Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

Hydrology and Earth System Sciences
Discussions

Estimating Daily Evapotranspiration Based on a Model of Evapotranspiration 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 evapotranspiration 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 bias 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 sites 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 $MJ \cdot m^{-2}$, and the mean bias error (MBE) decreased from 1.92 to 1.18 $MJ \cdot m^{-2}$.

Index Terms: Evapotranspiration; Heterogeneous surface; Temporal scale extrapolation; Evapotranspiration 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), A temperature-sharpening and flux aggregation scheme (TSFA) (Peng et al., 2016) have been developed to monitor land-atmosphere energy balance flux interactions; in most cases, spatially

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

Hydrology and Earth System Sciences
Discussions

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. These models can be used to assess surface fluxes from fine-resolution remote sensing (RS) data such as 30 m for Landsat, but larger biases will result when using data of much ~~coarser resolution~~ from sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) or the Advanced Very High Resolution Radiometer (AVHRR) about at 1 km resolution. These biases mainly result from non-linear operation model use and surface heterogeneities; of the two, scaling problems caused by surface heterogeneities are generally dominant (Garrigues et al., 2006; Jin et al., 2007; Tian et al., 2002; Gottschalk et al., 1999; Xin et al., 2012). To improve the accuracy of ET estimates, it is necessary to fully understand the scale effects of ET estimates and to develop ways of correcting these scale effects. To address the scale effect on energy fluxes, many studies have compared lumped calculations that aggregate fine resolution parameters to a coarser resolution with distributed calculations retrieved at fine resolutions and then aggregated to a coarser resolution; other studies have noted discrepancies between multi-sensor data aggregations. Moran et al. (1997) found a significant bias 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, whether the physical mechanism is suitable for application, what is the applicable condition and how to correct the scale effects are the three critical issues for the 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 biases 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 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. On the other hand, this method is difficulty to describe the surface information accurately due to land cover maps at coarse resolution.

Each of the above approaches reduces the bias 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 evapotranspiration fraction (EF) method (Nichols and Cuenca, 2010; Sugita and Brutsaert, 1991) and the reference evapotranspiration 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 evapotranspiration with the 24-hour trend in solar radiation; the function expresses the relationship between instantaneous evapotranspiration and daily evapotranspiration. The EF method, which is the most widely used, extrapolates the instantaneous EF to the daily EF based on the characteristics of EF which remain constant over one day. The reference evapotranspiration fraction method assumes that the instantaneous reference evapotranspiration fraction which is calculated as the ratio of the computed instantaneous ET at satellite overpass time from each pixel to reference crop's (such as alfalfa's) ET is the same as the average reference evapotranspiration fraction over the 24 h average; it then uses the reference crop's accumulated daily evapotranspiration to obtain the daily evapotranspiration. Chavez et al. (2008) compared different evapotranspiration temporal scale extrapolation methods and found that the EF method generates values most consistent with 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 EF of mixed pixels only needs area fraction (AF) of sub-pixels, which could be obtained from a high-resolution land-cover type map. The model was applied to an artificial oasis in the midstream of 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 carry on the key step of the model, i.e., correction of mixed pixels 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.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

Hydrology and Earth System Sciences
Discussions

2 Methodology

2.1 LE algorithm

The lumped calculation method (Peng et al., 2016) involves the following steps: (1) resampling finer resolution land surface parameters to coarser resolution and (2) estimating coarser resolution fluxes from parameters measured at a coarser resolution. The method does not consider scaling effects.

In this paper, a widely used one-source energy balance model was used to estimate the following energy flux components: net radiation (Rn), soil heat flux (G), sensible heat flux (H) and LE (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. Further details can be available in Peng et al. (2016).

2.2 The EF of mixed pixels

Various methods use EF to estimate daily ET based on RS data (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). The EF is the ratio of LE and AE (Rn-G), as follows:

$$EF = \frac{LE}{R_n - G} , \quad (4)$$

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

Hydrology and Earth System Sciences
Discussions

Research has shown that the EF remains constant over time and is thus well suited to denote the status of various components of the energy balance equation. Because the EF remains constant during the day (Nichols and Cuenca, 2010; Sugita and Brutsaert, 1991), it can be used to extrapolate the temporal scale of ET. Additionally, the EF can be used to correct scaling effects of the LE.

Due to land surface heterogeneity, the estimation of fluxes may be influenced by scale effects at a coarse resolution; these effects generate retrieval bias and require correction. 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:

$$LE = \sum s_i \, LE_i = \sum [s_i \cdot \frac{LE_i}{(Rn-G)_i} \cdot (Rn - G)_i] \,, \tag{5}$$

where LE denotes the accurate LE of mixed pixels, s_i denotes the AF of sub-pixel i, and LE_i denotes the LE of sub-pixel i. Eq. (4) and (5) can be combined as follows:

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

In Eq. (6), EF_i and $(Rn-G)_i$ denote the EF and AE of sub-pixel i in a certain mixed pixel respectively.

Here, Hypothesis 1 is proposed 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, Eq. (6) can be transformed in to the following expression:

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

where $\widehat{LE}$ denotes the latent heat flux in mixed pixels based on Hypothesis 1. There is a minor difference between $\widehat{LE}$ and accurate LE.

Rearranging Eq. (7) yields the following:

$$\frac{\widehat{LE}}{(Rn-G)} = \sum (s_i \cdot EF_i), \tag{8}$$

The EF_i values of different sub-pixels are required to perform the correction process. Thus, Hypothesis 2 is proposed:

the EF of each sub-pixel is equal to the EF of the nearest pure pixel(s) of the same land cover type.

This assumption 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". The EF of the sub-pixel i may be expressed using Hypothesis 2, because it ensures that the most similar conditions, phenological patterns and physical characteristics exist between the EF of the sub-pixel land class and the chosen EF.

Combining Eq. (7) and Eq. (8) yields the following equation:

$$\widehat{EF} = \sum (s_i \cdot EF_i), \tag{9}$$

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

where $\widetilde{EF}$ denotes the EF of the mixed pixel including scale effects based on Hypotheses 1 and 2.

Therefore, Eq. (7) may be reduced as above to the following:

$$\widetilde{LE} = (Rn - G) \cdot \widetilde{EF}, \tag{10}$$

The process described above suggests two conclusions: (1) the EF of a mixed pixel including scaling effects can be calculated from the area-weighted EF_i of its sub-pixels, and (2) the LE of a mixed pixel including scaling effects can be calculated by multiplying the $\widetilde{EF}$ in Eq. (9) with its AE.

2.2 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:

$$\frac{\mathrm{LE_{daily}}}{(\mathrm{R_n-G})_{\mathrm{daily}}} = \frac{\mathrm{LE_{inst}}}{(\mathrm{R_n-G})_{\mathrm{inst}}} = \mathrm{EF_{inst}}, \tag{11}$$

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

where the subscripts “daily” and “inst” indicate daily cumulative and instantaneous values, respectively. To calculate daily total evapotranspiration from Eq. (12), it is necessary to determine the EF and the daytime total available energy (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:

$$\mathrm{DANR} = 2 * \mathrm{Rn_{inst}}/\pi \sin\left[\left(\frac{\mathrm{t_{ovp}}-\mathrm{t_{rise}}}{\mathrm{t_{set}}-\mathrm{t_{rise}}}\right)\pi\right], \tag{13}$$

$$\mathrm{Rn_{daily}} = \int \mathrm{DANR}\, \mathrm{dt}, \tag{14}$$

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 (1) calculation of LE without scale effect, (2) calculation of the EF of mixed pixels, and (3) extrapolation of the temporal scale (Fig. 1).

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

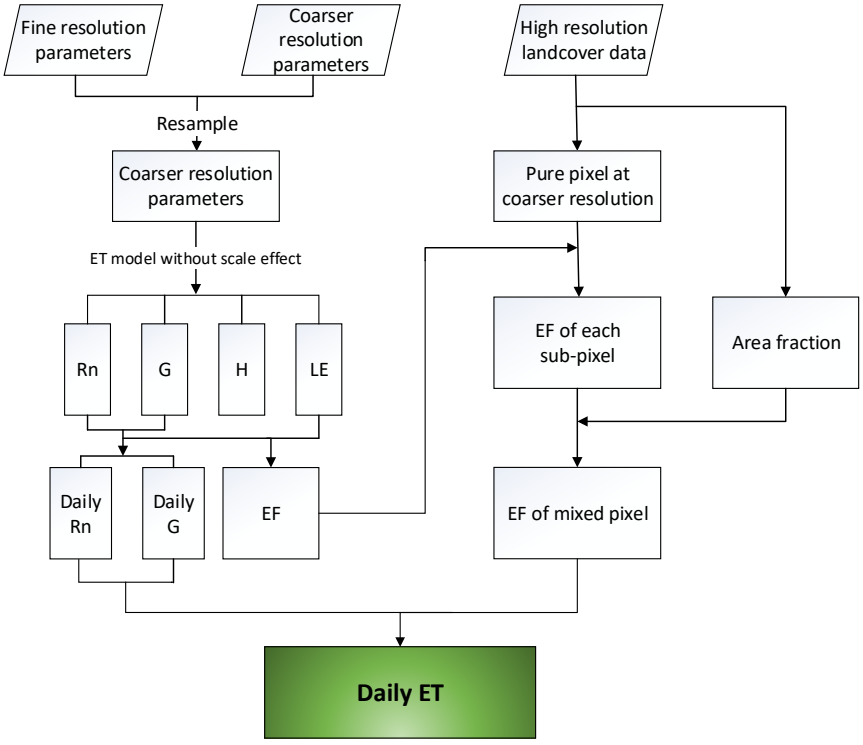

Figure 1: Flowchart of the EFAF, where trapezoids represent input variables or parameters, and rectangles represent variables or parameters. The inputs of EFAF encompass remotely sensed variables or parameters and meteorological forcing dataset. The abbreviations are defined as follows: Rn: net radiation; G: soil heat flux; H: sensible heat flux; LE: latent heat flux; EF: evapotranspiration 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. 2). The Heihe River watershed has a land surface area of approximately 128,000 km² and 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 km². 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 evapotranspiration 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).

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

Hydrology and Earth System Sciences
Discussions

The percentage of the numbers of land cover types (Yu et al., 2016) (Fig. 3) for the study area were extracted on a 300 m scale with 30 m land cover classifications. 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.

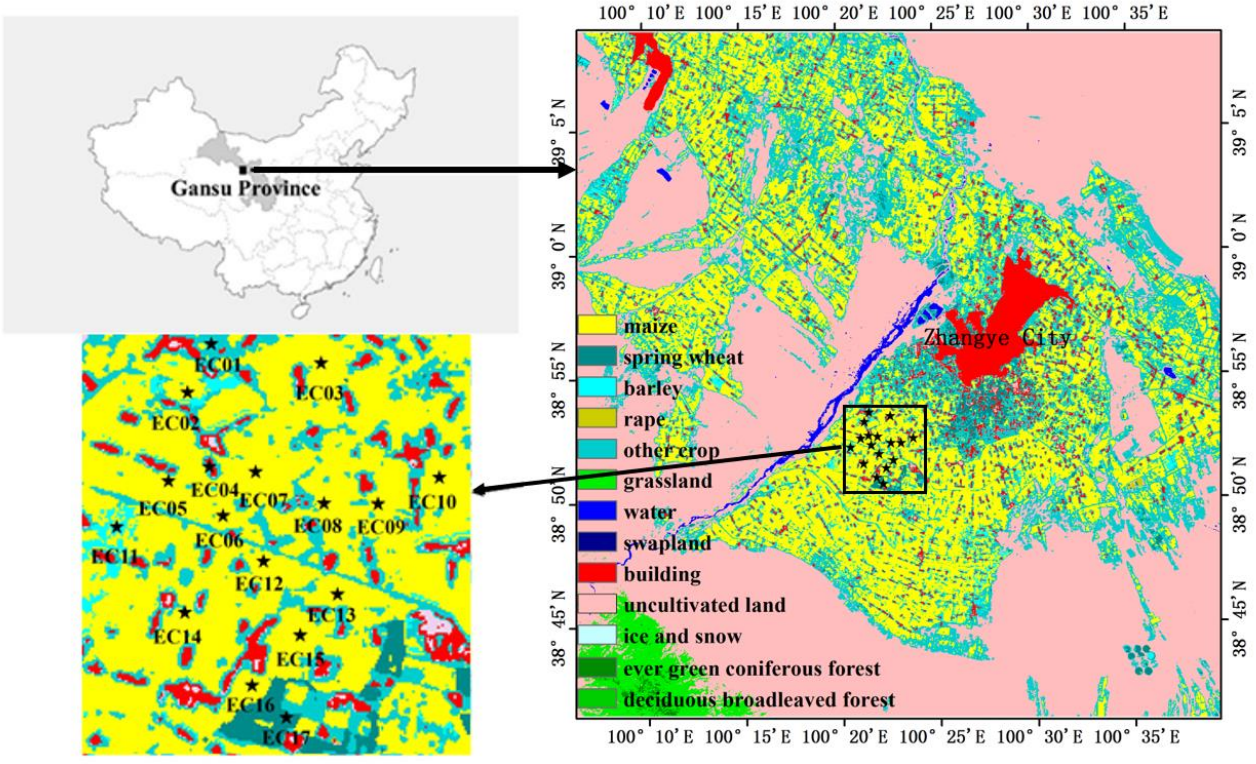

Figure 2: Distribution of in situ stations and land use classification at our study area.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

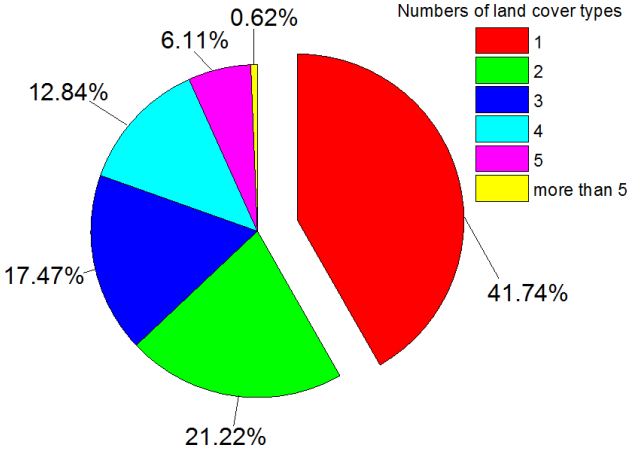

Figure 3: The percentage of the numbers of land cover types for the study at 300 m scale with 30 m land cover images.

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 evapotranspiration 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 of estimating ET for can only be applicable to the ~~sky-clear~~ conditions. Therefore, the satellite data selected for the study area were collected under clear or partly cloudy conditions based on data quality metrics. 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 evapotranspiration 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 vegetable (according to the experiential knowledge, but 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 datasete

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. 2.

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 was based on 30 min intervals; additional information can be found in previous reports (Liu et al., 2011; Liu et al., 2016; Xu et al., 2013). Soil heat fluxes were measured using three soil heat plates at a depth of 6 cm below the surface at each site. The surface soil heat flux was calculated using the method proposed by Yang and Wang (2008) and based on the soil temperature and moisture above the plates.

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 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. 4) and 22 August (Fig. 5), 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.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

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. 4 and 5). For example, Area I in Fig. 6 represents the city of Zhangye, and Area II in Fig. 6 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.

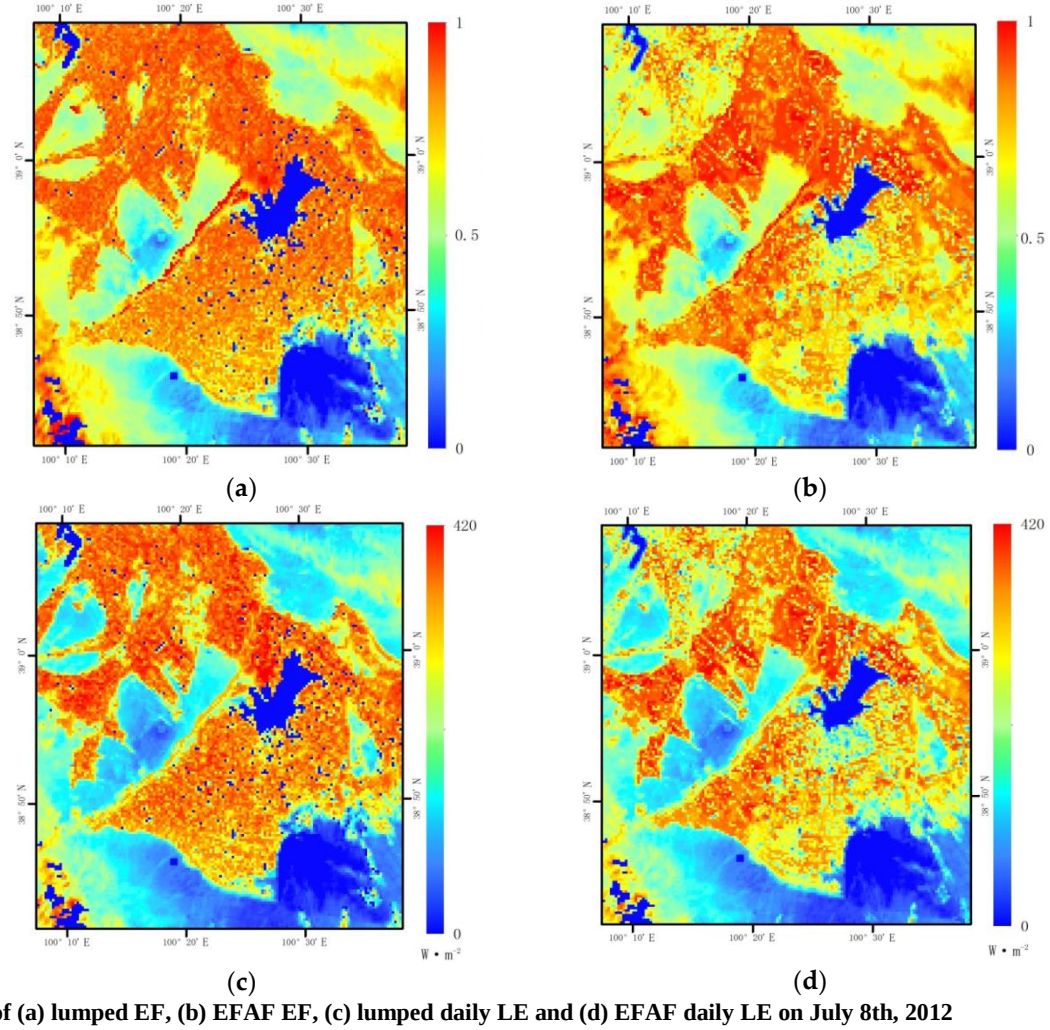

Figure 4. Maps of (a) lumped EF, (b) EFAF EF, (c) lumped daily LE and (d) EFAF daily LE on July 8th, 2012

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

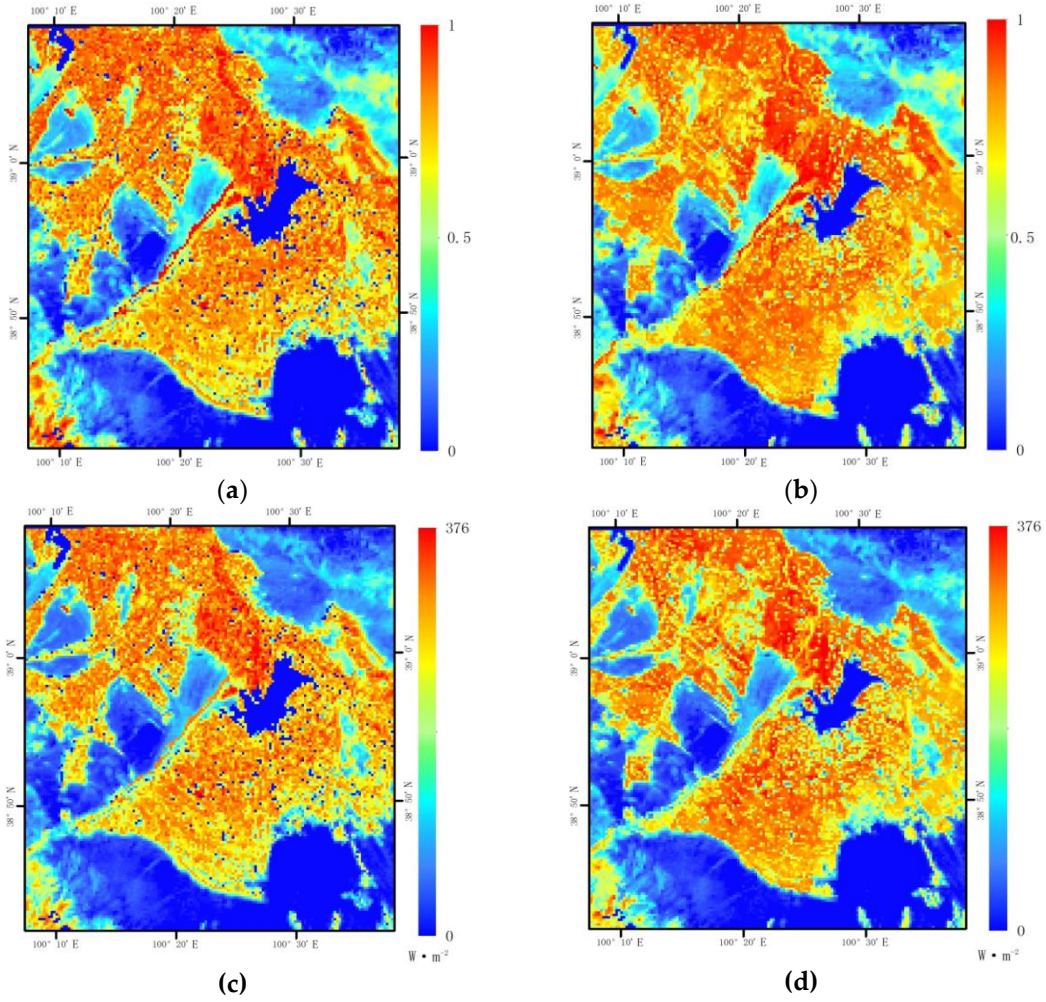

Figure 5. Maps of (a) lumped EF, (b) EFAF EF, (c) lumped daily LE and (d) EFAF daily LE on August 22nd, 2012

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. 6, 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. 6), 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.

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. 6), pixels dominated by buildings and including cropland and pixels dominated by cropland and including buildings account for 20 % and 80 %, respectively, and the spatially

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

averaged daily LE decreased from 8.98 to 7.39 MJ·m^{-2} on 8 July 2012. However, for pixels dominated by buildings, the spatially averaged daily LE increased from 0 to 4.70 MJ·m^{-2}.

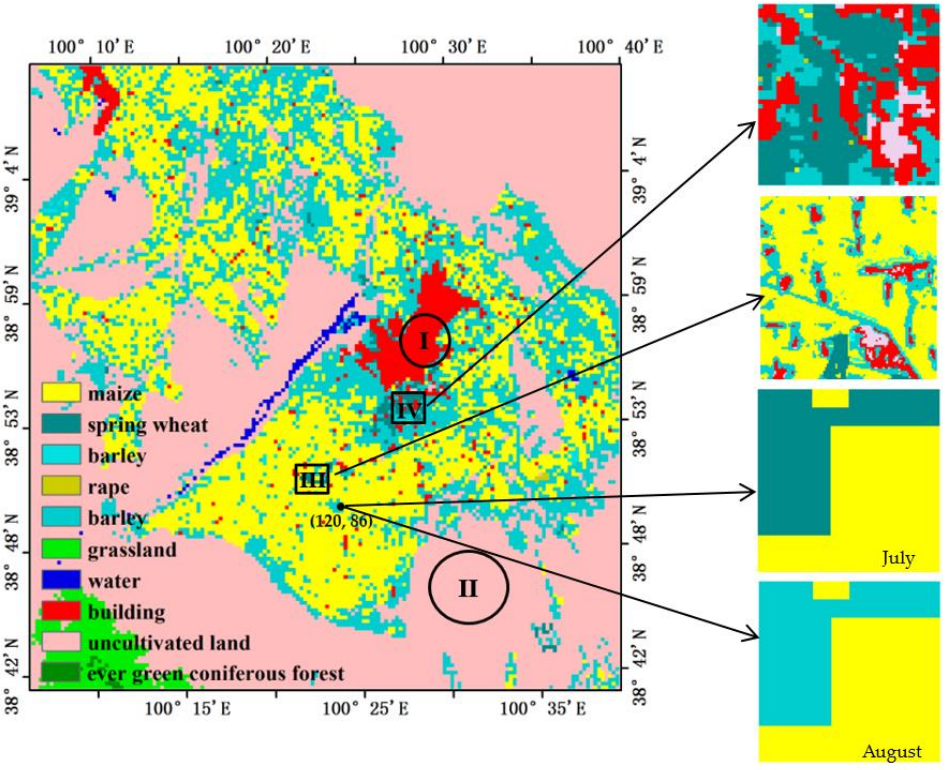

Figure 6. Land cover maps of this study in 300 m resolution and some region in 30 m resolution. Area I and II represent the area of pure pixels, Area III and IV represent the area of mixed pixels, and (120, 86) represents an example point of different land cover types in different months.

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. 6) 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}. 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}. 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. 5 (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}; and the mean bias error (MBE) decreased from 1.92 to 1.18 MJ·m^{-2}

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}, relative to the lumped LE, with an RMSE of 1.72 MJ·m^{-2}. 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}, relative to the lumped LE, which had an RMSE of 1.89 MJ·m^{-2}. 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.

Table 3. In situ validation results for daily LE

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

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-bias by ensuring that the retrieved LE can be assessed against each EC tower record individually (Fig. 7). The results (Fig. 7) 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 $MJ \cdot m^{-2}$ (an absolute decrease of 2.64 $MJ \cdot m^{-2}$); 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 $MJ \cdot m^{-2}$. Likewise, the difference of them against the EC measurements had also declined from 4.12 $MJ \cdot m^{-2}$ to 2.32 $MJ \cdot m^{-2}$. 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. 7 (d), with values of 8.36 $MJ \cdot m^{-2}$ on 27 July and 9.33 $MJ \cdot m^{-2}$ on 3 August. Even after the EFAF method was applied, these values were 5.20 $MJ \cdot m^{-2}$ and 4.59 $MJ \cdot m^{-2}$, 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 $MJ \cdot m^{-2}$, respectively, while the retrieved Rn of the pixels was 18.14 and 18.80 $MJ \cdot m^{-2}$, respectively.

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 on 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 $MJ \cdot m^{-2}$. Compared to the value of 16.52 $MJ \cdot m^{-2}$ found for EC, the EFAF LE was more accurate.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

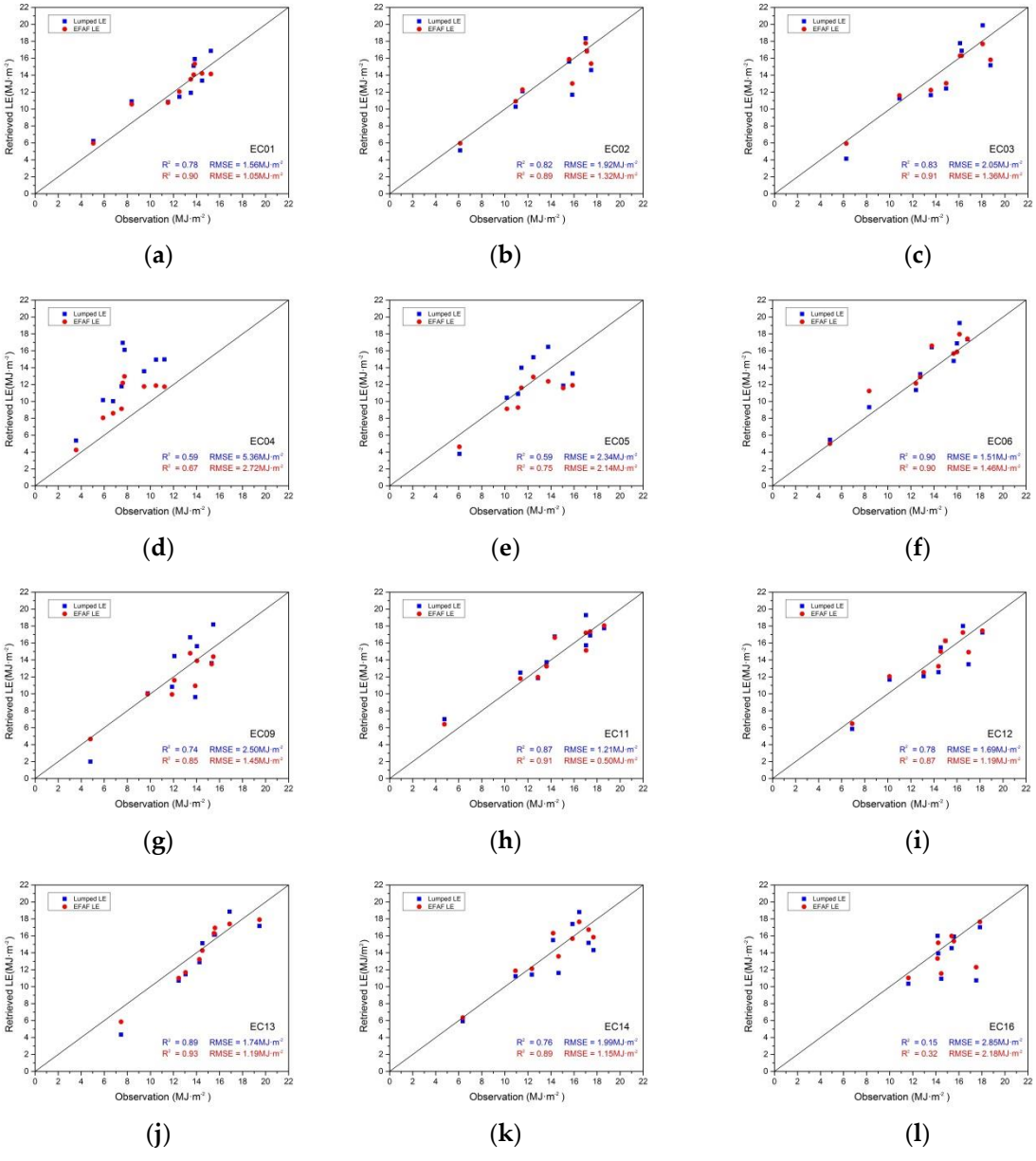

Figure 7. Scatter plots of the lumped LE (blue) and EFAF LE (red) against EC measurement LE at each site

The correction effects of certain sites were not found to be significant such as those of EC02, EC06, EC12 and EC14 and likely because surface heterogeneity was not found to be significant due to the presence of only two land cover types in the mixed pixels. These pixels also included a mixture of maize and other crops with similar EF values at the sites. However, the accuracy of estimating daily LE was improved when the impact of the sub-pixels for the EF was considered. For example, EC12 was measured as a maize-dominant pixel of maize and other crops with an AF of 74:26. On 27 July, maize cropland

presenting values nearest neighbouring EF of pure pixels generated a value of 0.97 while other crops presenting values nearest neighbouring EF of pure pixels generated a value of 0.84. The EF of this mixed pixel changed to 0.94 from 0.96 by EFAF and the daily LE decreased from 18.00 to 17.24 MJ/m^2 on that day. Compared to the value of 16.52 MJ/m^2 found for EC, the EFAF LE was found to be relatively more accurate for that day.

4.3 Error analysis

4.3.1 Error analysis of Hypothesis 1

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 bias 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. 8). 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. 8). The difference in frequency for values ±60 W·m^{-2} was extremely poor (less than 1%) and thus could be ignored.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

Hydrology and Earth System Sciences
Discussions

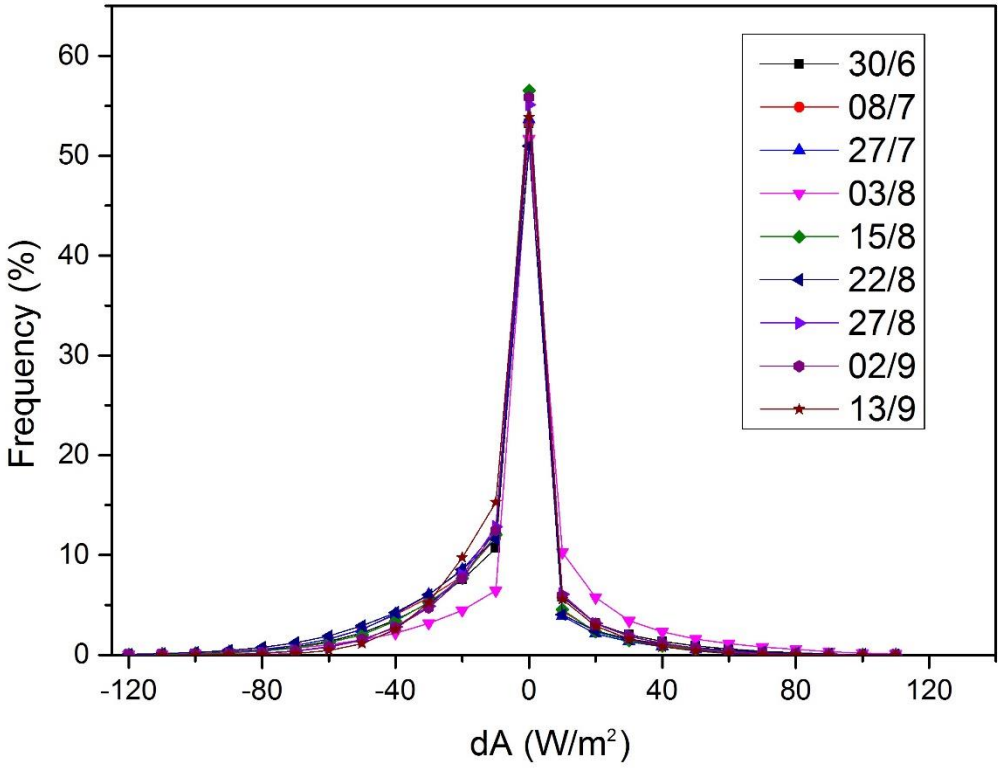

Figure 8. Distribution of difference AE (dA) and the frequency of difference for nine days

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.9)). This result indicated that Hypothesis 1 results in large biases in the transition zones between oasis areas and uncultivated land, but these biases often cancel one another because large negative and positive biases exist in a mixed pixel.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

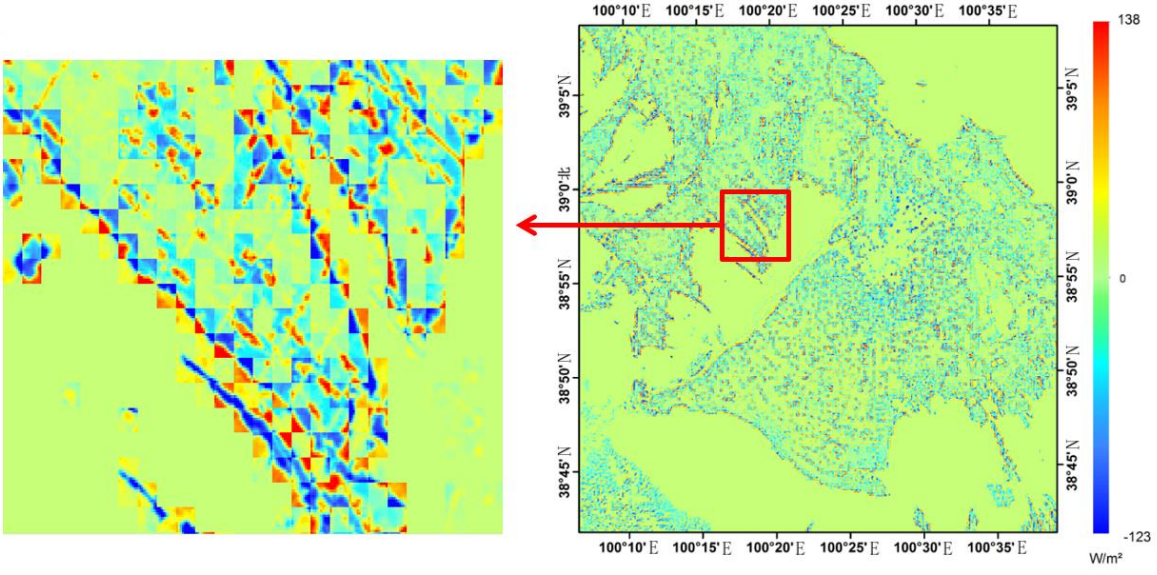

Figure 9. Spatial distribution of difference AE (dA) and a transition zone on 2nd September.

To evaluate the bias in the study area as a result of Hypothesis 1, the expected value (E(x)) of bias was measured based on the dA and its frequency. Fig. 10 shows the expected values of bias based on Hypothesis 1 (dA) for the nine days studied. Small expected values of less than 10 $W \cdot m^{-2}$ were observed when Hypothesis 1 was tested. A maximum bias value of -8.44 $W \cdot m^{-2}$ 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 biases resulting from Hypothesis 1 for maize, grass, bare soils and vegetables were approximately -6.50, -4.98, -1.86 and -6.84 $W \cdot m^{-2}$, respectively. We consider these biases to be acceptable.

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

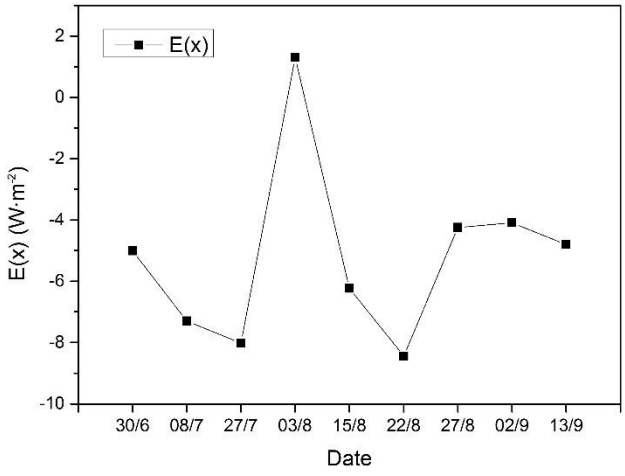

Figure 10. Expected values of bias based on Hypothesis 1 for the nine days

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 is equal to the EF of the closest 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. 11).

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 bias 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 $W \cdot m^{-2}$ on 8 July, 324.05 $W \cdot m^{-2}$on 27 July and 309.05 $W \cdot m^{-2}$on 22 August. This means that the maximum bias 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 $W \cdot m^{-2}$, respectively. Considering that most mixed pixels were closer to their nearest pure pixels than pure pixels were to their nearest pure pixels, the bias 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^2 value 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.

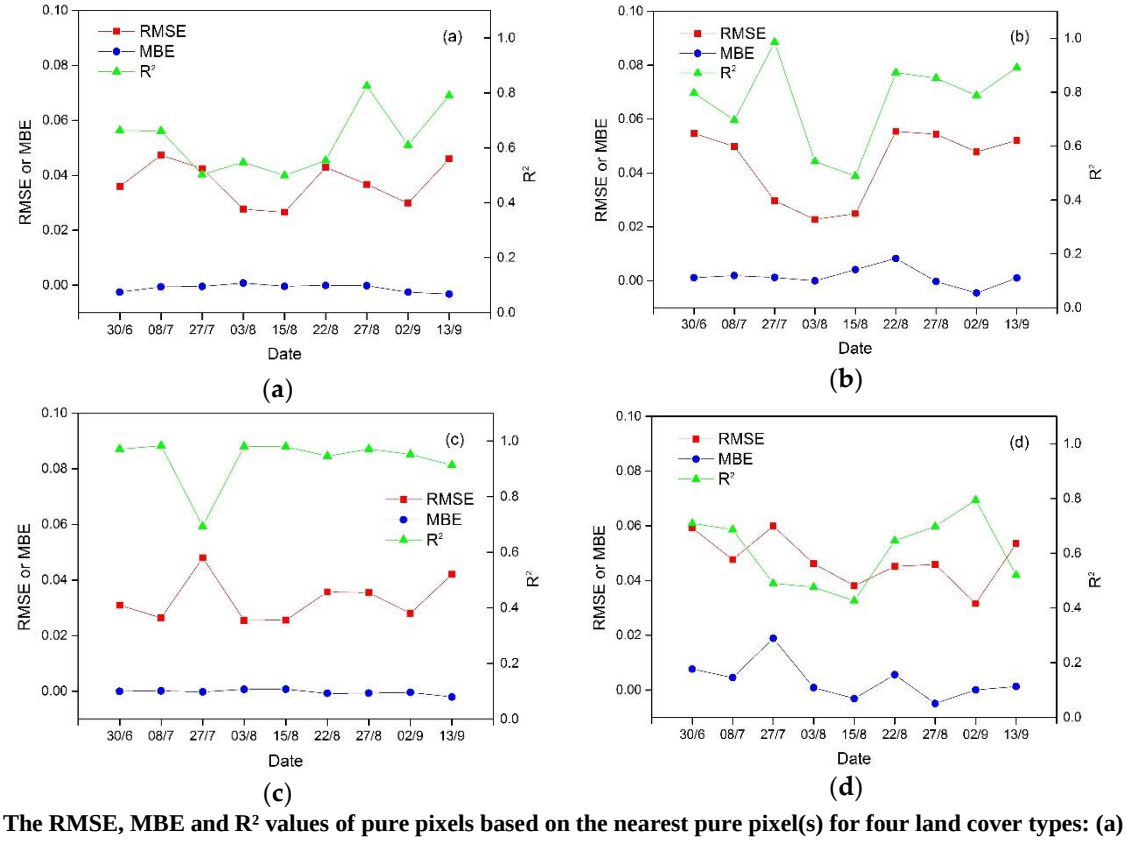

Figure 11. The RMSE, MBE and R^2 values of pure pixels based on the nearest pure pixel(s) for four land cover types: (a) maize, (b) grassland, (c) bare soils and (d) vegetables

In summary, Hypothesis 2 reproduces the EF of sub-pixels with an RMSE less than 0.06, resulting in a bias within 20 $W \cdot m^{-2}$ for LE estimation in this study. We consider this bias to be acceptable in surface flux estimation

4.4 Sensitivity analysis to 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.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-148
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 22 May 2018

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 bias 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 bias, 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 bias of 162.03 $W \cdot m^{-2}$ for the LE estimation. Additionally, incorrectly classifying bare soils as maize, grass or vegetables (or vice versa) also induced a greater bias; 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 $W \cdot m^{-2}$. However, incorrectly classifying bare soils as wheat (or vice versa) resulted in lower bias, with an absolute difference in EF of approximately 0.12.

Furthermore, while misclassifications between water bodies and bare soils could result in a higher bias 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 bias 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 $MJ \cdot m^{-2}$; without the EFAF, RMSE is 2.47 $MJ \cdot m^{-2}$. The two hypotheses result in lower bias, within 10 $W \cdot m^{-2}$ for Hypothesis 1 and 20 $W \cdot m^{-2}$ for Hypothesis 2. These results suggest that EFAF is reliable and has a great application field. 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 body, building 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 pixels estimates. As discussed in Sect. 4.4, relatively small biases resulted from the misclassification of vegetation with similar phenological conditions; however, larger biases 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 bias in the pure pixels over which the clouds cover.

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

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 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. 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 $MJ \cdot m^{-2}$, and the MBE decreased from 1.92 to 1.18 $MJ \cdot m^{-2}$, which is a quite significant improvements.

Error analysis suggests that the two key hypotheses of the model induce relatively little bias. The expected value of the absolute bias in AE due to Hypothesis 1 was within 7 $W \cdot m^{-2}$ 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. Yet, We note that the results from this study are probably limited and the model should be tested and validated in other areas.

Acknowledgments

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

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.

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.

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

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.

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, 2008, 477-481,

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, 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., 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.

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

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.

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 HiWATERmUSOEXE, 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.