Remote Sensing of Environment 158 (2015) 281–294

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Remote sensing of evapotranspiration over cotton using the TSEB and METRIC energy balance models ☆

Andrew N. French *, Douglas J. Hunsaker, Kelly R. Thorp

U.S. ALARC, 21881 North Cardon Lane, Maricopa, AZ 85138, USA

ARTICLE INFO

Article history:
Received 30 September 2013
Received in revised form 24 September 2014
Accepted 11 November 2014
Available online xxxx

Keywords:
Two source energy balance
METRIC
Maricopa
Arizona
Cotton
Irrigation
Remote sensing
Thermal infrared

ABSTRACT

Remote sensing of evapotranspiration (ET) can help detect, map and provide guidance for crop water needs in irrigated lands. Two remote sensing ET models based on thermal infrared (TIR), the Two-Source Energy Balance (TSEB) and the Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC), were tested for accuracy, and bias at fine (1 m) and moderate (30–120 m) spatial scales. Airborne and Landsat data were collected over Maricopa, Arizona in 2009 and 2011 as part of a cotton irrigation scheduling study. Based on soil moisture observations at 112 locations across 4.9 ha and image data spanning two growing seasons, TSEB and METRIC were found similarly accurate at both fine and moderate scales with average discrepancies no more than 1.9 mm/day. Tests at 1-m scales showed that TSEB and METRIC model sensitivities were seasonally correlated, with greater sensitivity modeled by METRIC in early growth and slightly greater sensitivity by TSEB at maturity. Time integration of flux estimates was done by assuming constant evaporative fraction and was also tested for 2011 data using ground-based TIR radiometers; this latter approach improved daily ET estimates by 0.8 mm/day or better in two cases. Time-series assessment of the utility of using evaporative fraction as a water-stress indicator was tested using Landsat data and both TSEB and METRIC. Two early season water depletion events were detected and none in mid-season. The impact of overpass frequency upon ET estimates was tested for the field as a whole and found that cumulative ET estimates were significantly affected, up to 200 mm out of ~1000 mm consumed. Results from this study showed that for ET accuracy, TSEB and METRIC perform similarly. METRIC is preferred when model ancillary data are sparse, while TSEB is preferred when support data are plentiful. Future ET modeling should consider implementing both to take advantage of their seasonally dependent sensitivities.

Published by Elsevier Inc.

1. Introduction

Accurate maps of evapotranspiration (ET) over crops are a way to improve detection of crop water stress, refine irrigation scheduling, and help manage scarce water supplies. Recently ET maps have begun to be incorporated within drought forecasting systems (Anderson et al., 2013), and thus are beginning to have major local and global impacts. ET mapping is also becoming important for management at watershed scales (Gibson, Münch, Engelbrecht, & Petersen, 2009; Kongo & Jewitt, 2006) and for water allocations (Consult, 2011). In recent times much has been written about ways to create maps using remote sensing data. These ways include use of vegetation indices derived from visible and near infrared bands (VNIR), predominately red (~670 nm) and near infrared (~790 nm), and the inclusion of thermal infrared (TIR) bands, predominated by bands over 10–13.5 μm. Using the normalized difference vegetation index (NDVI) in combination with crop coefficients has been shown by Glenn, Neale, Hunsaker, and Nagler (2011), Hunsaker, Fitzgerald, French, Clarke, and Pinter (2007), and others, to be an effective way to map ET over crops. When coupled with ancillary data and estimates of crop coefficients, ET can be reasonably estimated under standard, non-water-stressed conditions (Gonzalez-Dugo et al., 2009). The choice of modeling ET with VNIR data comes with a substantial advantage: satellite data at these wavelengths are readily available at <100 m resolution, often at no-cost (e.g. Landsat through landsat.usgs.gov). However, use of VNIR data alone also has a distinct disadvantage: short term onset of water stress signals from plants cannot be readily detected (Pinter et al., 2003) except at very fine resolution. Eventually over several days there will be changes in canopy architecture with consequent changes in reflectance, but for applications requiring near real-time information, VNIR-based ET maps will not be sufficient.

Using land surface temperatures (LST) derived from TIR data, in contrast, can provide the needed short-time information. Dehydrated plants are unable to transpire and lack of evaporative cooling results in elevated canopy temperatures. Water shortage in plant root zones is quickly represented by anomalous high plant canopy temperatures. The temperature changes exceed 1 K and are measurable from space. When combined with a surface energy balance model, LST data can be

☆ The USDA is an equal opportunity provider and employer.

* Corresponding author.
E-mail address: andrew.french@ars.usda.gov (A.N. French).

http://dx.doi.org/10.1016/j.rse.2014.11.003
0034-4257/Published by Elsevier Inc.

used to produce instantaneous ET estimates and plant stomatal conductance (Blonquist, Norman, & Bugbee, 2009), a direct indicator of plant stress. Preeminent energy balance models include one-source, contextual models such as SEBAL (Bastiaanssen, Menenti, Feddes, & Holtslag, 1998), its open-source variant, METRIC, (Allen, Tasumi, & Trezza, 2007) a time-integrated variant (Sun et al., 2009), VI/LST/resistance triangle approaches (Carlson, Capehart, & Gillies, 1995; Jiang & Islam, 1999), and the two-source biophysical approach, TSEB (Norman, Kustas, & Humes, 1995).

Questions about which model is best, or which to use, arise frequently. Model inter-comparisons (Gonzalez-Dugo et al., 2009; Timmermans, Kustas, Anderson, & French, 2007) help to highlight model benefits and shortcomings and guide future model development. Results from studies are equivocal, demonstrating good results in some seasons and poor results otherwise. For example, Chirouze et al. (2014) compared instantaneous ET results over crops in Northern Mexico and sometimes found good results with both contextual and dual source approaches and sometimes not.

For studies focused on ET from remote sensing over irrigated crops, the pathway ahead remains unresolved because high spatial resolution data (<100 m) in reflected and emission bands are required. They are not routinely available. Thus model performance at irrigation treatment scales is difficult to evaluate at time steps ranging from days to months. If remote sensing models cannot be shown to be robust and consistently more accurate than standardized weather-based ET models such as Penman-Monteith (Allen et al., 2005) then there is not good justification to implement them. Furthermore, existing model demonstrations typically utilize surface energy flux station measurements for validation (e.g., Byun, Liaqat, & Choi, 2014; Choi et al., 2009), an important but spatially blunt tool for measuring ET over discontinuous or patchy irrigated crops. For irrigation scheduling research at Maricopa, however, ET is obtained from intensive soil moisture monitoring and thus could provide more meaningful validation data for model assessment than would otherwise be possible. To that end a comparison study was conducted to evaluate two accessible but distinctly different models: TSEB and METRIC.

TSEB offers a physics-based approach: energy fluxes between the soil surface, plant canopy, and the overlying air are modeled and supported by physically meaningful parameterizations leading to distinct estimates of transpiration and evaporation from non-plant surfaces. This separation, part of crop ET models such as FAO56 Pereira, Allen, Smith, and Raes (2015), quantifies how much irrigation water is beneficial to plant growth compared with non-beneficial water loss at the soil surface. Implementation of TSEB, however, is complex, sensitive to LST observation errors and algorithmically incomplete without a constrained potential ET parameterization.

METRIC (and its parent model SEBAL) on the other hand, relies upon contextual LST data to model energy fluxes and makes no attempt to differentiate soil and canopy. 'Contextual' is used in the sense that LST values over target sites are modeled with respect to LST values observed at the same time and spatially nearby. This means that METRIC reference pixels can be applied to single or nearly simultaneous adjacent remote sensing scenes. Though criticized for physical simplifications, METRIC offers a major advantage over TSEB: its self-calibrating approach avoids difficult-to-resolve errors and uncertainties in LST data. METRIC enforces meaningful constraints on temperature end-members wherein the coldest pixels represent conditions close to potential ET and the hottest pixels represent conditions with minimal latent heat flux. Results from studies using METRIC/SEBAL worldwide are certainly encouraging but questions remain about how much local calibration is required and how to make the approach more objective and repeatable.

Although both TSEB and METRIC have been implemented to produce ET estimates, their underlying objectives are not the same and it makes little sense to compare the models in their entirety. Specifically, the salient features that need comparison are the turbulent flux components. While formulations for net radiation soil heat flux components are certainly different, their distinctiveness has little to do with LST data. For TSEB the emphasis lies with canopy geometry and separation of fluxes between plant and soil. The METRIC emphasis lies with atmospheric correction of satellite VNIR to obtain albedo estimates regardless of canopy structure. Thus the approach taken here is to conduct a TSEB/METRIC inter-comparison by providing both with the same net radiation and soil heat flux inputs and then evaluating ET outcomes. The inter-comparison is based on extensive observations over a cotton experiment in Maricopa, Arizona conducted 2009 and repeated in 2011. Remotely sensed data included airborne and Landsat observations. Companion papers describe geospatial modeling, crop simulations (Thorp et al., in review) and irrigation scheduling approaches (Hunsaker et al., in review). Considering that the TSEB algorithms incorporate biophysical soil and canopy properties not provided by METRIC, the TSEB net radiation and soil heat flux estimates were used as standard inputs for both. While this choice does mean that parameterization errors will propagate to both model outputs, differences in turbulent flux estimates will not be confounded.

Hence the presentation of the paper describes the methodology in Section 2, containing some mathematical modeling details in Section 2.1, followed by an overview of the experimental plan in Section 2.4. Results from model implementations are reported in Section 3, which are interpreted in Section 4 and summarized in Section 5.

2. Methods

2.1. Remote sensing of ET

Estimation of ET with the TSEB and METRIC approaches begins with energy balance:

$$LE = R_n - G - H \tag{1}$$

where LE is latent heat flux, R_n is net radiation, G is soil heat flux, and H is sensible heat flux (all computed in W/m^2). Photosynthetic and heat storage are neglected as minor components, being <5% of R_n, although in some instances the latter component may be important for full canopy during morning hours (Meyers & Hollinger, 2004). LE, the target term, cannot be measured with remote sensing and is computed as the residual from solving the other three terms in Eq. (1). To recover ET as a liquid water depth, LE values are divided by latent heat of vaporization and density of liquid water. Since remotely sensed LE values are an instantaneous observation not representative for the entire day, an extrapolation approach is needed. Here two approaches were investigated. In one, a constant evaporative fraction (EF) assumption (Lhomme & Elguero, 1999) is used, i.e., $EF = LE/[R_n - G]$ = constant. During mid-day hours EF is nearly constant, meaning that a single time of day observation could be sufficient for daily ET estimation. Constancy of EF, however, is not assured since it depends upon multiple factors, some of the more important being cloudiness, advected heat or moisture (Crago, 1996) and phase difference between net radiation and soil heat flux (Gentine, Entekhabi, Chehbouni, Boulet, & Duchemin, 2007). Daily ET (mm) is computed:

$$ET_{Daily} = 1000 \times \frac{EF}{\rho\lambda} \times (R_n - G) \times \sum_{t=0}^{t=n} \frac{R_{s,t}}{R_{s,o}} \Delta t \tag{2}$$

where $R_{s,o}$ is incoming solar radiation (W/m^2) at remote sensing time, $R_{s,t}$ is estimated incoming radiation (W/m^2) over the whole day, ρ is water density (kg/m^3), λ is heat of vaporization (J/kg), and Δt is the time sample interval (s). For this study at Maricopa, time step t was one hour, $n = 24$, and both R_s, o and $R_{s,t}$ were obtained from AZMET (Brown, 1989) observations (Table 1). The second approach used

Table 1
Field conditions at remote sensing times. Values are shown for day of year (DOY), 2 m air temperature (Air), relative humidity (RH), solar radiation (Solar).

Year	Date	DOY	Air (C)	RH %	Wind (m/s)	Solar (W/m^2)	Plant (m)
2009	27 May	147	31.7	1.6	444	969	0.42
2009	3 June	154	35.4	9.3	1.0	994	0.42
2011	30 July	211	34.9	28.9	2.2	947	1.35
2011	7 July	188	38.7	37.6	1.2	3.22	0.78
2011	21 July	202	35.2	38.1	1.2	3.13	1.0
2011	18 August	230	38.8	38.0	1.0	3.14	1.1
2011	8 September	251	37.7	27.0	0.8	3.10	1.2

ground-based infrared thermometers (described below) in an LST image extrapolation scheme to avoid the constant EF assumption. In a small-plot study (French, Hunsker, Clarke, Fitzgerald, & Pinter, 2010), LST values were estimated to 2.1 °C accuracy and 0.4 °C bias and thus the approach could be a viable approach to obtain ET estimates continually throughout the day. Other approaches (Colaizzi et al., 2014; Peters & Evett, 2004; Tang, Li, & Sun, 2013) use reference temperatures and reference ET at a fixed location to scale canopy temperatures and could be equally viable. Due to this extrapolation, the scheme cannot be used with the METRIC algorithm, but can be tested with TSEB. If spatially distributed LSTs over uniformly irrigated plots can be scaled in proportion to a mid-field reference observation point, then the energy balance models can be run at multiple times of day, results from which can then be summed directly:

$$LST_{pred} = LST_{obs} * [LST_{node} - LST_{\circ}] / [LST_{max} - LST_{\circ}] \quad (3)$$

where the predicted LST at any time of day (LST_{pred}) is a function of two terms: LST at overpass (LST_{obs}) and scaled LSTs observed by a ground-based radiometer (second term). This scale term consists of radiometer LST at the desired extrapolation time (LST_{node}), the pre-dawn LST (LST_{o}), and peak observed LST (LST_{max}). The potential advantage of this approach is more accurate modeling of daily ET on days with variable weather, cloudy skies, or when EF is not constant.

Models TSEB and METRIC are described by model authors Norman et al. (1995) and Allen et al. (2007), respectively, and are authoritative. Nevertheless, some details are repeated here to clarify their implementation in this study. TSEB and METRIC were invoked using code written in-house. Models were implemented in C (MinGW-w64 with GCC 4.8), file handling was done with Python V3.3, with data preparation and analyses done in Excel 2007 and R 3.0.1 (R Core Team, 2013).

R_n is the result of incoming and outgoing shortwave and longwave radiation:

$$R_n = (1-\alpha)R_s + \epsilon\, R_{L\downarrow} - \epsilon\sigma T^4 \quad (4)$$

where α is surface albedo, R_s is downwelling solar radiation, ϵ is surface longwave emissivity, R_L is downwelling longwave radiation, σ is the Stefan–Boltzmann constant, and T is surface temperature, raised to the fourth power to approximate blackbody radiation. For this study, albedo, short wave radiation, long wave radiation were all estimated following procedures outlined in Norman et al. (1995), in Campbell and Norman (1998), and incorporating local observations of weather, plant height, and leaf angle distribution (spherical used in 2009 and 2011).

Solar radiation was taken from hourly AZMET data (Table 1) for airborne overpass times, but verified against clear sky day estimates computed via procedures in Allen, Pereira, Raes, and Smith (1998). At this point, Eq. (4), is essentially the same formulation for TSEB and METRIC, absent atmospheric correction. In this study, none of the reflectance correction steps described in Allen et al. (2007) were used: for the airborne data they were not needed since data were calibrated against reference reflectance tarps, while the satellite data products included corrections.

G, a minor component of Eq. (1), cannot be observed remotely and was estimated indirectly. TSEB and METRIC approaches are not identical, but a common method was needed to ensure consistent estimates of available energy. Here the TSEB formulation was adopted:

$$G = c_G R_{n,s} \quad (5)$$

where a constant c_G was assumed, here set to 0.2, and $R_{n,s}$ is net radiation at the soil surface, as modeled within the TSEB multiscattering algorithm. Optimal values for c_G have diurnal variation (Friedl, 1996) but in this study no adjustment was provided.

The remaining terms in Eq. (1), turbulent fluxes H and LE, were solved in distinctly different ways for the two energy balance models.

2.2. TSEB turbulent flux computation

Estimation of ET by TSEB models the two sources of latent heat, transpiration and evaporation, distinctly. TSEB disaggregation accommodates the different energy transport mechanisms. Two TSEB variants exist, parallel and series network formulations, to explore the value of direct and indirect modeling of energy exchange between the soil and canopy. Results from either approach are similar. Here the series network was implemented because results from a row-crop study (Colaizzi et al., 2012) reported its greater accuracy. A key part of TSEB is incorporation of radiometric temperature, as this observation is indicative of plant water stress levels (Jackson, Reginato, & Pinter, 1981). For canopy latent heat flux (LE_C):

$$LE_C = R_{n,C} - H_C \quad (6)$$

where $R_{n,C}$ is net radiation at the canopy and H_C is sensible heat flux at the canopy. For the soil latent heat flux (LE_S):

$$LE_S = R_{nS} - G - H_S \quad (7)$$

Using multispectral image data in the VNIR and TIR, the 5 unknowns on the right hand sides of Eqs. (6) and (7) can be reduced to 1. The remaining unknown is solved by setting the Priestley–Taylor parameter α in the equation for canopy latent heat (LE_C):

$$LE_C = \alpha f_G R_{n,C} \left[\frac{\Delta}{\Delta+\gamma}\right] \quad (8)$$

α is nominally 1.26, but adjusted according to local conditions. f_G is the fraction of the vegetation that is green, Δ is the slope of the saturation vapor pressure vs. temperature function at ambient air temperature (kPA/K), and γ (kPA/K) is the psychrometric constant. Knowledge of α is critical for temperature disaggregation. Without this constraint (or some other one, e.g. observations provided at multiple view angles), Eqs. (6) and (7) are indeterminate.

The remaining TSEB equations, fully described in Norman et al. (1995), solve for soil sensible heat flux using radiometric temperature and subsequently soil latent heat flux by enforcing energy balance. To summarize, a TSEB innovation is to provide a way to estimate E and T separately. Provided accurate observed temperatures and reasonable estimates of α are used, both LE_S and LE_C can be accurately estimated. While doing this, it addresses the H over-estimation problem encountered in one-source models. In the latter case, excess resistance is introduced to offset effects caused when using radiometric temperature in place of aerodynamic temperature. Logical problems with this approach are discussed in Sun and Mahrt (1995).

2.3. METRIC turbulent flux computation

METRIC solves *H* from contextual image data and then *LE* by residual:

$$H = \rho_a c_p \frac{\Delta T}{r_{ah}} \quad (9)$$

where ρ_a is air density in kg/m^3, c_p is volumetric heat capacity in J/(kg K) and r_{ah} is resistance in s/m. The key term, ΔT, is the apparent difference between near surface air temperature and LST. It is solved in a different way from TSEB:

$$\Delta T = a + bT_s \quad (10)$$

where *Ts* is a best estimate of LST. The linear coefficients *a* and *b* are determined by using extreme LST values within the same remote sensing scene (Allen et al., 2007), using previously obtained values for R_n and *G*, and solving Eq. (1) for *H* for temperature extremes:

$$b = \frac{\Delta T_{Hot} - \Delta T_{Cold}}{T_{s,Hot} - T_{s,Cold}} \quad (11)$$

$$a = \Delta T_{Hot} - bT_{s,Hot} \quad (12)$$

where ΔT_{Hot} and ΔT_{Cold} are temperature gradients obtained from solving the energy balance equation at respectively hot and cold pixels. $T_{s,Hot}$ and $T_{s,Cold}$ are observed LST at the same hot and cold pixels. (Note the error in equation '50' in Allen et al. (2007), its solution returns *b*). At the chosen extreme cold value, *LE* was assumed 1.05 × alfalfa reference ET (Tasumi, 2003; Tasumi, Allen, Trezza, & Wright, 2005; Wright, 1982). For the comparable hot value *LE* was set to zero. LST at the hot pixel was obtained by correcting remotely sensed values for reflected downwelling sky radiance (Brutsaert, 1982) and a surface soil emissivity known from local experience to be ~0.97.

A major point of contention is how to select the extreme temperature values (Long & Singh, 2013), although research progress has been made to indicate that non-subjective pixel selection is feasible for first-order accuracy corrections (Morton et al., 2013). This study did not attempt to resolve the controversy. However, to help maintain objectivity and consistency within this study, we selected hot and cold reference pixels in a predetermined way. First, reference values were based on cluster means and never upon a single pixel. By necessity cluster sizes varied according to conditions and sensor resolution. For airborne data, sizes were in the thousands since many pixels fit within each irrigation treatment. For satellite data, sizes ranged from 2–8 depending upon Landsat resolution and positions of scan-line gaps. Second, cluster selection was based on LST statistics over the site and adjacent fields with cold and hot extremes chosen at 0.1% and 99.9% quantiles. For the airborne data pixel selection was limited to image mosaic extents, while for Landsat data the limitation was for cropland within the Maricopa Agricultural Center and immediately adjacent farms to the north and east. These constraints ensured that the bounding end members came close to spanning the full range of temperatures and occurred within the same local climate and soil types.

2.4. Maricopa cotton experiment 2009 & 2011

Experiments were conducted in March–October 2009 and repeated in March–October 2011 on a 4.9 ha cotton field (denoted 'F33') at the University of Arizona, Maricopa Agricultural Center (33° 04′ N, 111° 58′ W, 361 m MSL) in Maricopa, Arizona, USA (Fig. 1). Objectives for the experiment were to evaluate ways to improve irrigation scheduling in near real time (Hunsaker et al., in review), to monitor the spatial distribution of crop water use with remotely sensed image data, and to extend crop modeling in the geospatial domain (Thorp et al., in review). The field soil is a Casa Grande sandy clay loam (reclaimed fine–loamy, mixed, superactive, hyperthermic, Typic Natriargid; (Post, Mack, Camp, & Sulliman, 1988)). Experimental design was randomized complete block with 16 irrigation borders, each 12 m × 168 m consisting of four irrigation scheduling treatments in four blocks. Treatments were 1: 'FAO', a crop coefficient approach following FAO-56 methodology (Allen et al., 1998), 2: 'VI_A', a crop coefficient approach controlled by the remotely sensed vegetation index (NDVI) driven estimates of soil water depletion to 45% total available water (TAW) of the crop root zone as averaged for all zones (each border was analyzed in subdivisions of 4 m × 8 m zones) 3: 'VI_B', the same crop coefficient approach as for 'VI_A' but with a depletion threshold of 65% TAW for 5% of zones, and 4: 'MAC', a control irrigation approach wherein scheduling followed conventional on-farm practice. What constitutes 'on-farm practice' was an ill-defined and difficult-to-replicate treatment conducted by MAC farm managers using elapsed calendar days and their independent discretion. Despite this lack of control, 'MAC' was included in the design to allow representation of irrigation deliveries typical for central Arizona farmers. Following model estimates, water was delivered to each border plot from the south via a concrete-lined ditch and gated polypipe at 1.0 m spacing (but via siphon tubes in 'MAC' borders). Irrigation volumes were measured separately for each border using an in-line water meter in the irrigation pipe. Notable events are listed in Table 2.

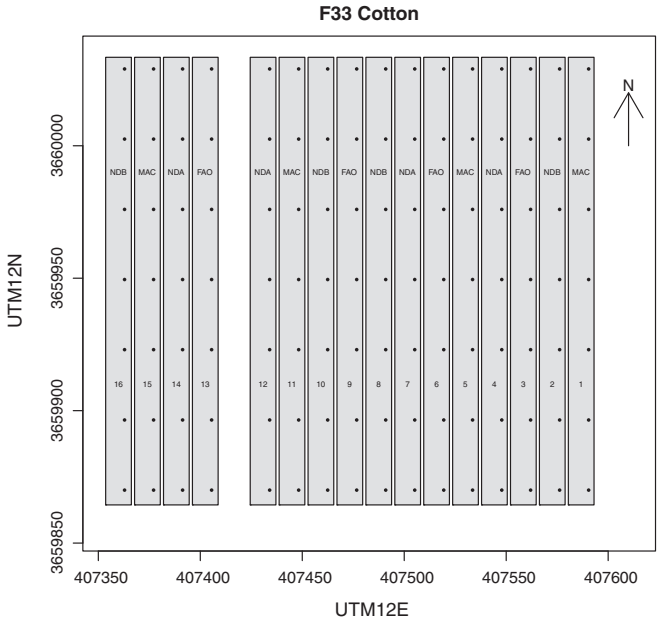

Fig. 1. Experimental layout for cotton grown at Maricopa, Arizona in 2009 and in 2011. Numbers denote plot identifiers, codes indicate their corresponding irrigation treatment type, dots indicate location of neutron probe access tubes (7 × 16) for soil moisture monitoring.

Experimental data collected included meteorological, soil, agronomic and hydrological observations at ground level. These observations monitored plant and water status throughout the growing season and

Table 2
Cotton experimental events 2009 & 2011.

Event	2009	2011
Planting	22 April (112)	19–20 April (109–110)
Emergence	28 April (118)	~26 April (116)
1st Irrigation	21 May (141)	12 May (132)
Defoliation	25 September–11 October (268–284)	4–5 October (277–278)
Harvest	27–28 October (300–301)	11–14 October (284–287)
Total rainfall (mm)	68	24
Total in-season irrigation (mm, est.)	840	850

provided data essential for parameterizing and running the surface energy balance models.

Meteorological data, including air temperature, humidity, wind speed and direction, and solar radiation, were collected from the nearby (~1700 m from F33) Maricopa AZMET station (ag.arizona.edu/azmet) and in 2011 from one in-field temperature–humidity sensor (HMP-45C-L, Vaisala) in a wireless network (Mention of a trade name, proprietary product, or specific equipment does not imply approval or endorsement by the USDA to the exclusion of others). The network was a ZigBee (www.zigbee.org) based system made by Crossbow (now MEMSIC Inc. Andover, MA) and consisted of 12 field nodes and one base node at the main USDA facility. A ZigBee network is a low-power, low-throughput, self-meshing system and thus well-suited for this agricultural study. In addition to air temperature/humidity data, the nodes transmitted radiometric temperatures of the plant canopy surface at 7–15 min intervals. These radiometers (SI-111, Apogee Instruments, Logan, UT) had a field of view of 44°, were mounted in selected borders (2, 4, 5, 6, 7, 8) on steel pipe at approximately 2 m height and close to soil moisture access tubes. They were inclined 45° from horizontal and towards the northeast. This orientation maximized plant canopy view fraction, avoided interception of sky temperature, and did not have self-shadowing effects except for brief (<20 min) afternoon intervals. Measurements were corrected for sky irradiance (Brutsaert, 1982) using an assumed 0.97 surface emissivity. All radiometers were factory calibrated to ~0.1 °C accuracy and checked in-house using our constant temperature room facility.

A key part of observations were soil moisture data collected with Model 503 neutron probes (Campbell Pacific Nuclear, Martinez, CA). Use of neutron probe observations is the most accurate and consistent way to determine in situ soil moisture profiles (Evett & Steiner, 1995). In this experiment they are superior to near surface flux observations from systems such as eddy covariance or scintillometry because observations represent the specific borders and do not depend upon fetch. They provided ET validation data by differencing observed volumetric soil water content on sequential dates (p. 80, Jensen, Burman, & Allen, 1990). Neutron probes detected soil moisture profiles from 0.1 m to 2.9 m in depth at 0.2 m intervals using access tubes installed with a tractor-mounted Giddings soil sampler (Model 25-TS, Giddings Machine Co., Windsor, CO). Moisture at the surface was not monitored. There were 112 tubes, with seven each per border arranged north–south (Fig. 1). Based on field calibrations volumetric soil moisture data were accurate to better than 0.02 m^3/m^3. Probe data were collected about every 7 days, with one time just prior to irrigations, and another 4–5 days after irrigation. Soil water balance calculations of ET, hearafter denoted 'SM', included measured irrigation amounts from the in-line water meter and measured precipitation from the AZMET station, when applicable for soil measurement intervals. Post-season monitoring of soil moisture content showed that deep percolation rates below the cotton root-zone to be <0.5 mm/week, an un-detectable flux when differencing soil moisture values over 7 day time steps.

2.4.1. Remote sensing data

Two sets of remote sensing observations were collected for the 2009 and 2011 study periods: airborne and Landsat satellite data. Initial plans were to use only airborne data; these would collect at data at fine scale (<1 m) and resolve the different irrigation treatments. As noted below, however, multiple instrument problems greatly reduced the number of acquired scenes to seven. While this was still meaningful for daily ET validation, a greater sample set was needed to evaluate model performance over the full cotton growing season. Although insufficient to resolve individual treatments, Landsat data could fill this gap on a field-wide basis since the study lacked major growth and irrigation discontinuities. Using both TM5 and TM7, spatial resolution was sufficient to observe the field sometimes as frequently as every 8 days with up to 45 pixels in the visible/near infrared bands 3 and 4, and up to 4 pixels in the thermal band 6.

Airborne remote sensing observations were collected multiple times in 2009 and 2011 (Table 3) using helicopter-based imaging platforms (Bell 46G in 2009, Hiller UH-12 in 2011) flying at ~800 m above ground level (AGL). Components were mounted on an aluminum plate, 6.35 mm × 370 mm × 305 mm, suspended underneath the airframe with a custom-made steel frame. Gross weight was less than 20 kg. The choice of a helicopter in these years, as opposed to a fixed-wing aircraft, was governed by camera frame rates needed to image small field sizes at high resolution and low airspeeds (since 2011, newer equipment has been acquired with 10+ Hz frame rate capabilities and lower-cost fixed-wing operations are now being conducted). The platform components varied over the experimental year, though data requirements did not: needed were <1 m image data in visible, near infrared bands (VNIR) to create vegetation index images, and data in the 8–13 μm thermal band to create land surface temperature maps. The components were a Duncan MS3100 3-band multispectral camera, 1392 × 1040 pixels (Geospatial Systems, now Optech, West Henrietta, NY), a FLIR 645C, 640 × 480 pixels, 25° × 18.8° field of view, 0.69 mrad, 7.5–13 μm, thermal infrared camera (FLIR Systems, AB, Danderyd, Sweden), laptop computer and portable 12 V battery. After failure in 2009, the Duncan MS3100 was replaced first with a pair of 8-bit EO machine vision cameras, individually filtered (filters acquired from Andover Corp., Salem, NH) and aligned 8-bit rolling shutter (IDS Imaging Development Systems GmbH, Obersulm, Germany), then in mid-2011 with a GSI MS4100.

Airborne red and NIR reflectance data were collected using 10 nm bandwidth filters centered at 670 nm and 790 nm. Co-registration quality of Duncan VNIR data was very good with only sub-pixel deviations between red and NIR image pixels. In contrast the machine vision cameras were not precision aligned and required intensive hand-selection of ground control points to achieve co-registration accuracies to 1 m. Ground reflectances were determined using 8 m × 8 m reference tarps (Group VII Technologies, Provo, UT), 4, 8, 48, and 64% reflectances, deployed at the field edge.

Airborne TIR data were collected over the 8–14 μm atmospheric window. After conversion into a FLIR public format using ImagineIR software, LST data from the FLIR camera were checked against the ground-based TIR data for temperature accuracy. They were found not significantly different within the camera specified accuracy of 2 °C. Prior experience at MAC showed that the fine texture soils had emissivities ~0.97 while cotton canopy emissivities exceed 0.98, thus uncertainties created by spatial emissivity variations were not expected to be significant.

Hand-held high-resolution oblique images were collected during flights using a Nikon D5000 camera. These images were used in conjunction with USGS National Map orthophotography to perform georegistration. In most cases resolution was <0.2 m.

As noted, multiple logistical and equipment problems interfered with sampling frequency of airborne data, necessitating changing camera configurations and eliminating some flight data from the analyses. Examples of problems encountered included excessive platform vibration, high ambient temperatures (>40 °C), serial communication failure, and image capture lags. Consequently only data from 3 days in 2009 (27 May, 3 June, 30 July) and 4 days in 2011 (7 July, 21 July, 18 August, and 8 September) were available for high spatial resolution energy

Table 3
Airborne remote sensing acquisition dates and platform components.

Year	Date	DOY	Time (MST)	Flight number	Equipment
2009	27 May	147	10:00	3	Duncan, FLIR
2009	3 June	154	10:10	4	Duncan, FLIR
2009	30 July	211	10:00	5	Duncan, FLIR
2011	7 July	188	11:00	3	EO, FLIR
2011	21 July	202	11:10	4	EO, FLIR
2011	18 August	230	11:50	6	GSI, FLIR
2011	8 September	251	12:00	7	GSI, FLIR

balance modeling. Considering that the realized airborne data would not be sufficient for testing TSEB and METRIC for seasonal scale water use, Landsat data were incorporated into the study. Though these data could only evaluate ET over the field as a whole, the collection turned out to be remarkably complete. In each year, 2009 and in 2011, 17 TM5/TM7 cloud-free and scan-gap free scenes during the cotton growing seasons. Considering an 8-day potential rate, successful acquisitions were 77%, a value confirming the suitability of Maricopa for testing remote sensing techniques.

Landsat data (Table 4) from WRS path 37 row 37, overpass time ~10:53 MST, were downloaded from the USGS site earthexplorer.usgs.gov and processed in two ways. For NDVI images, the LEDAPS reflectance processing (Schmidt, Jenkerson, Masek, Vermote, & Gao, 2013) was used since that product included atmospheric correction. For LST images, L1 data were collected, then atmospherically corrected, using MODTRAN5 (Berk et al., 1998) and NOAA ESRL radiosondes (www.esrl.noaa.gov/raobs/) launched from the Tucson, AZ site 'TWC' (32.23° N, 110.96° W). An alternative correction tool was considered (Barsi, Schott, Palluconi, & Hook, 2005), but was not used; its profile resolution was not as good as radiosonde data and its spatial resolution (NCEP 2.5° grid cell size) was significantly greater than the 132 km distance between Maricopa and 'TWC'. Comparison between Landsat and ground-based LST values for 6 overpass dates in 2011 showed good correction accuracy with a mean difference of 1.8 °C. Pixel selections for all scenes were done using the previously described quantile protocol by first defining customized regions of interest for each scene; this ensured that mean NDVI and LST values represented only the cotton field site or nearby reference sites.

2.4.2. ET modeling implementation

Having assembled remote sensing image data, computation of F33 ET values were done by running TSEB code, then METRIC code. Air temperature, humidity, solar radiation and wind speed were assumed to have no spatial variability across F33. Georegistered NDVI, LST and plant height maps were combined with meteorological data and input to the series-formulated TSEB. Output soil and canopy LE fluxes were summed, then time integrated to daily values using the constant EF approach. EF was computed from TSEB estimates of R_n, G and LE. Resulting daily ET values were then averaged over each of the 16 irrigation border treatments. Using the same NDVI, LST images data, plus the TSEB-derived R_n and G terms, the METRIC code was run. For pixel end member selection, quantiles from the LST data were generated with end-member selections made at the 0.1% and 99.9% probability levels.

Table 4
Satellite remote sensing acquisitions.

2009			2011		
Date	DOY	TM	Date	DOY	TM
May	134	7	12 May	132	5
May	150	7	20 May	140	7
June	158	5	28 May	148	5
June	166	7	5 June	156	7
July	190	5	13 June	164	5
July	198	7	29 June	180	5
August	214	7	7 July	188	7
August	222	5	15 July	196	5
August	230	7	31 July	212	5
August	238	5	24 August	236	7
September	254	5	1 September	244	5
September	262	7	17 September	260	5
September	270	5	25 September	268	7
October	278	7	3 October	276	5
October	286	5	11 October	284	7
October	294	7	19 October	292	5
October	302	5	27 October	300	7
Acquired	17		17		
Potential	22		22		
Success rate	77%		77%		

Output LE fluxes were then converted to daily ET estimates by averaging and integrating in the same way as for TSEB results. 7 airborne flight data sets were processed in these ways. To model Landsat observations, 17 each from 2009 and 2011, a different data assembly procedure was needed due to the relatively coarse resolution. For these instances, 3-pixel maps containing NDVI, LST and plant heights for three targets: F33, a reference bare soil patch, and a reference cool, vegetated patch. TSEB runs only utilized the F33 target pixel, while METRIC runs utilized all three. Not all Landsat overpasses could be used due to cloud cover or scan line gaps from Landsat 7. To obtain validation data, measured ET values were based on soil moisture changes observed at 112 neutron probe access tube sites. Water depth equivalents were first determined for each neutron probe sample level by multiplying the fractional volumetric water content by the sampling depth intervals, then summed over the whole interval to obtain a total depletion for the time between samples. Average daily ET was then assumed constant for the mid-point of the computed time interval. To obtain ET at overpass times, the mid-point values were linearly interpolated. Using all seven soil moisture data sets for each border, the interpolated ET values were aggregated to determine mean and standard deviations of daily ET in mm/day.

3. Results

3.1. Comparing airborne modeled ET with observations

Accuracy of remotely sensed ET was assessed for experimental data collected on 3 dates in 2009 — 27 May, 3 June, 30 July, days of year (DOY), 147, 154, 211 — and for data collected on 4 dates in 2011 — 7 July, 21 July, 18 August, and 8 September, DOY 188, 202, 230, 251. Fig. 2 shows average daily ET values for all 16 borders plotted against soil moisture based ET.

Remote sensing ET estimates agreed with SM ET to ~2 mm/day. Generally good agreement is verified in summary results listed in Table 5, where ET estimates are averaged over all borders. SM estimates ranged from 4.25 to 10.8 mm/day based on weekly depletions. Remote sensing ET estimates had a similar range, 5.77 to 10.75 mm/day, based on in mid-day instantaneous observations. Model-SM differences were less than 2.07 mm/day. For 2–4 instances for TSEB and METRIC difference were less than 1.0 mm/day.

While ET estimates from TSEB and METRIC were similar to each other in both years, the ranges of estimates over the cotton canopy in the different treatments showed even greater similarity, i.e. relative model sensitivities were correlated (Fig. 3). Correlation between TSEB and METRIC was moderate to strong, with R^2 ranging 0.6–0.9 for the seven airborne survey days. Significantly, the variability had a seasonal dependence. For early season data with sparse cover, variability of METRIC-derived ET was greater than TSEB-derived ET, while for mid-late season full cover conditions, variability of METRIC was about equal to or slightly less than ET variability estimated with TSEB. As shown in Fig. 3, the linear fit to each survey data set had steep slopes for DOY 147 and 202 and shallow slopes for DOY 230 and 251 in 2011 (too little variability existed for DOY 211 in 2009 to generate a trend line for that data set). Combining linear models from both years into one display (Fig. 4) shows a seasonal trend, where the range of METRIC-based ET estimates over vegetation was as much as 4x that estimated by TSEB. Whiskers at each sample represents the standard error for the ET model ratio estimate. In the later growing season, the range of METRIC ET became less than TSEB ET. Why this systematic relationship exists appears to be due to the relative significance of LST differences considered in each model (Fig. 5). METRIC relies on LST values at cold and hot reference pixels to create an apparent LST/air temperature gradient by scaling. When the ranges of vegetation LST values seen in all 16 borders were compared with the range spanned by the cold and hot references, a similar pattern observed in Fig. 4 is observed in Fig. 5A. Thus in early growth stages, where LST values over canopy contain a soil temperature component, the amount of LST variation constitutes a major

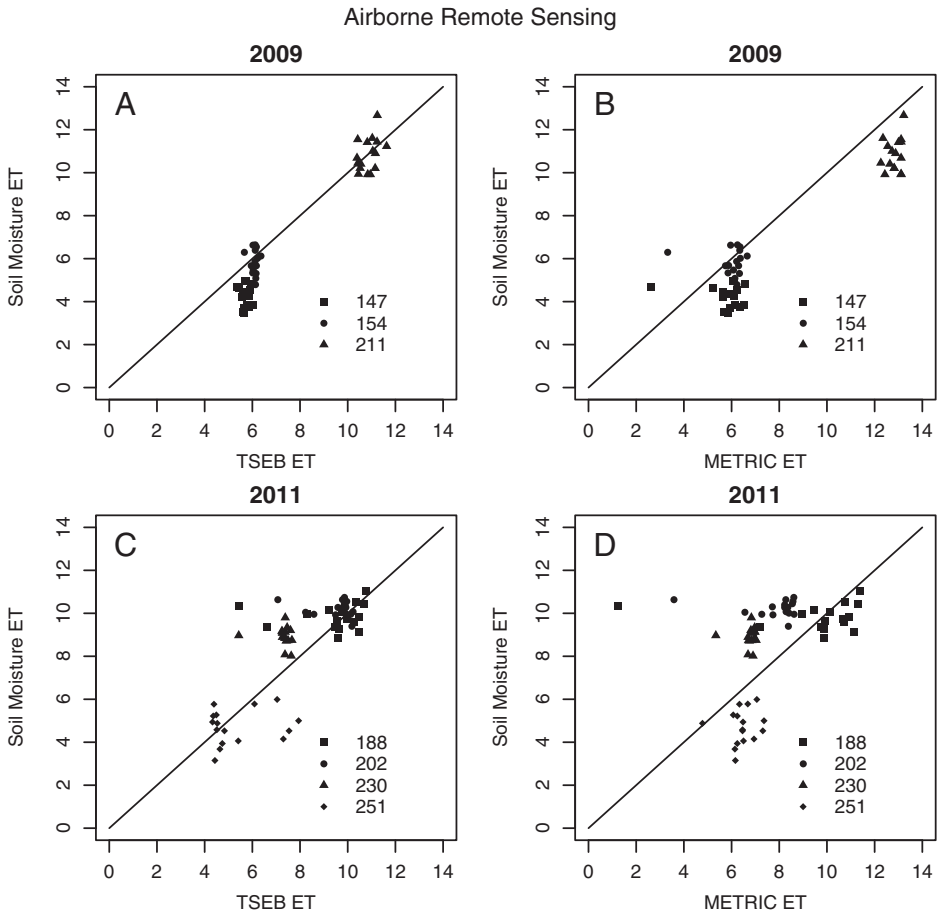

Fig. 2. Modeled and reference daily ET derived from airborne remote sensing in 2009 and 2011. Border-averaged ET estimated with TSEB (left column) and METRIC (right column) are compared with soil moisture derived ET.

portion (~0.6) of total LST variability in the scene. In contrast, TSEB relies entirely upon LST/air temperature gradients and uses neither reference end members nor scaling. This means that any ET variability is represented—assuming constant air temperature over the site—by corresponding unscaled LST values (Fig. 5B). In this experiment, LST variability was ~4.5–7 °C, with higher ranges corresponding to greater ET values as shown in Fig. 2. At full-cover both models were found to be equally sensitive to changes in LST observations, the timing for which is indicated by the intersection of the horizontal line with ratio trend.

Results at border scales (i.e. 12 m × 168 m) supported field-wide observations. To illustrate, ET results from FAO-treatment border 9 are shown for early, mid and late season growth stages in 2009 and 2011 (Fig. 6). Selection of border 9 has no special significance—other borders showed similar results. The dashed line indicates mean SM daily ET, the bracketing solid lines the standard deviation of SM ET. Viewing trends for the full growing season, ET values meet expectations as previously reported by Hunsaker, Barnes, Clarke, Fitzgerald, and Pinter (2005) and Erie, French, Bucks, and Harris (1982). At early growth stages cotton ET was 4 mm/day with standard deviations less than 1 mm/day. Daily ET rapidly increased with plant growth, with peak values reached by early July of 10 mm/day with standard deviations up to 2 mm/day. The large variability of ET was consistent with the surface irrigation method, where large areas of open water were exposed for several hours during floods. In the late season, daily ET rapidly declined to 4–5 mm/day due to termination of irrigations prior to desiccant applications.

Table 5
Airborne-based, field-averaged, daily ET (mm/day).

Year	DOY	TSEB	METRIC	Integrated	Soil moisture
2009	147	5.96/1.71	6.21/1.96	–	4.25
2009	154	6.21/0.36	6.32/0.47	–	5.85
2009	211	10.75/−0.10	12.83/1.98	–	10.85
2011	188	9.92/0.08	10.4/0.57	9.27/−0.56	9.83
2011	202	9.90/−0.30	8.23/−1.97	8.68/−1.52	10.20
2011	230	7.52/−1.39	6.86/−2.05	7.85/−1.06	8.91
2011	251	5.77/1.05	6.63/1.91	4.97/0.25	4.72

While agreement to better than 1.0 mm/d is desired, that outcome is unrealistic given experimental constraints. SM data were extensive in time and profile depths, but spatially limited. Remote sensing data, on the other hand, were a really comprehensive but time-limited. An example of these resolution differences is shown in Fig. 7, where a South–North transect was taken to compare SM-based ET against LST data. Because NDVI data did not show a significant South–North trend (mean: 0.56, standard deviation: 0.03, corresponding to cover variation ~49–55%), an SM ET trend would be expected to be reflected by a corresponding LST trend. Such a trend did exist at large scale. For the 7 SM access points, ET decreased from 12 to 8 mm/d (Fig. 7A), matching expectations due to irrigation practice. Irrigation gates lay to the south, meaning that water preferentially infiltrated at the source where there was greater opportunity time. This ET trend was matched in a general sense with border averaged LSTs (Fig. 7B). Higher temperatures (38 °C) corresponded approximately to SM-ET of 8 mm/day and lower temperatures (34 °C) to 12 mm/day. Hence a 4 °C change accompanied an ET change of 4 mm/day for a ratio of 1 °C:1 mm/day.

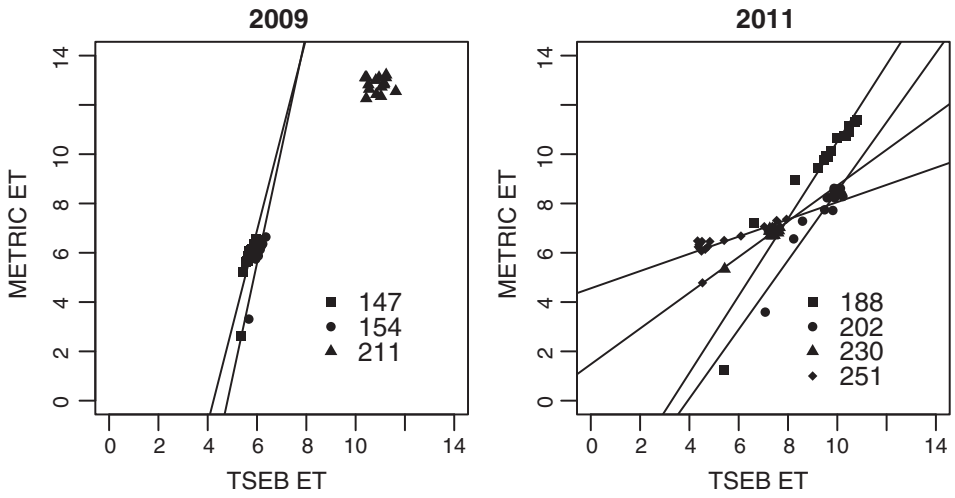

Fig. 3. Daily ET from TSEB and METRIC compared for 2009 and 2011. Each symbol represents one of 7 particular days. The lines are linear models representing the relative model sensitivities of ET over cotton canopy. No linear fit was created for DOY 211 due to lack of linear trend.

However, at higher spatial frequencies the correspondence was less consistent. A 3 °C temperature spike (3,659,900.0 m North UTM) correlates with an ET drop of ~1.5 mm/day, for a ratio of 2 °C:1 mm/day. The difference in this ratio from the ratio of the overall trend illustrates that remotely sensed variability of LST at the fine spatial scales may not represent weekly SM ET variations (Fig. 8). Mapping of instantaneous airborne LST data at 1 m resolution showed strong coherent turbulent patterns, related to southeast wind at 2 m/s, had variations >8 °C. The patterns were independent of the experimental design and thus likely provide no corresponding fine spatial scale daily ET information.

Results from 2011 airborne data were used to assess an alternate conversion of instantaneous fluxes to daily ET. If a way exists to estimate the spatial distribution of LST at all times of day based on time-continuous observations at a few ground locations, then it could be feasible to obtain improved daily ET estimates when skies are not uniformly clear. LST observations collected at 7–15 min intervals from 12 wireless nodes, in combination with Eq. (3) and the TSEB model, were used to estimate ET at all hours on four separate days. This time synthesis approach did improve ET results on two days (DOY 230 and 251), one of which corresponded to partly cloudy skies following the airborne survey (Table 5). On DOY 251 (8 September 2011), total solar radiation was 23.2 MJ/m^2, approximately 90% of expected for a clear sky day. The synthesized LST-based TSEB ET was 86% of a clear sky model, consistent with the reduced incoming shortwave radiation for that day. In this case the sample size was too small to be conclusive. Nevertheless the improvement suggests additional evaluation would be worthwhile.

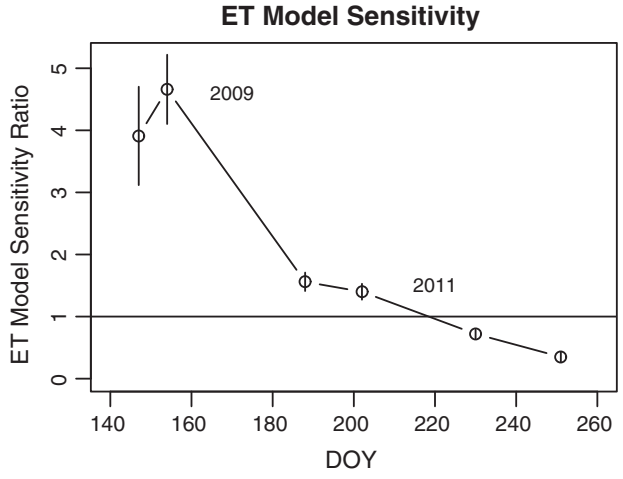

Fig. 4. Daily ET model sensitivity over cotton. Slopes derived from fits shown in Fig. 3 are plotted vs. time to show the decreasing relative sensitivity of METRIC estimates to TSEB estimates. At DOY 220, model sensitivities are approximately equal.

3.2. Comparing Landsat modeled ET with observations

Following procedures similar to those used for airborne data, ET estimation accuracy was assessed at the ~200 m field-wide scale for the full cotton growth cycles in 2009 and 2011 using Landsat 5 and Landsat 7 data. Although the satellite data were too coarse to resolve treatment differences in the cotton experiment, they provided full-season observations not provided by the airborne data and thus were able to assess model performance at all growth stages.

Model input data highlighted the differences in cotton plant growth in 2009 and 2011 (Fig. 9A) where peak NDVI were 0.802 and 0.742 corresponded with mid-August plant heights 1.40 and 1.08 m respectively. Seasonal LST patterns (Fig. 9B) showed a distinguishing feature of irrigated summer crops, where LST values peak early, then drop throughout the growing season, and then once again during senescence and fall weather. In contrast, non-irrigated lands follow the pattern shown the gray lines where LST values track air temperature patterns. LST-air temperature differences (ΔT, Fig. 9C), represented complementary patterns for daily ET. In this case, ΔT tracked increasing ET in early and mid-growing seasons and decreasing ET during senescence.

ET model outputs are shown in Fig. 10, where the left column plots Landsat-based daily ET results for 2009 and 2001. The right column shows cumulative ET representing seasonal crop water used. Each plot contains three ET data sets: SM observations (open circles) interpolated to Landsat overpass dates for the daily displays, TSEB estimates (triangles) and METRIC estimates (diamonds). Irrigation events, scaled to 1/50th of applied depth in mm, are plotted at the bottom of the ET daily displays. By comparing these events with daily ET one can seek correlations. In 2009 TSEB and METRIC ET values were nearly the same on a seasonal basis with ±0.2 mm/day bias (Table 6), but close agreement was consistent only for early season overpass times. In 2011 TSEB ET values were more biased than METRIC ET (−1.6 vs. −0.7 mm/day) though in this year positive was more dominant than in 2009. The result of these differences was that cumulative water use estimates in 2009 were similar for both models, differing by 34 mm, whereas in 2011 TSEB was 93 mm less than METRIC.

An alternative way of assessing remotely sensed ET estimates was to consider changes in evaporative fraction instead of changes in total ET.

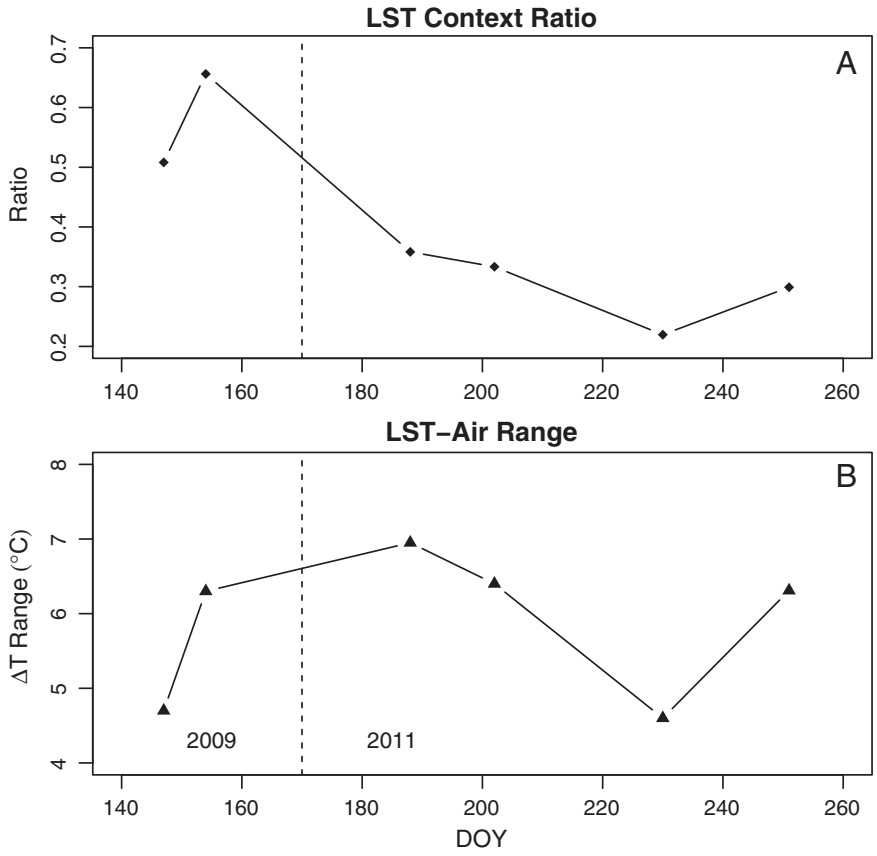

Fig. 5. Factors controlling ET model sensitivity. METRIC ET values are determined by scaling LST observations with respect to cold and hot pixels; its sensitivity of ET estimates to LST changes is controlled by the relative range of vegetation LSTs to the range spanned by cold to hot reference pixels. Shown are ratios between these ranges for the different airborne flight days (A). TSEB ET values, on the other hand, are controlled by LST/air temperature differences (ΔT) and do not use reference pixels. Thus where for sites with a nearly constant air temperature, TSEB model sensitivity is controlled by absolute vegetation LST (B). The dashed line separates 2009 and 2011 data.

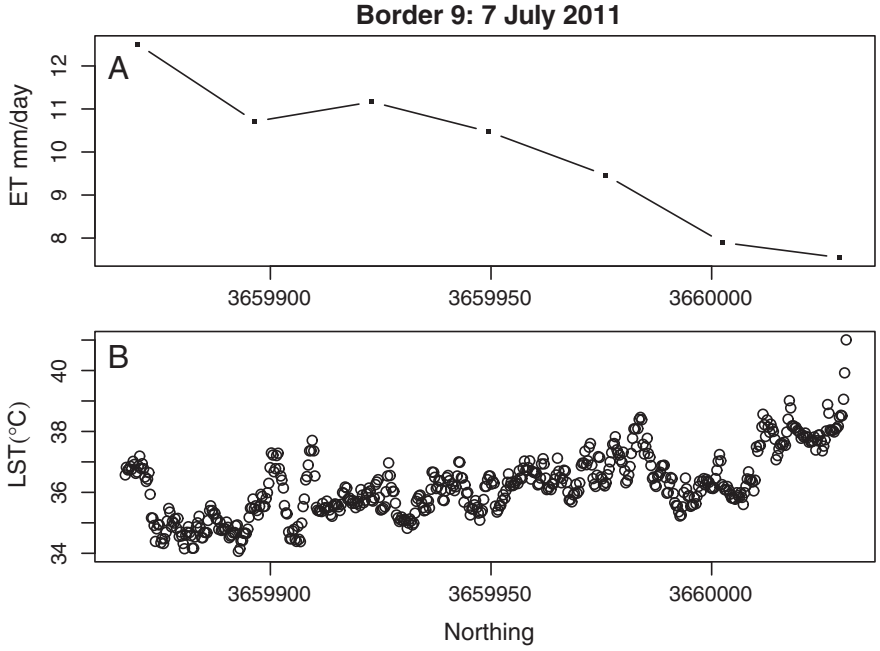

Fig. 6. Modeled and observed ET over the 2009/2011 cotton experiment at Maricopa, AZ for the irrigated border 9 (FAO treatment). The dashed line denotes soil moisture based ET, and the solid lines the corresponding standard deviation for those observations. Triangles indicate TSEB mean daily ET estimates obtained from using the constant evaporative fraction method, diamonds METRIC estimates, and whiskers standard deviations for values computed within border 9.

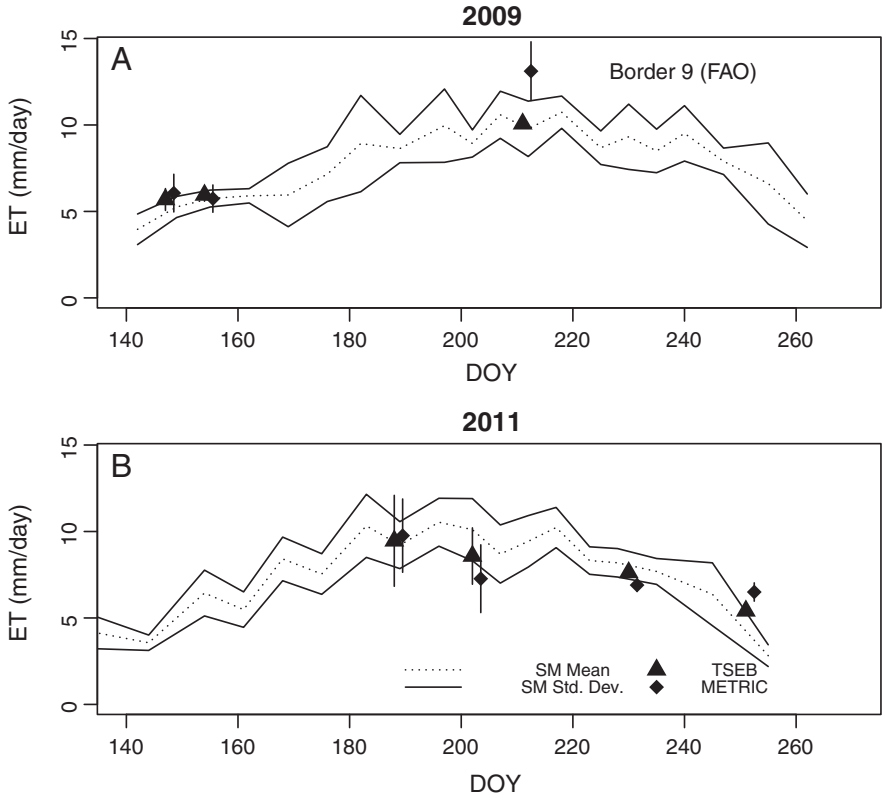

Fig. 7. Profile of ET along border 9 on 7 July 2011 obtained from soil moisture data (A), and the corresponding profile for LST data obtained from the airborne TIR instrument (B).

In equilibrium and well-watered conditions EF would be close to 1.0, whereas for episodes with significant root-zone depletion EF will decrease below 1.0. What value of EF corresponds to crop water stress or a signal to irrigate would need to be determined experimentally. As suggested by Zipper and Loheide (2014), a measure of relative ET could be an effective water-stress indicator. Fig. 11 illustrates results from 2011 (results from 2009 are similar). Mean mid-field water content was mapped over time for 15 depths, two of which are displayed in Fig. 11A: the near-surface (10 cm) and main root zone (90 cm). Irrigation events, on the order of 100 mm, are indicated at the bottom of Fig. 11A. Rainfall was mainly unimportant — the only significant rainfall was 30 mm on DOY 190. By comparing irrigation events at multiple depths on can monitor rooting depth: if water content decreases more rapidly than observed for drainage, then it can be concluded that those depths lie within the active root zone. Irrigation inputs have immediate impact on shallow soil moisture for the whole season, where the water content promptly increase on the order of 10%. In contrast, moisture fluctuations at 90 cm only become significant

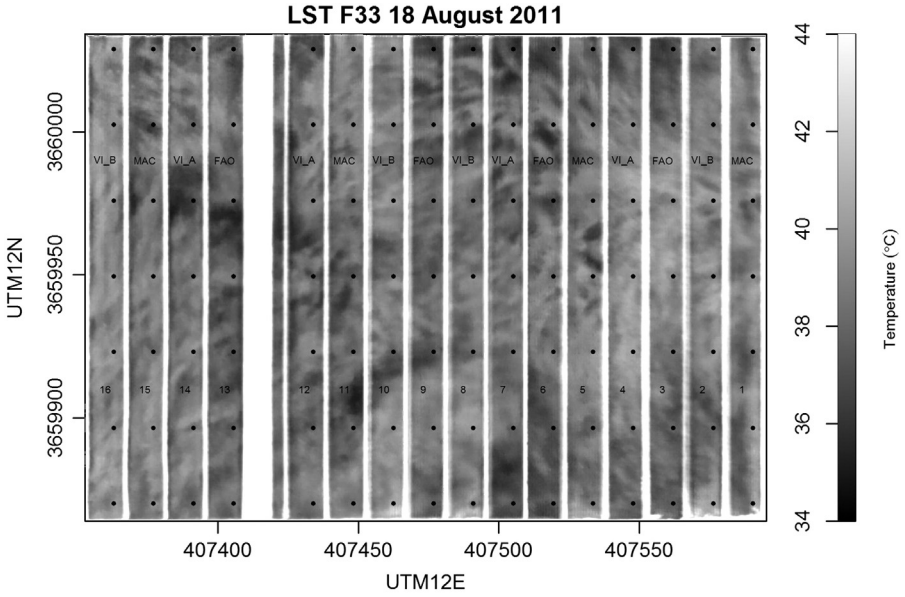

Fig. 8. LST patterns observed on 21 July 2011 near local noon across all irrigated borders. The swirl patterns represent LST variations on the order of 5 °C induced by changing wind speeds across the field.

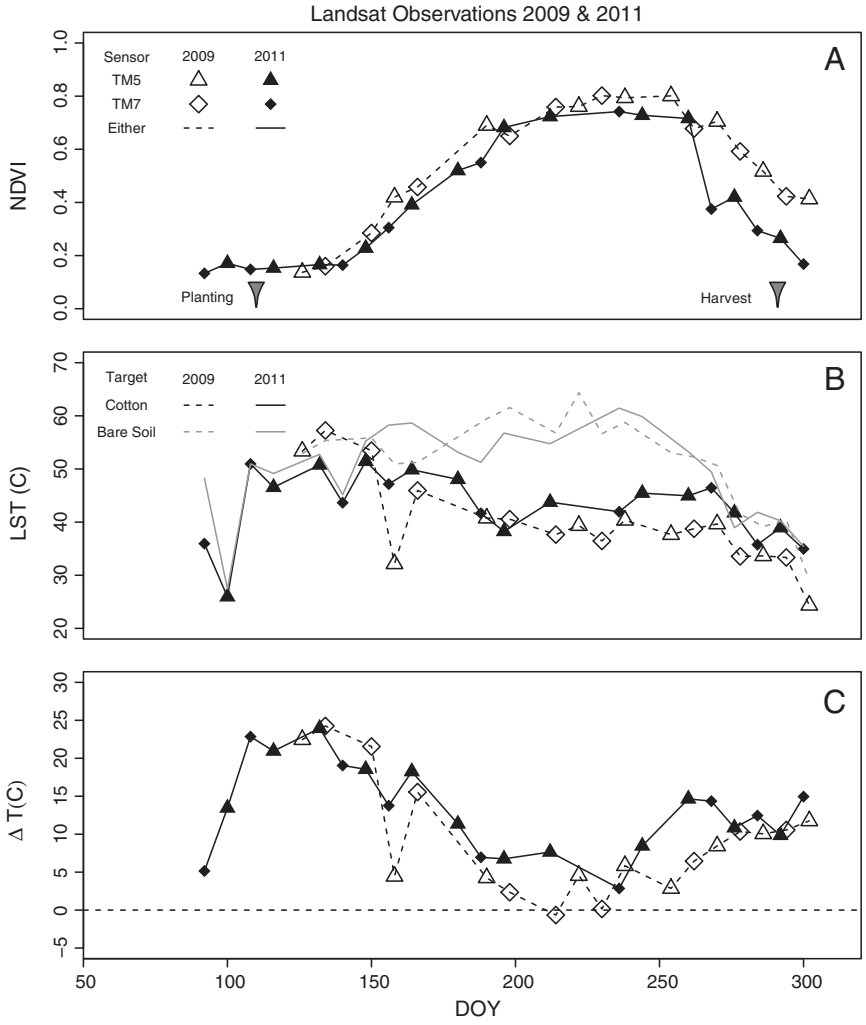

Fig. 9. Landsat-derived surface observations of NDVI, LST and ΔT for the 2009 and 2011 cotton experiments. The 2009 crop was denser and matured earlier than the 2011 crop.

mid-season (>DOY 170), reflecting rooting depth development time. Corresponding changes in remotely sensed EF are shown in Fig. 11B. Using the combined Landsat 5 and 7 data collection for 2011 showed two early season EF drops on DOY 150 and 175 that appear to be correlated with soil moisture depletions at both 10 and 90 cm depths. Note however that from mid-season (~DOY 200) onward, no soil moisture depletion events appeared to be detected by Landsat-based EF estimates. Because the cotton crop was irrigated to avoid severe stress — with a few exceptions irrigations were applied to the four treatments when soil moisture depletion exceeded ~45% — the negative result for mid-season data could indicate that the EF approach is sensitive to observation timing.

Having evaluated Landsat-based ET estimates from TSEB and METRIC using 17 overpasses in each year, it was also possible to evaluate the impact of remote sensing image frequency. Clearly for irrigation applications weekly observations are higher spatial resolution are important, else timely updates would not be available to match on-farm scheduling practice. Furthermore, frequent overpasses are needed to reduce very long time gaps due to cloud cover (Gao, Masek, Schwaller, & Hall, 2006). However, if the objective is seasonal ET estimation, lower overpass frequencies could be acceptable. In this study the impact of overpass frequency on cotton ET was assessed by comparing model results using Landsat 5 or 7 or both (Table 7). Regardless of ET model, TSEB or METRIC, results showed similar impact in both 2009 and 2011. Seasonal ET based on nominal 16-day overpass samples was commonly 100–200 mm different from estimates generated when both satellite data sets were available. While this result does not mean that combing Landsat 5 and 7 data returned best agreement with SM data, it does show that ET forecasts can be strongly affected by overpass frequencies. In this study the effect ranged up to 20% of SM observations, a substantial deviation.

4. Discussion

Outcomes from a multiscale model study compared TSEB and METRIC and reference SM ET observations. The goal was to determine ET estimation accuracy over an irrigated agricultural crop and to evaluate their relative performance.

In contrast to other model comparison studies (e.g. Consoli & Vanella, 2014), this study evaluated, estimated, and validated, ET at daily and seasonal time scales using locally constrained soil moisture depletion estimates instead of surface flux observations with temporally and spatially varying flux footprints. The comparisons focused on core estimates of ET model differences directly related to remotely sensed LST; other differences were removed by standardizing *Rn* and *G* estimates. The study did not incorporate model specific optimizations, a practice that would have had beneficial results for both approaches. For example, the Priestley–Taylor parameter α used in TSEB was set to a nominal value, 1.26, instead of some greater value specifically applicable to irrigated crops. For METRIC, subjective pixel selection was replaced with an objective one, despite model guidelines (Allen et al., 2007) that judgment is needed on a case-by-case basis.

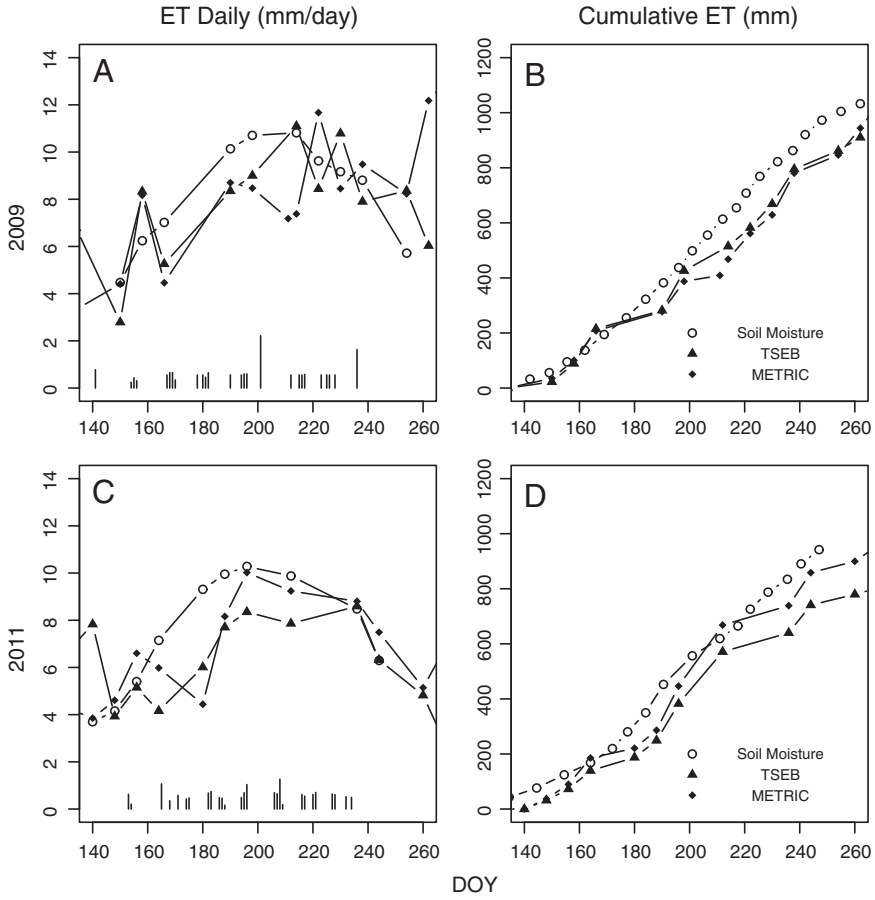

Fig. 10. Landsat-derived daily and cumulative ET for 2009 and 2011 using TSEB and METRIC.

When TSEB and METRIC were compared at fine and moderate spatial scales, on average no large differences in daily ET estimation accuracy were observed. Agreement with SM data was 1.9 mm/day or better. Considering that details of water flow within the soil profile were not known, much better agreement, e.g. to less than 1.0 mm/day, may not be feasible. When daily ET was seasonally integrated, some differences did appear. This was most notable in 2011, where METRIC showed lower bias in cumulative ET than TSEB. Differences up to 20% were found, an outcome possibly due to better early season estimates provided by METRIC.

Comparing TSEB and METRIC at 1 m resolution using airborne data showed their model sensitivities were strongly correlated and that the correlation had a seasonal dependency. TSEB and METRIC results showed strong linear positive covariation (Fig. 3), meaning that while the ET themselves were not necessarily closely correlated, their relative model sensitivities were. For sparse cover, ET variability derived from METRIC was much greater than observed with TSEB. For full cover, this ET variability diminished to values slightly less than TSEB ET. Since the same LST and NDVI data were used for both, this covariation appears to be a direct representation of the relationship between METRIC's contextual scaling to obtain an apparent LST/air temperature gradient vs. TSEB's explicit link with air temperature as demonstrated in Fig. 5. Should other studies show replication of this phenomenon, this correlated model sensitivity finding could be used for seasonal-dependent model selection, where sensitivity covariance could be used to weight relative importance of ET estimates while running both models in parallel. Such an approach could improve ET estimation accuracy, for example, by assigning higher confidence levels to ET estimates that show verifiable accurate higher sensitivities.

Table 6
Seasonal ET, 15 May to 22 September 2009 and 2011 using Landsat and soil moisture data.

Year	TSEB			METRIC			Soil moisture
	RMSE	Bias	Cumulative ET	RMSE	Bias	Cumulative	
2009	1.6	−0.2	910	1.9	0.2	944	1033
2011	0.5	−1.6	881	0.5	−0.7	972	918

Availability of results from two models allowed multiple tests of an EF-based stress indicator. Results showed some signaling of water shortfall in the root zone in two early season events shortly before and after planting, however later season indicators did not appear. Circumstances that could explain this negative finding are inadequate spatial resolution needed to discriminate irrigation borders, management practice that maintained maximum soil water depletion to <45%, and sub-optimal timing of remote sensing acquisitions.

Results from estimating daily ET both remote sensing and ground-based LST observations showed in two instances that time synthesis or hourly ET fluxes could improve upon results from the constant EF approach. However, degree of improvement and justification for the additional field effort is needed.

Lastly, model estimates enabled testing of the impact of overpass frequency. Weekly overpasses are close to essential for irrigation scheduling and decision support systems. They are highly preferred in other water applications such as water district accounting, but a question to answer is if improvements in ET are substantial enough to warrant the extra costs of additional satellite platforms. This study showed that for two years of irrigation that there were significant differences. Modeling estimates at daily and seasonal time scales depended upon the frequency of data availability. Landsat-based ET estimates, using either TSEB or METRIC, showed similar results. Differences up to 20% of seasonal water used, 100–200 mm, occurred when nominal 16-day overpass frequencies were used compared with 8-day frequencies.

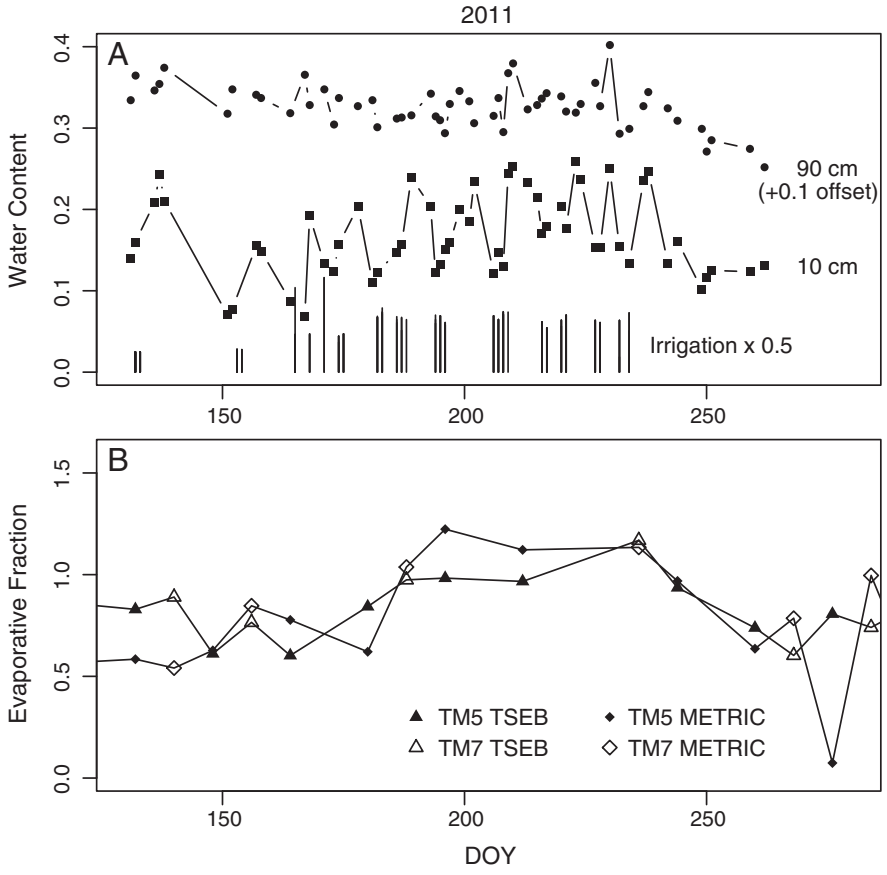

Fig. 11. Testing remote sensing detection of crop water stress. Soil moisture content at two depths (A) and EF derived from TSEB and METRIC for 2011 observations (B) are plotted for the entire cotton growing season.

This difference is large enough to have an impact upon irrigation decision making at farm and irrigation district scales.

5. Conclusions

A two-year remote sensing ET study was undertaken at Maricopa, Arizona to evaluate approaches to monitoring ET at daily to seasonal time scales. The study evaluated two prominent ET models, TSEB and METRIC. The structure of the comparison was constructed to remove confounding model components not germane to remote sensing of LST. Both models were found similarly accurate to 1.9 mm/day. TSEB and METRIC performed well at 1-m and 30–120 m scales and resolved episodic and seasonal ET changes. Model sensitivity was found to be correlated in a seasonally systematic way where METRIC had greater sensitivity to LST changes over less dense cotton cover while TSEB had greater sensitivity during mature phases. An alternative time scaling method to the commonly used constant EF approach was tested and worked well for two instances, but more evidence is needed under partly cloudy conditions to verify its consistency. Tests were conducted with matching temporal variations of EF to soil moisture depletions to evaluate its utility as a simple-to-implement crop water stress detection approach. Results indicated some signaling in the early growth stage. However, EF variations for most of the seasons did not track water depletions observed in the root zone. A test on the impact of overpass frequency upon ET accuracy was conducted and showed significant benefit to an 8-day overpass frequency compared with 16-day observation intervals. When considering model selection, model complexity and availability of ancillary data need to be considered. In this cotton study neither factors significantly hindered implementation, but where surface conditions are not well known the METRIC approach is recommended. When surface conditions are well constrained, the TSEB approach is recommended for its ability to adapt to known biophysical conditions.

Table 7
Impact of overpass frequency upon seasonal ET. Values shown are cumulative seasonal ET (mm) for field F33 as a whole.

	TSEB			METRIC			Soil moisture
Year	TM5	TM7	Both	TM5	TM7	Both	
2009	929	1073	910	993	1127	944	1033
2011	789	1072	881	899	1087	972	918

References

Allen, R.G., Pereira, L.S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration, guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56 Food and Agriculture Organization of the United Nations.*

Allen, R.G., Tasumi, M., & Trezza, R. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-model. *Journal of Irrigation and Drainage Engineering, 133*, 380–394.

The ASCE standardized reference evapotranspiration equation. Allen, R., Walter, I., Elliott, R., Howell, T., Itenfisu, D., & Jensen, M., et al. (Eds.). (2005). American Society of Civil Engineers (Prepared by the Task Committed on Standardization of Reference Evapotranspiration of the Environmental and Water Resources Institute).

Anderson, M., Hain, C., Otkin, J., Zhan, X., Mo, K., Svoboda, M., et al. (2013). An intercomparison of drought indicators based on thermal remote sensing and NLDAS-2 simulations with U.S. drought monitor classifications. *Journal of Hydrometeorology, 14*, 1035–1056.

Barsi, J.A., Schott, J.R., Palluconi, F.D., & Hook, S.J. (2005). Validation of a web-based atmospheric correction tool for single thermal band instruments. In J.J. Butler (Ed.), *Earth Observing Systems X. Bellingham, WA: SPIE volume 5882 of Proceedings of the SPIE.*

Bastiaanssen, W., Menenti, M., Feddes, R., & Holtslag, A. (1998). A remote sensing surface energy balance algorithm for land (SEBAL) 1. Formulation. *Journal of Hydrology, 212–213*, 198–212.

Berk, A., Bernstein, L., Anderson, G., Acharya, P., Robertson, D., Chetwynd, J., et al. (1998). MODTRAN cloud and multiple scattering upgrade with application to AVIRIS. *Remote Sensing of Environment*, *65*, 367–375.

Blonquist, J., Norman, J., & Bugbee, B. (2009). Automated measurement of canopy stomatal conductance based on infrared temperature. *Agricultural and Forest Meteorology*, *149*, 2183–2197.

Brown, P. (1989). Accessing the Arizona Meteorological Network (AZMET) by computer. *Technical Report Ext. Rep. No. 8733*. University of Arizona.

Brutsaert, W. (1982). *Evaporation into the atmosphere: Theory, history and applications*. Dordrecht, Holland: D. Reidel Pub. Co.

Byun, K., Liaqat, U., & Choi, M. (2014). Dual-model approaches for evapotranspiration analyses over homo- and heterogeneous land surface conditions. *Agricultural and Forest Meteorology*, *197*, 169–187.

Campbell, G.S., & Norman, J.M. (1998). *An introduction to environmental biophysics* (2nd ed.). Springer-Verlag.

Carlson, T., Capehart, W., & Gillies, R. (1995). A new look at the simplified method for remote sensing of daily evapotranspiration. *Agricultural and Forest Meteorology*, *54*, 161–167.

Chirouze, J., Boulet, G., Jarlan, L., Fieuzal, R., Rodriguez, J., Ezzahar, J., et al. (2014). Intercomparison of four remote-sensing-based energy balance methods to retrieve surface evapotranspiration and water stress of irrigated fields in semi-arid climate. *Hydrology and Earth System Sciences*, *18*, 1165–1188.

Choi, M., Kustas, W., Anderson, M., Allen, R., Li, F., & Kjaersgaard, J. (2009). An intercomparison of three remote sensing-based surface energy balance algorithms over a corn and soybean production region (Iowa, U.S.) during SMACEX. *Agricultural and Forest Meteorology*, *149*, 2082–2097.

Colaizzi, P., Agam, N., Tolk, J., Evett, S., Howell, T., Gowda, P., et al. (2014). Two-source energy balance model to calculate E, T, and ET: Comparison of Priestley–Taylor and Penman–Monteith formulations and two time scaling methods. *Transactions of the ASABE*, *57*, 479–498.

Colaizzi, P., Evett, S., Howell, T., Gowda, P., OShaughnessy, S., Tolk, J., et al. (2012). Two-source energy balance model-refinements and lysimeter tests in the Southern High Plains. *Transactions of the ASABE*, *55*, 551–562.

Consoli, S., & Vanella, D. (2014). Comparisons of satellite-based models for estimating evapotranspiration fluxes. *Journal of Hydrology*, *513*, 475–489.

Consult, W. (2011). Operational monitoring product for planning and water allocation in the international Incomati Basin (WATPLAN). *Technical Report Stakeholder analysis report WE Consult Maputo, Mozambique*.

Crago, R.D. (1996). Conservation and variability of the evaporative fraction during the daytime. *Journal of Hydrology*, *180*, 173–194.

Erie, L., French, O., Bucks, D., & Harris, K. (1982). Consumptive use of water by major crops in the Southwestern United States, Conservation Research Report 29. *Technical Report USDA, Agricultural Research Service*.

Evett, S., & Steiner, J. (1995). Precision of neutron scattering and capacitance type water content gauges from field calibration. *Soil Science Society of America Journal*, *59*, 961–968.

French, A., Hunsker, D., Clarke, T., Fitzgerald, G., & Pinter, P.J., Jr. (2010). Combining remotely sensed data and ground-based radiometers to estimate crop cover and surface temperatures at daily time steps. *Journal of the Irrigation and Drainage Division*, *136*, 232–239.

Friedl, M. (1996). Relationships among remotely sensed data, surface energy balance, and area-averaged fluxes over partially vegetated land surfaces. *Journal of Applied Meteorology*, *35*, 2091–2103.

Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 2207–2218.

Gentine, P., Entekhabi, D., Chehbouni, A., Boulet, G., & Duchemin, B. (2007). Analysis of evaporative fraction diurnal behaviour. *Agricultural and Forest Meteorology*, *143*, 13–29.

Gibson, L., Münch, Z., Engelbrecht, J., & Petersen, N. (2009). Remote sensing as a tool towards resource assessment and determination of the legal compliance of surface and groundwater use. *Technical Report WRC Report No. 1690/1/09 Water Research Commission Pretoria, South Africa*.

Glenn, E., Neale, C., Hunsaker, D.J., & Nagler, P. (2011). Vegetation index-based crop coefficients to estimate evapotranspiration by remote sensing in agricultural and natural ecosystems. *Hydrological Processes*, *25*, 4050–4062.

Gonzalez-Dugo, M., Neale, C., Mateos, L., Kustas, W., Prueger, J., Anderson, M., et al. (2009). A comparison of operational remote sensing-based models for estimating crop evapotranspiration. *Agricultural and Forest Meteorology*, *149*, 1843–1853.

Hunsaker, D.J., Barnes, E.M., Clarke, T.R., Fitzgerald, G.J., & Pinter, Paul J., Jr. (2005). Cotton irrigation scheduling using remotely-sensed and FAO-56 basal crop coefficients. *Transactions of ASAE*, *22*, 1395–1407.

Hunsaker, D.J., Fitzgerald, G.J., French, A.N., Clarke, T.R., & Pinter, Paul J., Jr. (2007). Wheat irrigation management utilizing multispectral crop coefficients: 1. Crop evapotranspiration prediction. *Transactions of the ASABE*, *50*, 2017–2033.

Hunsaker, D., French, A., Waller, P., Bautista, E., Thorp, K., & Andrade-Sanchez, P. (2014w). Irrigation scheduling using spatial information in real-time: An evaluation with cotton grown under surface irrigation. *Agricultural Water Management*, *1–44* (in review).

Jackson, R., Reginato, R., & Pinter, P.J., Jr. (1981). *Canopy temperature as a crop water stress indicator*.

Jensen, M., Burman, R., & Allen, R. (1990). Evapotranspiration and irrigation water requirements. *ASCE Manuals and Reports on Engineering Practice 70 American Society of Civil Engineers*.

Jiang, L., & Islam, S. (1999). A methodology for estimation of surface evapotranspiration over large areas using remote sensing observations. *Geophysical Research Letters*, *26*, 2773–2776.

Kongo, V., & Jewitt, G. (2006). Preliminary investigation of catchment hydrology in response to agricultural water use innovations: A case study of the Potshini catchment- South Africa. *Physics and Chemistry of the Earth*, *31*, 976–987.

Lhomme, J. -P., & Elguero, E. (1999). Examination of evaporative fraction diurnal behaviour using a soil-vegetation model coupled with a mixed-layer model. *Hydrology and Earth System Sciences*, *3*, 259–270.

Long, D., & Singh, V.P. (2013). Assessing the impact of end-member selection on the accuracy of satellite-based spatial variability models for actual evapotranspiration estimation. *Water Resources Research*, *49*, 2601–2618.

Meyers, T., & Hollinger, S. (2004). An assessment of storage terms in the surface energy balance of maize and soybean. *Agricultural and Forest Meteorology*, *125*, 105–115.

Morton, C., Huntington, J., Pohll, G., Allen, R., McGwire, K., & Bassett, S. (2013). Assessing calibration uncertainty and automation for estimating evapotranspiration from agricultural areas using METRIC. *Journal of the American Water Resources Association*, *49*, 549.

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

Pereira, L.S., Allen, R.G., Smith, M., & Raes, D. (2015). Crop evapotranspiration estimation with FAO56: Past and future. *Agricultural Water Management*, *147*, 4–20.

Peters, R.T., & Evett, S.R. (2004). Modeling diurnal canopy temperature dynamics using one-time-of-day measurements and a reference temperature curve. *Agronomy Journal*, *96*, 1553–1561.

Pinter, P., Jr., Hatfield, J., Schepers, J., Barnes, E., Moran, M., Daughtry, C., et al. (2003). Remote sensing for crop management. *Photogrammetric Engineering and Remote Sensing*, *69*, 647–664.

Post, D., Mack, C., Camp, P., & Sulliman, A. (1988). Mapping and characterization of the soils on the University of Arizona agricultural center. *Proceedings Hydrology and Water Resources in Arizona and the Southwest* (pp. 49–60). Tucson, Arizona: University of Arizona.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (URL: http://www.R-project.org/).

Schmidt, G., Jenkerson, C., Masek, J., Vermote, E., & Gao, F. (2013). Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description. *Open-File Report 1057 USGS*.

Sun, J., & Mahrt, L. (1995). Determination of surface fluxes from the surface radiative temperature. *Journal of the Atmospheric Sciences*, *52*, 1096–1106.

Sun, Z., Wang, Q., Matsushita, B., Fukushima, T., Ouyang, Z., & Watanabe, M. (2009). Development of a Simple Remote Sensing EvapoTranspiration model (Sim-ReSET): Algorithm and model test. *Journal of Hydrology*, *376*, 476–485.

Tang, R., Li, Z. -L., & Sun, X. (2013). Temporal upscaling of instantaneous evapotranspiration: An intercomparison of four methods using eddy covariance measurements and MODIS data. *Remote Sensing of Environment*, *138*, 102–118.

Tasumi, M. (2003). *Progress in operational estimation of regional evapotranspiration using satellite imagery*. (Ph.D. thesis). ID: Univ. of Idaho Moscow.

Tasumi, M., Allen, R., Trezza, R., & Wright, J. (2005). *Satellite-based energy balance to assess within-population variance of crop coefficient curves131. (pp. 94–109), 94–109*.

Thorp, K., Hunsaker, D., French, A., Bautista, E., & Bronson, K. (2014w). Integrating geospatial data and cropping system simulation within a geographic information system to analyze spatial seed cotton yield, water use, and irrigation requirements. *Precision Agriculture*, 1 (in review).

Timmermans, W., Kustas, W., Anderson, M., & French, A. (2007). An intercomparison of the Surface Energy Balance Algorithm for Land (SEBAL) and the Two-Source Energy Balance (TSEB) modeling schemes. *Remote Sensing of Environment*, *108*, 369–384.

Wright, J. (1982). New evapotranspiration crop coefficients. *Journal of the Irrigation and Drainage Division*, *108*, 57–74.

Zipper, S., & Loheide, S. (2014). Using evapotranspiration to assess drought sensitivity on a subfield scale with HRMET, a high resolution surface energy balance model. *Agricultural and Forest Meteorology*, *197*, 91–102.