative values. The simulated EC measurements were averaged over the estimated upwind source area for each flux tower. The results (Table 3) indicate that in general, the EFAF LE values are more consistent with the EC measurements than the lumped LE values. Comparing the lumped and EFAF methods shows that the coefficient of determination (R^2) increased from 0.62 to 0.82; the root mean square error (RMSE) decreased from 2.47 to 1.60 MJ m^{-2} (0.99 to 0.64 mm), a decreased of approximately 35.22%; and the mean bias error (MBE) decreased from 1.92 to 1.18 MJ m^{-2} (0.77 to 0.47 mm).

Table 3. In situ validation results for the daily LE

	Lumped LE ($MJ \cdot m^{-2}$)			EFAF LE ($MJ \cdot m^{-2}$)			RMSE decreased from
Date	R^2	MBE	RMSE	R^2	MBE	RMSE	lumped LE to EFAF LE (%)
30 June	0.16	-1.42	2.59	0.59	-1.20	1.95	24.71%
8 July	0.16	0.40	1.99	0.63	-0.32	1.38	30.65%
27 July	0.24	2.49	3.37	0.65	0.53	1.62	51.93%
3 August	0.50	1.37	3.09	0.87	0.53	1.78	42.39%
15 August	0.39	1.48	1.87	0.72	0.95	1.32	29.41%
22 August	0.01	-1.70	3.18	0.54	-1.43	2.19	31.13%
29 August	0.43	-0.73	1.72	0.63	-0.73	1.38	17.77%
2 September	0.18	0.72	1.72	0.52	0.87	1.48	13.95%
13 September	0.01	-0.64	1.89	0.32	-0.08	0.90	52.38%
Total	0.63	0.21	2.47	0.82	-0.10	1.60	35.22%

Table 3 also presents the lumped LE and EFAF LE results against the EC measurements for each day. The EFAF LE better reproduced the EC measurements than the lumped LE on all nine days. Combining the EFAF LE with EC data on 29 August resulted in a slightly more accurate LE estimate, with an RMSE of 1.38 MJ m^{-2} (0.55 mm), relative to the lumped LE, with an RMSE of 1.72 MJ m^{-2} (0.69 mm), the accuracy increased by approximately 13.95% according to the RMSE. This difference is likely related to the fact that the slight heterogeneity in land surface temperature decreased the scale error that resulted from thermal dynamics. In addition, the EFAF LE results for 13 September were more accurate, yielding an RMSE of 0.90 MJ m^{-2} (0.36 mm), relative to the lumped LE, which had an RMSE of 1.89 MJ m^{-2} (0.76 mm), the RMSE decreased by approximately 52.38%. This improvement may result from the greater landscape heterogeneity, which created obvious scale effects in the LE results; ripe maize, growing vegetables, withered grass, and bare soils coexisted in the study area on that day.

However, uncertainties resulting from scale mismatches between RS data and the EC footprint could reduce the confidence and skill of the EFAF method. A unique aspect of the present study is that the EC data are consistent across the simulations on all nine days; this feature minimizes tower-uncertainties by ensuring that the retrieved LE can be assessed against each EC tower record individually (Fig. 8). The results (Fig. 8) show that the EFAF LE had smaller RMSE values and higher R^2 values than the lumped LE for all EC sites, indicating that the EFAF method improved the accuracy of daily LE estimates. However, this improvement in accuracy differed across sites.

The correction effect of the EFAF method was most distinct at the EC04 site, and the RMSE at EC04 decreased from 5.36 to 2.72 MJ m^{-2} (2.15 to 1.09 mm) (decreased by approximately 49.25%); this improvement stemmed from the fact that EC04 had the highest complexity of all sites. Maize-dominated pixels in EC04 included maize, vegetables, buildings and bare soil, at a ratio of 53:26:19:2, respectively. We conclude that maize and vegetables were land cover types with a high EF, while bare soil had a low EF. For buildings, the EF value was 0 in this study. For example, on 30 June, the EF of mixed pixels in EC04 was 0.81. However, the average EF values of the pure pixels positioned closest to maize and vegetables among the sub-pixels were 0.88 and 0.88, respectively and that of bare soil was 0.65. Therefore, when scale effects were taken into consideration, the EF of the mixed pixels was 0.70. Using the EFAF method, the daily LE of the mixed pixel where EC04 was located decreased from 13.57 to 11.78 MJ m^{-2} (5.45 to 4.73 mm). Similarly, the difference between these estimates and the EC measurements also declined from 4.12 MJ m^{-2} to 2.32 MJ m^{-2} (1.67 to 0.93 mm) (decreased by approximately 43.3%). Additionally, there were large discrepancies between the observed and retrieved LE values at EC04. Specifically, there are two points far from the 1:1 line in Fig. 8 (d), with values of 8.36 MJ m^{-2} (3.36 mm) on 27 July and 9.33 MJ m^{-2} (3.75 mm) on 3 August. Even after the EFAF method was applied, these values were 5.20 MJ m^{-2} (2.09 mm) and 4.59 MJ m^{-2} (1.84 mm), respectively, because EC04 was positioned in a maize-dominated pixel and the EC tower was located in a built-up area, thus generating errors associated with temperature retrieval that would create further errors in estimating Rn. For example, on 27 July and 3 August, the Rn observed by AWS for the EC station was 15.95 and 15.35 MJ m^{-2}, respectively, while the retrieved Rn of the pixels was 18.14 and 18.80 MJ m^{-2}, respectively. The remaining larger errors in such pixels are a reminder that this method has limitations under certain extreme conditions. More complex models should be built for such circumstances and more information other than land cover should be included when considering subsurface heterogeneity to obtain results that are as accurate as those obtained for homogeneous sites.

The correction effect was not significant for sites such as EC02, EC06, EC12, and EC14; these sites had minimal surface heterogeneity, with only two land cover types present in the mixed pixels. These pixels also included a mixture of maize and other crops with similar EF values. However, the accuracy of daily LE was improved based on the effects of mixed pixels on EF. For example, EC12 was a maize-dominated pixel, with a 74:26 ratio of maize to other crops in July. On 27 July, the mean EF of the pure pixels closest to the maize area was 0.97; for the other crops, the EF of the pure pixels was 0.84. The EF of this mixed pixel changed from 0.96 to 0.94 when the EFAF method was used, and the daily LE decreased from 18.00 to 17.24 MJ m^{-2} (7.23 to 6.92 mm). Compared to the value of 16.52 MJ m^{-2} (6.63 mm) found for EC, the EFAF LE was more accurate.

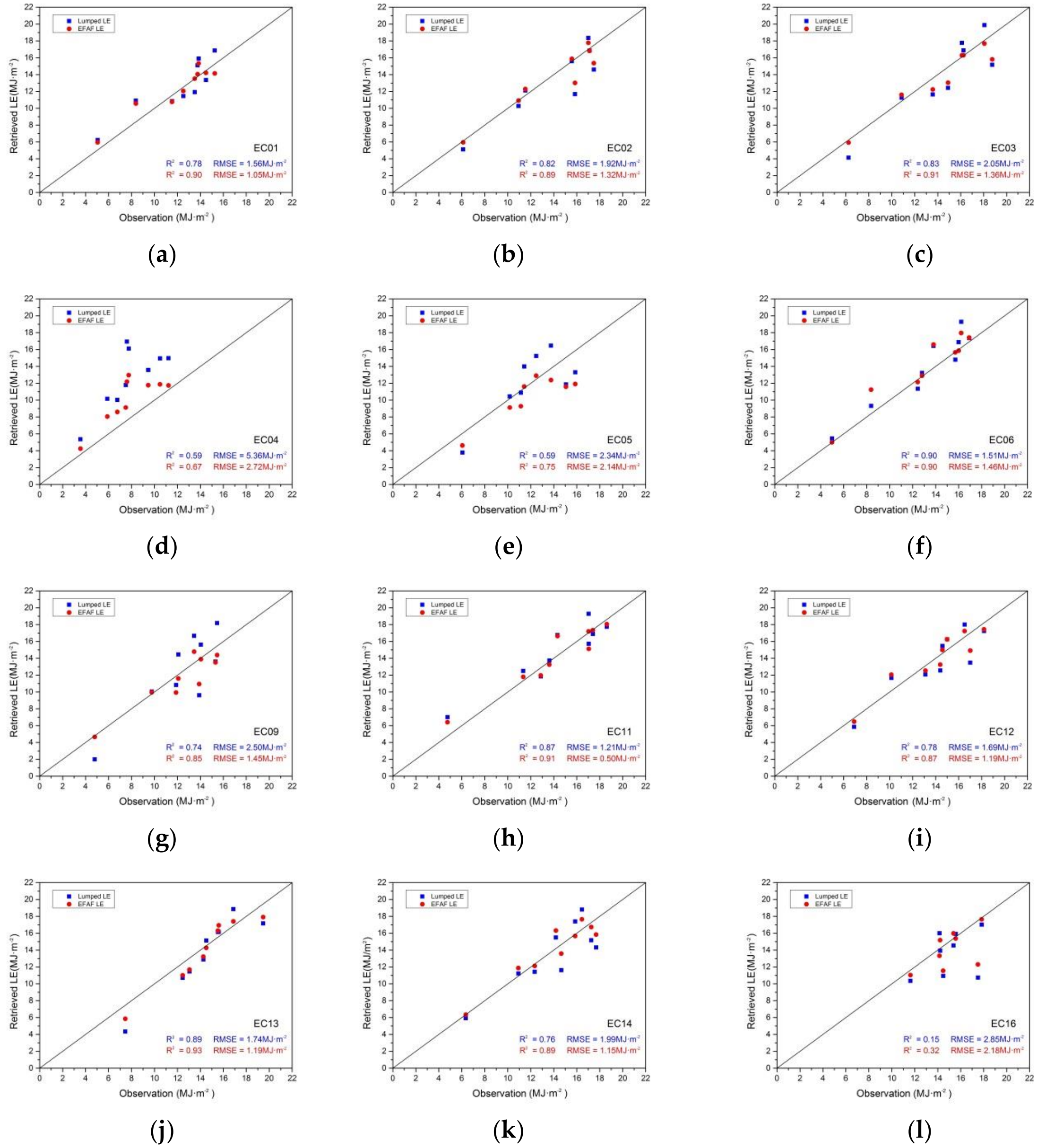


Figure 8. Scatter plots of the lumped LE (blue) and EFAF LE (red) against the EC measurement LE at each site

4.3 Error analysis

4.3.1 Error analysis of Hypothesis 1

Hypothesis 1 states that the AE of each sub-pixel is approximately equal to that of any other sub-pixels in the same mixed pixel within an acceptable margin of errors (e.g. $50\ \mathrm{W \cdot m^{-2}}$ (Seguin B et al.,1999; Kustas and Norman, 2000;

Sánchez et al., 2007)) and is equivalent to the AE of the mixed pixel. To quantify the error associated with Hypothesis 1 for ET estimation, each lumped AE (Rn-G) was compared to the original 30 m pixel located within it, i.e., the pixel values of a lumped 300 m resolution were compared to the 10 × 10 set of 30 m pixels that they were drawn from. The difference AE (dA) and percent frequency of difference were measured from the 30 m resolution sub-pixels ($\boldsymbol{A_{sub}}$) with the same values as the lumped AE measured at a 300 m resolution from each mixed pixel, relative to the original 30 m of distributed AE ($\boldsymbol{A_d}$) for the nine days.

$$\mathbf{dA} = \boldsymbol{A_{sub}} - \boldsymbol{A_d}, \tag{15}$$

$$\mathbf{f} = \frac{\boldsymbol{dA}}{\sum \boldsymbol{dA}}, \tag{16}$$

In all cases, the peak of the histogram is positioned at approximately 0 W m^{-2} (Fig. 9). This result indicates that the differences between the lumped and distributed AE range from -5 to 5 W m^{-2}, so the errors caused by Hypothesis 1 were minor for the AE estimations of most of the mixed pixels.

Furthermore, the frequency distribution of the difference in AE follows a generally symmetric distribution approximately 0 W m^{-2} at a range of ±120 W m^{-2}, though the frequency was low when the differences in AE were greater than 10 W m^{-2} or less than -20 W m^{-2} (less than 10%) (Fig. 9). The difference in frequency for values ±60 W m^{-2} was extremely poor (less than 1%) and thus could be ignored.

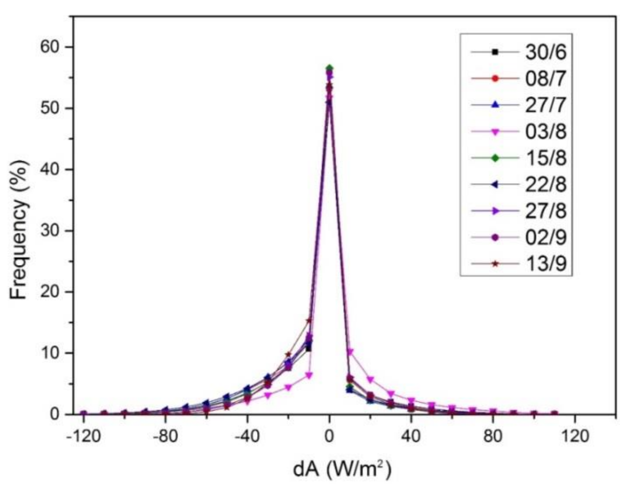


Figure 9. Distribution of the difference AE (dA) and the frequency of the difference for nine days

In addition, larger dA values mainly occurred at the transition zones between oasis areas and uncultivated land, and where large positive and negative dA values existed in a mixed pixel (for example, the dA value on 2 September (Fig. 10)). This result indicated that Hypothesis 1 results in large errors in the transition zones between oasis areas and uncultivated land, but these errors often cancel one another because large negative and positive errors exist in a mixed pixel.

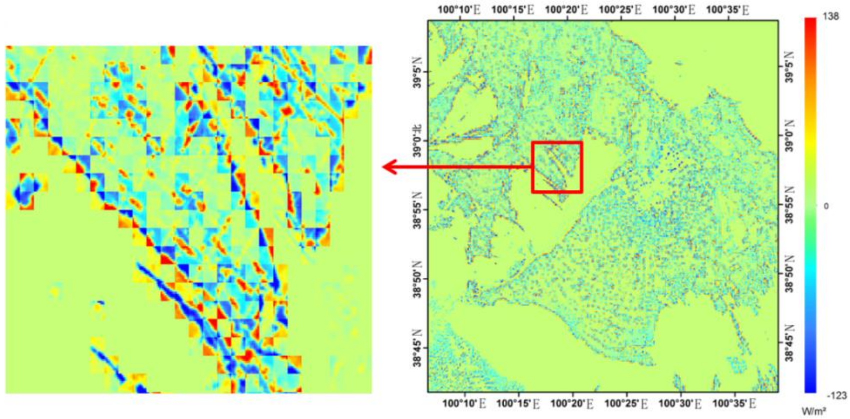


Figure 10. Spatial distribution of the difference AE (dA) and a transition zone on 2nd September.

To evaluate the errors in the study area as a result of Hypothesis 1, the expected value (E(x)) of error was measured based on the dA and its frequency. Fig. 11 shows the expected values of error based on Hypothesis 1 (dA) for the nine days studied. Small expected values of less than 10 W m^{-2} were observed when Hypothesis 1 was tested. A maximum error value of -8.44 W m^{-2} was found on 22 August. The mean EF of pure pixels for maize, grass, bare soils and vegetables was 0.77, 0.59, 0.22 and 0.81, respectively, on the same day. This result suggests that the LE estimation errors resulting from Hypothesis 1 for maize, grass, bare soils and vegetables were approximately -6.50, -4.98, -1.86 and -6.84 W m^{-2}, respectively. We consider these errors to be acceptable (Seguin B et al., 1999; Kustas and Norman, 2000; Sánchez et al., 2007).

$$\mathrm{E(x)} = \int_{-\infty}^{\infty} dA(x) f(x) dx, \tag{17}$$

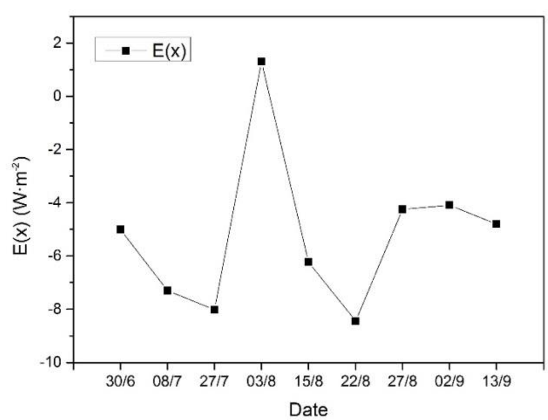


Figure 11. Expected error values based on Hypothesis 1 for the nine days

4.3.2 Error analysis of Hypothesis 2

To evaluate the errors associated with Hypothesis 2, which 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, the EF for each pure pixel, which is regarded as the correct value, was compared to the mean EF of its nearest pure pixel(s). The RMSE, MBE and R^2 values were calculated for each maize, grass, bare soil and vegetable land cover type (Fig. 12).

The EF of pure pixels appears to be well reproduced by Hypothesis 2; the overall RMSE is less than 0.06, indicating that Hypothesis 2 results in little error in the EF of sub-pixels estimations. For each land cover type, the maximum RMSEs were 0.047 for maize on 8 July, 0.055 for grass on 22 August, 0.048 for bare soils on 27 July and 0.059 for vegetables on 27 July, respectively. The simple averaged AE for the entire study area was 315.46 W m^{-2} on 8 July, 324.05 W m^{-2}on 27 July and 309.05 W m^{-2}on 22 August. This means that the maximum error in the LE estimates caused by Hypothesis 2 for maize, grass, bare soil and vegetables was approximately 14.83, 17.00, 15.55 and 19.12 W m^{-2}, respectively. Considering that most mixed pixels were closer to their nearest pure pixels than pure pixels were to their nearest pure pixels, the error in LE estimation caused by Hypothesis 2 might actually be lower.

The MBEs of EF for four land cover types were less than 0.01. These low values indicate that using Hypothesis 2 does not have adverse effects on calculating the EF of sub-pixels. Greater MBEs were observed in vegetables, ranging from -0.0050 to 0.019, and in grasslands, ranging from -0.0045 to 0.0083; in comparison, the MBE of maize ranged from -0.0037 to 0.00076 and the MBE of bare soil ranged from -0.0020 to 0.00075. These differences are likely related to the accuracy of classification. Areas with vegetables and grasses may include different species with various phenological patterns; in contrast, the phenological patterns of maize varied less and the bare soils were relatively homogeneous.

However, the R^2value differed between maize, grassland and vegetables.

The lower correlations were mainly caused by the uncertainty associated with positive or negative differences between the EF of a pure pixel and the mean EF of its nearest pure pixel(s); this uncertainty arises because of the heterogeneity in surface roughness and other variables among vegetation land cover types. For bare soils, there was a lower R^2 value on 27 July. This value can be attributed to the higher RMSE, which may have been caused by a brief cloudy period on that day that was not properly identified in the cloud detection process over uncultivated land.

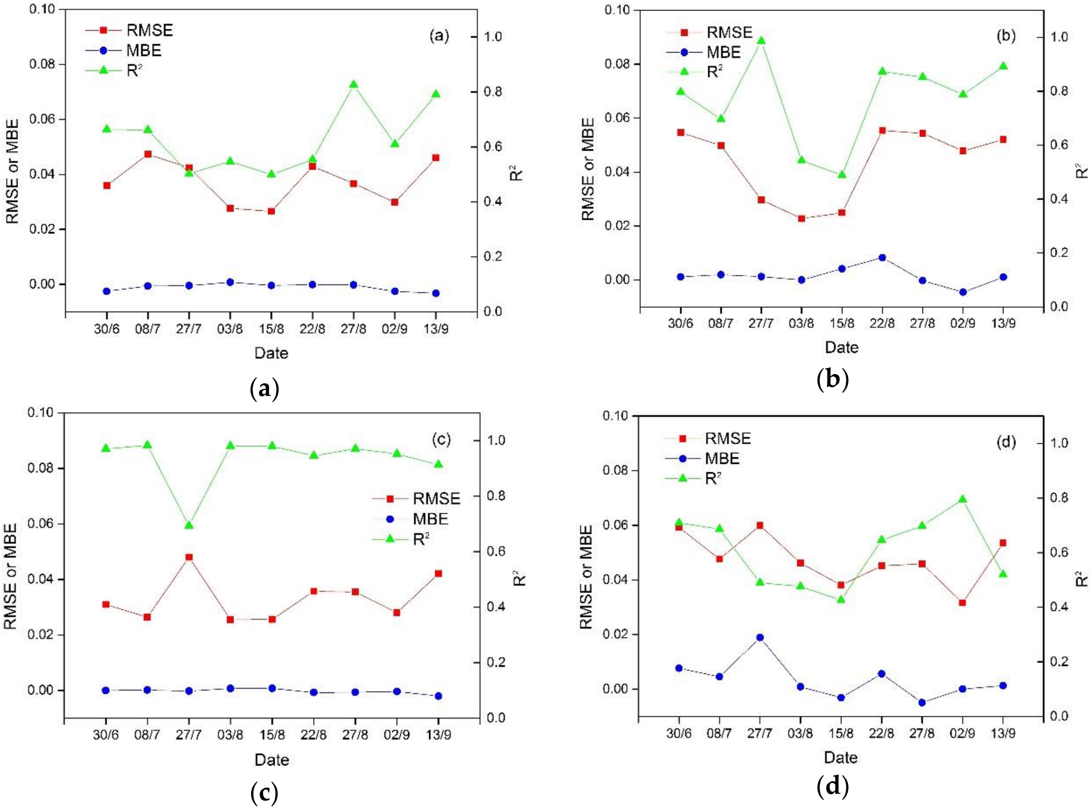


Figure 12. The RMSE, MBE and R^2 values of pure pixels based on the nearest pure pixel(s) for four land cover types: (a) maize, (b) grassland, (c) bare soils and (d) vegetables

In summary, Hypothesis 2 reproduces the EF of sub-pixels with an RMSE less than 0.06, resulting in errors within 20 W m^{-2} for LE estimation in this study. We consider such errors to be acceptable in surface flux estimation (Seguin B et al., 1999; Kustas and Norman, 2000; Sánchez et al., 2007).

4.4 Sensitivity analysis of the land cover map

An accurate high-resolution map of land cover types is essential when calculating the mixed pixel EF using EFAF. Incorrect specification of the underlying land cover is particularly critical because the EF and AF of sub-pixels are based on the land cover map.

To assess the sensitivity of the land cover map and AE, reference values were obtained from the retrieved data set on 27 July; these values indicate a wider range of phenological conditions and thermal dynamics. Other days had relatively homogeneous phenology conditions and thermal dynamics; at these times, the sensitivity analysis is conservatively estimated. The simple averaged pure-pixel EF was calculated to investigate the sensitivity of the seven main land cover types in the study area, i.e., maize, grass, bare soil, wheat, vegetables, buildings and water bodies. Of these, the EFs of buildings and water bodies were defined as 0 and 1, respectively.

Table 4. Differences in EF and LE caused by incorrect classification

Incorrect classification	EF or LE ($W \cdot m^{-2}$)	Correct classification						
		Maize	Grass	Bare soils	Wheat	Vegetables	Water bodies	Buildings
Maize	EF	0	0.07	0.38	0.5	0.07	-0.06	0.94
	LE	0	22.68	123.14	162.03	22.68	-19.44	304.62
Grass	EF	-0.07	0	0.31	0.43	0	-0.13	0.87
	LE	-22.68	0	100.46	139.35	0	-42.13	281.93
Bare soils	EF	-0.38	-0.31	0	0.12	-0.31	-0.44	0.56
	LE	-123.14	-100.46	0	38.89	-100.46	-142.59	181.47
Wheat	EF	-0.5	-0.43	-0.12	0	-0.43	-0.56	0.44
	LE	-162.03	-139.35	-38.89	0	-139.35	-181.47	142.59
Vegetables	EF	-0.07	0	0.31	0.43	0	-0.13	0.87
	LE	-22.68	0	100.46	139.35	0	-42.13	281.93
Water bodies	EF	0.06	0.13	0.44	0.56	0.13	0	1
	LE	19.44	42.13	142.59	181.47	42.13	0	324.06
Buildings	EF	-0.94	-0.87	-0.56	-0.44	-0.87	-1	0
	LE	-304.62	-281.93	-181.47	-142.59	-281.93	-324.06	0

The "+" and "-" symbols indicate overestimation and underestimation, respectively.

The average AE was 324.05 $W \cdot m^{-2}$ over the entire study area.

Table 4 shows the difference in EF between the correct and incorrect classifications; the "+" and "-" symbols indicate overestimation and underestimation, respectively. The results demonstrate that little errors was introduced by misclassifications among maize, grass and vegetables, because they have similar phenological conditions during the period of high water use efficiency, which is especially true of grass and vegetables because of their similar roughness length.

Conversely, a greater error, with an absolute difference of 0.5 in EF, occurred because of misclassification between wheat and other vegetation types. As ripe wheat changes colour from green to yellow or brown, its water use efficiency decreases; this resulted in a error of 162.03 W m^{-2} for the LE estimation. Additionally, incorrectly classifying bare soils as maize, grass or vegetables (or vice versa) also induced a greater error; the absolute difference in EF ranged from 0.31 to 0.38 and the absolute difference in LE ranged from 100.46 to 123.14 W m^{-2}. However, incorrectly classifying bare soils as wheat (or vice versa) resulted in lower error, with an absolute difference in EF of approximately 0.12.

Furthermore, while misclassifications between water bodies and bare soils could result in a higher error in LE estimation, this rarely occurred because of the unique spectral characteristics of water and bare soils. Similarly, misclassification between buildings and other land cover types would induce a greater error because the EF of buildings was set to 0 in this study.

5 Discussion

The most significant contribution of EFAF is related to its capacity to correct spatial scale errors in the EF of mixed pixels; it can be used to calculate daily ET from daily AE data based on two hypotheses. This attribute could be beneficial in global ET mapping and water resources management compared to models that do not consider spatial scale effects. Validation of the EFAF results against EC measurements across the HiWATER experimental sites demonstrates that EFAF can reproduce the LE of mixed pixels with an RMSE of 1.60 MJ m^{-2} (0.64 mm); without the EFAF, RMSE is 2.47 MJ m^{-2} (0.99 mm). The two hypotheses result in lower error, within 10 W m^{-2} for Hypothesis 1 and 20 W m^{-2} for Hypothesis 2. These results suggest that EFAF is reliable and has a considerable application potential. In particular, EFAF has the following advantages:

(1) EFAF is uniquely able to identify the ET values of different land cover types in mixed pixels. This represents an improvement relative to single-source models that assume homogeneous land cover and two-source models that only distinguish bare surfaces from vegetated surfaces. Single-source models generate significant errors when applied to partially vegetated surfaces because they represent the surface as a single uniform layer (Timmermans et al., 2007). Two-source models are influenced by the characteristics of different vegetation species including canopy height and phenological conditions and can not distinguish other land cover types including water bodies, buildings and ice. In contrast, EFAF functions over heterogeneous surface can identify different land cover types (e.g., maize, grass, bare soil, vegetables, water bodies and buildings) from high resolution land cover images.

(2) EFAF reduces the uncertainties associated with both spatial scale and temporal scale. The EFAF method is based on the EF model, which is widely accepted for temporal extrapolation between data collected a satellite overpass time and daily ET. In the EFAF, the algorithm used to calculate the EF of mixed pixels is based on two hypotheses. The case study results presented in Sect. 4.1 and Sect. 4.2 demonstrate that the EFAF could significantly reduce the errors caused by the heterogeneous surfaces in a watershed located in north-western China, as well as reproducing the daily LE, particularly the spatial distribution of daily LE. Therefore, EFAF can be used for regional, continental or even global applications.

(3) EFAF is easy to apply. In EFAF, calculating the mixed pixel EF only involves determining the AF of sub-pixels, which can be obtained from a high resolution map of land cover types. Furthermore, the module for inhomogeneous surfaces is independent and easy to embed in traditional RS algorithms of heat fluxes; these algorithms were mainly designed to calculate LE or ET under unsaturated conditions and did not consider heterogeneities in the land surface.

(4) EFAF is robust in terms the mechanism of ET, especially through its two hypotheses. Hypothesis 1 is based on the theory of low spatial scale effects for AE. Hypothesis 2 is based on TFL, which ensures the maximum likelihood estimation of ET in land cover, phenology, surface topography and roughness length.

(5) EFAF requires relatively few inputs, at most two or three. The first type of input is remotely sensed ET or LE images with no consideration of the spatial scale effect. These images can be obtained from ET products or calculated using RS algorithms of heat fluxes that were mainly designed to calculate LE or ET under unsaturated conditions and do not

consider heterogeneities in the land surface (including single-source and two-source models). The second type of input is high spatial resolution land cover images, which are readily available. For example, GlobeLand30 is a global land cover data with a 30 m resolution, which can be downloaded free of cost from the following website: https://glc30.tianditu.com. The third type of input is daily AE, which is available directly from LE products in the first type of input or can be calculated using forcing data and heat flux algorithms.

However, similar to other remotely sensed ET models, EFAF has several limitations:

(1) Incorrect classifications directly impact the EF of mixed pixel estimates. As discussed in Sect. 4.4, relatively small errors resulted from the misclassification of vegetation with similar phenological conditions; however, larger errors resulted from the misclassification of vegetation with different phenological conditions and misclassification between vegetation and water bodies. Major errors resulted from the misclassification of buildings, bare soils and vegetation and of buildings, bare soils and water, though this was less common.

(2) LE and EF retrievals are limited to clear-sky conditions. Clouds limit thermal infrared (TIR) observations of land surface temperatures and of the downward shortwave radiation, which control energy partitioning and ET (Bastiaanssen et al., 1998; Allen et al., 2007a; Ershadi et al., 2013). For example, TIR measurements within 1 k uncertainty allow ET estimates to have a relative error within 10% (Hook et al., 2004; Blonquist et al., 2009; Cammalleri et al., 2012; Hulley et al., 2012; Fisher et al., 2013a). If a cloud covers a mixed pixel area, the EFAF can reduce the effects of the cloud, but there will be a large error in the pure pixels covered by clouds.

(3) Mismatch between the footprint of the EC measurement mismatches and the satellite image pixels is likely to increase the uncertainties in validation and create discrepancies between the retrieved LE and EC measurements, which are especially relevant the LE or ET of heterogeneous surfaces. This problem is beyond the scope of this study and should be addressed in future work.

(4) The underlying assumption and starting point of this method is that the pure pixel is really the actual “purity” of the pure pixels; therefore, the EF of pure pixels is representative at least to surrounding the mixed pixels. Only land cover information was used to define pure pixels; therefore, subsurface heterogeneity in pure pixels caused by other aspects (such as variations in the surface variables) may have certain influences on the results. Including additional features in the definition of pure pixels may increase the complexity of the model and the difficulties of its application significantly.

6 Conclusions

This study aimed to develop an operational model for estimating the daily ET of heterogeneous surfaces that is capable of reproducing daily ET with reasonable accuracy but easy to apply. A simple model (EFAF) was developed to calculate the ET of mixed pixels based on the EF and AF from a high-resolution map of land cover types. Temporal scale extrapolation of the instantaneous latent heat flux (LE) at satellite overpass time to daily ET depends on the widely accepted EF model. For heterogeneous surfaces, an equation was derived to calculate the EF of mixed pixels based on two key hypotheses.

Hypothesis 1 states that the AE of each sub-pixel is approximately equal to that of any other sub-pixels in the same mixed pixel within an acceptable margin of error and is equivalent to the AE of the mixed pixel. Hypothesis 2 states that the EF of each sub-pixel is equal to the EF of the nearest pure pixel(s) of the same land cover type. Determination of the EF of mixed pixels also depends on high-resolution land cover data to calculate the AF and the position of pure pixels. Daily ET is calculated by combining the EF of mixed pixels and the daily AE, which can be obtained from energy flux products or retrieved using forcing data.

The EFAF method was applied to an artificial oasis in the midstream of the Heihe River using HJ-1B satellite data at a spatial resolution of 300 m. The results show that the EFAF can improve the accuracy of daily ET estimation relative to the lumped method. Validations at 12 sites with EC systems during 9 days of HJ-1B overpass showed that the R^2 increased from 0.62 to 0.82, the RMSE decreased from 2.47 to 1.60 MJ m^{-2} (0.99 to 0.64 mm), and the MBE decreased from 1.92 to 1.18 MJ m^{-2} (0.77 to 0.47 mm), which are a significant improvements.

Error analysis suggests that the two key hypotheses of the model induce relatively little error. The expected value of the absolute error in AE due to Hypothesis 1 was within 7 W m^{-2}, and the maximum RMSE of the EF for each land cover type due to Hypothesis 2 was 0.047 for maize, 0.055 for grass, 0.048 for bare soil and 0.059 for vegetables. However, we note that the results from this study are probably limited and the model should be tested and validated in other areas.

In brief, the estimated LE of pure pixels is considered accurate and used to calculate its EF. Based on this parameter, the equation for the EF of mixed pixels was established with two key hypothesises. A finer resolution land cover map is needed to search for “pure pixels” as well as to calculate area ratio of each land cover in mixed pixels. This process can derive the daily ET from coarse resolution remote sensing data with acceptable accuracy, and no other finer resolution data are needed in the EFAF method. Thus, this method may be applicable on a daily basis with daily coarse resolution imagery, such as MODIS, and only one finer resolution land cover map for a certain length of time, i.e., a week, month or season, as long as the land cover change is not extreme in that period. It is quite convenient for regional applications that need long-term running. This method can also be used as a correcting technique for LE estimations or remote sensing products since calculating the EF of mixed pixels is carried out after calculating heat fluxes that could be based on an energy balance equation or other methods at the very beginning. The application of the EFAF could be limited with very coarse resolution data since the probability of pure pixels becomes very low. In these circumstances, a compromise may have to be made between the “purity” of pure pixels and the searching distance for the pure pixels. Additional investigations are needed to evaluate the performance of this method with different remote sensing products.

Acknowledgments

We thank all of the scientists and engineers who took part in the HiWATER experiment. We also thank the editors and reviewers for their generous help in revising the paper. This study was jointly supported by the Chinese Natural Science

Foundation Project (grant no. 41871252 and no. 41371360) and the Special Fund from the Chinese Academy of Sciences (KZZD-EW-TZ-18).

References

Allen, R. G., Tasumi, M., Morse, A., Trezza, R., Wright, J. L., Bastiaanssen, W., Kramber, W., Lorite, I., and Robison, C. W.: Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-applications, Journal of Irrigation & Drainage Engineering, 133, 395-406, 2007a.

Allen, R. G., Tasumi, M., and Trezza, R.: Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Model, Journal of irrigation and drainage engineering, 133, 380-394, 2007b.

Anderson, M., Kustas, W., Norman, J., Hain, C., Mecikalski, J., Schultz, L., González-Dugo, M., Cammalleri, C., d'Urso, G., and Pimstein, A.: Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery, Hydrology and Earth System Sciences, 15, 223-239, 2011.

Anderson, M. C., Allen, R. G., Morse, A., and Kustas, W. P.: Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources, Remote Sensing of Environment, 122, 50-65, 2012.

Bastiaanssen, W. G., Menenti, M., Feddes, R., and Holtslag, A.: A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation, Journal of hydrology, 212, 198-212, 1998.

Bisht, G., Venturini, V., Islam, S., and Jiang, L.: Estimation of the net radiation using MODIS (Moderate Resolution Imaging Spectroradiometer) data for clear sky days, Remote Sensing of Environment, 97, 52-67, 2005.

Bin, L., and Roni, A.: The Impact of Spatial Variability of Land-Surface Characteristics on Land-Surface Heat Fluxes, Journal of Climate, 7, 527-537, 1994.

Blonquist Jr, J., Norman, J., and Bugbee, B.: Automated measurement of canopy stomatal conductance based on infrared temperature, Agricultural and Forest Meteorology, 149, 2183-2197, 2009.

Blyth, E. M., and Harding, R. J.: Application of aggregation models to surface heat flux from the Sahelian tiger bush, Agricultural & Forest Meteorology, 72, 213-235, 1995.

Bonan, G. B., Pollard, D., and Thompson, S. L.: Influence of Subgrid-Scale Heterogeneity in Leaf Area Index, Stomatal Resistance, and Soil Moisture on Grid-Scale Land–Atmosphere Interactions, Journal of Climate, 6, 1882-1897, 1993.

Bonan, G. B., Levis, S., Kergoat, L., and Oleson, K. W.: Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models, Global Biogeochemical Cycles, 16, 5-1-5-23, 2002.

Cammalleri, C., Anderson, M., Ciraolo, G., D'urso, G., Kustas, W., La Loggia, G., and Minacapilli, M.: Applications of a remote sensing-based two-source energy balance algorithm for mapping surface fluxes without in situ air temperature observations, Remote Sensing of Environment, 124, 502-515, 2012.

Cammalleri, C., Anderson, M. C., Gao, F., Hain, C. R., and Kustas, W. P.: A data fusion approach for mapping daily evapotranspiration at field scale, Water Resources Research, 2013.

Carlson, T.: An overview of the" triangle method" for estimating surface evapotranspiration and soil moisture from satellite imagery, Sensors, 7, 1612-1629, 2007.

Chávez, J. L., Neale, C. M. U., Prueger, J. H., and Kustas, W. P.: Daily evapotranspiration estimates from extrapolating instantaneous airborne remote sensing ET values, Irrigation Science, 27, 67-81, 2008.

Chen, J. M.: Spatial scaling of a remotely sensed surface parameter by contexture, Remote Sensing of environment, 69, 30-42, 1999.

El Maayar, M., and Chen, J. M.: Spatial scaling of evapotranspiration as affected by heterogeneities in vegetation, topography, and soil texture, Remote Sensing of environment, 102, 33-51, 2006.

Ershadi, A., Mccabe, M. F., Evans, J. P., and Walker, J. P.: Effects of spatial aggregation on the multi-scale estimation of evapotranspiration, Remote Sensing of Environment, 131, 51-62, 2013.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., and Han, D.: Gap filling strategies for defensible annual sums of net ecosystem exchange, Agricultural & Forest Meteorology, 107, 43-69, 2001.

Fisher, J., Mallick, K., Lee, J., Hulley, G., Hughes, C., and Hook, S.: Uncertainty in evapotranspiration from uncertainty in land surface temperature, American Meteorological Society, 2013.

Garrigues, S., Allard, D., Baret, F., and Weiss, M.: Quantifying spatial heterogeneity at the landscape scale using variogram models, Remote Sensing of Environment, 103, 81-96, 2006.

Gottschalk, L., Batchvarova, E., Gryning, S. E., Lindroth, A., Melas, D., Motovilov, Y., Frech, M., Heikinheimo, M., Samuelsson, P., and Grelle, A.: Scale aggregation — comparison of flux estimates from NOPEX, Agricultural & Forest Meteorology, s 98–99, 103-119, 1999.

Gu, J., Li, X., and Huang, C.: Land Cover Classification in Heihe River Basin with Time Series - MODIS NDVI Data, International Conference on Fuzzy Systems and Knowledge Discovery, 477-481, 2008.

Ha, W., Gowda, P. H., and Howell, T. A.: A review of downscaling methods for remote sensing-based irrigation management: part I, Irrigation Science, 31, 831-850, 10.1007/s00271-012-0331-7, 2013.

He, L., Chen, J. M., Pisek, J., Schaaf, C. B., and Strahler, A. H.: Global clumping index map derived from the MODIS BRDF product, Remote Sensing of Environment, 119, 118-130, 2012.

Hong, S. H., Hendrickx, J. M. H., and Borchers, B.: Up-scaling of SEBAL derived evapotranspiration maps from Landsat (30 m) to MODIS (250 m) scale, Journal of Hydrology, 370, 122-138, 2009.

Hook, S. J., Chander, G., Barsi, J. A., Alley, R. E., Abtahi, A., Palluconi, F. D., Markham, B. L., Richards, R. C., Schladow, S. G., and Helder, D. L.: In-flight validation and recovery of water surface temperature with Landsat-5 thermal infrared data using an automated high-altitude lake validation site at Lake Tahoe, IEEE Transactions on Geoscience and Remote Sensing, 42, 2767-2776, 2004.

Hu, Z. L., and Islam, S.: A framework for analyzing and designing scale invariant remote sensing algorithms, Geoscience and Remote Sensing, IEEE Transactions on, 35, 747-755, 1997.

Hu, G., and Jia, L.: Monitoring of evapotranspiration in a semi-arid inland river basin by combining microwave and optical remote sensing observations, Remote Sensing, 7, 3056-3087, 2015.

Hulley, G. C., Hughes, C. G., and Hook, S. J.: Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and MODIS thermal infrared data, Journal of Geophysical Research: Atmospheres, 117, 2012.

Jackson, R. D., Hatfield, J. L., Reginato, R. J., Idso, S. B., and Jr, P. J. P.: Estimation of daily evapotranspiration from one time-of-day measurements, Agricultural Water Management, 7, 351-362, 1983.

Jiang, B., Liang, S., Townshend, J. R., and Zan, M. D.: Assessment of the Radiometric Performance of Chinese HJ-1 Satellite CCD Instruments, IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 6, 840-850, 2013.

Jiao, J., Xin, X., Shanshan, Y. U., Zhou, T., and Peng, Z.: Estimation of surface energy balance from HJ-1 satellite data, Journal of Remote Sensing, 18, 1048-1058, 2014.

Jin, Z., Tian, Q., Chen, J. M., and Chen, M.: Spatial scaling between leaf area index maps of different resolutions, Journal of Environmental Management, 85, 628, 2007.

Kato, S., and Yamaguchi, Y.: Analysis of urban heat-island effect using ASTER and ETM+ Data: Separation of anthropogenic heat discharge and natural heat radiation from sensible heat flux, Remote Sensing of Environment, 99, 44-54, 2005.

Kimball, J. S., Running, S. W., and Saatchi, S. S.: Sensitivity of boreal forest regional water flux and net primary production simulations to sub‐grid‐scale land cover complexity, Journal of Geophysical Research Atmospheres, 104, 27789-27801, 1999.

Kormann, R., and Meixner, F. X.: An Analytical Footprint Model For Non-Neutral Stratification, Boundary-Layer Meteorology, 99, 207-224, 2001.

Kustas, W. P., and Norman, J. M.: Evaluating the effects of subpixel heterogeneity on pixel average fluxes, Remote Sensing of Environment, 74, 327-342, 2000.

Kustas, W. P., Moran, M. S., and Meyers, T. P.: The Bushland Evapotranspiration and Agricultural Remote Sensing Experiment 2008 (BEAREX08) Special Issue, Advances in Water Resources, 50, 1-3, 2012.

Kustas, W. P., Norman, J. M., Anderson, M. C., and French, A. N.: Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship, Remote Sensing of Environment, 85, 429-440, 2003.

Li, H., Liu, Q., Zhong, B., Du, Y., Wang, H., and Wang, Q.: A single-channel algorithm for land surface temperature retrieval from HJ-1B/IRS data based on a parametric model, Geoscience and Remote Sensing Symposium, 2010, 2448-2451,

Li, L., Xin, X. Z., Su, G. L., and Liu, Q. H.: Photosynthetically active radiation retrieval based on HJ-1A/B satellite data, Science China Earth Sciences, 53, 81-91, 2011.

Li, X., Cao, C., and Chang, C.: The first law of geography and spatial temporal proximity, Chinese Journal of Nature, 29, 69-71, 2007.

Li, X., Cheng, G., Liu, S., Xiao, Q., Ma, M., Jin, R., Che, T., Liu, Q., Wang, W., and Qi, Y.: Heihe Watershed Allied Telemetry Experimental Research (HiWATER): Scientific Objectives and Experimental Design, Bulletin of the American Meteorological Society, 94, 1145-1160, 2013.

Li X and W. Y.: Prospects on future developments of quantitative remote sensing, Acta Geographica Sinica, 68, 1163-1169, 10.11821/dlxb201309001, 2013.

Li, Z. L., Tang, B. H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A.: Satellite-derived land surface temperature: Current status and perspectives, Remote Sensing of Environment, 131, 14-37, 2013.

Liang, S., Stroeve, J., and Box, J. E.: Mapping daily snow/ice shortwave broadband albedo from Moderate Resolution Imaging Spectroradiometer (MODIS): The improved direct retrieval algorithm and validation with Greenland in situ measurement, Journal of Geophysical Research Atmospheres, 110, -, 2005.

Liu, D., Li, J., Yu, Q., Tong, X., and Ouyang, Z.: Energy balance closure and its effects on evapotranspiration measurements with the eddy covariance technique in a cropland, Acta Ecologica Sinica, 32, 5309-5317, 2012.

Liu, Q., Wang, L., Qu, Y., Liu, N., Liu, S., Tang, H., and Liang, S.: Preliminary evaluation of the long-term GLASS albedo product, International Journal of Digital Earth, 6, 69-95, 2013.

Liu, S., Xu, Z., Song, L., Zhao, Q., Ge, Y., Xu, T., Ma, Y., Zhu, Z., Jia, Z., and Zhang, F.: Upscaling evapotranspiration measurements from multi-site to the satellite pixel scale over heterogeneous land surfaces, Agricultural & Forest Meteorology, 230, 97-113, 2016.

Liu, S. M., Xu, Z. W., Wang, W. Z., Jia, Z. Z., Zhu, M. J., Bai, J., and Wang, J. M.: A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy balance closure problem, Hydrology & Earth System Sciences, 15, 1291-1306, 2011.

Long, D., and Singh, V. P.: A two-source trapezoid model for evapotranspiration (TTME) from satellite imagery, Remote Sensing of Environment, 121, 370-388, 2012.

Ma, M., and Veroustraete, F.: Interannual variability of vegetation cover in the Chinese Heihe River Basin and its relation to meterological parameters, International Journal of Remote Sensing, 27, 3473-3486, 2006.

McCabe, M. F., and Wood, E. F.: Scale influences on the remote estimation of evapotranspiration using multiple satellite sensors, Remote Sensing of Environment, 2006.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E., and Franz, T. E.: The future of Earth observation in hydrology, Hydrology and Earth System Sciences, 21, 3879, 2017.

Miller, H. J.: Tobler's First Law and Spatial Analysis, Annals of the Association of American Geographers, 94, 284–289, 2004.

Moran, M. S., Humes, K. S., and Pinter Jr, P. J.: The scaling characteristics of remotely-sensed variables for sparsely-vegetated heterogeneous landscapes, Journal of Hydrology, 190, 337-362, 1997.

Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W.: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sensing of Environment, 111, 519-536, 2007.

Mu, Q., Zhao, M., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sensing of Environment, 115, 1781-1800, 2011.

Nichols, W. E., and Cuenca, R. H.: Evaluation of the evaporative fraction for parameterization of the surface energy balance, Water Resources Research, 29, 3681-3690, 2010.

Nilson, T.: A theoretical analysis of the frequency of gaps in plant stands, Agricultural Meteorology, 8, 25-38, 1971.

Norman, J. M., Kustas, W. P., and Humes, K. S.: Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature, Agricultural & Forest Meteorology, 77, 263-293, 1995.

Pan, Z., and Liu, G.: Evapotranspiration Research of Yellow River Delta Using Remote Sensing Method, Geo-information Science, 3, 91-96, 2003.

Peng, Z., Xin, X., Jiao, J. J., Zhou, T., and Liu, Q.: Remote sensing algorithm for surface evapotranspiration considering landscape and statistical effects on mixed pixels, Hydrology & Earth System Sciences, 20, 4409-4438, 2016.

Sánchez, J., Kustas, W., Caselles, V., and Anderson, M.: Modelling surface energy fluxes over maize using a two-source patch model and radiometric soil and canopy temperature observations, Remote sensing of Environment, 112, 1130-1143, 2008.

Seguin, B., Becker, F., Phulpin, T., Gu, X., Guyot, G., Kerr, Y., King, C., Lagouarde, J., Ottlé, C., and Stoll, M.: IRSUTE: a minisatellite project for land surface heat flux estimation from field to regional scale, Remote Sensing of Environment, 68, 357-369, 1999.

Sharma, V., Kilic, A., and Irmak, S.: Impact of scale/resolution on evapotranspiration from Landsat and Modis images, Water Resources Research, 52, 1207-1221, 2016.

Su, Z.: The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes, Hydrology & Earth System Sciences, 6, 85-99, 2002.

Sugita, M., and Brutsaert, W.: Daily evaporation over a region from lower boundary layer profiles measured with radiosondes, Water Resources Research, 27, 747-752, 1991.

Tian, Y., Woodcock, C. E., Wang, Y., Privette, J. L., Shabanov, N. V., Zhou, L., Zhang, Y., Buermann, W., Dong, J., and Veikkanen, B.: Multiscale analysis and validation of the MODIS LAI product: I. Uncertainty assessment, Remote Sensing of Environment, 83, 414-430, 2002.

Tobler, W.: On the First Law of Geography: A Reply, Annals of the Association of American Geographers, 94, 304-310, 2004.

Valor, E., and Caselles, V.: Mapping land surface emissivity from NDVI: application to European, African, and South American areas, Remote Sensing of Environment, 57, 167-184, 1996.

Wu, J., Ding, Y., Wang, G., Yamazaki, Y., and Kubota, J.: Evapotranspiration of intercropping field in an artifical oasis in arid region, Transactions of the Chinese Society of Agricultural Engineering, 22, 16-20, 2006.

Xin, X., Liu, Y., and Liu, Q.: Spatial-scale error correction methods for regional fluxes retrieval using MODIS data, Journal of Remote Sensing, 16, 207-231, 2012.

Xu, Z., Liu, S., Li, X., Shi, S., Wang, J., Zhu, Z., Xu, T., Wang, W., and Ma, M.: Intercomparison of surface energy flux measurement systems used during the HiWATER㎝USOEXE, Journal of Geophysical Research Atmospheres, 118, 13-13,157, 2013.

Yang, K., and Wang, J. M.: A temperature prediction-correction method for estimating surface soil heat flux from soil temperature and moisture data, Science China Earth Sciences, 51, 721-729, 2008.

Yu, W., Li, J., Liu, Q., Zeng, Y., Yin, G., Zhao, j., and Xu, B.: Extraction and Analysis of Land Cover Heterogeneity over China, Advances in Earth Science, 31, 1067-1077, 2016.

Zhang, L., and Lemeur, R.: Evaluation of daily evapotranspiration estimates from instantaneous measurements, Agricultural & Forest Meteorology, 74, 139-154, 1995.

Zhang, X., Zhao, X., Liu, G., Qian, K., and Wu, D.: Radioactive Quality Evaluation and Cross Validation of Data from the HJ-1A/B Satellites' CCD Sensors, Sensors, 13, 8564-8576, 2013.

Zhong, B., Ma, P., Nie, A., Yang, A., Yao, Y., Lü, W., Zhang, H., and Liu, Q.: Land cover mapping using time series HJ-1/CCD data, Science China Earth Sciences, 57, 1790-1799, 2014a.

Zhong, B., Zhang, Y., Du, T., Yang, A., Lv, W., and Liu, Q.: Cross-Calibration of HJ-1/CCD Over a Desert Site Using Landsat ETM+ Imagery and ASTER GDEM Product, IEEE Transactions on Geoscience & Remote Sensing, 52, 7247-7263, 2014b.