Available online at www.sciencedirect.com
ScienceDirect
Remote Sensing of Environment 105 (2006) 271 – 285
Remote Sensing of Environment
www.elsevier.com/locate/rse

Scale influences on the remote estimation of evapotranspiration using multiple satellite sensors

Matthew F. McCabe [a,b,*], Eric F. Wood [a]

[a] Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA
[b] Los Alamos National Laboratory, PO Box 1663, MS-D436, Los Alamos, NM, 87545, USA

Received 8 April 2005; received in revised form 12 July 2006; accepted 15 July 2006

Abstract

There is considerable interest in using remote sensing to characterize the hydrologic behavior of the land surface on a routine basis. Information on moisture fluxes between the surface and lower atmosphere reveals linkages and land–atmosphere feedback mechanisms, aiding our understanding of energy and water balance cycles. Techniques that combine information on land and atmospheric properties with remotely sensed variables would allow improved prediction for a number of hydrological variables. Over the last few decades, there has been a focus on better determining evapotranspiration and its spatial variability, but for many regions routine prediction is not generally available at a spatial resolution appropriate to the underlying surface heterogeneity. Over agricultural regions, this is particularly critical, since the spatial extent of typical field scales is not regularly resolved within the pixel resolution of satellite sensors. Understanding the role of landscape heterogeneity and its influence on the scaling behavior of surface fluxes as observed by satellite sensors with different spatial resolutions is a critical research need. To attend this task, data from Landsat-ETM (60 m), ASTER (90 m), and MODIS (1020 m) satellite platforms are employed to independently estimate evapotranspiration. The range of the satellite sensor resolutions allows analyses that span scales from (point-scale) in-situ tower measurements to the MODIS kilometer-scale. Evapotranspiration estimates derived at these multiple resolutions were assessed against eddy covariance flux measurements collected during the 2002 Soil Moisture Atmospheric Coupling Experiment (SMACEX) over the Walnut Creek watershed in Iowa. Together, these data allow a comprehensive scale intercomparison of remotely sensed predictions, which include intercomparisons of the evapotranspiration products from the various sensors as well as a statistical analysis for the retrievals at the watershed scale. A high degree of consistency was observed between the retrievals from the higher-resolution satellite platforms (Landsat-ETM and ASTER). The MODIS-based estimates, while unable to discriminate the influence of land surface heterogeneity at the field scale, effectively reproduced the watershed average response, illustrating the utility of this sensor for regional-scale evapotranspiration estimation.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Evapotranspiration; Remote sensing; Scale; Hydrology; ASTER; MODIS; Landsat; SMACEX; SMEX; Catchment hydrology; Satellite

1. Introduction

Evapotranspiration is one of the most important components of the hydrological cycle. Combined with rainfall and runoff, it controls the availability and distribution of water at the Earth's surface, and for this reason, is of significance to a number of water-related research and application areas. Quantifying the spatial variability in hydrological and land surface variables is important to water resource management, particularly in agricultural regions. There are also strong interconnections and feedbacks between hydrological variables, like soil moisture and evapotranspiration, with regional hydrometeorology (Betts et al., 1997; Hipps & Kustas, 2000), which broadens the interest in accurately determining energy partitioning of surface fluxes. The need for effective regional measurements of the surface energy balance, surface temperature and soil moisture, has resulted in research into the use of satellite remote sensing. However, the problem of "point perfection but regional presumption" (Prince et al., 1998) is an issue, where much of the forcing data and validation information is spatially sparse at discrete locations.

Issues of scale are of critical interest to hydrological investigations (Grayson & Blöschl, 2000; Kalma & Sivapalan,

* Corresponding author. Los Alamos National Laboratory, PO Box 1663, MS-D436, Los Alamos, NM, 87545, USA. Tel.: +1 505 606 1800; fax: +1 505 664 0362.
E-mail address: mmccabe@lanl.gov (M.F. McCabe).

0034-4257/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2006.07.006

1995; Wood, 1994). Often, there is a gap between the desired resolution and the practical scale at which processes can be represented, constrained to a large degree by the availability of forcing and evaluation data. Methods to address spatial and temporal disparities between landscape heterogeneity and sensor and model resolution are limited, since an adequately developed theory of scale dependence or scaling in hydrology does not yet exist (Beven & Fisher, 1996). Yet, it is an issue that continues to perplex many aspects of hydrological modeling, particularly surface flux estimation (Raupach & Finnigan, 1995). While the patchiness of landscapes is difficult to measure on the ground, satellite remote sensing can observe these variations across a number of scales (up to continental scales), with spatial coverage that is unattainable from ground-based measurements. Despite the challenges of obtaining skillful retrievals, satellites observations are the only viable method for measuring relatively small-scale variations in surface fluxes and states at regional scales, and to examine the same region at multiple scales from a variety of available sensors with different spatial, spectral, and temporal characteristics. Determining effective techniques to make use of data from various sensors has been the focus of considerable research (Jackson, 1997; Kustas, 1990; Wan & Dozier, 1996), but little actual intercomparison between sensors has been undertaken, with most analysis focused on individual satellite platforms.

The work presented here extends recent research in estimating surface heat fluxes from satellite observations. The work uses remote sensing platforms and field data collected during the 2002 Soil Moisture Atmospheric Coupling Experiment (SMACEX) (Kustas et al., 2005), conducted over the Walnut Creek catchment in central Iowa. French et al. (2005b) examined two different modeling approaches to illustrate the utility of employing the Advance Spaceborne Thermal Emission and Reflection Radiometer (ASTER) for surface flux prediction. They show that ASTER data holds much promise for high-resolution flux mapping, even though the models used illustrated variability in individual surface flux prediction. Su et al. (2005) presented results for the retrieval of evapotranspiration over the Walnut Creek watershed using a Landsat-ETM scene. Considering the complex spatial structure of the land surface in the catchment, good agreement between the Landsat-based retrievals and in-situ tower measurements was observed. Kustas et al. (2004) present one methodology to assess the effects of pixel resolution on surface flux variability using high-resolution (60 m) Landsat-ETM data (Li et al., 2004). By aggregating the high-resolution data, they mimicked lower-resolution sensors having pixel resolutions ranging from 120 m, 240 m, and 960 m; the later close to the 1-km pixel scale of the thermal band on the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor. From the simulated lower-resolution sensors, they attempt to understand the influence of sensor resolution on flux prediction. The authors found good agreement in comparing modeled fluxes with in-situ and aircraft data, although at MODIS-like resolutions, it was not possible to distinguish variation in fluxes between soybean and corn fields.

One limitation in the Kustas et al. (2004) aggregation study is that they focused on a single sensor, progressively degrading the resolution to mimic expected results from coarser scale sensors, a typical approach for such analyses. The approach in this study differs from previous studies in that data from three different sensors, Landsat, ASTER, and MODIS, are used independently to study the effect of sensor resolution on land surface flux estimates. As discussed in Section 3, the Landsat image was collected within 30 min of the ASTER and MODIS image. By assembling these data, the study provides for the first time an intercomparison of flux estimates from different sensor systems relative to the high-quality SMACEX ground-based measurements and assesses the consistency in predictions across satellite platforms and resolutions. The surface fluxes were estimated using the Surface Energy Balance Systems (SEBS) (Su, 2002), with vegetation and surface temperature information derived separately from each sensor. Meteorological forcing data is determined from an average of the in-situ measurements, coincident with different satellite overpass times. To the authors' knowledge, this represents the first detailed intercomparison of these satellite systems for coincident surface flux estimation.

2. SEBS land surface model

The Surface Energy Balance System (SEBS) was developed for the estimation of surface heat fluxes using satellite earth observation data in combination with routinely available meteorological forcing. Detailed description of the model is provided in Su (2002) and Su et al. (2005). The SEBS model consists of (1) a set of algorithms to determine land surface physical parameters such as albedo, emissivity, temperature, vegetation coverage, etc. from spectral reflectance and radiance; (2) an extended model to determine the roughness length for heat transfer (Su, 2002); and (3) a formulation to determine the evaporative fraction on the basis of the energy balance at limiting wet and dry cases. The following presents a brief overview to SEBS, with further details available in Su (2002) and Su et al. (2005).

SEBS requires as inputs three sets of information. The first set consists of land surface albedo, emissivity, temperature, fractional vegetation coverage and leaf area index, and the height of the vegetation (or roughness height). When vegetation information is not explicitly available, the Normalized Difference Vegetation Index (NDVI) is used as a surrogate. The second set includes air pressure, temperature, humidity, and wind speed at a reference height, derived from either in-situ measurement or alternatively as variables estimated by large-scale meteorological models. The third set of data include downward solar radiation and downward longwave radiation, which can either be measured directly, from atmospheric model output or through simple parameterization schemes (e.g. Brutsaert, 1991).

SEBS consists of several separate modules to estimate the net radiation, soil heat flux, and the partitioning of available energy into sensible and latent heat fluxes. The surface energy balance equation is expressed as:

$$R_n = G_0 + H + \lambda E \quad (1)$$

where R_n is the net radiation, G_0 is the soil heat flux, H is the sensible heat flux, and λE is the latent heat flux. In SEBS, the net radiation is estimated from components of the radiative energy balance, expressed as:

$$R_n = (1-\alpha)R_{swd} + \varepsilon R_{lwd} - \varepsilon\sigma T_s^4 \tag{2}$$

where α is the broadband albedo in the visible and near infrared band, ε is the broadband emissivity in the thermal infrared band, R_{swd} is the incident solar radiation, R_{lwd} is the downward longwave radiation, T_s is the satellite derived surface temperature, and ó is the Stephan–Boltzman constant. In this study, radiation components are primarily determined from in-situ observations, as discussed further in Section 3.

For satellite-based applications, estimates of the soil heat flux are not generally available, so empirical formulations based on relationships with the net radiation and vegetation fraction are often used to estimate the total soil heat flux for the area. If G_0 is defined as a simple fraction of the net radiation, then following Choudhury et al. (1987), it can be parameterized as:

$$G_0 = R_n \cdot (\Gamma_v f_v + \Gamma_s(1-f_v)) \tag{4}$$

where Γ_v and Γ_s are the soil heat flux ratios for full vegetation canopy (0.05) (Monteith & Unsworth, 1990) and bare soil (0.315) (Kustas & Daughtry, 1990), and f_v is the vegetation fraction. Interpolation can be performed between these limiting cases using distributed estimates of the fractional coverage. Here, spatial variation in the ground heat flux is calculated using an empirical relationship between the NDVI and fractional cover (Baret et al., 1995):

$$f_v = 1-\left(\frac{NDVI-NDVI_\infty}{NDVI_s-NDVI_\infty}\right)^k \tag{5}$$

where $NDVI_s$ represents the value for a bare soil (0.2013), $NDVI_\infty$ the value for a full canopy (0.8986), and k is 0.6175, all of which were experimentally determined. NDVI data was derived from both Landsat and ASTER reflectance data and aggregated up to their respective thermal band resolutions.

The sensible heat flux H can be solved using a combination of the three equations below, derived in part from Monin–Obukhov similarity theory (Monin & Obukhov, 1954):

$$u = \frac{u_*}{k}\left[\ln\left(\frac{z-d_0}{z_{0m}}\right) - \Psi_m\left(\frac{z-d_0}{L}\right) + \Psi_m\left(\frac{z_{0m}}{L}\right)\right] \tag{6a}$$

$$\theta_0-\theta_a = \frac{H}{ku_*\rho C_p}\left[\ln\left(\frac{z-d_0}{z_{0h}}\right) - \Psi_h\left(\frac{z-d_0}{L}\right) + \Psi_h\left(\frac{z_{0h}}{L}\right)\right] \tag{6b}$$

$$L = -\frac{\rho C_p u_{*3}\theta_v}{kgH} \tag{6c}$$

where u is the wind speed, u_* is the friction velocity, ρ is the air density, C_p is the specific heat of air at constant pressure, k is the von Karman's constant, d_0 is the zero plane displacement, z is the height above the surface, z_{0m} and z_{0h} are the roughness height for momentum and heat transfer, respectively, θ_0 and θ_a are the potential temperatures at surface and at height z, respectively, Ψ_h and Ψ_m are the stability correction functions for sensible heat and momentum transfer, respectively, L is the Obukhov length, g is the acceleration due to gravity, and θ_v is the virtual temperature near the surface. When near surface wind speed and vegetation parameters (height and leaf area index) are available (as used here), the within canopy turbulence model proposed by Massman (1987) is employed to estimate aerodynamic parameters, the displacement height (d_0) and the roughness height for momentum (z_{0m}). If only the height of the vegetation is available (e.g. through vegetation classification information) the relationships proposed by Brutsaert (1991, 1999) are used.

In SEBS, model output can be constrained by considering dry-limit and wet-limit conditions, which give the upper and lower boundary of sensible heat flux estimation. A summary of the approach follows, but the reader is directed to Su (2002) for complete details. Under the dry-limit, latent heat is assumed zero due to the limitation of soil moisture and the sensible heat flux is at its maximum value, which is the available energy. Under the wet-limit, where the evaporation takes place at potential rate (λE_{wet}) (i.e. the evaporation is only limited by the available energy under the given surface and atmospheric conditions), the sensible heat flux takes its minimum value (H_{wet}) and can be estimated using a Penman–Monteith parameterization (Monteith, 1981):

$$H_{wet} = \left((R_n-G_0) - \frac{\rho C_p}{r_{ew}}\cdot\frac{e_s-e}{\gamma}\right)\Big/\left(1+\frac{\Delta}{\gamma}\right) \tag{7}$$

where e and e_s are actual and saturation vapor pressure, respectively, γ is the psychrometric constant, Δ is the rate of change of saturation vapor pressure with temperature (i.e. $\partial e_{sat}(T)/\partial T$), and r_{ew} is the bulk surface external or aerodynamic resistance estimated under the assumption that the bulk internal resistance is zero. The dry- and wet-limit constraints allow an expression for the actual evapotranspiration to be formulated from knowledge of these limits, resulting in the following two equations for the relative evaporation (8) and the evaporative fraction (9).

$$\Lambda_r = 1-\frac{H-H_{wet}}{H_{dry}-H_{wet}} \tag{8}$$

$$\Lambda = \frac{\Lambda_r\cdot\lambda E_{wet}}{R_n-G_0} \tag{9}$$

All the observation data requirements for the SEBS model as implemented herein are listed in Table 1. In addition to tower-based flux measurements used to evaluate SEBS, other observations including detailed vegetation parameters and surface meteorological data were collected during the field experiment. The vegetation parameters (vegetation height, vegetation fraction, and LAI) and wind speed are used to

Table 1
Meteorological forcing data and sensor characteristics

	Landsat	ASTER	MODIS
Meteorological variables			
Incoming solar radiation (W .m^{-2})	859.0	914.0	914.0
Air temperature (°C)	29.35	30.1	30.1
Vapour pressure (kPa)	2.2	2.2	2.2
Wind speed (m .s^{-1})	5.32	5.4	5.4
Wind direction (°)	225	225	225
Atmospheric pressure (kPa)	98.2	98.2	98.2
Satellite and other data			
Overpass time (local)	10:40 a.m.	11:12 a.m.	11:12 a.m.
SEBS modelling resolution	60 m	90 m	1020 m
LST resolution	60 m	90 m	1020 m
NDVI resolution	30 m	30 m	30 m Landsat data
Albedo	Aggregated from ASTER (15 m) surface reflectance		
Fractional cover	Based on approach of Kustas and Daughtry (1990)		
Emissivity	Fractional cover mixing model Sobrino et al. (2001)		
LAI	Based on approach of Xavier and Vettorazzi (2004)		
Land cover	Aggregated Landsat data from Jackson et al. (2004)		

estimate the roughness height for heat and momentum transfer, used in (6) and elsewhere.

3. Data description and methodology

The Soil Moisture Atmospheric Coupling Experiment (SMACEX) was conducted in Iowa from June 19 through July 9 in 2002. The primary purpose of the field campaign was to understand how horizontal heterogeneities in vegetation cover, soil moisture, and other land surface variables influence surface flux exchanges with the atmosphere. In conjunction with the Soil Moisture Experiment (SMEX) (see http://hydrolab.arsusda.gov/), the SMACEX data set provides a detailed and comprehensive observation of hydrological interactions over Walnut Creek, a small watershed in central Iowa. Approximately 95% of the watershed (and much of the surrounding region) is used for row crop agriculture, with corn and soybean representing nearly 80% in relatively equal proportions. The experiment compiled a variety of measurement sources and techniques, including aircraft flux and tower-based eddy covariance observations (at 30-min intervals) (Prueger et al., 2005), and in-situ measurement of a number of land surface variables. In all, 14 towers were available for meteorological and flux measurement (see Fig. 1), although only 12 of these provided continuous flux measurements during the course of the campaign. Quality control of the flux data occasionally limited the number of towers. Further information and a detailed description of the experiment can be found in Kustas et al. (2005).

3.1. Landsat and ASTER variables

Li et al. (2004) used Landsat-based radiances to calculate land surface brightness temperatures (LST) with a mean absolute difference of 0.98 K compared to in-situ measurements. The data used in this study (from Li et al., 2004) was collected at 10:40 a.m. July 1, 2002 (path-26, row-31). For more general applications, where quality in-situ data is not available, achieving such accurate retrieval of the surface temperature is not likely, given the radiometric resolution of the data. Likewise, the ASTER-based LST (11:12 a.m. July 1, 2002 overpass) (from French et al., 2005b), calculated LST using a temperature emissivity separation approach following Gillespie et al. (1998). Atmospheric corrections, required as part of the analysis, were performed using MODTRAN4, with data obtained from a radiosonde release at Kelley, Iowa (42.00°N, 93.61°W). The radiosonde was released at 16:07 UT- near coincident with the satellite overpasses.

Albedo estimates for all satellite platforms were calculated from the ASTER Level 2 surface reflectance product. The ASTER product (from French et al., 2005b) was aggregated from the original 15-m resolution to the sensor specific surface temperature resolution. Emissivity was calculated using a fractional cover mixing model (see Sobrino et al., 2001), which describes the composite emissivity as a combination of fixed component (vegetation and bare soil) emissivity values, determined here from literature data (Chen et al., 2004; Li et al., 2004; Sobrino et al., 2001).

3.2. MODIS variables

In this study, use has been made of the MOD11A1 LST operational product (Wan et al., 2002), resampled to a nominal resolution of 1020 m using the MODIS Reprojection Tool (http://edcdaac.usgs.gov/). It was necessary to reproject the LST data to ensure consistency between the other data sources used to estimate the surface fluxes with the MODIS LST product. The same meteorological forcing data were used for MODIS and ASTER, given that these sensors operate from the same platform (EOS-TERRA). The emissivity, NDVI, and land surface cover were aggregated from the high-resolution (30 m) Landsat data, while the albedo was determined from the ASTER product, as described above and summarized in Table 1. While some error inevitably arises through this process, these aggregated data are superior to equivalent operational products at MODIS resolutions. For instance, MODIS-based land classification of the region indicates a single homogeneous land unit of type ‘cropland’.

3.3. Land cover classification

A key area of uncertainty lies in the specification of land surface cover type. Such details are critical for correctly characterizing related vegetation and atmospheric properties (e.g. vegetation height and roughness length)—particularly over a heterogeneous surface. Generally, the required level of information is not adequately resolved within operational products (e.g. MOD12 1-km Land Cover Product), with classifications limited to generic descriptions. Broad land cover classifications increase the uncertainty in flux retrievals, particularly in the case of the Walnut Creek watershed, where the dominant vegetation types (corn and soybean) exist as patchwork across the catchment.

To counter the lack of detail at MODIS-type resolutions, it is necessary to make effective use of available Landsat classifications. Supervised classifications of the Walnut Creek watershed, determined from 30-m resolution Landsat overpasses, were aggregated to the 1020-m MODIS resolution. In theory, aggregation is a relatively straightforward concept, but in practice can cause difficulties—particularly in strongly heterogeneous environments where no single vegetation unit dominates at large scales. The technique employed here calculates a simple majority of vegetation elements within a pixel and assigns the large-scale unit that value. Fig. 2 illustrates the result of this analysis, comparing high-resolution data with the resulting MODIS-equivalent product. Inevitably, some misclassification occurs even when using the best available high-resolution data. However, the effects of this are not as significant as specifying a uniform land cover, as is the case with operational products.

Overall, there is generally good agreement between the high-resolution classification and the MODIS equivalently aggregated product. Dominant features are retained, although the mix of corn, soybean, and grass on the eastern portion of the catchment complicates the specification of a single dominant unit. Regardless, the derived product represents the best available classification for land surface features at MODIS-type resolutions, relative to operational products, which only indicates a single broad classification type of cropland.

4. Multi-scale–multi-sensor flux estimation

4.1. Cross compatibility of flux retrievals

Surface flux maps were determined for the Walnut Creek watershed for each of the Landsat, ASTER, and MODIS satellite images using the SEBS model and available forcing data (see Fig. 3). These maps represent the spatial characterization of the flux patterns at successively increasing scales, offering insight into the ability of MODIS resolution sensors to retrieve surface fluxes over heterogeneous terrain. Table 3 compares the estimated fluxes from the three sensors to in-situ flux observations extracted from the SMACEX tower data at 11:00 a.m. (local time), as this approximated the mid-point of the ASTER/MODIS and Landsat satellite overpass times.

A strong correlation between Landsat and ASTER flux estimates and the underlying surface is immediately evident, with patterns of evapotranspiration strongly linked to vegetation type (see Fig. 2). As expected, the visual agreement between these images (Fig. 3) is also strong, due to having similar spatial data resolutions (60 m and 90 m). Contrast between soybean (yellow) and corn fields (blue) is apparent across the catchment, highlighting the heterogeneous nature of the study area. Interestingly, Fig. 3 is not completely dominated by these colors, with distinct hues between the soybean and corn. The gradation of evaporative response possibly reflects different stages of vegetation growth, or alternatively, could indicate gradients in soil moisture conditions across the landscape (e.g. Fig. 1, McCabe et al., 2005). Tower data (see Table 3) clearly show that on this day corn transpires at a higher rate relative to soybean. Previous analysis of flux tower data using SEBS (Su et al., 2005) indicated that the flux difference between soybean and corn varies between 100 and 150 W/m^2 during the SMACEX campaign, a difference consistent with the results in Table 2 and in Kustas et al. (2004).

In comparison, the MODIS-based flux retrievals, while displaying some limited spatial variability, do not exhibit the same heterogeneity that characterize the higher-resolution imagery. Given that land cover types were aggregated to 1 km (Fig. 2) for the MODIS-based retrievals, it might be expected that some pattern related to their classification would be reflected. However, the MODIS-based estimates lay between soybean

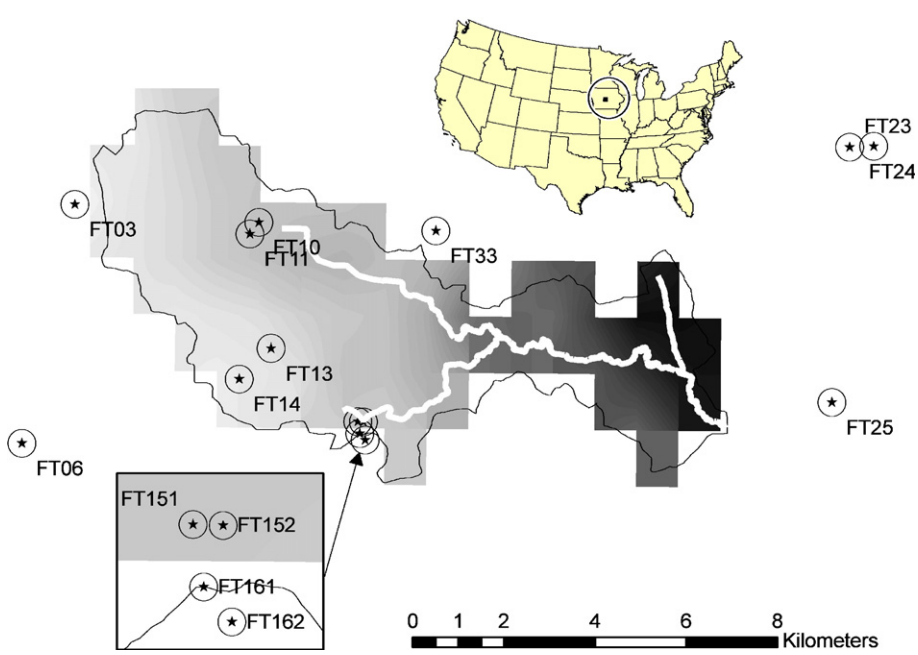

Fig. 1. Location of Walnut Creek watershed in Iowa and the position of flux towers distributed across the watershed. Gradients within the catchment represent the topographic relief, representing a range from 270 m (black) to 325 m (white). Streamlines identifying the main Walnut Creek branch are also included.

Table 2
Statistics for catchment-scale evapotranspiration

	Landsat	ASTER	MODIS
Samples, *n*			
Corn	5533	2623	26
Soybean	4774	2369	25
Catchment	11,711	5738	51
Mean (W/m^2)			
Corn	442.05	457.60	388.53
Soybean	283.97	322.12	372.53
Catchment	367.48	392.3	380.97
Standard deviation (W/m^2)			
Corn	65.30	96.72	45.8
Soybean	37.90	59.81	25.0
Catchment	97.21	105.3	35.3

Values for both the catchment average and crop specific estimates are presented, along with number of samples and the standard deviation. For ASTER and Landsat, component samples do not sum to the total number of pixels. This is a result of additional land cover types being present at high resolutions. Values are consistent with observations of soybean and corn comprising nearly 90% of the watershed.

and corn retrievals, suggesting that heterogeneity effects from surface temperature and vegetation decrease as scale increases.

The lack of spatial variability in the MODIS-based evapotranspiration retrievals is evident in Fig. 3, and when

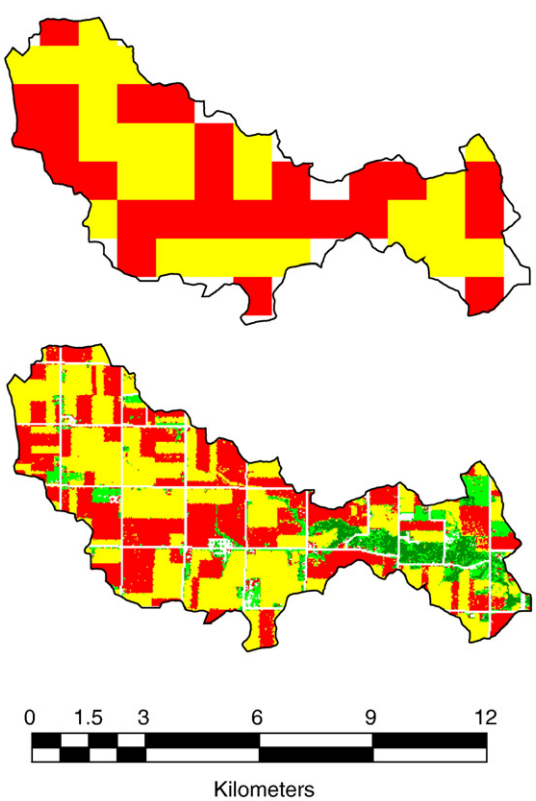

Fig. 2. Land cover classifications determined by supervised classification of Landsat data at 30-m resolution (bottom) (Jackson & Cosh, 2003) and the aggregated product at MODIS 1-km resolution (top). Although the higher detail imagery included grassland and tree classifications (green), only soybean (red) and corn (yellow) crop types were retained at the 1-km scale.

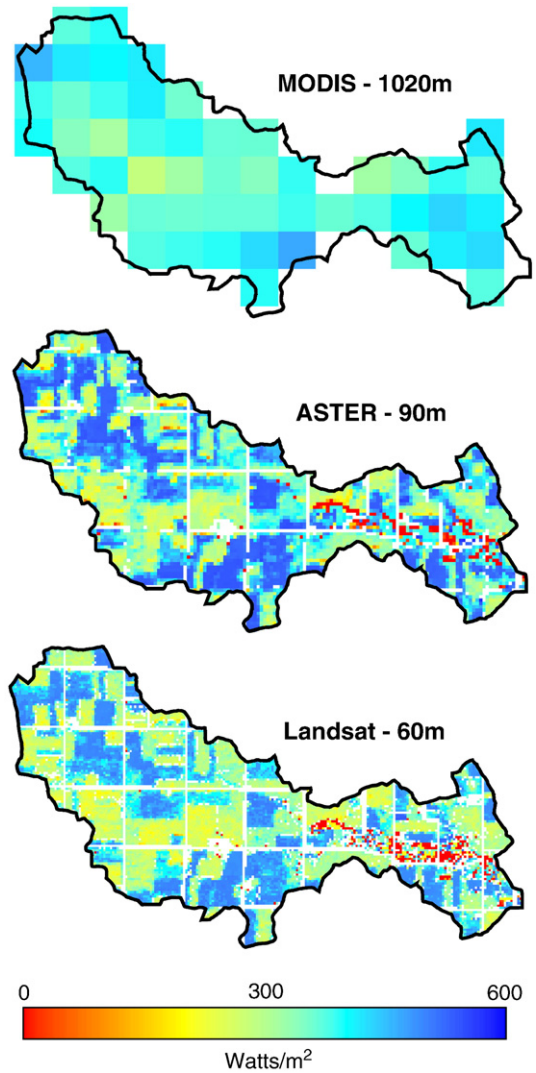

Fig. 3. Estimates of evapotranspiration for MODIS, ASTER, and Landsat-ETM sensors over the Walnut Creek watershed for July 1, 2002. Satellite overpasses are near coincident, with Landsat measurements at 10:40 a.m. and the MODIS/ASTER overpass near 11:10 a.m. (local time). Generally, corn fields have higher evapotranspiration (blue), while soybean fields have lower values (yellow-green).

the frequency and cumulative distribution plots are displayed, as shown in Fig. 4. The ASTER and Landsat estimates show marked similarity, displaying a near-bimodal separation of flux patterns from the corn and soybean fields and a wide range in the distribution, while the MODIS-based evaporative fluxes display a comparatively narrow distribution about the mean of the two crop types. The three distinct peaks within the Landsat predictions for corn (higher ET) may be indicative of some relationship to crop condition, stages of growth, or water availability within the watershed, but there is insufficient data to explore this extensively. From Fig. 4, both Landsat and ASTER distributions indicate a clear distinction between soybean and corn, although the ASTER results present a smoother transition between the two crop types with a single peak about 530 W/m^2. The absence of peaks in the ASTER data (vis-à-vis Landsat) is likely related to either its marginally lower resolution (90 m) or

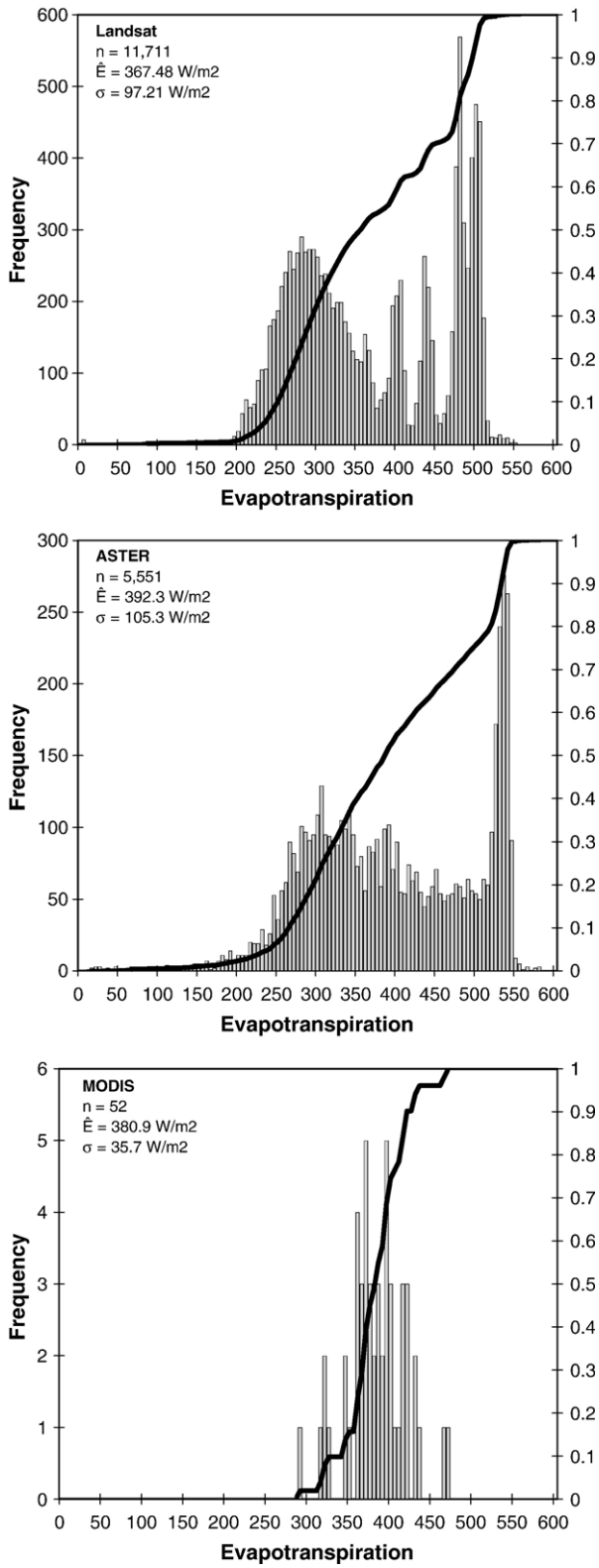

Fig. 4. Frequency (bar) and cumulative (line) distributions plots for evapotranspiration across the Walnut Creek watershed for Landsat, ASTER, and MODIS sensors. A bin size of 5 W/m^2 was specified. Statistics illustrate the number of samples within the catchment (n), the spatial mean (Ê), and standard deviation (σ). Further details can be found in Table 2.

differences resulting from the independently derived NDVI distributions (see Table 1).

A recent investigation by Kustas et al. (2004) of flux estimation in Walnut Creek employed histograms derived from high-resolution Landsat imagery to examine the influence of scale on ET (see Fig. 3 in Kustas et al., 2004). Although a different retrieval model was employed to estimate their ET (Norman et al., 1995), a general bimodal distribution is evident in both studies. The smaller bin-size in Fig. 4 (5 W/m^2 as opposed to 20 W/m^2 used in Kustas et al., 2004) more clearly discriminates differences in crop type and scale effects. Fig. 4 also displays a greater difference between the 60-m Landsat scale and the 90-m ASTER scale than is evident between the 60-m and 120-m Landsat intervals used by Kustas et al. (2004), probably because they aggregated data from Landsat rather than compare different sensors. Both analyses do agree on the decreased ability to determine flux variation as scale increases, with the distributions at MODIS scales unable to differentiate fluxes between land cover type.

It is expected that flux predictions based on Landsat should be slightly lower than estimates from ASTER and MODIS, given its 32-min earlier overpass time. This is found, with the Landsat-based spatial average (367.5 W/m^2) approximately 25.0 W/m^2 less than the later overpasses. This is also consistent with both the observation data and expected trends in the diurnal cycle. Interestingly, the averages for the individual vegetation units fall either side of this value, with the differences between corn and soybean in the ASTER image approximately 15 W/m^2 and 38 W/m^2 greater than in Landsat.

ASTER-based results indicate a slightly larger, but probably insignificant, difference in the catchment-averaged evapotranspiration relative to the coincident MODIS-based estimates (392.3 W/m^2 compared to 380.9 W/m^2). This suggests that the heterogeneity effects from vegetation, which are significant at small scales (high resolutions, see Table 2 and Fig 3), become less so with increased spatial scale, and calls for more studies on the spatial scaling of evapotranspiration over heterogeneous landscapes.

4.2. Spatial scaling of surface fluxes

4.2.1. Catchment-scale flux aggregation

To understand consistency in retrievals between scales, the Landsat- and ASTER-based latent heat flux predictions were aggregated to MODIS type resolutions. The aggregation approach is restricted to integer multiples of the original resolution; hence, the aggregated ASTER data is 990 m (90 × 11 m) and the Landsat overpass 1020 m (60 × 17 m). Moran et al. (1997) suggest caution on the ad-hoc aggregation of flux data, particularly over patchy vegetation. However, given that MODIS results are inherently aggregating the surface condition, it is instructive to assess the degree of accord between aggregated high-resolution flux data.

Fig. 5 illustrates the resulting Landsat and ASTER evapotranspiration patterns after aggregating the retrieval model output to an equivalent MODIS resolution. Initial comparison indicates a relatively consistent agreement among sensors.

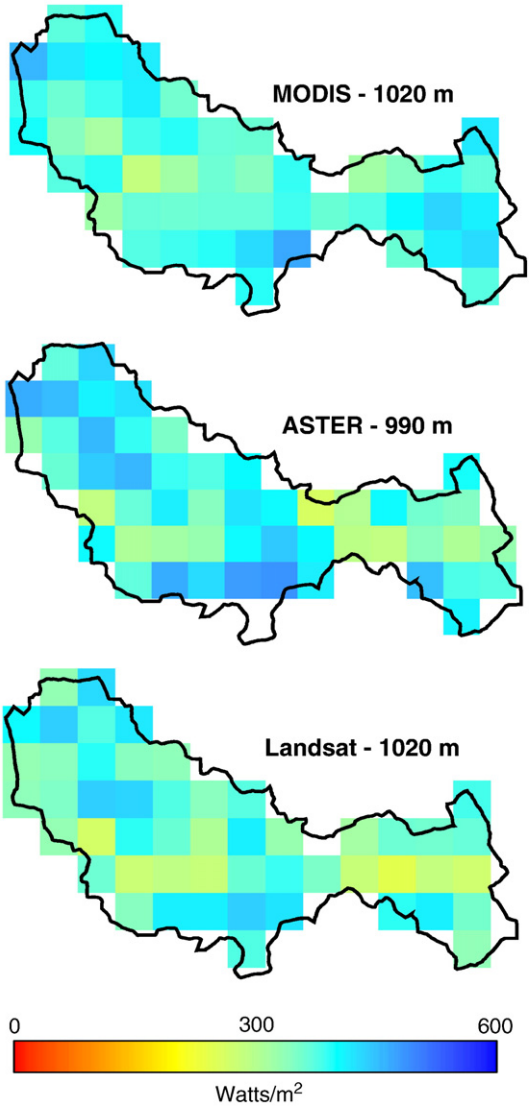

Fig. 5. Aggregated estimates of evapotranspiration for ASTER and Landsat-ETM compared to MODIS values. Spatial aggregates are calculated as integer multiples of the sensor resolution, hence the disparity between ASTER (90 × 11 m) and the Landsat (60 × 17 m) images.

There are subtle land cover influences in the aggregated high-resolution flux retrievals, most noticeably in the central western region where large fields of corn are separated by a horizontal band of soybean. The effect is not evident in the MODIS retrieval, even though the pre-defined vegetation units (see Fig. 2) might be expected to impose some structure on flux response.

As expected, the areal average of the evaporative flux maintains consistency across scales (since it is linear aggregation). However, a statistical analysis between the aggregated Landsat and spatially equivalent MODIS image reflects a low level of agreement, with a correlation coefficient (r) of 0.073. The root mean squared error (RMSE) between collocated pixels is 64 W/m^2, higher than the 'useful' accuracy limit of 50 W/m^2 as suggested by Kustas and Norman (2000). Error between measurements cannot be attributed entirely to differences in overpass times, since applying the estimated 25 W/m^2 offset between Landsat and MODIS (see Section 4.1) makes only a minor difference to the RMS error. Since the ASTER data is not an integer multiple of the MODIS data, a similar analysis is not trivial. The pixel shift between the western and eastern sides of the catchment (~15 km) approaches 500 m. Considering this disparity, the correlation estimate of 0.27 and associated RMS error of 56.45 W/m^2 when compared with the MODIS data should be viewed with some balance.

4.3. Tower-scale flux aggregation

While general application of aggregation approaches serve to identify broad-scale results and catchment response, aggregating data at specific locations allows a more equivalent comparison, particularly with available ground-based flux data. Landsat and ASTER retrievals were aggregated around each of the 14 flux tower locations (see Fig. 1) to the equivalent MODIS resolution. Comparison of these retrievals across the 14 sites indicates good agreement between Landsat and ASTER

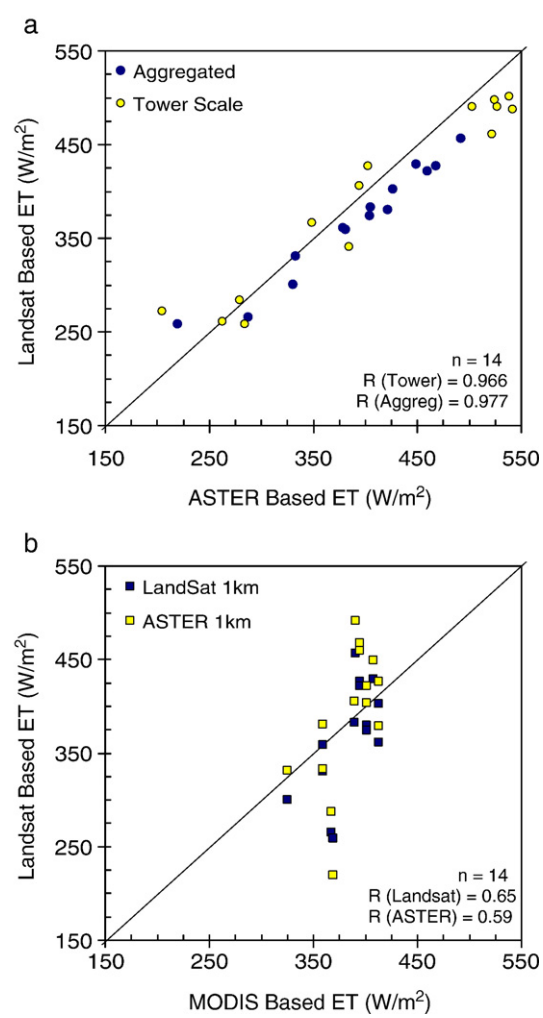

Fig. 6. Satellite-based flux measurements for (a) Landsat plotted against ASTER at tower locations (see Fig. 1) for both aggregated (closed) and tower-scale (open) values. Data are linearly averaged over the equivalent MODIS resolution, centred around the location of the flux tower. Flux measurements are also compared with (b) the spatially equivalent MODIS evapotranspiration for ASTER (open) and Landsat (closed). Not all towers were available for measurements.

retrievals (see Fig. 6a), with correlation coefficients for both aggregated and single pixel comparisons greater that 0.96. In contrast, the relative homogeneity of the MODIS retrieved flux distribution is evident in Fig. 6b, with the scaled Landsat and ASTER retrievals plotted against MODIS retrievals from the same tower locations. In four cases, two towers are located within the same MODIS pixel, but all available towers are included in the analysis. While the MODIS retrievals have a relatively narrow distribution at the tower sites (325–412 W/m^2), there is generally reasonable agreement with the aggregated high-resolution retrievals. In most instances, values agree to within 25% of the MODIS estimates, which are within the accuracy expected over heterogeneous landscapes.

There is a notable difference in the correlation when comparing Landsat and MODIS imagery at all points across the catchment (r=0.071) and to values at the selected tower locations (r=0.65). One explanation might relate to the positioning of the flux towers. Due to the high degree of surface heterogeneity in the catchment, if towers were randomly placed in the catchment and the aggregated Landsat and MODIS retrievals compared, then the correlations for these sites should be close to the catchment average presented above. However, the flux towers are not sited randomly. Rather, they are located in the least heterogeneous areas to maximize the spatial representativeness of measurements. Further analysis on the representativeness of the tower locations is examined in the following section with respect to statistical variation from aggregation.

4.4. Spatial variability of surface fluxes

As noted above, the spatial mean of the MODIS retrievals (380.9 W/m^2) agrees well with the spatial mean of both the Landsat and ASTER retrievals (367.5 W/m^2 and 392.3 W/m^2, respectively). Kustas et al. (2004) aggregated Landsat retrievals across four distinct spatial resolutions, all having the Walnut Creek average latent heat flux of approximately 391 W/m^2 (see Kustas et al., 2004, Table 3). However, examining the average retrieval of the 50-km^2 watershed provides less insight into aggregation effects than examining distinct locations within the catchment and the spatial variability among these locations. Identifying suitable sub-areas for the analysis within the catchment is challenging, given the need to representatively sample the dominant vegetation types against the background of surface heterogeneity. The task of selecting focus areas is constrained by the predetermined distribution of tower sites (see Fig. 1), which provide the data to evaluate the satellite retrievals.

Table 3 presents the individual satellite retrieved evaporative fluxes at each of 11 available tower locations. For comparison, flux tower observations using a footprint weighted approached (Schuepp et al., 1990) are also available in Table 1 of French et al. (2005a). French et al. (2005b) presented a comparison of two energy balance models: TSEB (Norman et al., 1995) and SEBAL (Bastiaanssen et al., 1998) against the footprint corrected flux observation data at selected tower locations. The results in Table 3 for SEBS agree well with the tower flux measurements, with an average deviation of −6 W/m^2 for Landsat and 18 W/m^2 for ASTER at the pixel scale (approaching the 25 W/m^2 bias estimated in Section 4.1).

Fig. 7a compares the pixel-scale ASTER and Landsat data from Table 3 with observed fluxes (not footprint corrected) at 11:00 a.m. (local time) for the 11 towers reporting valid measurements at this time. Since tower data are only reported at 30-min increments, the two satellite overpasses straddle the 11:00 a.m. reporting time by approximately ±15 min. While there

Table 3
Latent heat fluxes (W/m^2) derived from MODIS, ASTER, and Landsat data at both pixel scale and an aggregated MODIS scale (~1000 m)

Tower ID	Land type	Tower obs. (W/m^2)	Adjusted tower obs. (W/m^2)[a]	Pixel-scale flux estimates (W/m^2)			Aggregated (MODIS-scale) flux estimates (W/m^2)			
				MODIS-based	ASTER-based	Landsat-based	ASTER	σ_{ASTER}	Landsat	$\sigma_{Landsat}$
FT03	S(C)[b]	272	354	391	503	490	492	58	456	58
FT06	C	448	506	408	527	490	449	107	429	90
FT10	S	–	–	413	349	367	379	113	361	100
FT11	C	–	–	413	524	497	427	94	403	94
FT13	S	184	–	367	263	261	288	53	265	46
FT14	S	256	–	325	284	258	331	64	300	57
FT151	C	410	488	394	542	487	468	99	427	96
FT152	C	381	475	394	538	501	460	105	422	97
FT161	S	242	299	401	384	341	422	98	380	101
FT162	S	–	–	401	394	406	404	120	374	105
FT23	S	315	427	359	279	284	333	87	331	73
FT24	C	415	533	359	522	461	381	87	359	81
FT25	C	294	–	369	205	272	219	107	259	75
FT33	C	360	472	390	403	449	405	108	383	90
$\bar{\Delta}$*				−57	18	−6	−18		−45	

Data are compared with eddy covariance-based latent heat flux at each of the available tower locations throughout the SMACEX study and also footprint average fluxes derived in French et al. (2005a). Standard deviations (σ) for the aggregated predictions are calculated from the pixel averaged data centred on the tower locations. Average deviation ($\bar{\Delta}$) from the tower data is also shown.

[a] Values taken from French et al. (2005a) for 'footprint weighted' tower measurements, which were adjusted for wind speed and fetch. Estimates were determined at the ASTER/MODIS overpass time.

[b] Land cover classified as soybean in the field notes, but located in a corn field from provided coordinates.

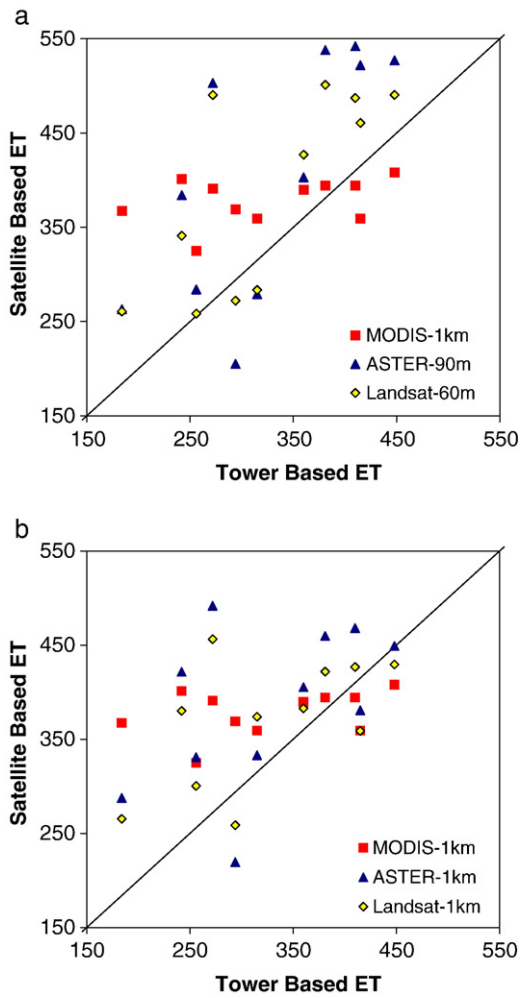

Fig. 7. Satellite-based flux measurements comparing MODIS (1 km), ASTER (90 m), and Landsat (60 m) for (a) the pixel scale and (b) the MODIS equivalent aggregated response, relative to tower-based flux measurements at 11:00 a.m. (local time). Eleven flux towers were available for comparison at this time, comprising six soybean and five corn sites.

seems a tendency for over-prediction in the satellite-based estimates, the correlation between the data sets is relatively strong (0.71 and 0.74 for ASTER and Landsat, respectively).

The eddy covariance approach estimates an areal average evapotranspiration, with its source area determined by the height of the tower and the associated fetch (Stannard, 1997). Thus, a comparison of larger-scale flux measurements with the tower-based measurements might be better than comparisons based on a single pixel. Fig. 7b presents ASTER and Landsat retrievals aggregated to 1 km (from Section 4.2), plotted against the tower-based eddy covariance measurements. While it is unlikely that the fetch actually extends to the 1-km range, results do illustrate a general improvement in agreement, even though the correlation coefficients for both series are reduced with respect to the pixel comparisons. Excluding tower (FT03), classified as soybean in the SMACEX field notes but classified as a corn field in the Landsat data, improves considerably the overall statistics for both pixel and aggregated comparisons. Upon its removal, the correlation coefficients (r) for ASTER and Landsat improved to 0.81 and 0.87 respectively at the pixel scale, and by over 30% to 0.63 and 0.75 for the 1-km scale. This reconfirms that land cover misclassification is an important issue with remote sensing data. In this specific analysis, validation of the satellite-based retrieval is corroborated by surface temperature measurements, which indicate a considerably lower value for FT03 than an adjacent soybean field. This lower temperature is consistent with temperatures from the cornfields, thereby justifying the removal of this spurious data point (or reprocessing the data using vegetation parameters consistent with corn).

The root mean squared error (RMSE) for both the pixel and aggregated retrievals, when compared to the tower-based fluxes (see Fig. 7), illustrate some interesting trends. For the 10 tower sites (FT03 excluded), a reasonable level of agreement was observed, with the Landsat having lower RMSE for the pixel-scale comparisons than ASTER. Consistent with trends in Fig. 7, RMSE decreased for both retrievals when aggregated to the coarser 1-km MODIS equivalent resolution, with Landsat displaying a smaller RMSE (61.94 W/m^2) than ASTER (82.04 W/m^2) when compared to the tower measurements.

4.5. Scaling characteristics in heterogeneous environments

In determining whether MODIS is able to represent flux characteristics over heterogeneous surfaces, a common approach is to analyze the scaling behavior derived from higher-resolution data (e.g. Kustas et al., 2004). Due to heterogeneity, one intuitively expects that the limit of flux discrimination should approach the dimensions of the average field. To examine this proposition, Landsat data were successively aggregated to increasing scales, centered at each of the 14 flux towers. Fig. 8 presents these results, providing the areally averaged patterns of evapotranspiration and the resulting standard deviation for soybean and corn vegetation units, averaged across the flux towers. Minimum and maximum values are given, along with ±0.5 standard deviations. The mean value is represented as the mid-point of the distribution.

Some caution in interpreting these results is necessary since the statistics are based on a single instance in time. Still, the result does allow some interpretation of the spatial characteristics of the retrievals, which is not possible using tower flux measurements alone. As the pixel scale is increased, two trends emerge upon comparison of the individual vegetation units. While the average evapotranspiration decreases for corn, there is a similar, although not equivalent, rise in the soybean. Between the 120-m and 1020-m scales, the average evapotranspiration for the corn crop decreases by 66 W/m^2 while the soybean increases by 26 W/m^2. Further, while the change in the standard deviation amongst the tower-averaged evapotranspiration is smaller for corn sites (50–30 W/m^2) than for soybean (60–50 W/m^2), there is a decrease in its absolute value with increasing pixel size (i.e. variation in evapotranspiration between towers is smaller at low resolutions).

The averaged MODIS-based retrievals (1 km) at the available tower locations were approximately 390 W/m^2 and 380 W/m^2 for corn and soybean respectively (see Table 3). These contrast with the equivalently scaled Landsat aggregations from Fig. 8 of

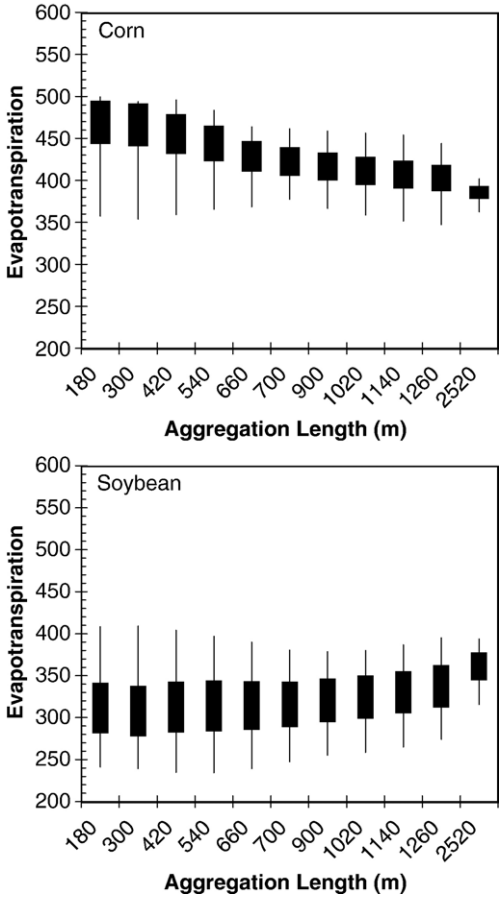

Fig. 8. Variation of the evapotranspiration (W/m^2) and standard deviations at the 14 flux tower sites. Values are determined from Landsat data with increasing aggregation over (top) corn and (bottom) soybean field sites. Lines on each plot represent the average of the flux tower values for the maximum and minimum estimates and the boxes bound the ±0.5 standard deviations (σ/2), with the centre being the mean prediction across tower sites.

410 W/m^2 and 330 W/m^2, which more clearly distinguish between the dominant vegetation units. While the vegetation specific scaling trends from the Landsat image (Fig. 8) tend towards convergence, the scale at which this might occur does not coincide with the MODIS type resolution. In this instance, it appears that simple aggregation studies of high-resolution surface fluxes do not explain sufficiently the spatial characteristics observed here of the retrievals using lower-resolution measurements.

Given the results in Fig. 8, the in-situ flux measurements should be useful in evaluating larger-scale flux estimates. Results from Fig. 8 show that flux averages at the Landsat-5 scale (120 × 120 m) and MODIS scale (1020 × 1020 m) are within 10–15%; however, there is an equivalent rise in the tower-averaged standard deviation occurring within the sampling footprint, suggesting an increase of within field uncertainty with increasing sampling area, which converges towards a maximum value approaching 80 W/m^2. So, while the mean values of the tower-based evapotranspiration converge with increasing scale, the degree of variability within individual fields containing these increases, making accurate estimation at large scales difficult. The multi-sensor estimation (Table 3) also indicates that the Landsat–ASTER predictions are not regularly within 10% of the MODIS values. Again, there seems to be an incompatibility between averaging a single image of high-resolution retrievals and subsequent analyses based on multiple resolutions. Additional observations of field-scale data and scaled evaporative fluxes are needed to better understand the efficacy of tower flux measurements for MODIS-scale retrieval evaluation.

4.6. Influence of land surface classification on flux retrieval

The influence of land cover misclassification and its impact on flux retrievals is not well-understood or appreciated at all sensor scales. Even at Landsat resolutions, there can be misclassification, even though at its high resolution the distinction between soybean and cornfields is clear. However, assessing the effect of misclassification has not been well documented—an issue of some importance given the increased likelihood of occurrence in the mixed agricultural fields typical of the Walnut Creek catchment. To quantify this effect, SEBS-

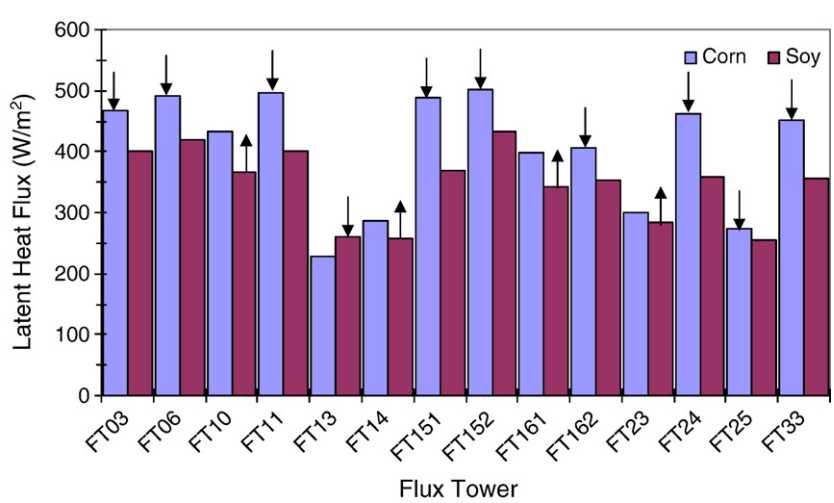

Fig. 9. Changes in SEBS predicted latent heat flux due to variation in land cover type. The position of the arrows identifies the original land cover unit and also denote the direction of change (±) induced by changing the land cover classification.

based Landsat predictions of latent heat flux were re-run, specifying either homogeneous corn or soybean land cover. In both cases, satellite surface temperature measurements remained consistent with the heterogeneous case, with the change in classification affecting only the vegetation height and subsequent changes in the aerodynamic properties and parameters used in the retrieval model.

Fig. 9 details the differences in the SEBS predicted latent heat flux for the changes in the land cover classification. The arrows are positioned above the original land cover type and denote the direction of change, being positive or negative. For corn, all of the nine tower sites within the Landsat scene indicate a decrease in retrieved evaporative flux, with the average decrease resulting from misclassifying corn to soybean being 76 W/m^2. For soybean, four of five tower sites record an increase in retrieved flux, with the average for all towers of 27 W/m^2. For one location (FT13), the change from original soybean to misclassified corn resulted in an estimated flux that was lower by 32 W/m^2.

Even considering the cursory nature of this analysis, it is clear that correctly specifying vegetation type is critical in obtaining accurate flux predictions. While it would seem that high-resolution remote sensing information (i.e. NDVI, Ts, albedo) is sufficient to estimate fluxes in heterogeneous environments (even with application of broad-scale homogeneous land unit classifications such as the MOD12 1-km land cover product), the retrieval accuracy of these predictions will depend strongly on correctly specifying the surface type.

5. Summary and discussion

There is an increasing ability to monitor hydrological variables from space with a combination of sensors offering multi-spatial and multi-temporal estimation. Unfortunately, increased spatial resolution comes at the expense of lower temporal coverage. The possibility of resolving these competing considerations is complicated by the lack of an adequately developed scaling theory for hydrology (see Blöschl & Sivapalan, 1995), which limits an effective synthesis of multi-scale remotely sensed observations. As a result, routine prediction of evapotranspiration is constrained to a great extent by issues of measurement scale, i.e. the scale of the underlying land surface unit.

Following is a brief summary and discussion of this study, presented here to understand and characterize the competing roles of land surface heterogeneity and remote sensing resolution in the retrieval of evaporative flux. Attention is directed towards assessing the potential of multi-sensor evapotranspiration retrievals from space, with particular focus on the consistency between higher-resolution satellite sensor retrievals and those by MODIS, which includes comparisons of spatial mean retrievals over large aggregating areas and the ability of MODIS to capture the spatial variability of surface fluxes.

5.1. Multi-sensor flux estimation

There are a number of difficulties associated with using remote sensing to retrieve the spatial and temporal variation in surface fluxes. The primary difficulty rests with the situation that sensors with sufficiently high spatial resolution to be useful for high-resolution (less than 100 m) retrievals are inevitably coupled to relatively low orbiting platforms with narrow swath-widths, precluding a high temporal repeat rate. Also, these systems are polar orbiting, and thus are sun-synchronous, resulting in overpasses occurring at the same time of the day and limiting the capacity to examine important diurnal effects. Despite these limitations, satellites remain the only viable alternative to measuring relatively small-scale variations, over regional domains, on a regular basis. While better high-resolution characterization of surface fluxes, in the temporal domain, is limited by a simple lack of platforms, there is an increasing ability to examine multiple spatial resolutions from a variety of Earth observing sensors.

The multi-sensor/multi-resolution comparison of evaporation retrievals in this paper indicated good consistency between the high-resolution Landsat and ASTER retrievals. Both the spatial characterization and frequency distribution of the results demonstrate the ability of these sensors to retrieve detail on the land surface flux. Potentially, knowledge of crop health, growth stage, and soil moisture patterns can be inferred from these evaporative patterns, with such information useful in regional irrigated agricultural studies. As shown in Sections 4.1 and 4.2, the agreement between aggregated ASTER and Landsat retrievals was high ($r=0.96$), demonstrating consistency within these two independent data sets. A relatively consistent bias of approximately 25 W/m^2 was attributed to the 32-min difference in overpass times between the two sensors.

Additionally, the surface fluxes estimated from all sensors showed good agreement when aggregated to larger scales, particularly in the mean spatial flux. Maintaining this consistency in the areal average is important at MODIS level resolutions, since the ability to identify crop specific distributions is limited when the length scale of surface heterogeneity is less than the sensing resolution.

5.2. Spatial flux variation

In their study of flux aggregation over Walnut Creek using Landsat data, Kustas et al. (2004) suggested that MODIS type data would be unable to resolve flux variations between soybean and corn. While their conclusion is supported by the work presented here, there is comparatively little knowledge on the ability of the MODIS sensor to accurately characterize fluxes at larger scales. Although our results indicate that MODIS has limited capacity in capturing the spatial variability in fluxes where significant sub-pixel heterogeneity is present, the results also suggest that at large scales it successfully estimates the spatial average flux. The MODIS-based retrievals fell between the bimodal responses of the higher-resolution ASTER and Landsat retrievals, which were shown to differentiate variations in flux due to vegetation differences.

Linearly aggregating Landsat and ASTER data to MODIS equivalent scales disregards the non-linear scaling behavior of surface variables and parameters, and resulted in a mixed level of agreement between the high- and low-resolution sensors.

While ASTER and Landsat retrievals show spatial and statistical consistency, the concurrent MODIS retrievals displayed limited agreement. This lack of consistency within the scaled results raises issues regarding assessments of sensor resolution using aggregated high-resolution data, due to an incomplete understanding of how constituent variables (e.g. surface temperature, albedo, emissivity, NDVI, etc.) scale (e.g. Anderson et al., 2004; Garratt & Prata, 1996; Moran et al., 1997; Njoku et al., 1996). While it is known that surface temperature scales non-linearly, the effects of disregarding this when calculating an areally averaged surface temperature, and the subsequent influence on flux prediction is not well quantified. Techniques to disaggregate lower-resolution surface temperature (i.e. MODIS) to higher resolutions (ASTER/Landsat) are likely to introduce considerable uncertainty into flux prediction, particularly over heterogeneous surfaces, since scaling relationships are not well understood. Further analysis of the scaling characteristics of variables influencing the surface temperature is required to understand the impact of scaling on surface fluxes, and the relationship among retrievals from different resolution sensors.

Results from the aggregation of higher-resolution imagery also indicated substantial spread relative to the essentially point-scale tower data. Generally, the remote sensing retrievals were higher than the corresponding in-situ flux estimates. Although there are considerations with estimating the areal flux as an average of the pixels surrounding the tower, as opposed to calculating the shape of the effective source area (Cooper et al., 2003) and weighting the pixel responses accordingly (see Table 3), the results also demonstrate significant variation within individual agricultural fields, and subsequently within the MODIS pixel. Given this, it is difficult to evaluate thoroughly the larger-scale remote sensing results, particularly over heterogeneous terrain, unless there is sufficient tower data whose locations have been carefully selected.

Techniques to disaggregate surface fluxes from MODIS level resolutions offer possibilities (Kustas et al., 2003; Norman et al., 2003), but additional investigations over longer time periods for a variety of heterogeneous landscapes are required to better assess various approaches. As evaporative flux responds as a complex function of surface temperature, soil moisture, and vegetation condition, it is likely that scaling relationships will vary considerably in response to changes in these variables.

5.3. Other problems and issues

5.3.1. Land cover influences

Accurate flux retrieval requires accurate surface and atmospheric forcing. Correctly specifying the underlying land cover is particularly critical, since many of the aerodynamic and plant-physiological properties (vegetation height, roughness length, displacement height) are related to the surface type. In this paper, scale equivalent surface cover information was determined by aggregating high-resolution (30 m) Landsat classifications. In principal, this appears to be the most accurate approach, but in practice, such detail is not generally available. The highest resolution routinely available for surface information is 1 km, which is unsuitable for identifying unique land units or distinguishing between vegetation classes (i.e. classified as cropland as opposed to corn and soybean). The error induced by assuming a homogeneous cover was computed to range from a decrease of 75 W/m^2 for corn classified as soybean to an increase of 27 W/m^2 for soybean fields classified as corn.

5.3.2. Validation–evaluation data

Comparison of satellite-based retrievals with field-based evaluation data, regardless of their quality, offers an imperfect assessment of satellite retrievals due to a variety of factors, including: site representativeness to sensor pixel scale, sensor accuracy, and the physical equivalence of measured versus observed variables (time averaged atmospheric boundary-layer measurements compared with instantaneous, largely empirically based predictions). Differences between raw flux tower measurements and those adjusted by closure techniques (Twine et al., 2000) or footprint weighting (Schuepp et al., 1990) further influence comparisons to satellite retrievals. In spite of these considerations, the available eddy-correlation data represent the best available means of evaluating the remotely sensed predictions over the Walnut Creek watershed.

To ameliorate the impacts of heterogeneity, flux tower placement becomes a critical consideration, particularly where source areas extend beyond the scale of the field being measured. While the positioning of towers is not arbitrary, there are inevitably non-scientific constraints placed on their location; access to fields by landowners and access to the equipment for servicing two primary examples.

Analysis of the spatial consistency between flux tower data for July 1, 2002, identified a high correlation for corn sites ($r=0.8$), while the soybean fields exhibited a more varied level of agreement. Coupled with results from Fig. 8, both greater inter- and intra-field variability is evident for soybean relative to corn. Across all scales, corn exhibits a reduced range and standard deviation. The most obvious explanation for these trends is that soybean is more closely linked to the surface condition. With a smaller rooting depth and incomplete cover at that time, soybean should naturally be more coupled to the soil wetness condition, whereas corn is likely to be less influenced. While this might seem a rather arbitrary observation, it does prompt some thought on the placement of limited numbers of towers for evaluation purposes and also the concurrent measurement of soil moisture (as was carried out in the SMACEX-SMEX campaign).

6. Conclusion

The investigated SMACEX landscape represents a challenging satellite-based flux retrieval exercise, primarily because of the sharply defined boundaries and patchwork distribution of vegetation and its varied state. While understanding the scale effects is important in the Walnut Creek watershed given the level of heterogeneity, the presented results are applicable to regions where the field scale is greater than the spatial resolution of the satellite sensor. Indeed, it is expected that more

homogeneous landscapes (at least at the pixel scale) should display an increased level of agreement.

The usefulness of satellite-based flux retrievals is tempered by the frequency of their return rate. For practical purposes (i.e. irrigation scheduling/water allocation) the (at best) bimonthly temporal observations offered by high-resolution platforms is inadequate. On the other hand, an estimate of a daily catchment-average evapotranspiration from MODIS would provide an extremely valuable data source for assessing current land and vegetation states—useful in weather prediction, flood forecasting, and other water resource management areas.

While it is expected that different energy balance models will vary in their ability to characterize the surface flux at defined model scales, the results presented here should be representative of flux prediction across scales, irrespective of the model employed. From comparisons of high-resolution satellite data available as part of NASA's Earth Observing System (EOS), a consistent level of agreement in evaporative flux across a small watershed was obtained. Estimates of the latent heat flux at high resolutions agreed well with spatially distributed flux towers, indicating that the SEBS model can correctly partition available energy into surface heat flux components. MODIS-based retrievals also compared well with estimated catchment average mean evaporative flux, indicating considerable potential in extending such retrievals to other landscapes. Further investigation on the scaling behavior of surface temperature and the impact of disaggregating MODIS type resolutions to intermediate scales, particularly in heterogeneous environments, is required to bridge the temporal and spatial divide between operational high-resolution platforms, and provide more useful predictive measures of surface fluxes.

Acknowledgements

The authors would like to thank Dr. Fuqin Li (USDA) and Dr. Andy French (USDA) for providing the high quality surface temperature data from Landsat and ASTER sensors, respectively. Without this information, little progress on the scaling mechanisms in surface flux retrieval would be possible, and their assistance is appreciated. The flux data from the 2002 Soil Moisture Atmospheric Coupling Experiment (SMACEX) over Iowa depended on the participation of the numerous volunteers who collected the data and the organizers who managed the campaign, particularly Dr. Bill Kustas (USDA) and Dr. Tom Jackson (USDA). The assistance of Hongbo Su (Princeton University) in the compilation of surface flux maps is greatly appreciated, along with the many discussions with him on aspects of the analysis. The research is supported through NASA grant NNG04GQ32G: A Terrestrial Evaporation Product Using MODIS Data.

References

Anderson, M. C., Neale, C. M. U., Li, F., Norman, J. M., Kustas, W. P., Jayanthi, H., et al. (2004). Upscaling ground observations of vegetation water content, canopy height, and leaf area index during SMEX02 using aircraft and Landsat imagery. *Remote Sensing of Environment*, *92*(4), 447−464.

Baret, F. J., Clevers, G. P. W., & Steven, M. D. (1995). The robustness of canopy gap fraction estimates from red and near-infrared reflectances: A comparison of approaches. *Remote Sensing of Environment*, *54*, 141−151.

Bastiaanssen, W. G. M., Pelgrum, H., Wang, J., Ma, Y., Moreno, J. F., Roerink, G. J., et al. (1998). A remote sensing energy balance algorithm for land (SEBAL): 2. Validation. *Journal of Hydrology*, *212–213*, 213−229.

Betts, A. K., Chen, F., Mitchell, K. E., & Janjic, Z. I. (1997). Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta Model using FIFE data. *Monthly Weather Review*, *125*(11), 2896−2916.

Beven, K. J., & Fisher, J. (1996). Remote sensing and scaling in hydrology. In J. B. Stewart, E. T. Engman, R. A. Feddes, & Y. Kerr (Eds.), *Scaling up in hydrology using remote sensing.* Chichester, UK: John Wiley & Sons.

Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling—a review. *Hydrological Processes*, *9*(3–4), 251−290.

Brutsaert, W. (1991). *Evaporation into the atmosphere—theory, history and applications.* Netherlands: Kluwer Academic Publishers.

Brutsaert, W. (1999). Aspects of bulk atmospheric boundary layer similarity under free-convective conditions. *Reviews of Geophysics*, *37*, 439−451.

Chen, L. F., Li, Z. -L., Liu, S., Tang, Y., & Zhong, B. (2004). Definition of component effective emissivity for heterogeneous and non-isothermal surfaces and its approximate calculation. *International Journal of Remote Sensing*, *25*, 231−244.

Choudhury, B. J., Idso, S. B., & Reginato, R. J. (1987). Analysis of an empirical model for soil heat flux under a growing wheat crop for estimating evaporation by an infrared-temperature based energy balance equation. *Agricultural and Forest Meteorology*, *39*, 283−297.

Cooper, D. I., Eichinger, W. E., Archuletaa, J., Hipps, L. E., Kao, J., Leclerc, M. Y., et al. (2003). Spatial source-area analysis of three-dimensional moisture fields from lidar, eddy covariance, and a footprint model. *Agricultural and Forest Meteorology*, *114*, 3−4.

French, A. N., Jacob, F., Anderson, M. C., Kustas, W. P., Timmermans, W., Gieske, A., et al. (2005a). Corrigendum to "Surface energy fluxes with the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) at the Iowa 2002 SMACEX site (USA)". *Remote Sensing of Environment*, *99*(4), 471.

French, A. N., Jacob, F., Anderson, M. C., Kustas, W. P., Timmermans, W., Gieske, A., et al. (2005b). Surface energy fluxes with the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) at the Iowa 2002 SMACEX site (USA). *Remote Sensing of Environment*, *99*(1–2), 55−65.

Garratt, J. R., & Prata, A. J. (1996). Surface radiation budget: Scaling up from local observations. In J. B. Stewart, E. T. Engman, R. A. Feddes, & Y. Kerr (Eds.), *Scaling up in hydrology using remote sensing* (pp. 77−91). John Wiley & Sons.

Gillespie, A., Rokugawa, S., Matsunaga, T., Cothern, J., Hook, S., & Kahle, A. (1998). A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) images. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 1113−1126.

Grayson, R. B., & Blöschl, G. (2000). *Spatial patterns in catchment hydrology.* Cambridge: Cambridge University Press.

Hipps, L. E., & Kustas, W. P. (2000). Surface evaporation and its spatial variations. In R. B. Grayson & G. Blöschl (Eds.), *Spatial patterns in catchment hydrology.* Cambridge, UK: Cambridge University Press.

Jackson, T. J. (1997). Soil moisture estimation using special satellite microwave imager satellite data over a grassland region. *Water Resources Research*, *33*(6), 1475−1484.

Jackson, T. J., Chen, D., Cosh, M. H., Li, F., Anderson, M. C., Walthall, C., et al. (2004). Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sensing of Environment*, *92*, 475−482.

Jackson, T. J., & Cosh, M. H. (2003). *SMEX02 regional Landsat imagery for Ames, Iowa. Boulder, CO: National Snow and Ice Data Center. Digital media.*

Kalma, J. D., & Sivapalan, M. (1995). *Scale issues in hydrological modelling.* London, UK: John Wiley & Sons.

Kustas, W. P. (1990). Estimates of evapotranspiration with a one- and two-layer model of heat transfer over partial canopy cover. *Journal of Applied Meteorology*, *29*, 704−715.

Kustas, W. P., & Daughtry, C. S. T. (1990). Estimation of the soil heat flux/net radiation ratio from spectral data. *Agricultural and Forest Meteorology*, *49*, 205−223.

Kustas, W. P., Hatfield, J. L., & Prueger, J. H. (2005). The Soil Moisture Atmosphere Coupling Experiment (SMACEX): Background, hydrometeorological conditions and preliminary findings. *Journal of Hydrometeorology*, *6*(6).

Kustas, W. P., Li, F., Jackson, T. J., Prueger, J. H., MacPherson, J. I., & Wolde, M. (2004). Effects of remote sensing pixel resolution on modeled energy flux variability of croplands in Iowa. *Remote Sensing of Environment*, *92* (4), 535−547.

Kustas, W. P., & Norman, J. M. (2000). Evaluating the effects of subpixel heterogeneity on pixel average fluxes. *Remote Sensing of Environment*, *74*, 327−342.

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

Li, F. Q., Jackson, T. J., Kustas, W. P., Schmugge, T. J., French, A. N., Cosh, M. H., et al. (2004). Deriving land surface temperature from Landsat 5 and 7 during SMEX02/SMACEX. *Remote Sensing of Environment*, *92*(4), 521−534.

Massman, W. J. (1987). An analytical one-dimensional model of momentum transfer by vegetation of arbitrary structure. *Boundary - Layer Meteorology*, *83*, 407−421.

McCabe, M. F., Gao, H., & Wood, E. F. (2005). An evaluation of AMSR-E derived soil moisture retrievals using ground based, airborne and ancillary data during SMEX 02. *Journal of Hydrometeorology*, *6*(6), 864−877.

Monin, A. S., & Obukhov, A. M. (1954). Basic laws of turbulent mixing in the atmosphere near the ground. *Tr. Akad. Nauk SSSR Geofiz. Inst.*, *vol. 24 (151)*. (pp. 163−187).

Monteith, J. L. (1981). Evaporation and surface temperature. *Quarterly Journal of the Royal Meteorological Society*, *107*, 1−27.

Monteith, J. L., & Unsworth, M. H. (1990). Principles of environmental physics. London, UK, Routledge, Chapman and Hall.

Moran, M. S., Humes, K. S., & Pinter, P. J. (1997). The scaling characteristics of remotely-sensed variables for sparsely-vegetated heterogeneous surfaces. *Journal of Hydrology*, *179*, 337−362.

Njoku, E. G., Hook, S. J., & Chehbouni, A. (Eds.). (1996). *Effects of surface heterogeneity on thermal remote sensing of land parameters. Scaling up in hydrology using remote sensing*. West Sussex, UK: John Wiley & Sons.

Norman, J. M., Anderson, M. C., Kustas, W. P., French, A. N., Mecikalski, J. R., Torn, R., et al. (2003). Remote sensing of surface energy fluxes at 10(1)-m pixel resolutions. *Water Resources Research*, *39*(8), 1221. doi:10.1029/2002WR001775.

Norman, J. M., Kustas, W. P., & Humes, K. S. (1995). A two-source approach for estimating soil and vegetation energy fluxes from observations of radiometric surface temperature. *Agricultural and Forest Meteorology*, *77*, 263−293.

Prince, S. D., Goetz, S. J., Dubayah, R. O., Czajkowski, K. P., & Thawley, M. (1998). Inference of surface and air temperature, atmospheric precipitable water and vapour pressure deficit using advanced very high resolution radiometer satellite observations: comparison with field observations. *Journal of Hydrology*, *212–213*, 230−249.

Prueger, J. H., Hatfield, J. L., Kustas, W. P., Hipps, L. E., Macpherson, J. I., Neale, C. M. U., et al. (2005). Tower and aircraft eddy covariance measurements of water vapor, energy and carbon dioxide fluxes during SMACEX. *Journal of Hydrometeorology*, *6*(6), 954−960.

Raupach, M. R., & Finnigan, J. J. (1995). Scale issues in boundary-layer meteorology: Surface energy balance in heterogeneous terrain. *Hydrological Processes*, *9*, 589−612.

Schuepp, P. H., Leclerc, M. Y., MacPherson, J. I., & Desjardins, R. L. (1990). Footprint prediction of scalar fluxes from analytical solutions of the diffusion equation. *Boundary - Layer Meteorology*, *50*, 355−373.

Sobrino, J. A., Raissouni, N., & Li, Z. (2001). A comparative study of land surface emissivity retrieval from NOAA data. *Remote Sensing of Environment*, *75*, 256−266.

Stannard, D. I. (1997). A theoretically based determination of Bowen ratio fetch requirements. *Boundary - Layer Meteorology*, *83*, 375−406.

Su, B. (2002). The surface energy balance system (SEBS) for the estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences*, *6*(1), 85−99.

Su, H., McCabe, M. F., Wood, E. F., Su, Z., & Prueger, J. H. (2005). Modeling evapotranspiration during SMACEX02: Comparing two approaches for local and regional scale prediction. *Journal of Hydrometeorology*, *6*(6), 910−922.

Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P. R., Meyers, T. P., et al. (2000). Correcting eddy-covariance flux underestimates over a grassland. *Agricultural and Forest Meteorology*, *103*(3), 279−300.

Wan, Z., & Dozier, J. (1996). A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Transactions on Geoscience and Remote Sensing*, *34*, 832−905.

Wan, Z., Zhang, Y., Zhang, Q., & Li, Z. -L. (2002). Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. *Remote Sensing of Environment*, *83*, 163−180.

Wood, E. F. (1994). Scaling, soil moisture and evapotranspiration in runoff models. *Advances in Water Resources*, *17*, 1−2.

Xavier, A. C., & Vettorazzi, C. A. (2004). Mapping leaf area index through spectral vegetation indices in a subtropical watershed. *International Journal of Remote Sensing*, *25*(9), 1661−1672.