


Article

Remote Sensing of Evapotranspiration over the Central Arizona Irrigation & Drainage District, USA

Andrew N. French ^{1,†,*} , Douglas J. Hunsaker ¹, Lahouari Bounoua ², Arnon Karnieli ³, William Luckett ¹, and Robert Strand ⁴

¹ USDA Agricultural Research Service, U.S. ALARC, Maricopa, AZ USA

² NASA Goddard Space Flight Center, Greenbelt, MD USA

³ Ben Gurion University of the Negev, Israel

⁴ LemnaTec Corporation, Research Triangle Park, NC

* Correspondence: andrew.french@ars.usda.gov; Tel.: +01-520-316-6371

† Current address: USDA, Agricultural Research Service, U.S. ALARC, 21881 North Cardon Lane, Maricopa, AZ 85318 USA

Abstract: A remote sensing-based evapotranspiration (ET) study was conducted over the Central Arizona Irrigation and Drainage District (CAIDD), an Arizona agricultural region. ET was assessed means for 137 wheat plots, 183 cotton plots, and 225 alfalfa plots. The remote sensing ET models were the Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC), the Two Source Energy Balance (TSEB), and Vegetation Index ET for the US Southwest (VISW). Remote sensing data were principally Landsat 5, supplemented by Landsat 7, MODIS Terra, MODIS Aqua, and ASTER. The models produced similar daily ET for wheat, with 6–8 mm/d mid-season. For cotton and alfalfa daily ET showed greater differences, where TSEB produced largest daily ET, METRIC the least, and VISW in the midrange. Modeled cotton ET at mid-season ranged from 9.5 mm/d (TSEB), to 8 mm/d (VISW), and 6 mm/d (METRIC). For alfalfa ET, values at peak cover ranged from 8 mm/d (TSEB), 6 mm/d (VISW), and 5 mm/d (METRIC). Model bias ranged -10% to +18%. Relative to potential ET, FAO-56 ET, and USDA-SW gravimetric-ET, model variability ranged from negligible to 35% of annual crop water use. Model averaging was found a useful way to consider and reconcile all ET estimates.

Keywords: evapotranspiration; remote sensing; TSEB; METRIC; Landsat; Arizona; wheat; cotton; alfalfa

0. Introduction

As one of the world's largest users of fresh water resources, irrigated agriculture has a profound effect on the water cycle, water availability, distribution of water, and water quality. Irrigated croplands are essential contributors— they comprise 15% of arable land yet provide 36% of food (FAOSTAT Statistical Database, <http://apps.fao.org>). However they are also large consumers of fresh water resources, where 42% of fresh water withdrawals are used for irrigation purposes [1]; amounts that are ~65–69% when a footprint analysis is considered [2] or when water withdrawals for thermo-electric power generation are excluded [3,4]. Despite high yields and relatively low risk (relative to rainfed lands), the large amounts of water used for irrigation often conflict with competing demands by urban populations [5]. In arid lands these conflicts are exacerbated by already strained resources, climate change, and population growth. In Arizona, the reality of limits to Colorado Rivers resources have led to the development of a 'Lower Basin Drought Contingency Plan' (<https://new.azwater.gov/lbdcp>), an initiative that potentially has major implications (i.e. reductions in

supply) for agricultural water. Addressing these problems require assessments of irrigation practices to ensure that farms are currently using water efficiently and the development of tools to facilitated additional water saving strategies at farm to irrigation district scales. These assessments and tools can be realized by using remote sensing data, which are being increasingly used to monitor, manage, and forecast agriculture water use at fine spatial scales, examples include efforts such as OpenET to make remote tools widely available (<https://etdata.org>), operational ET mapping in California's Central Valley [6], Climate Engine, an on-line modeling resource [7], and agricultural monitoring by the European Commission's Joint Research Centre (<https://ec.europa.eu/jrc/en/research-topic/agricultural-monitoring>).

Despite the strong interest in remote sensing for ET, the ideal choices for which ET model, and under what circumstances the ought to be use, are not resolved. Choices are driven by the kind and quality of remote sensing and ground-based data sets, by the monitoring and frequency requirements, timeliness of results, and by the economic value of the ET maps. In recent years model inter-comparison studies have been conducted and reported ([8–10]), each showing reasons for adopting one particular approach or another. However, trends in other disciplines such as climate assessment are highlighting that for forecast accuracy it is equally important, or possibly more important, to simulate processes with multiple models. Ensemble modeling quantifies uncertainties and leads to better estimates, either because model averaging tends to improve estimates, or because employing multiple models increases the likelihood of adding new information to the analysis [11–13].

This study was undertaken to evaluate ET with the aim of assessing modeling uncertainties when using differing but established remote sensing paradigms. The research was initiated as part of a joint US-Israeli effort to improve inverse biophysical modeling of ET, for example with approaches such as the Simple Biosphere Model (SiB2, [14]), which can provide a way to use remote sensing to diagnose the minimum crop water requirement. Ideally, many models would be included, however for practical reasons—available resources, time, and modeling expertise—three were selected: a contextual surface energy balance model, a biophysically focused surface energy balance model, and a strictly empirical vegetation index model. Each have strengths and complementary weaknesses. The contextual model was the Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC; [15]), a thermal infrared approach that provides ET estimates over stressed and un-stressed vegetation and requires minimal ancillary correction, but lacks the capacity to disaggregate transpiration from evaporation. The biophysical model was the Two Source Energy Balance (TSEB; [16]), also a thermal model capable of diagnosing ET under wide ranges of water stress with capacity to estimate both transpiration and evaporation, a capability that comes at increased model complexity and sometimes difficult-to-parameterize inputs. The third model was the Vegetation Index for the US Southwest (VISW; [17,18]) a strictly visible near infrared (VNIR) empirical approach. VISW benefits are its model simplicity, use of high spatial and high temporal resolution remote sensing data, while its chief drawbacks are need for crop-specific calibrations and relative insensitivity to rapid changes in vegetation water status.

Site selection to test the models was also based on practical matters, namely identifying sites growing familiar crops, managed using practices familiar to the authors, and encompassing an area large enough to include 10's to 100's of fields. The chosen site was the Central Arizona Irrigation and Drainage District (CAIDD) an ~ 35 km x 24 km, 35450 ha, agricultural region in South Central Arizona, USA. Typical crops at CAIDD have been cotton, small grains, forage, and vegetables. The useful attributes for this site are long growing seasons, non-deficit irrigation, hot and mostly dry summers and ample clear skies for remote sensing. Crops assessed for ET for the year 2008 were durum wheat, cotton, and alfalfa in plots ranging between 1.4 and 313.4 ha.

1. Results

In order to evaluate ET results, data quality of input weather and remote sensing needed to be assessed as they have a direct impact on the potential accuracy and variability of ET.

1.1. CAIDD air temperature

Accurate near surface air temperature is required by all three ET models, either as a constraint defining ET_o or as a boundary condition for estimating sensible heat flux values. In 2008 no weather station was deployed within CAIDD, where the nearest station- Coolidge [19]- lay 40 km to the North. To supply representative air temperature values to the models, cool, well-watered vegetated canopy LST values directly over CAIDD, validated by adiabatically adjusted Coolidge temperatures. Although it was not possible to directly verify accuracy of this approach, the aim is to achieve air temperature estimates within $\pm 2^\circ\text{C}$ for every modeled satellite overpass time. Considering all available TM5 and TM7 data (respective spatial resolutions of 120 m and 60 m) and median values of automatically selected cold pixels (Fig. 1), a linear model showed close correspondence between cool canopy temperatures obtained on clear sky days and projected AZMET Coolidge air temperatures. Exceptions were helpful in identifying remote sensing data sets to be excluded from modeling. These were observed for four days- DOY 77, 285, 333, and 357 - each confirmed to be partly cloudy days when TIR data were strongly affected by atmospheric conditions. Considering the high correlation between temperatures (RMSE $\sim 2.4^\circ\text{C}$), the cold-pixel air temperature proxy was adopted for use with inputs needed for TSEB and METRIC, though not for VISW, where modified AZMET air temperatures were used for ET_o estimation.

1.2. Land Surface Temperature Assessment

Land surface temperature (LST) data, critical to performance of the TSEB and METRIC models, were atmospherically corrected using the radiative transfer model, MODTRAN (Spectral Sciences, Inc.), and atmospheric temperature and humidity profiles from Tucson radiosonde data. Considering 44 overpass times for TM5 and TM7 in 2008, results showed $\sim 4.5^\circ\text{C}$ temperature differences between the land surface and at-sensor values. Exceptions occurred on the same 4 overpass times as previously noted for estimating near surface air temperatures. Three relevant facts about atmospherically correction thermal infrared data are illustrated in Fig. 2A. First, MODTRAN results show that requirements for accurate profiles are not demanding when LST values were $\sim 22^\circ\text{C}$, which means that for these conditions standard profiles could be used to estimate corrections. On the other hand, accurate estimates of atmospheric conditions become increasingly important as LST values increase beyond $\sim 35^\circ\text{C}$, in such cases correction errors could approach $8\text{--}10^\circ\text{C}$ when using inaccurate atmospheric profile input data. These Tucson model results show that at-sensor temperatures are almost always cooler than LST and that the range of observed temperatures is substantially less than actual surface conditions. This latter point is illustrated in Fig. 2B, where the impact of temperature span reduction is seasonal. For 2008 during monsoonal conditions, satellite observations of LST are reduced by $\sim 6^\circ\text{C}$. This means that atmospheric correction of satellite-based TIR data is important for any ET model using LST observations.

Having performed atmospheric corrections to the TIR data, relative accuracy of TM5 data were assessed by comparing them with TM7 and ASTER data. Fig. 3 shows the seasonal pattern for temperatures in 754 CAIDD plots: symbols represent mean values, bars represent ± 1 standard deviation. None of the three sensors had coincidental overpass times so direct comparisons were not possible. However a distinct LST bias was observed between TM5 and the other two sensors. Considering ASTER to be the most accurate, TM5 bias for 2008 was 3.8°C (Tab. 1). This offset was applied to all TM5 images prior to running TSEB, though not for METRIC since one of its features is insensitivity to LST bias.

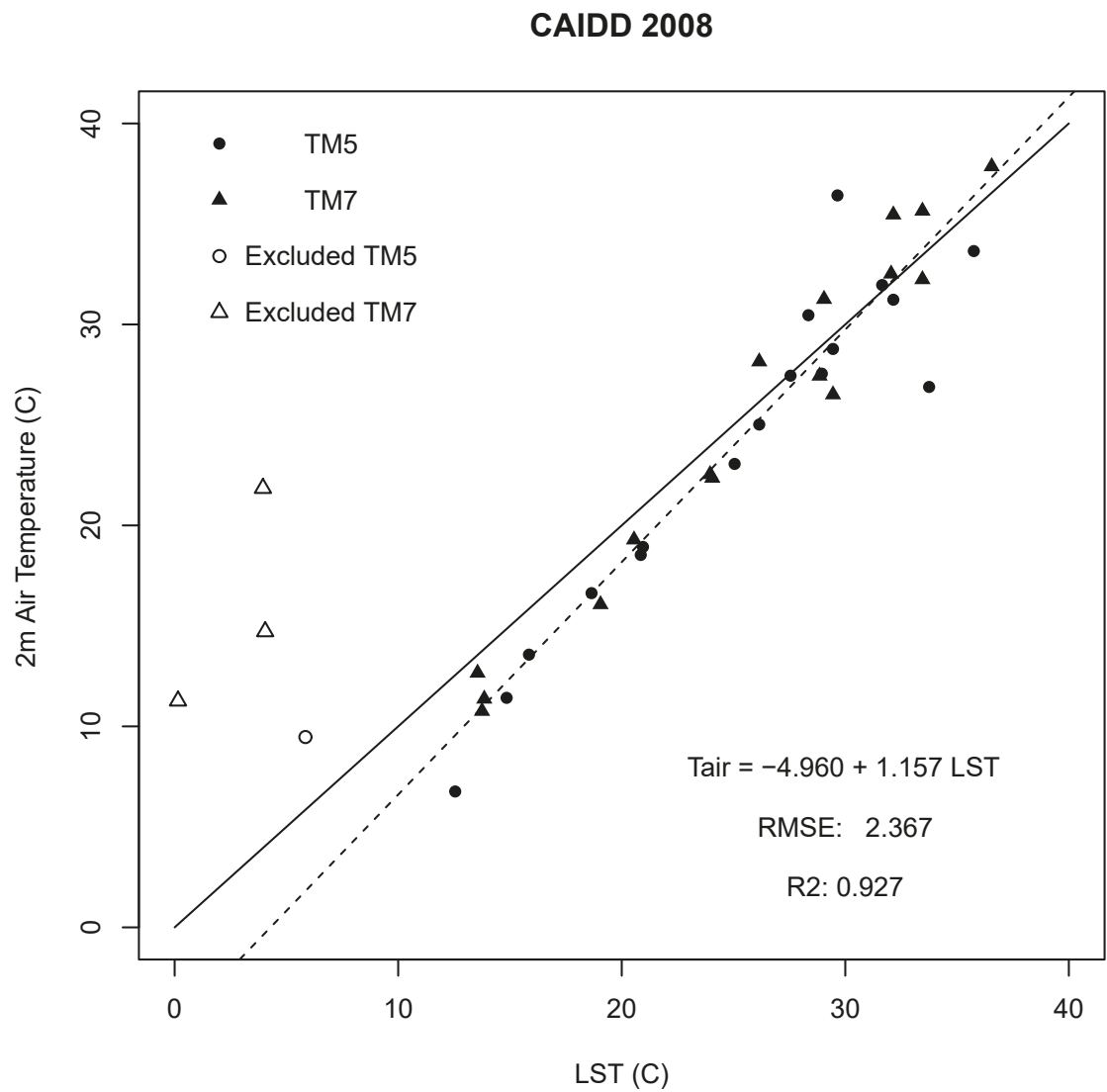


Figure 1. Cool surface temperatures compared with modeled 2 m air temperature at CAIDD.

Table 1. Comparing LST estimates from ASTER, TM7, and TM5 over CAIDD in 2008.

Sensor	DOY								Bias
	51	60	140	179	220	236	268	275	
ASTER	23.9	31.0	42.5	45.7	42.6	45.5	39.6	41.2	-
TM7	29.4	29.4	42.8	41.1	39.2	42.8	40.4	37.8	1.1
TM5	20.1	23.5	38.2	40.8	41.3	38.1	40.5	39.5	3.8

TWC 2008

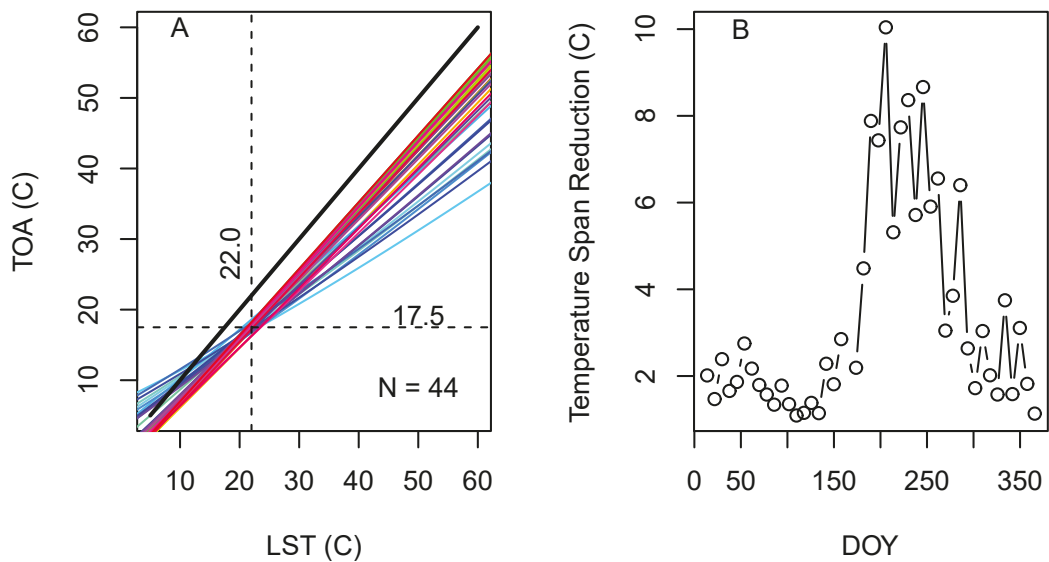


Figure 2. Thermal infrared corrections over South Central Arizona in 2008.

Table 2. Cumulative ET results

Crop	ET Model						
	ET_o	FAO-56	USDA-SW	METRIC	TSEB	VISW	Model Mean
Wheat	770.6 (28.6)	566.8 (-175.2)	655.6 (-86.4)	726.0 (-16.0)	842.8 (100.8)	657.1 (-84.9)	742.0 (-84.9)
Cotton	1234.6 (252.0)	972.9 (-9.7)	1046.5 (63.9)	803.6 (-179.0)	1142.8 (160.2)	1001.3 (18.7)	982.6 (18.7)
Alfalfa	1733.7 (306.7)	-	1887.7 (460.7)	1317.4 (-109.6)	1746.1 (319.1)	1217.3 (-209.7)	1427.0 (-209.7)

1.3. Evapotranspiration Results: Wheat

Remotely sensed ET estimates for durum wheat plots were assessed for 137 plots across CAIDD with planting dates spanning late 2007 to early 2008. Fig. 4 shows daily estimates made from the three models, METRIC (red), TSEB (blue), and VISW (green), linearly interpolated from 8 overpass days (days indicated by solid squares) to daily time steps. Values were smoothed with Savitzky-Golay polynomials to approximate weekly ET patterns. Average of the models, using equal weighting is shown as a heavy purple line. Standardized ET estimates are shown as symbols: FAO-56 (solid triangles), and USDA-SW ET values (open squares). Rainfall ($\times 0.1$ mm) is shown in light blue along the bottom.

Results show good agreement over the ~ 150 growing days with peak ET of 8 mm/d reached by all models and supported by FAO-56 estimates. METRIC and TSEB estimates generally closely tracked within 1 mm, indicating no discrepancies between using contextual LST values vs. using best estimates of LST and near surface air temperatures. Discrepancies existed between thermal models and the VISW between late January and mid-February, where the latter produced low daily ET values, ~ 1 mm/d, while METRIC and TSEB estimates ranged 2-3 mm/d, this difference could be due to the ability of the LST approaches to detect surface evaporation while VISW was modeled solely for transpiration. Considering consistency of estimates and their average, the USDA-SW model appears to be the outlier.

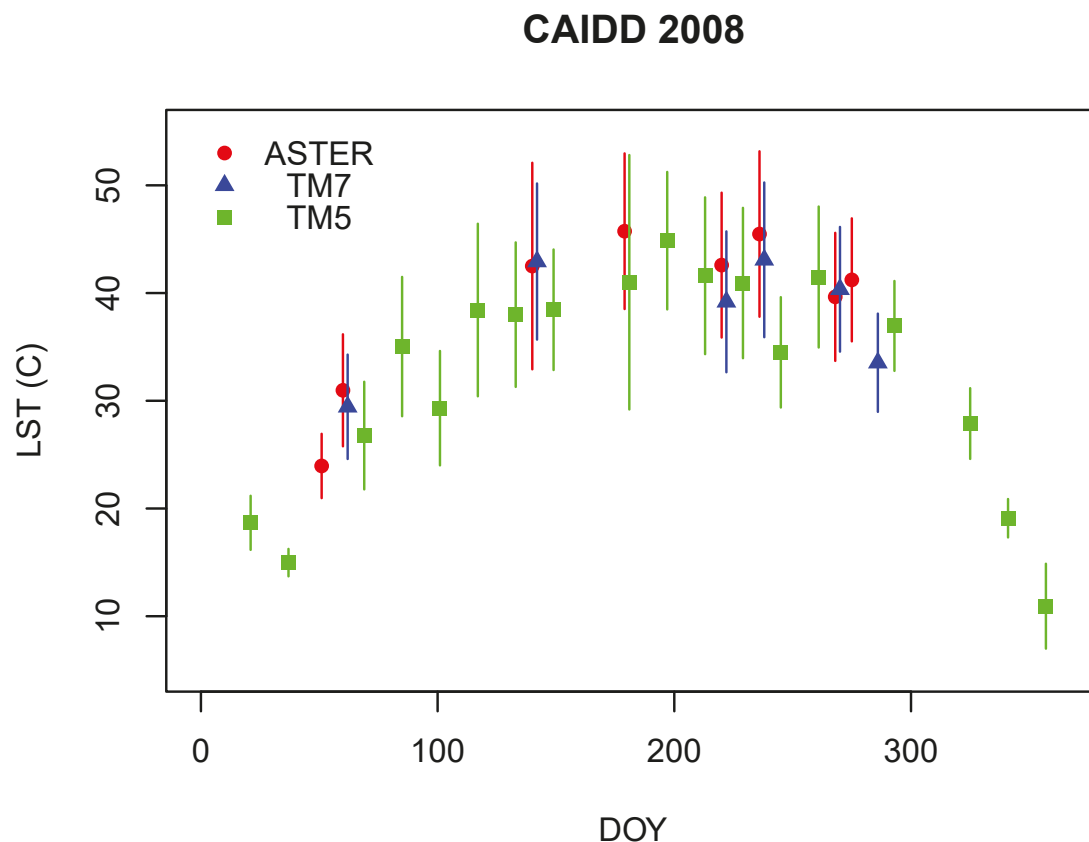


Figure 3. Comparing LST time series over CAIDD as observed by ASTER, TM7, and TM5

When daily ET values are accumulated, modeling consistency is confirmed, seasonal transpiration amounts were estimated to be 726.0, 842.8, 657.1 mm respectively for METRIC, TSEB, and VISW (Tab. 2). These values roughly agree with general estimates for durum wheat crop water requirements in Central Arizona to be 610 mm plus another 300-450 mm of water to accommodate irrigation inefficiencies [21]. If soil evaporation were included, VISW estimates would likely exceed 750 mm. The cumulative display also highlights that model variability for METRIC and VISW is about twice that of TSEB, indicating their respective sensitivity to end-member pixel selections and local weather.

1.4. Evapotranspiration Results: Cotton

In contrast to winter-time estimates, results for ET modeling over cotton showed broad agreement with seasonal growth but significant inter-model biases ranging between 1-4 mm/d (Fig. 6). For days shortly after planting METRIC, VISW indicate 2 mm/d, while TSEB returned 4-5 mm/d, an outcome indicating that TSEB estimates high surface evaporation in April. Mid-season ET estimates were unexpectedly low, ~ 6 mm/d, and much less than the 8-10 mm/d estimates provided by other models. Seasonal total ET results (Fig. 7) confirm estimation biases, with transpiration values of 803.6, 1142.8, 1001.3, respectively for METRIC, TSEB, and VISW. These values fall within cotton ET obtained under detailed controlled experiments 2014 and 2015 in Maricopa, AZ, where irrigation amounts ranged between 757- 939 mm [22]. As for wheat, plot-to-plot variability of ET over cotton for METRIC was about twice that of TSEB. Plot variability for the VISW, however, was about the same as for TSEB, possibly indicative of close coupling of NDVI with LST values.

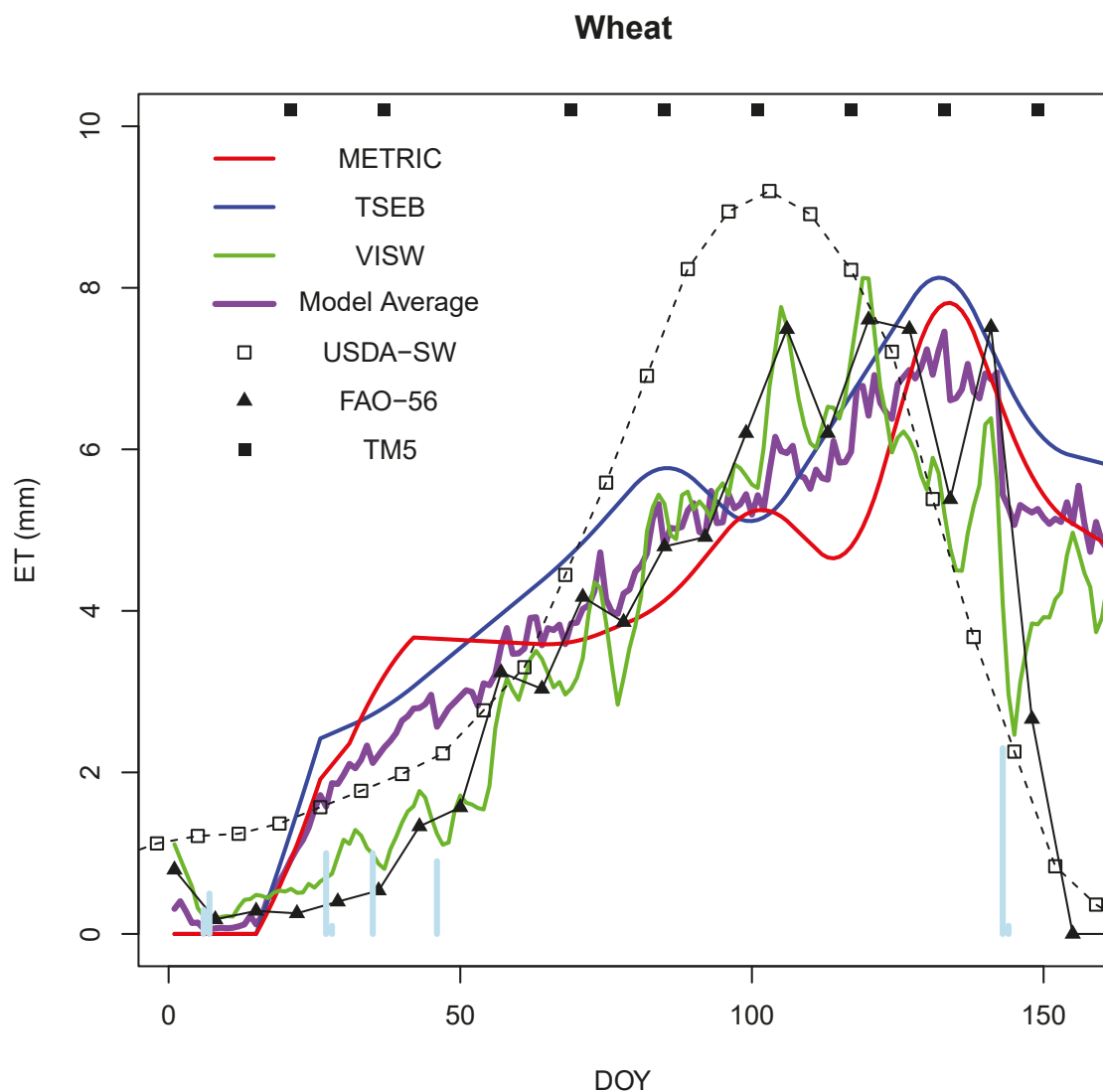


Figure 4. Wheat Daily ET estimates vs. Day of Year (DOY) from averages of 137 CAIDD plots in between January 1 and June 4, 2008. Shown are METRIC (red), TSEB (blue), VISW (green), USDA nominal estimates from [20] (open squares), and FAO-56 (solid triangles). Ensemble average ET for METRIC, TSEB, and VISW is shown as a heavy purple line. Overpass times for TM5 are denoted along the top margin by solid squares. Rainfall amounts are shown in light blue (x0.1 mm) along the bottom margin.

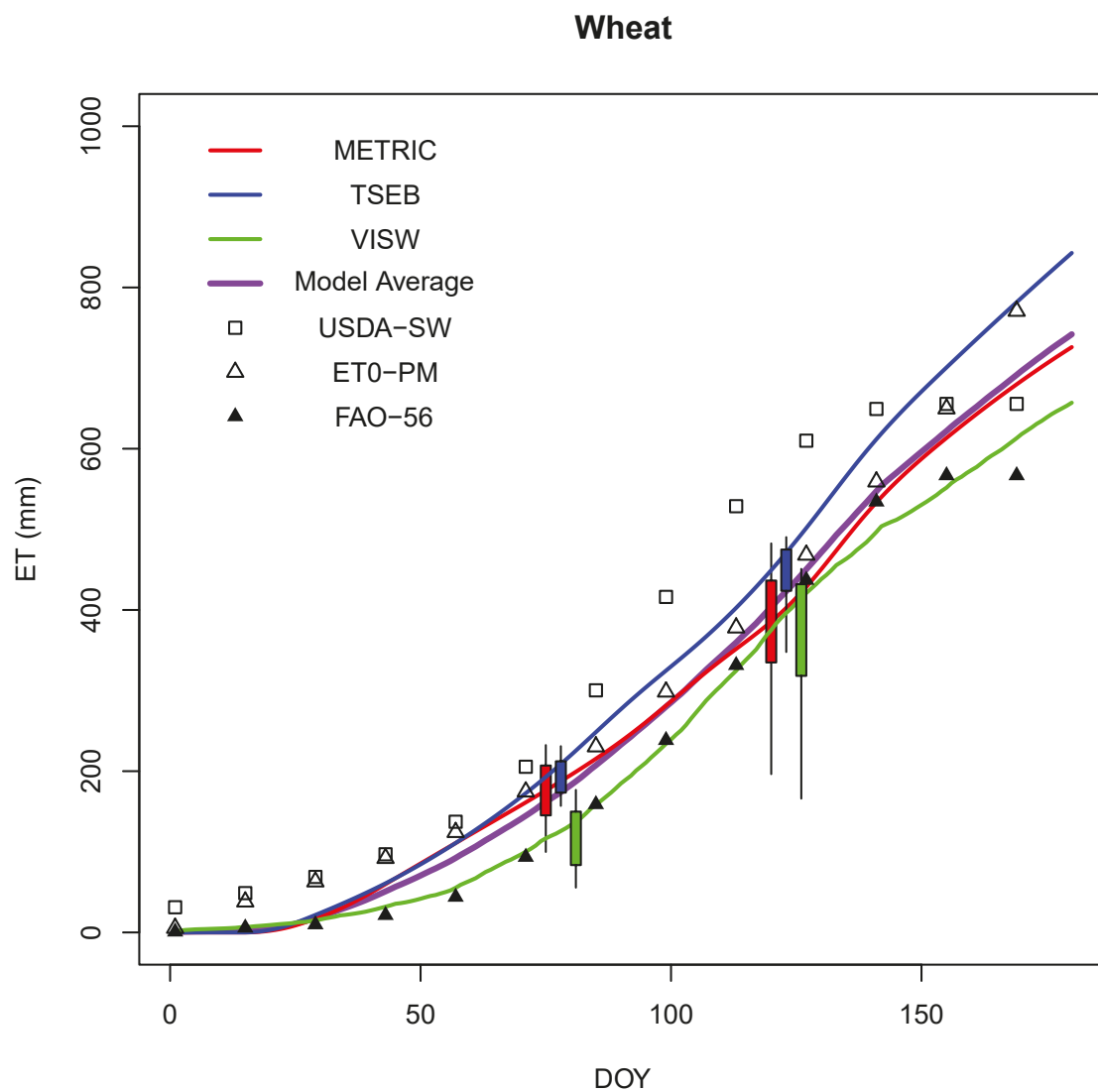


Figure 5. Wheat cumulative ET, derived from data shown in Fig. 4. ASCE potential ET for a tall crop are shown as open triangles. The box and whiskers—offset by 2 and 4 days for TSEB and VISW to help visualization—display the standard deviation and range of ET considering all 137 wheat plots on two days, 15 March and 30 April.

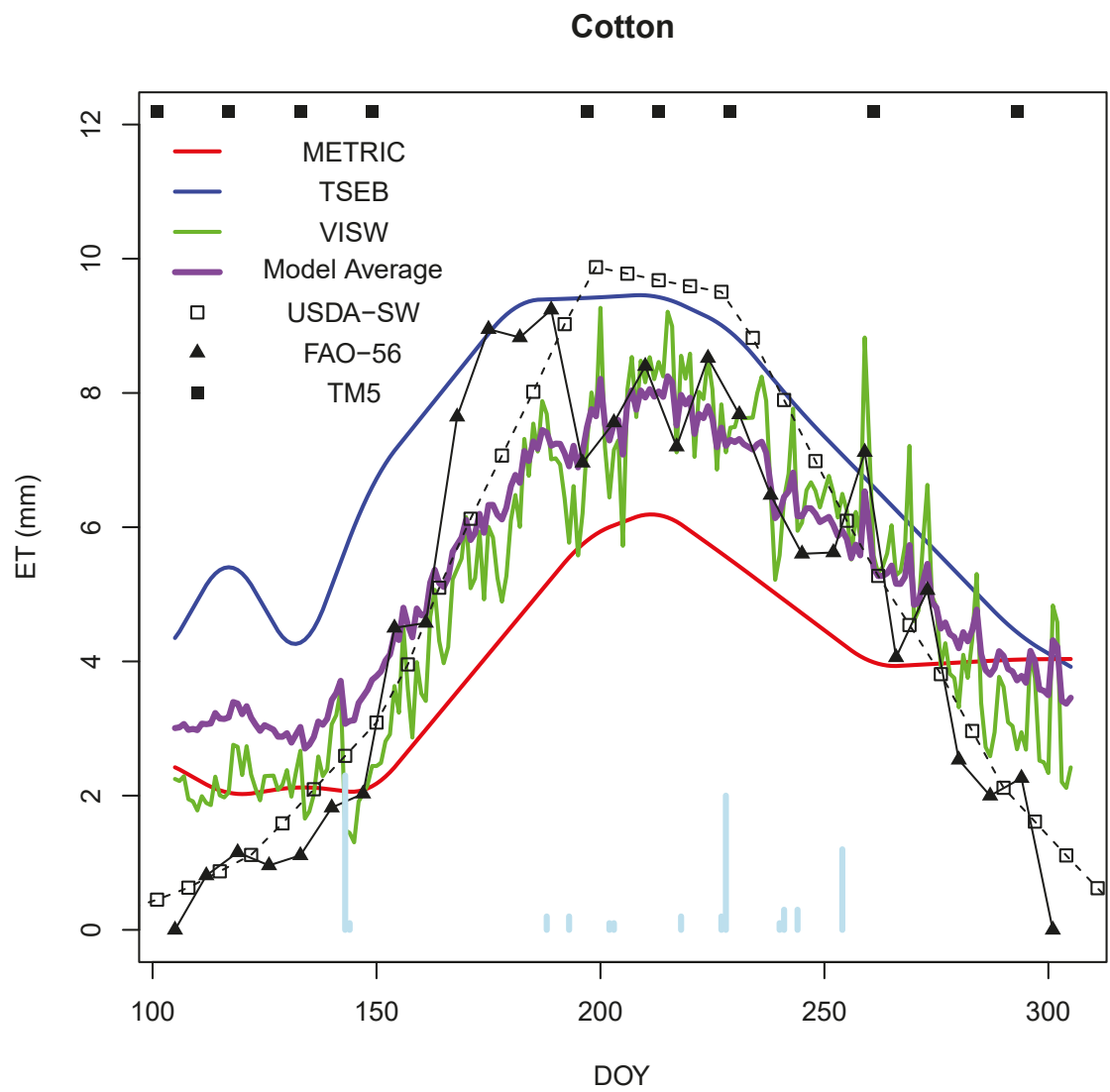


Figure 6. Cotton Daily ET for 183 plots at CAIDD.

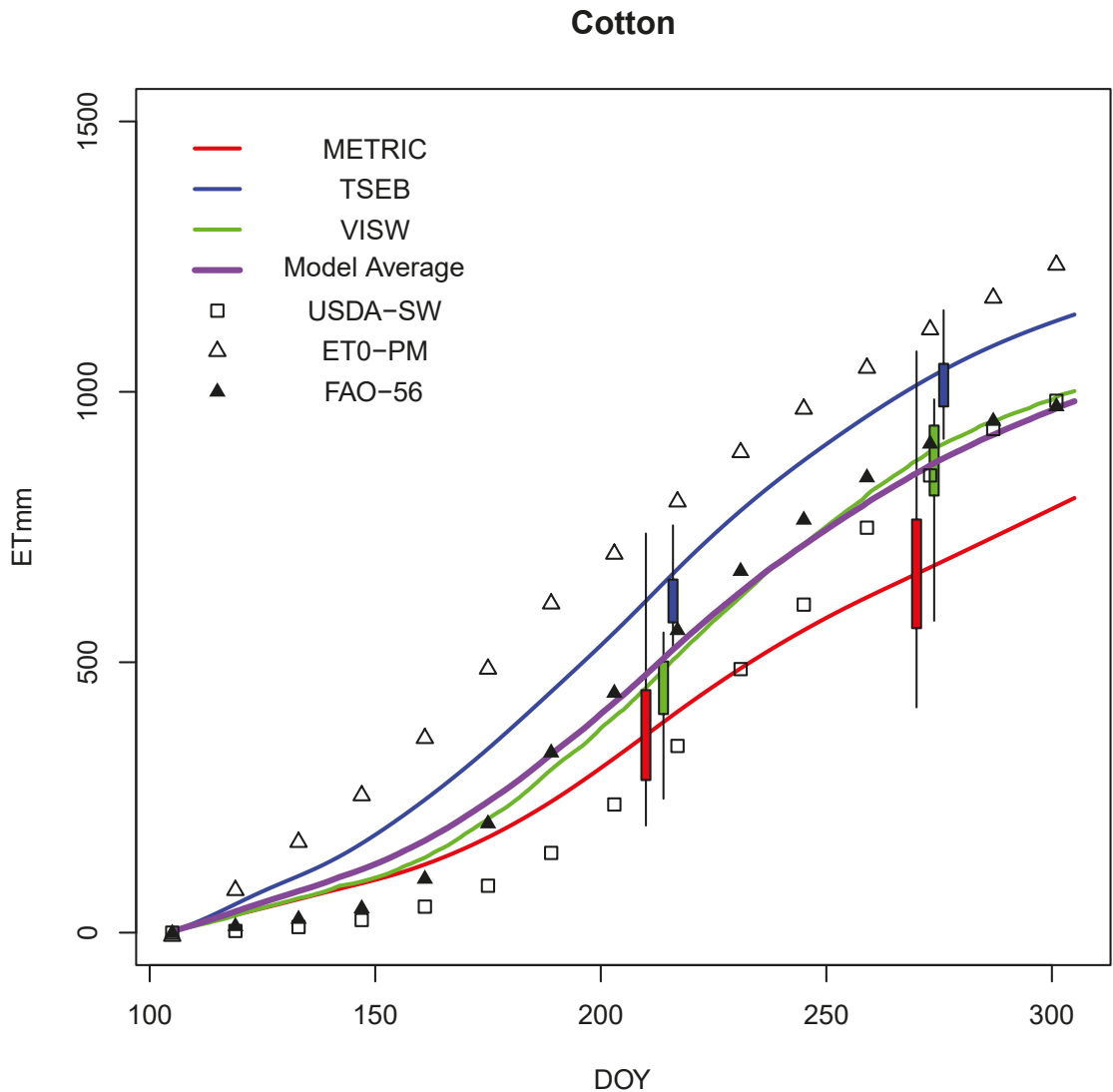


Figure 7. Cotton Cumulative ET derived from data shown in Fig. 6 for 183 plots. Box and whisker displays show standard deviation and range of cotton ET for 28 July and 26 September 2008.

Because a greater number of cloudy days occurred for summer months, a background test using weekly MODIS Terra and Aqua data was conducted. Although the spatial resolution, 250 m for red and NIR bands, is questionable for fields typically less than 180 m in extent at CAIDD, the test could help answer the question about potential ET modeling improvement with more frequent satellite overpasses. Using the MOD09Q1 and MYD09Q1 climatological products, VISW was implemented, but without daily interpolation steps. Resulting ET estimates closely match higher spatial resolution TM5 VISW data, but with a dynamic range compromise: over sparse vegetation, MODIS data over-estimates ET while over full cover tends to under-estimate ET.

1.5. Evapotranspiration Results: Alfalfa

ET estimates over alfalfa reflect a combination of the daily patterns observed over wheat and cotton, namely that there was close agreement between models in the cooler months, and significant biases in the summertime. Alfalfa is a year-round crop in Central Arizona with cuttings occurring approximately every two months, diagrammatically illustrated in the hypothetical FAO-56

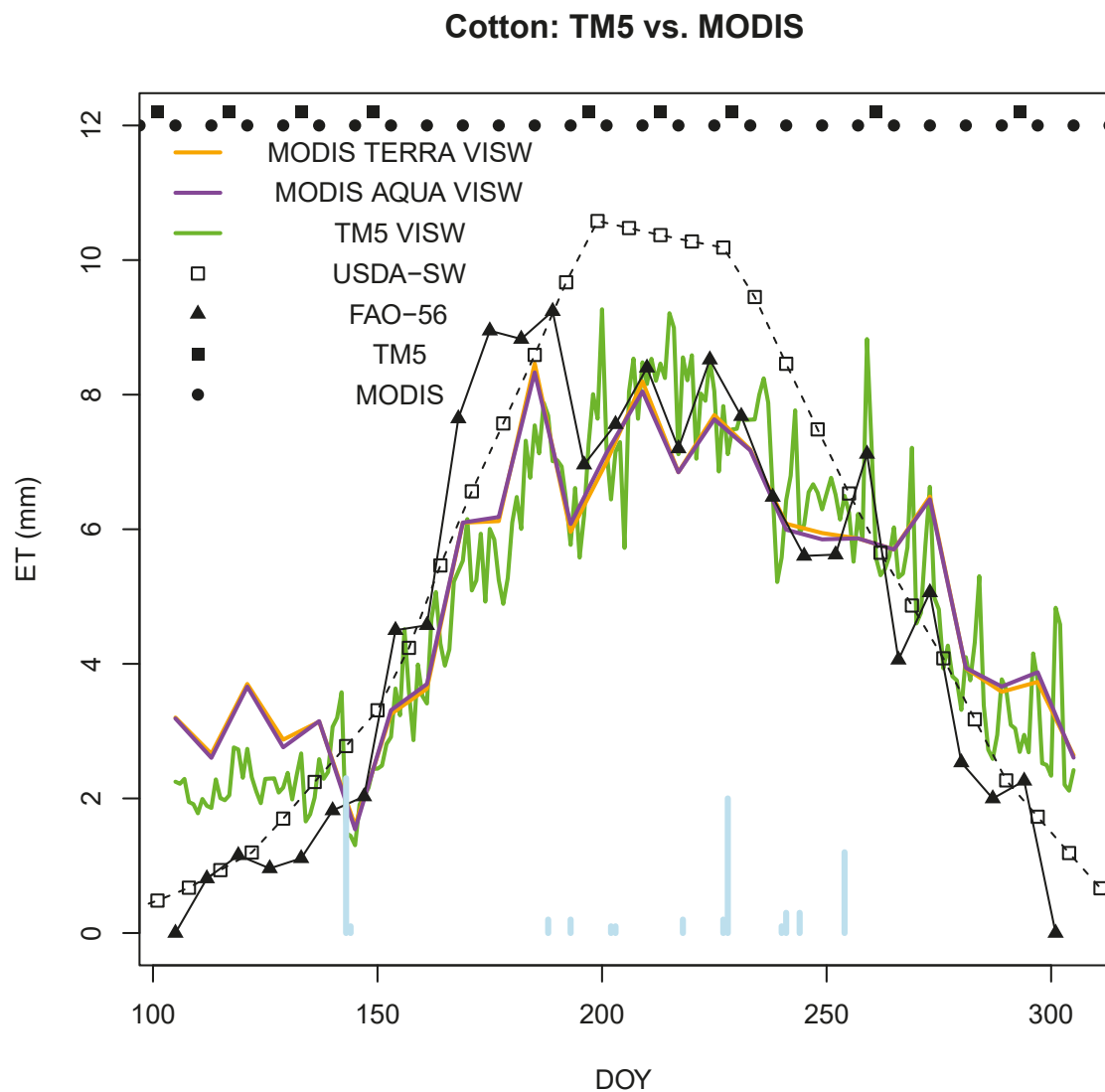


Figure 8. Cotton daily ET estimated with VISW using TM5 LEDAPS VNIR (green) MOD09Q1 (orange) and MYD09Q1 (purple). TM5 data were observed at nominally 30 m, 16-day intervals (squares), while MOD09Q1 (Terra, morning overpass) and MYD09Q1 (Aqua, afternoon overpass) are weekly composites of the best available 250 m VNIR. There was virtually no difference in MODIS ET estimates.

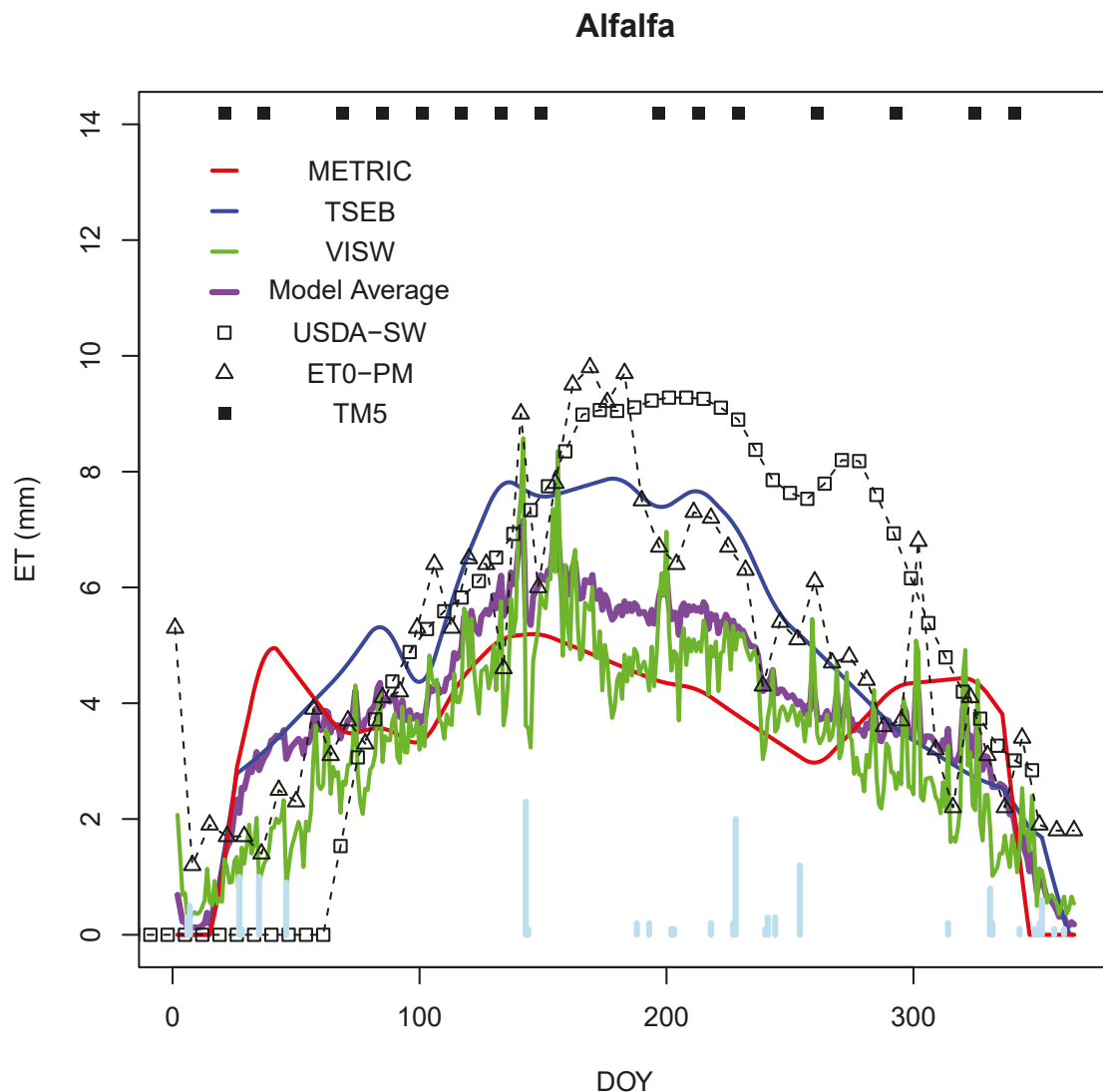


Figure 9. Alfalfa Daily ET for 225 plots on all days in 2008 at CAIDD. Due to the inability to accurately capture cutting days, no FAO-56 model was computed.

crop coefficient chart (Fig. 14). Early and late year ET values ranged between 2 and 5 mm/d, while mid-summer values ranged 5–8 mm/d, less than seen in a canopy resistance study, 8–9 mm/d [23], and compatible with values obtained between September and November in 1983: a range of 4.2–7.4 mm/d reported from a Coolidge, AZ study [24]. The USDA-SW model indicating peak ET values over 9 mm/d were not confirmed by remote sensing estimates. The USDA-SW model also predicts a late September ET depression from 9 to 7.5 mm/d on ~DOY 270, a feature possibly detected by METRIC, but not confirmed by any of the other models. In general, remote sensing data overpasses were insufficiently frequent to resolve rapid ET fluctuations, meaning that the fluctuations present in the display are representations of local weather. With frequent overpasses and high spatial resolution—as observed by the Ven μ s satellite (<https://www.theia-land.fr/en/products/venus>)—peak cover and harvest days could be detected and crop coefficients accurately mapped over time. However, TM periodicity was insufficient to resolve these patterns, which could mean a low-bias remote sensing of ET from all VI-based models.

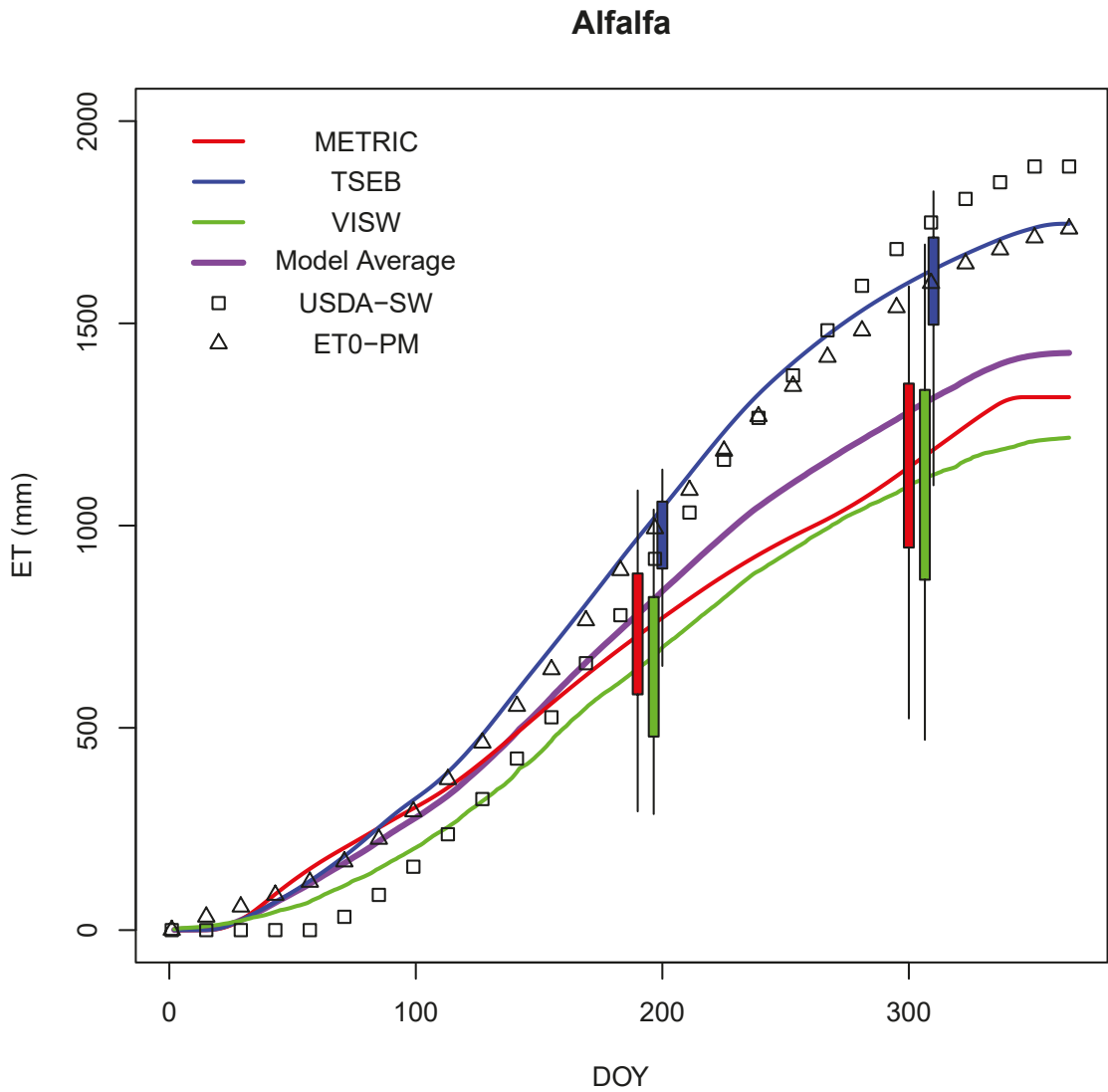


Figure 10. Alfalfa Cumulative ET

Cumulative ET estimates for alfalfa were the highest of the three examined crop types: 1317.4, 1746.1, and 1217.3 mm/d for METRIC, TSEB, and VISW respectively, in broad agreement with lysimeter data collected in Phoenix, AZ [25] and greater than observed at higher elevations in New Mexico, ~1190 mm [26]. As for wheat results, plot-to-plot variability of ET, as estimated by METRIC and VISW were approximately twice that of TSEB estimates.

2. Discussion

Results from assessing multiple remote sensing data sets demonstrated strategies for handling sparse air temperatures, biased TIR satellite observation (critical for TSEB), automated LST end-member selection for the METRIC approach, and quantile selection of NDVI extremes to help standardize implementation of VI-based models. The study found that LST data over well-watered, cool canopies were reasonable proxies for near surface air temperatures, easing ET modeling problems when no local weather station is available; this addresses the problem of unavailable nearby air temperature data needed for modeling the surface energy balance. Eventually consistent, accurate, and finely gridded weather data may make this approach unnecessary. The study implemented a

cross-comparison analysis of sensors- TM5, TM7, and ASTER- to improve radiometric accuracy. This improvement is similar to findings reported in [27], where multiple sensors were utilized for data fusion. By analyzing LST trends over all of 2008 observed by TM5 with respect to TM7 and ASTER, we were able to justify a 3.7 °C LST correction to TM5, a change that significantly reduced TSEB's ET bias. Using atmospherically corrected TM5 data, we used an LST/NDVI quantile selection method to identify so-called 'cold' pixels, then used those anchor points to search the neighborhood of each for 'hot' pixels. Very large (tens of thousands) samples were generated, greatly exceeding capabilities of manual end-member picking. Lastly, a related quantile selection approach for vegetation index normalization was implemented and tested. The approach is initially subjective in setting quantiles, but is objective and replicable for subsequent NDVI normalization. Unresolved in this study is an understanding of the impact of spatial resolution between sensors used to calibrate k_{cb} estimates and sensors used to employ them. Data underpinning calibrations in this study are based on footprints less than 2 m in diameter with NDVI values exceeding 0.9, while satellite-based footprints were 30 m where NDVI values uncommonly exceeded 0.65. As new satellites with finer spatial and temporal sampling, as for example provided by the *Venus* and ECOSTRESS missions, the scale divergence problem will be greatly reduced.

Results from this remote sensing study at CAIDD showed ET biases and uncertainties were affected by model and crop. As shown in Table 2, relative to model averaging results for cumulative ET values for all crops, METRIC was biased low ~100 mm, TSEB biased high ~190 mm, and VISW biased low ~90 mm. These respectively represent uncertainties of -10%, +18%, and -9%, and suggest that either METRIC or VISW could be used in preference to TSEB, despite prior experience showing TSEB bias is more on the order of 10% [28]. However, when biases were considered by crop type model preferences change. For wheat, a winter crop at CAIDD, the METRIC model had superior performance with a -16 mm bias, while TSEB returned +100 mm and VISW -85 mm. For cotton, VISW came closest to ensemble mean with a 19 mm bias, while METRIC and TSEB returned biases of -180 mm and +160 mm respectively. For alfalfa, all models diverge from the ensemble mean of 1427 mm by over 100 mm with biases consistent with those observed for the other two crops, i.e. METRIC and VISW were relatively low ET estimators while TSEB was relatively high. For METRIC in particular, this result is at odds with findings reported by [29], who report anomalously high ET values for a sorghum crop. Absent ground validation data, the significance of the model differences is unclear, but may be indicative of problems underlying model assumptions and sensitivities to inaccurate or infrequent inputs. In the case of METRIC, having seasonal dependency was not expected, suggesting that the horizontal temperature differencing methodology be checked for LST dependency. For TSEB, efforts to reduce TM5 LST bias effects may have been insufficient, possibly meaning that bias corrections should be adjusted on a scene-by-scene basis: bias LST corrections of 1-2 °C could readily remove any ET bias between TSEB and model averages.

Considering model biases exhibited in this study were larger than desired- one could seek differences in the range of 10%- future implementation strategies that could be considered are adopting ensemble modeling as a standard practice and creating shared standardized data sets to be used for model verification. Inputs, outputs, parameter settings, and intermediate results from runs against such data sets could be provided as a part of the publication and documentation. Currently, ET models results are reported based on the author's own implementation with no apparent means to verify consistency with other implementations of the same model namesake. Tracking down inconsistencies due implementation differences could significantly improve community ET modeling results. In this study a simple to implement, equally weighted model averaging approach, using only 3 models, showed a benefit: outlier estimates were reduced and total ET estimates appear to track ET values known from experience. An additional benefit for averaging is a shift away from model-intercomparisons, with an emphasis on accepting that each model has different information to contribute to the overall ET estimates. There are many documented remote sensing ET models that could be added to an ensemble framework. Such a framework would likely be impractical

on single-user platforms, but with rapidly improving IT technology and cloud-based computing, as being proposed by the OpenET organization, model sample sizes could be readily increased to 10 or more. The other strategy to address ET bias, model verification, needs community support, cooperation, and extensive sharing of common and well-documented data sets.

3. Materials and Methods

3.1. The Central Arizona Irrigation & Drainage District (CAIDD)

The CAIDD district is located in South Central Arizona and is a triangularly shaped region ~35 km North-South and ~24 km East-West at its southern boundary Fig. 11. CAIDD lies midway between Phoenix, 90 km to the northwest, and Tucson, 80 km to the southeast. CAIDD nominally comprises 35400 ha and delivers ~ 339207000 m³/year (ed4.biz). Irrigation water is supplied in approximately equal portions from ground water wells (350), and Colorado River water via the Central Arizona Project (CAP). Water deliveries within CAIDD are conveyed by three main canals: Santa Rosa, Central Main, and South Main. On-farm deliveries are provided by canal turnouts; in 2008 there were 327 active turn-outs. Per agreement with CAIDD directors, water volumes and release dates at each turn-out during 2008 were provided to the authors, but with the condition that proprietary water use information not be disclosed. The nearest weather station data for CAIDD was provided by the Coolidge AZMET [30] station, approximately 50 km to the North. Using a CAIDD-provided map and imagery from Google EarthTM (mention of trademarked or commercial products in this manuscript is for reader convenience, USDA provides no endorsement or recommendations for these products), CAIDD canals and turn-out locations were digitized and incorporated into an in-house GIS database. Using Google Earth imagery and TM5 scenes, boundaries of 754 active cropping areas were digitized as shape files within the ArcGIS editor environment. Considering the objective to assess average crop evapotranspiration across the district, a small sample (~30) would likely be sufficient for analysis, however a larger sample size was needed to offset errors in a multi-cropping environment where planted areas change shape and size seasonally. Thus this study identified defined crop areas for most of the 2008 season at CAIDD. Using the USDA Cropland map for Arizona 2008 [31], crop mask files for wheat, cotton, and alfalfa were generated at its source 30 m resolution. Estimated crop classification accuracy exceed 80%. Dominant crops at CAIDD include wheat, cotton, alfalfa, corn, and sorghum, with several combinations occurring within a year. Using the total area of digitized cropping areas, ~22000 ha (54305 acres), and provided water delivery for 2008 of 374506700 m³ (303618 ac-ft), average irrigation water depth was 1700 mm (5.6 feet). Crop varieties grown at CAIDD in 2008 were not available.

3.2. Remote sensing data

This study evaluated ET over the CAIDD considering 5 main sources of remote sensing data collected in 2008: Landsat TM5, Landsat TM7, ASTER, MODIS Terra and Aqua. Additionally, the study implicitly used Resourcesat-1 (IRS-P6) and AWiFS via USDA Cropland crop cover maps.

The core data for the study were 22 scenes from TM5 (Table 4). Data were downloaded from the USGS Explorer website (<https://earthexplorer.usgs.gov>) as geotiff files; geo-positioning errors relative to well-defined ground-reference structures such as irrigation canals were small and usually less than the nominal 30 m pixel size. 23 TM7 were also assessed, but due to missing scan lines and inconsistent reflectances relative to TM5, their use was limited to assessment of land surface temperatures. Normalized Difference Vegetation index (NDVI) maps were derived from 30 m red and near infrared (bands 3 and 4 respectively) obtained from the USGS LEDAPS data collection [32]. LEDAPS data are atmospherically corrected land surface reflectance, obviating the need to independently perform atmospheric corrections in the visible and near infrared bands. Land surface temperature (LST) maps were derived from 120 m TM5 band 6 and 60 m TM7 band 6. LST values

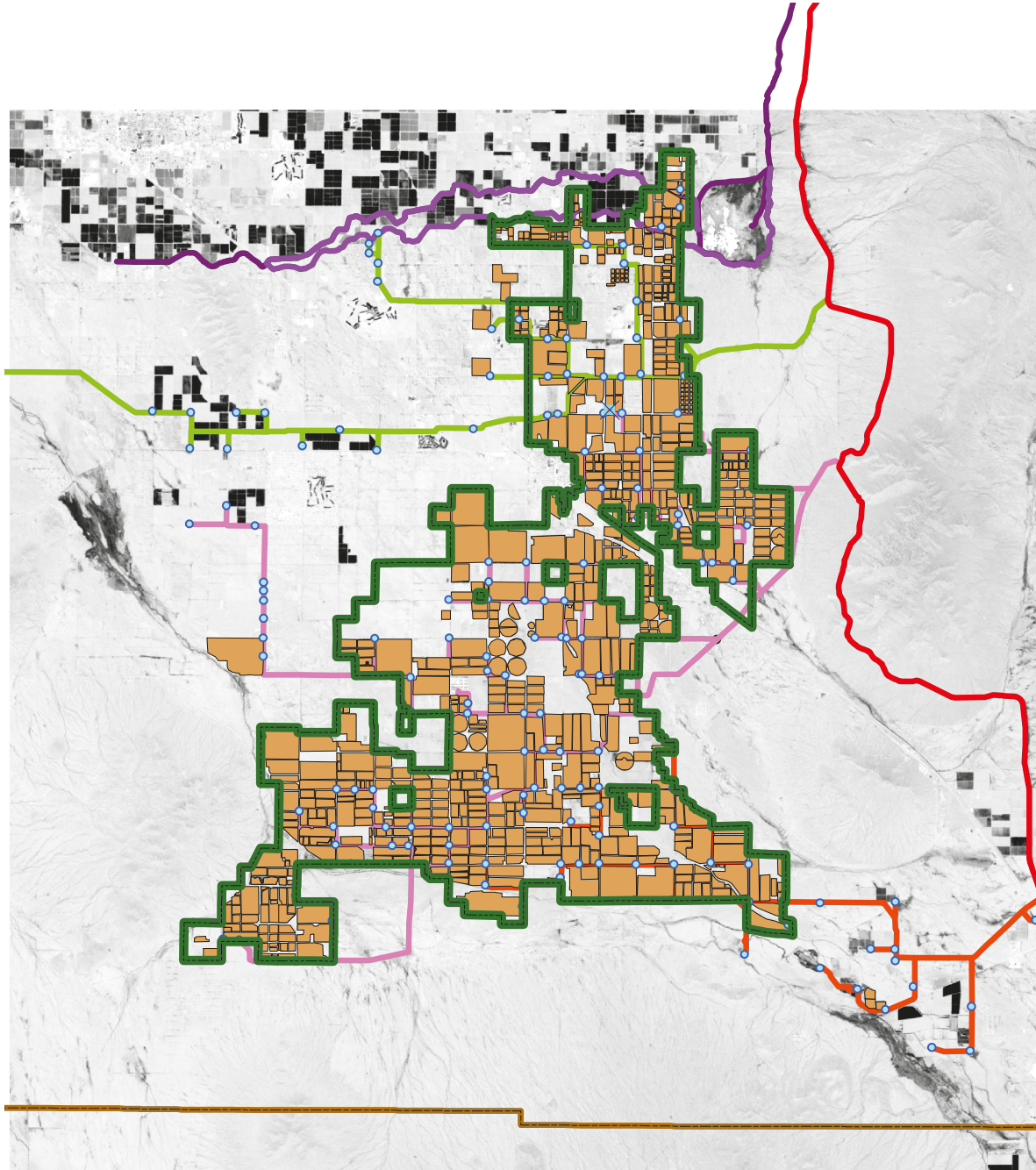


Figure 11. CAIDD regional map. The dark green line denotes the CAIDD extent, tan polygons the modeled plots, and blue circles irrigation turn-outs. The remaining bold lines denote the Central Arizona Project canal (Red), the Casa Grande Canal (purple), the Santa Rosa Canal (light green), the Central Main Canal (pink), and the South Main Canal (orange).

Table 3. ASTER observations over CAIDD in 2008.

Index	Date	DOY	Notes
1	21 February 2008	51	NW CAIDD
2	1 March 2008	60	few clouds, all of CAIDD
3	20 May 2008	140	small coverage of east CAIDD
4	28 June 2008	179	patchy clouds
5	24 August 2008	236	west 1/2 of CAIDD
6	25 September 2008	268	west 1/2 of CAIDD
7	2 October 2008	275	NW CAIDD

were computed by converting TM top-of-atmosphere radiances ($L_{sensor,\lambda}$) to land surface radiances ($L_{surf,\lambda}$) via Eq. 1:

$$L_{surf,\lambda} = \frac{L_{sensor,\lambda} - L \uparrow_{\lambda}}{\tau_{\lambda}} - (1 - \epsilon_{\lambda})L \downarrow_{\lambda} \tag{1}$$

The spectrally dependent atmospheric correction terms are τ_{λ} (transmissivity), $L \uparrow_{\lambda}$ (upwelling atmospheric radiance), $L \downarrow_{\lambda}$ (downwelling radiance), ϵ_{λ} (surface emissivity, 0.97 assumed), all obtained by using radiosonde atmospheric profile data from the NOAA FSL Tucson site [33] and the radiative transfer modeling package MODTRANv5 [34] (Spectral Sciences, Inc., Burlington, MA). To convert L_{surf} into LST, a temperature-spectral radiance lookup table was used. This table was created by solving the Planck blackbody function for 0.2 μm steps, each weighted by the corresponding TM5 spectral response function F :

$$\bar{L} = \frac{\sum_{\lambda_{ll}}^{\lambda_{ul}} L_{\lambda} F_{\lambda} R_{\lambda} \Delta\lambda}{\sum_{\lambda_{ll}}^{\lambda_{ul}} F_{\lambda} R_{\lambda} \Delta\lambda} \tag{2}$$

Both TM5 and TM7 were downloaded, merged and geopositioned. Although TM data are nominally georegistered, on occasion large positioning errors occurred and had to be corrected. Use of both TM5 and TM7 would provide 8-day samples under clear sky conditions. Unfortunately, two factors made use of TM7 difficult: inconsistent reflection patterns with respect to TM5, and the well-known gaps due to the broken scan line corrector. While the gaps could be accommodated when developing plot averages, reflection inconsistencies were troublesome. For relatively slowly developing crops such as cotton, reflection anomalies could be readily identified, but for highly dynamic alfalfa cropping systems, where cuttings occur at intervals ~ 6 weeks, uncertainties were high as to whether the TM7 variations were real or artifacts.

To help validate LST data obtained from TM sensors, 7 clear to mostly clear, L1B ASTER [35] scenes covering all or parts of CAIDD were downloaded. Data from its 5 thermal infrared bands were processed in a similar way to that shown in Eqs. 1 and 2, but with additional steps to refine the surface temperature and emissivity estimates via an approach known as ‘normalized emissivity’ (NEM, [36]). By considering temperatures observed in different parts of the 8–11 μm thermal infrared window, one is able to improve LST accuracy over estimates provided by a single thermal band such as for TM5 and TM7 [37]. ASTER is mounted on the sun-synchronous TERRA platform and has similar overpass times of day as for TM5 (18:20 vs. 17:46 UTC), though with different overpass days. This means that while direct LST validations were not possible, it is reasonable to expect temperature differences to be on the order a few $^{\circ}\text{C}$.

Lastly, 250 m red/near infrared MODIS from both the morning overpass platform Terra, and the afternoon overpass platform Aqua were downloaded and incorporated into the ET analysis routines. The relatively coarse spatial resolution of MODIS would normally preclude its consideration for this study—typical field sizes at CAIDD are ~ 180 m—these data sets have superior temporal resolution relative to TM5 and may provide improved time precision of cropping activities such as planting and harvesting. 47 MOD09Q1 and 46 MYD09Q1, weekly composited, red/NIR surface reflectance data were downloaded and included in the CAIDD data set. Map tiles, sourced in sinusoidal grid, were

re-projected to UTM zone 12 coordinates as geotiff files using the MODIS Reprojection Tool (MRT, [38]).

3.3. Evapotranspiration Models

To assess the efficacy of different remote sensing approaches for estimating spatially distributed ET, three different models were implemented: the Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC [39]), the Two-Source Energy Balance (TSEB [16]), and Evapotranspiration Mapping with Vegetation Index ([17,18]). Summaries of, and deviations from their standards are described below; complete details for METRIC and TSEB can be found in the source citations. The chief characteristics distinguishing the models are modeling ET using the surface energy balance (TSEB, METRIC) vs. using crop-specific ET calibrations (VI), consideration of canopy vs. soil biophysics (TSEB), ET calibration using contextual data from remote sensing (METRIC), and adaptability to non-standard ET conditions (TSEB, METRIC).

3.3.1. METRIC

ET estimation with METRIC is based on the assumption of surface energy balance:

$$LE = R_n - G - H \quad (3)$$

where ET is represented in its energy flux form, latent heat (LE), and equal to net radiation (R_n), minus soil heat flux (G) and sensible heat flux (H). Units for all components in W/m^2 ; sign convention is positive R_n for incoming radiation, and positive G , H , LE for fluxes away from the soil/canopy surface. Conversion to mass water flux is:

$$ET = \frac{LE}{\lambda \rho_l} \quad (4)$$

where λ is the latent heat of vaporization, and ρ_l is the density of liquid water. METRIC estimates R_n using satellite-based reflectance and thermal infrared data [39,40]:

$$R_n = R_{\downarrow s} - \alpha R_{\downarrow s} + R_{\downarrow l} - R_{\uparrow l} - (1 - \epsilon_{surf}) R_{\downarrow l} \quad (5)$$

where the radiation components— all estimated using remotely sensed observations— are downwelling shortwave radiation ($R_{\downarrow s}$), surface albedo (α), downwelling longwave radiation ($R_{\downarrow l}$), upwelling longwave radiation ($R_{\uparrow l}$), and land surface emissivity (ϵ_{surf}). METRIC estimates of G were derived from a SEBAL empirical relationship with surface albedo and NDVI [41]:

$$G = R_n \left[T_{surf} (0.0038 + 0.0074\alpha) (1 - 0.98NDVI^4) \right] \quad (6)$$

where T_{surf} is surface temperature in Celsius.

The remaining term on the right-hand side of Eq. 3, H , is solved using METRIC's distinctive feature of representing vertical temperature gradients between the surface soil/vegetation canopy and the overlying atmosphere in terms of horizontal temperature gradients readily accessible from thermal remote sensing at resolutions better than 100 m. The usual resistance equation for H is described:

$$H = \rho_a c_p \frac{\Delta T}{r_a} \quad (7)$$

where ρ_a is density of moist air, c_p is specific heat of the moist air, ΔT is the air temperature gradient between the surface and overlying air, and r_a is a resistance term representing the effectiveness of heat transport between the surface and overlying air. The crux of remotely sensed surface energy balance models is how they resolve ΔT and r_a . In some models, such as TSEB (described below), ΔT is determined from radiometric surface temperature and ambient air temperature. The r_a term

is estimated from models of surface roughness, wind speed, and atmospheric stability. Following the SEBAL method, METRIC solves for these terms indirectly by selection of so-called ‘cold’ and ‘hot’ pixels that are presumed to represent reference H end-members. Knowing H for the extreme conditions allows estimation of H fluxes at all other pixels. These reference pixels are respectively denoted H_{cold} and H_{hot} , and are used together to compute the coefficients needed to convert LST observations into an apparent ΔT . The equation are:

$$\Delta T = a + bT_s \quad (8)$$

where the parameters a and b are solved used the selected hot and cold pixels:

$$b = \frac{\Delta T_{hot} - \Delta T_{cold}}{T_{s,hot} - T_{s,cold}} \quad (9)$$

$$a = \Delta T_{hot} - bT_{s,hot} \quad (10)$$

ΔT_{hot} and ΔT_{cold} are temperature gradients derived from solutions of Eq. 3 using reference conditions and observations. This approach simplifies, and potentially increases accuracy in ET computations because biases due to errors in LST estimation are canceled by the hot/cold differencing. Thus a reference cold pixel is assumed to represent ET at a standardized rate, $1.05 * ET_r$, a rate typically corresponding to either a short (0.12 m) or tall (0.5 m) crop ET as defined in [42]; the additional 5% above reference ET accommodates conditions with evaporation at the soil surface. The surface energy balance equation for cold pixels becomes:

$$H_{cold} = (R_n - G)_{cold} - LE_{cold} \quad (11)$$

where LE_{cold} is the energy flux equivalent of $1.05ET_r$. At the ‘hot’ pixel, METRIC allows residual evaporation based on antecedent surface moisture, and follows the pattern in Eq. 11. For this study, however, antecedent moisture was considered unknown and difficult to accurately model. Thus the traditional SEBAL approach, where LE was assumed to be zero:

$$H_{hot} = (R_n - G) \quad (12)$$

A significant obstacle for wide deployment of METRIC is the stated need for trained experts knowledgeable in surface energy balance, this requirement not only limits the number of practitioners, but also introduces subjectivity and non-reproducibility of ET results [43]. This study assessed the need by implementing an alternative objective and reproducible approach based on the hypothesis that quantile selection of hot and cold pixels can return meaningful ET reference locations. The method implemented is similar to that previously described in [43] who implemented the approach over almonds in a California study. Given a 30 m resolution TM scene, choose as cold reference pixels all those within 95-99% NDVI quantiles and also with 1-10% LST quantiles as illustrated in Fig. 12. While the quantile selection can be automated, the choice of quantile levels remains subjective as there are as of yet no established criteria. For the 2008 TM data we sought levels that were not outliers yet wished to include thousands of sample points for subsequent steps—experience at multiple sites could be used to formalize optimal quantiles. Next, for each selected cold pixel, scan all neighboring pixels within a fixed search radius and choose the maximum LST found. In this study the radius was set to 300 m, a value approximately double a typical CAIDD field size to ensure high probability of finding bare soil pixels. Third, compute the median cold and hot pixel values based on all previously selected cold/hot value pairs. Lastly, used the median values to compute the a and b coefficients in Eqs. 10 and 9.

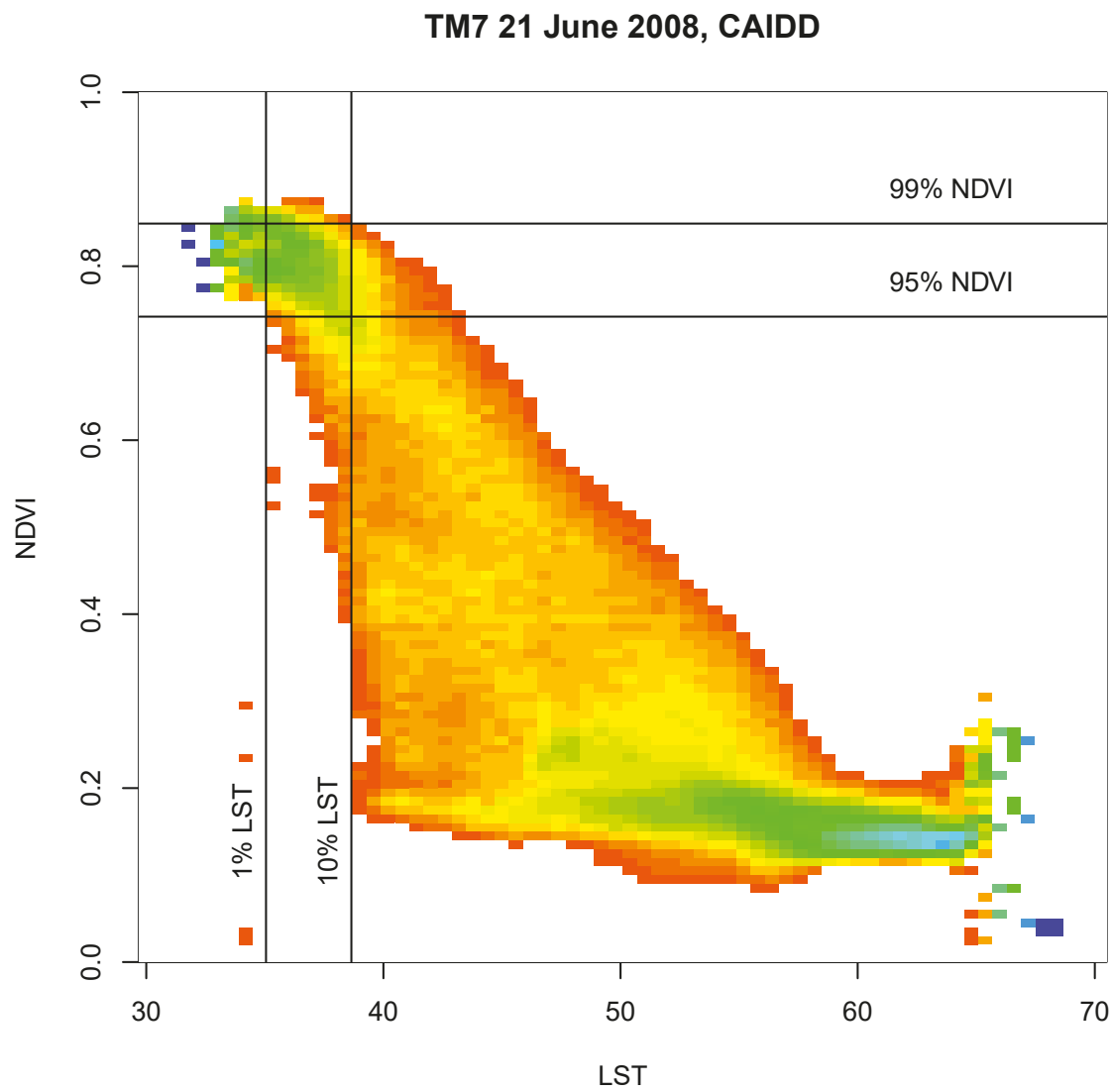


Figure 12. NDVI vs. LST histogram

344 3.3.2. TSEB

The TSEB model is a thermal-based remote sensing ET approach that is distinguished by its partitioning of surface fluxes into separate transpiration and evaporation components. This partitioning allows for differing transport resistances between soil, canopy, and atmosphere and accommodates use of radiometric instead of aerodynamic temperature in the estimation of ΔT by modifying Eq. 7:

$$H_{soil} = \rho c_p \frac{\Delta T}{r_a + r_{soil}} \quad (13)$$

where a single resistance term (used by METRIC) is replaced by two: one for the canopy-air transport, another for soil-air transport. A third resistance term (r_x) is added to the formulation when TSEB is implemented as a series network, the approach adopted for this study. Thus for the soil evaporative heat flux (LE_{soil}), one computes:

$$LE_{soil} = R_{n,soil} - G - H_{soil} \quad (14)$$

TSEB does not use radiometric temperature observations to solve the canopy energy flux—a constraint imposed by the inability to uniquely decompose composite temperature into soil and canopy components. Instead, the canopy is modeled to transpire at close to potential rates as estimated by the Priestley-Taylor parameter α :

$$LE_{canopy} = \alpha f_G R_{n,canopy} \left[\frac{\Delta}{\Delta + \gamma} \right] \quad (15)$$

When solutions of TSEB lead to negative LE_{canopy} values, an indicator of condensation, the α term is adjusted until LE_{canopy} equals or exceeds zero. Having obtained component latent fluxes, the total surface LE is:

$$LE = LE_{soil} + LE_{canopy} \quad (16)$$

Conversion of LE to ET in mm/hour is:

$$ET = \frac{LE}{\lambda \rho_l} \times 3.6 \times 10^6 \quad (17)$$

345 To obtain ET, LE fluxes from the soil and canopy sources are summed. For the CAIDD study, the series
 346 network TSEB was implemented using methods in [16] with two exceptions. The first was to adopt
 347 median cold pixels (as computed for METRIC) for the near surface air temperature. This approach
 348 has been used in other contexts [44] and helps accommodate the fact that the nearest air temperature
 349 observations lie 40 km away from CAIDD and might have been significantly different. The second
 350 amendment was to set plant heights for wheat, cotton, and alfalfa to nominal heights as indicated in
 351 Table 5 and surface roughness over bare soil of 0.02 m:

352 These simplifications were imposed because heights were unknown but approximations are
 353 needed to run TSEB.

354 3.3.3. VISW

The Vegetation Index for the South Western US (VISW) model is an empirical ET estimation approach based on experimental data collected in Arizona [17,18,45–47], where wheat, cotton, and alfalfa crops were grown under controlled irrigation regimes and frequently monitored with proximal sensing of NDVI. The approach aims to use remotely sensed vegetation indices as a proxy for the crop coefficient k_{cb} , thereby replacing standardized values with real-time observations. Use of empirically constrained vegetation indices to derive crop coefficients is a concept pursued previously by several researchers, including [14,48–52]. The original concept for a crop coefficient is a single scaling factor:

$$ET = k_c ET_o \quad (18)$$

FAO-56 Wheat & Cotton

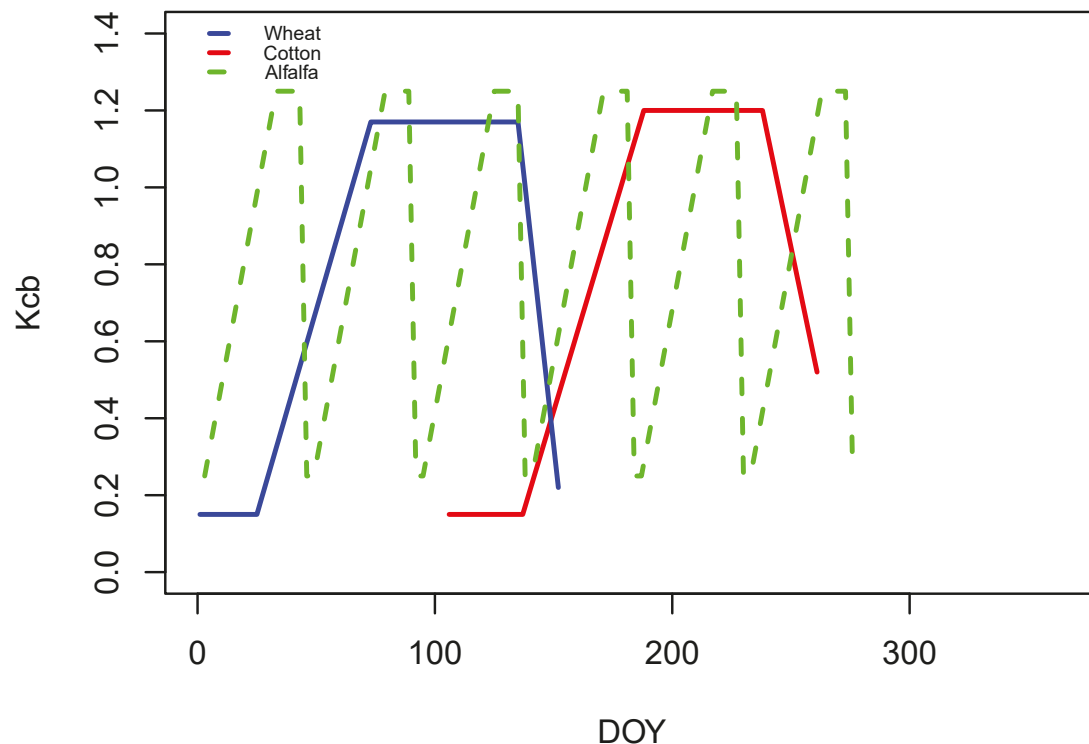


Figure 13. Basal crop coefficients for wheat, cotton, and alfalfa

where ET_o is potential evapotranspiration. In the FAO-56 [15] approach, the crop coefficient is re-partitioned into transpiration and evaporation components:

$$ET = (k_{cb} + k_e)ET_o \quad (19)$$

where k_{cb} , a basal crop coefficient, represents transpiration, while k_e , represents surface evaporation component.

Modeling of soil evaporation under the VI paradigm is not addressed in this study. The simplicity and efficacy for using crop coefficients to model, forecast, and manage crop water requirements has been noted for many years, notably documented in [15,48,53]. Examples of k_{cb} coefficients are illustrated for wheat, cotton, and a hypothetical alfalfa crop with 6 cuttings:

A chief shortcoming of the crop coefficient methodology is its need for local calibration [15,53], which has led to suggestions by [50,51,54,55] that remote sensing observation of vegetation, via an index such as NDVI, could greatly alleviate this constraint. Adoption of VI-based ET estimation is becoming well-known and to a limited extent, commercialized or facilitated by governmental agencies at local (Crop Circle (Holland Scientific, Lincoln, NE); Greenseeker (Trimble, Sunnyvale, CA)) and regional scales (CIMIS California TOPS Satellite Irrigation Management Support (SIMS), [6], <https://ecocast.arc.nasa.gov/simsi>). Vegetation indices are inherently slow reactors to changing surface conditions such drought or delayed irrigation and hence are not ideal for monitoring, forecasting or managing crop water use under water deficits. Thus while future water shortfalls may change operations, we believe few farms in US Southwest practice deficit irrigation. This means that the VI approaches calibrated under standard conditions is a viable water management tool.

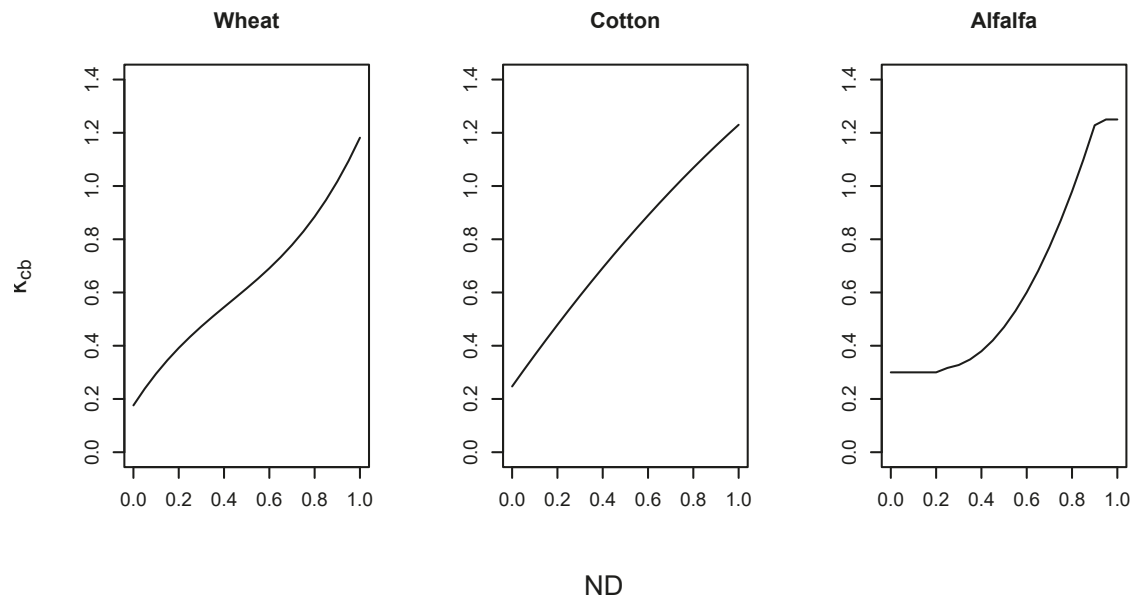


Figure 14. Crop coefficient transformation functions for wheat, cotton, and alfalfa.

In this study ET_o is derived from the tall-crop (0.5 m) standardized ET equation as presented in [42]:

$$ET_o = \frac{[0.408\Delta(R_n - G)] + \left[\gamma \frac{C_n}{T+273} u_2 (e_s - e_a)\right]}{\Delta + \gamma(1 + C_d u_2)} \quad (20)$$

where the notation is changed solely for convenience in this manuscript from ' ET_{SZ} ' in [42] to the potential ET notation ' ET_o '. The new terms are γ for the psychrometric 'constant', e_s for saturation vapor pressure, e_a for ambient vapor pressure, u_2 for 2-m wind speed, and numerator/denominator constants C_n and C_d accommodating different computational time steps and crop heights (citation Table 1 is partially reproduced here as Table 6). Computation of ET_o in Eq. 20 is done by following the procedures in [42] for hourly time steps.

To provide a basis for comparison of ET results between models, the FAO-56 parameters for wheat, cotton, and alfalfa were defined based on its documentation and common planting date practices in Central Arizona. The FAO-56 methodology is illustrated for wheat and cotton in Fig. 13). (but not for alfalfa: it is a multi-year crop with multiple and difficult-to-predict cutting times).

Transformation of vegetation indices to k_{cb} is done empirically based on replicated field experiments of different crops and reflectance measurements with accurately calibrated radiometers or imagers. Due to variations in sensor spatial resolution, spectral band placement, sensor calibration, atmospheric clarity, and soil moisture conditions, raw NDVI values are not consistent between experiments without a normalization procedure. For the calibrations used in this study, normalization re-scales original NDVI values to lie between 0 and 1:

$$ND_* = [NDVI - NDVI_l] / [NDVI_u - NDVI_l] \quad (21)$$

where TM5 NDVI values are re-scaled to new values (ND_*) that lie between empirically chosen minimum (L_l) and maximum (L_u) limits, roughly defined as bare-soil and full-cover.

For the three crops considered in this study the transformation equations used are polynomials as shown in Fig. 14.

For wheat a cubic equation [18,47] is used to accommodate significantly increased ET while NDVI approaches saturation:

$$k_{cb}[\text{wheat}] = \min \begin{cases} 0.15 \\ 1.75913 + 1.325341ND_* - 1.46574ND_*^2 + 1.146003ND_*^3 \end{cases} \quad (22)$$

For cotton, a quadratic was adequate for experimental results [17]:

$$k_{cb}[\text{cotton}] = \max \begin{cases} 0.15 \\ 0.2471 + 1.2012ND_* - 0.2183ND_*^2 \end{cases} \quad (23)$$

For alfalfa, a quadratic was also used (Hunsaker, personal communication, based on experiments reported in [45]):

$$k_{cb}[\text{alfalfa}] = \begin{cases} 0.3 & \text{if } ND_* < 0.25 \\ \min \begin{cases} 1.25 \\ 0.40925 - 0.86315ND_* + 1.97011ND_*^2 \end{cases} & \text{otherwise} \end{cases} \quad (24)$$

NDVI normalization is needed to ensure that the full range of crop coefficients can be represented while viewing a range of different background soil colors and moisture [18]. To avoid biasing results based on local conditions, this study adopted a quantile selection approach, where all plot-mean NDVI values were aggregated by crop type– Wheat, Cotton, Alfalfa– and their resulting distributions were modeled with Beta functions. The Beta function is specified by two shape parameters, by convention identified as α and β . Because the Beta distribution domain spans 0-1, NDVI values were transformed from a nominal range found for the NDVI 2008 CAIDD data set, 0.2 to 0.85, to 0.0 to 1.0. This transformation step simplifies the function fitting process. Using empirical histograms for each crop, Beta shape parameters were determined in multiple steps. Beta shape parameters were computed by first determining sample means and variances, and then estimating α and β terms via the method of moments (Eq. 25).

$$\begin{aligned} \alpha &= \bar{X} [(\bar{X}(1 - \bar{X}) / X_{var}) - 1] \\ \beta &= (1 - \bar{X}) [(\bar{X}(1 - \bar{X}) / X_{var}) - 1] \end{aligned} \quad (25)$$

These estimates initialized their final determination with the maximum likelihood method, implemented via the R 'fitdistr' function found in the MASS package. Quantiles at the 10% and 90% probability levels were then computed and transformed back to source NDVI equivalents. Selection of these levels was subjective but not arbitrarily chosen: probabilities spanning a larger range of 5% and 95% were found to produce anomalously low crop coefficient values for wheat and alfalfa, while probabilities spanning a lesser range fail to reproduce representative crop coefficients for sparse cover. The normalization procedure is illustrated for wheat crops (Fig. 15). Using instances with early January planting (green lines, Fig. 15A), an empirical histogram was generated (Fig. 15B), bins were re-scaled to lie between 0 and 1, a Beta distribution function was fit, quantiles at 10% and 90% probabilities identified, then indices transformed back to original NDVI space by inverting Eq. 21. After the inversion, the recomputed quantiles are respectively the values for $NDVI_l$ and $NDVI_u$.

Resulting NDVI normalization parameters are shown in Tab. 8.

CAIDD Wheat 2008

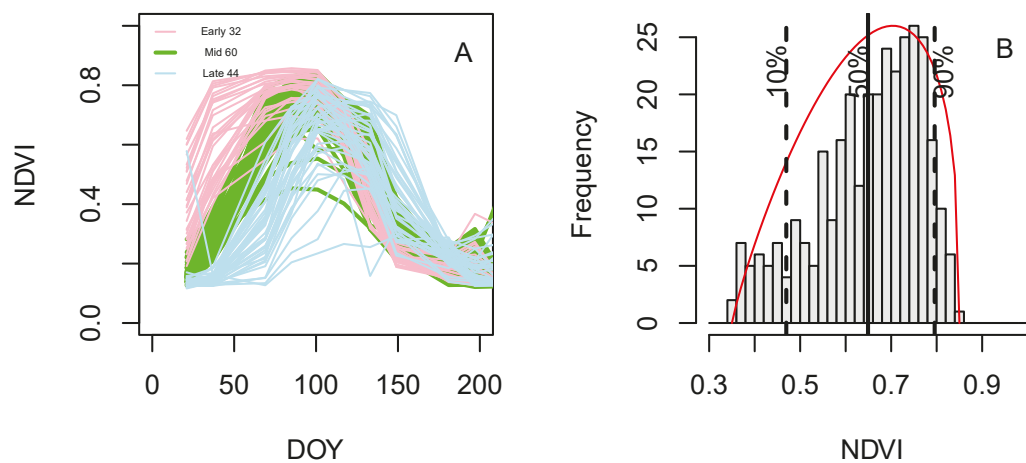


Figure 15. Wheat NDVI normalization procedures. To convert surface reflectance-based NDVI to normalized NDVI, the NDVI time series for each crop was assembled. For wheat (A), there were three different planting intervals; the normalization procedure considered the distribution from all intervals as one sample distribution. From the generated histogram, a maximum-likelihood Beta distribution (red curve) was fit based on selected end-members (B). Using the modeled NDVI distribution, upper and lower quantiles were selected, then applied all individual plot NDVI traces to create normalized NDVI values.

3.4. Vegetation Index Transformations

4. Conclusions

An ET study using remote sensing models was performed over irrigated farms in South Central Arizona and quantified daily and cumulative water use for wheat, cotton, and alfalfa. Multiple tests to evaluate input data quality were performed. Actions were taken to adjust biases, and standardized procedures establish to remove arbitrariness. Three models were implemented, two surface energy balance approaches, TSEB and METRIC, and one strictly empirical vegetation index approach, VISW. When compared with each other and against standardized estimates it was found that TSEB produced the highest, METRIC the lowest, and VISW middle level ET values. Differences had a seasonal dependence, where all models agreed within 18% for winter durum wheat plots, while differed up to 35% during summertime cotton and alfalfa trials.

Author Contributions: conceptualization, A.French, D. Hunsaker, L. Bounoua, and A. Karnieli; methodology, A.French, D. Hunsaker, L. Bounoua, and A. Karnieli; software, A.French; validation, A.French, D. Hunsaker; formal analysis, A.French, and D. Hunsaker; A.French, D. Hunsaker, L. Bounoua, and A. Karnieli; resources, A. French, D. Hunsaker, L. Bounoua, A. Karnieli, W. Lockett, and R. Strand; data curation, A.French and D. Hunsaker; writingoriginal draft preparation, A. French; writingreview and editing, A. French, D. Hunsaker, L. Bounoua, and A. Karnieli; visualization, A. French; supervision, A.French; project administration, A. French and A. Karnieli; funding acquisition, A.French, L. Bounoua, and A. Karnieli

Funding: “This research was funded in part by BARD project IS-448012 and by the USDA”.

Acknowledgments: The CAIDD kindly shared maps and water use data used in this study.

Conflicts of Interest: “The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results”.

Abbreviations

The following abbreviations are used in this manuscript:

ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AWiFS	Advanced Wide Field Sensor
AZMET	The Arizona Meteorological Network
CAIDD	Central Arizona Irrigation and Drainage District
ET	evapotranspiration
IRS-P6	Resourcesat-1
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LST	Land Surface Temperature
METRIC	Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration
MODIS	Moderated Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index
NOAA FSL	National Oceanic and Atmospheric Administration Forecast Systems Library
TIR	Thermal Infrared Radiation
TM5	Landsat 5 Thematic Mapper
TM7	Landsat 7 Thematic Mapper
TSEB	Two Source Energy Balance Model
USDA-SW	U.S. Department of Agriculture crop water use for Arizona, Conservation Report 29
USGS	United States Geological Survey
VISW	Vegetation Index Southwest
VNIR	Visible Near Infrared Radiation

- Dieter, C.; Maupin, M.; Caldwell, R.; Harris, M.; Ivahnenko, T.; Lovelace, J.; Barber, N.; Linsey, K. Estimated use of water in the United State in 2015. In *U.S. Geological Survey Circular*; Number 1441, 2018.

2. Zhan-Ming, C.; Chen, G. Virtual water accounting for the globalized world economy: National water footprint and international virtual water trade. *Ecological Indicators* **2013**, *28*, 142–149.
3. Hutson, S.; Barber, N.; Kenny, J.; Linsey, K.; Lumia, D.; Maupin, M. Estimated water use of water in the United States in 2000. In *U.S. Geological Survey Circular*; Number 1268, 2004.
4. Clemmens, A.; Allen, R.; Burt, C. Technical concepts related to conservation of irrigation and rainwater in agricultural systems. *Water Resour. Res.* **2008**, *44*, 1–16.
5. Howell, T. Enhancing water use efficiency in irrigated agriculture. *Agron. J.* **2001**, *93*, 281–289.
6. Melton, F.S.; Johnson, L.F.; Lund, C.P.; Pierce, L.L.; Michaelis, A.R.; Hiatt, S.H.; Guzman, A.; Adhikari, D.D.; Purdy, A.J.; Rosevelt, C.; Votava, P.; Trout, T.J.; Temesgen, B.; Frame, K.; Sheffner, E.J.; Nemani, R.R. Satellite Irrigation Management Support With the Terrestrial Observation and Prediction System: A Framework for Integration of Satellite and Surface Observations to Support Improvements in Agricultural Water Resource Management. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2012**, *5*, 1709–1721.
7. Huntington, J.; Hegewisch, K.; Daudert, B.; Morton, C.; Abatzoglou, J.; McEvoy, D.; Erickson, T. Computing of climate and remote sensing data for advanced natural resource monitoring and process understanding **2017**. 10.1175, 2397–2409.
8. Timmermans, W.; Kustas, W.; Anderson, M.; French, A. An intercomparison of the Surface Energy Balance Algorithm for Land (SEBAL) and the Two-Source Energy Balance (TSEB) modeling schemes. *Remote Sens. Environ.* **2007**, *108*, 369–384.
9. Gonzalez-Dugo, M.; Neale, C.; Mateos, L.; Kustas, W.; Prueger, J.; Anderson, M. A comparison of operational remote sensing-based models for estimating crop evapotranspiration. *Agric. For. Meteorol.* **2009**, *149*, 1843–1853.
10. French, A.; Hunsaker, D.; Thorp, K. Remote sensing of evapotranspiration over cotton using the TSEB and METRIC energy balance models. *Remote Sens. Environ.* **2015**, *158*, 281–294.
11. Gosling, S.; Bretherton, D.; Haines, K.; Arnell, N. Global hydrology modelling and uncertainty: running multiple ensembles with a campus grid. *Phil. Trans. R. Soc. A* **2010**, *368*, 4005–4021.
12. Schellekens, J.; Dutra, E.; Dutra, A.; Torre, A.M.; Balsamo, G.; v. Dijk, A.; Weiland, F.S.; Minvielle, M.; Calvet, J.C.; Decharme, B.; Eisner, S.; Fink, G.; Flörke, M.; Pessenteiner, S.; v. Beek, R.; Polcher, J.; Beck, H.; Orth, R.; Calton, B.; Burke, S.; Dorigo, W.; Weedon, G. A global water resources ensemble of hydrological models: the earthH2Observe Tier-1 dataset. *Earth Syst. Sci. Data* **2017**, *9*, 389–413.
13. Dirmeyer, P.; Gao, X.; Zhao, M.; Guo, Z.; Oki, T.; Hanasaki, N. GSWP-2: Multimodel analysis and implications for our perception of the land surface. *B. Am. Meteorol. Soc.* **2006**, *87*, 1381–1397.
14. Bounoua, L.; Imhoff, M.; Franks, S. Irrigation requirement estimation using vegetation indices and inverse biophysical modeling. Geoscience and Remote Sensing Symposium (IGARSS), 2010, pp. 1823–1826.
15. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. Crop evapotranspiration, guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, Food and Agriculture Organization of the United Nations, 1998.
16. Norman, J.; Kustas, W.; Humes, K. A two-source approach for estimating soil and vegetation energy fluxes from observations of directional radiometric surface temperature. *Agric. For. Meteorol.* **1995**, *77*, 263–293.
17. Hunsaker, D.J.; Barnes, E.M.; Clarke, T.R.; Fitzgerald, G.J.; Paul J. Pinter, J. Cotton irrigation scheduling using remotely-sensed and FAO-56 basal crop coefficients. *Transactions of the ASAE* **2005**, *48*, 1395–1407.
18. Hunsaker, D.J.; Fitzgerald, G.J.; French, A.N.; Clarke, T.R.; Paul J. Pinter, J. Wheat irrigation management utilizing multispectral crop coefficients: II. Irrigation scheduling performance, grain yield, and water use efficiency. *Transactions of the ASABE* **2007**, *50*, 2035–2050.
19. Brown, P. Accessing the Arizona Meteorological Network (AZMET) by computer. Ext. repo no. 8733, University of Arizona, Tucson, AZ, 1989.
20. Erie, L.; French, O.; Bucks, D.; Harris, K. Consumptive use of water major crops in the southwestern United States. Technical Report Conservation Research Report Number 29, USDA, Agricultural Research Service, 1982.
21. Husman, S.; Ottman, M. Irrigation of small grains in Arizona. In *The University of Arizona Cooperative Extension*; Number AZ1345, 2015.

22. Thorp, K.; Hunsaker, D.; Bronson, K.; Andrade-Sanchez, P.; Barnes, E. Cotton irrigation scheduling using a crop growth model and FAO-56 methods: field and simulation studies. *J. ASABE* **2017**, *60*, 2023–2039.
23. Bavel, C.V. Changes in canopy resistance to water loss from alfalfa induced by soil water depletion. *Agr. Meteorol.* **1967**, *4*, 165–176.
24. Ottoni, T.; Matthias, A.; Warrick, A. Field estimates of alfalfa evapotranspiration by energy balance-resistance modelling. *Theor. Appl. Climatol.* **1987**, *38*, 15–23.
25. Hunsaker, D.; P.J. Pinter, J. Measured and calculated evapotranspiration of alfalfa in weighing lysimeters. ASAE Annual International Meeting, 2000, pp. 1755–1770.
26. Saz, A.; Bawazir, S.; Samani, Z.; Skaggs, R. Alfalfa evapotranspiration in Albuquerque's South Valley. In *New Mexico State University Research Report*; Number 787, 2014.
27. Semmens, K.; Anderson, M.; Kustas, W.; Gao, F.; Alfieri, J.; McKee, L.; Prueger, J.; Hain, C.; Cammalleri, C.; Yang, Y.; Xia, T.; Sanchez, L. Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach. *Remote Sens. Environ.* **2016**, *185*, 155–170.
28. Song, L.; Liu, S.; Kustas, W.; Zhou, J.; Xu, Z.; Xia, T.; Li, M. Application of remote sensing-base two-source energy balance model for mapping field surface fluxes with composite and component surface temperatures. *Agric. For. Meteorol.* **2016**, *230–231*, 8–19.
29. Wagle, P.; Bhattarai, N.; Gowda, P.H.; Kakani, V. Performance of five surface energy balance models for estimating daily evapotranspiration in high biomass sorghum. *ISPRS J. Photogrammetry and Remote Sensing* **2017**, *128*, 192–203.
30. of Arizona, T.U. AZMET: The Arizona Meteorological Network. <https://cals.arizona.edu/azmet/>.
31. United States Department of Agriculture (USDA), National Agricultural Statistics Service (NASS), R.; Development Division (RDD), Geospatial Information Branch (GIB), S.A.R.S.S. 2008 Arizona Cropland Data Layer | NASS/USDA. [USDA, NASS Marketing and Information Services Office, Washington, D.C. https://www.nass.usda.gov/Research_and_Science/Cropland/metadata/metadata_az08.htm](https://www.nass.usda.gov/Research_and_Science/Cropland/metadata/metadata_az08.htm), 2008.
32. U.S. Geological Survey. *Product Guide, Landsat 4-7 Surface Reflectance (LEDAPS) Product*, version 8.3 ed., 2018.
33. National Oceanic and Atmospheric Administration-Earth System Research Laboratory. <https://ruc.noaa.gov/raobs>, 2018. archived 12-hourly radiosonde data.
34. Berk, A.; Bernstein, L.; Anderson, G.; Acharya, P.; Robertson, D.; Chetwynd, J.; Adler-Golden, S. MODTRAN cloud and multiple scattering upgrade with application to AVIRIS. *Remote Sens. Environ.* **1998**, *65*, 367–375.
35. Abrams, M. The Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER): Data products for the high spatial resolution imager on NASA's Terra platform. *International Journal of Remote Sensing* **2000**, *21*, 847–859.
36. Kahle, A.B.; Alley, R.E. Separation of temperature and emittance in remotely sensed radiance measurements. *Remote Sens. Environ.* **1992**, *42*, 107–111.
37. Hulley, G.C.; Hughes, C.G.; Hook, S.J. Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and MODIS thermal infrared data. *Journal of Geophysical Research: Atmospheres* **2012**, *117*. doi:10.1029/2012JD018506.
38. MODIS reprojection tool, user's manual. https://lpdaac.usgs.gov/sites/default/files/public/mrt41_usermanual_032811.pdf, release 4.1 ed., 2011.
39. Allen, R.; Tasumi, M.; Trezza, R. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-model **2007**. *133*, 380–394.
40. Tasumi, M.; Allen, R.G.; Trezza, R. Estimation of at-surface reflectance and albedo from satellite for routine, operational calculation of land surface energy balance. *J. Hydro. Eng.* **2008**, *13*, 51–63.
41. Bastiaanssen, W. SEBAL-based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey. *J. Hydrol.* **2000**, *229*, 87–100.
42. on Standardization of Reference Evapotranspiration, T.C. The ASCE Standardized Reference Evapotranspiration Equation. Technical report, Environmental and Water Resources Institute, 2005.
43. He, R.; Jin, Y.; Kandelous, M.M.; Zaccaria, D.; Sanden, B.L.; Snyder, R.L.; Jiang, J.; Hopmans, J.W. Evapotranspiration Estimate over an Almond Orchard Using Landsat Satellite Observations. *Remote Sensing* **2017**, *9*.

- 532 44. Czajkowski, K.; Goward, S.; Ouaidrari, H. Impact of AVHRR filter functions on surface temperature
533 estimation from the split window approach. *Int. J. Remote Sens.* **1998**, *19*, 2007–2012.
- 534 45. Hunsaker, D.; P.J. Pinter, J.; Cai, H. Alfalfa basal crop coefficients for the FAO-56 procedures in the desert
535 regions of the southwestern U.S. *Trans. ASAE* **2002**, *45*, 1799–1815.
- 536 46. Hunsaker, D.; Jr., P.P.; Fitzgerald, G.; Clarke, T.; Kimball, B.; Barnes, E. Tracking spatial and temporal
537 cotton ET patterns with a normalized difference vegetation index. Irrigation association exposition and
538 technical conference, San Diego, CA, Nov. 18-20, 2003, pp. 126–137.
- 539 47. Hunsaker, D.; P.J. Pinter, J.; Kimball, B. Wheat basal crop coefficients determined by normalized difference
540 vegetation index. *Irrig. Sci.* **2005**, *24*, 1–14.
- 541 48. Wright, J. New evapotranspiration crop coefficients. *J. Irrig. Drain. Div., ASCE* **1982**, *108*, 57–74.
- 542 49. Heilman, J.; Heilman, W.; Moore, D. Evaluating the crop coefficient using spectral reflectance. *Agron. J.*
543 **1982**, *74*, 967–971.
- 544 50. Bausch, W. Remote sensing of crop coefficients for improving the irrigation scheduling of corn. *Agric.*
545 *Water Mgmt.* **1995**, *27*, 55–68.
- 546 51. Neale, C.; Jayanthi, J.; Wright, J. Crop and irrigation water management using high-resolution airborne
547 remote sensing. Proc. ICID Workshop Remote Sensing of ET for Large Regions; International Commission
548 on Irrigation and Drainage, , 2003. CD-ROM.
- 549 52. Bronson, K.; White, J.; Conley, M.M.; Hunsaker, D.; Thorp, K.; French, A.; Mackey, B.; Holland, K. Active
550 optical sensors in irrigated Durum wheat: nitrogen and water effects. *Agronomy J.* **2017**, *109*, 1060–1071.
- 551 53. Doorenbos, J.; Pruitt, W. *Crop Water Requirements*; Number FAO Irrigation and Drainage Paper 24, Food
552 and Agriculture Organization of the UN: Rome, Italy, 1977.
- 553 54. Paul J. Pinter, J.; Hatfield, J.L.; Schepers, J.S.; Barnes, E.M.; Moran, M.S.; Daughtry, C.S.; Upchurch, D.R.
554 Remote Sensing for Crop Management. *Photogrammetric Engineering and Remote Sensing* **2003**, *69*, 647–664.
- 555 55. Johnson, L.; Scholasch, T. Remote sensing of shaded areas in vineyards. *Hort Tech.* **2005**, *15*, 859–863.

Table 4. Landsat 5 Data

Index	Date	DOY	Quality	Clouds	Sky Cover	Solar Elev.	Solar Azi.	Tair	RH	U	Rsolar	Rel. Rsolar
1	1/21/2008 10:49	21	9	-	51	31.20	150.60	11.03	38.97	1.12	1.567	0.795
2	2/6/2008 10:48	37	9	-	1	34.60	147.50	6.50	76.36	0.70	1.738	0.813
3	2/22/2008 10:48	53	9	-	94	39.40	144.20	13.98	62.22	1.84	1.208	0.509
4	3/9/2008 10:48	69	9	-	2	45.00	140.80	16.54	28.72	0.76	2.084	0.793
5	3/25/2008 10:48	85	9	-	0	50.80	137.00	22.84	27.04	0.56	2.438	0.849
6	4/10/2008 10:46	101	9	-	0	56.30	132.20	18.70	27.19	4.23	2.845	0.929
7	4/26/2008 10:46	117	9	-	1	61.00	126.30	24.60	17.35	0.42	3.047	0.952
8	5/12/2008 10:46	133	9	-	0	64.30	119.30	27.41	18.28	3.62	2.735	0.835
9	5/28/2008 10:46	149	9	-	0	65.90	112.70	27.46	19.69	1.00	3.043	0.923
10	6/13/2008 10:48	165	NA	-				33.52	13.36	1.52	3.182	0.964
11	6/29/2008 10:46	181	9	-	7	65.40	107.30	36.92	16.31	1.71	2.971	0.909
12	7/15/2008 10:45	197	9	-	10	63.90	110.10	33.48	38.58	1.50	2.723	0.844
13	7/31/2008 10:43	213	9	-	2	61.80	115.90	36.09	21.95	0.43	2.693	0.849
14	8/16/2008 10:43	229	9	-	2	59.10	123.50	31.72	49.84	0.98	2.711	0.873
15	9/1/2008 10:43	245	9	-	19	55.70	131.70	30.30	51.79	0.96	2.663	0.887
16	9/17/2008 10:43	261	9	-	0	51.50	139.40	30.99	33.60	1.93	2.553	0.893
17	10/3/2008 10:43	277	9	-	65	46.80	145.80	30.16	31.58	2.52	2.169	0.811
18	10/19/2008 10:41	293	9	-	0	41.80	150.60	27.66	24.53	1.24	2.196	0.895
19	11/4/2008 10:41	309	9	-	49	37.00	153.60	21.82	40.40	2.83	1.946	0.871
20	11/20/2008 10:40	325	9	-	0	32.90	155.10	18.00	25.20	1.11	1.818	0.895
21	12/6/2008 10:40	341	9	-	0	30.00	155.10	12.89	54.38	1.10	1.648	0.874
22	12/22/2008 10:41	357	9	-	51	28.50	153.70	8.93	69.78	0.84	1.476	0.810

Table 5. Nominal plant heights for TSEB

Crop	Planting	Mid-Season	Height at Mid-Season
Wheat	January 1	March 15	0.5
Cotton	April 15	July 15	0.5
Alfalfa	-	-	0.5 (before cutting)

Table 6. Constants for ASCE Standardized ET computation

Time Step	Short Crop		Tall Crop	
	C_n	C_d	C_n	C_d
Daily	900	0.34	1600	0.38
Hourly, daytime	37	0.24	66	0.25
Hourly, nighttime	37	0.96	66	1.7

Table 7. FAO-56 crop parameters

Crop		DOP	Init	Dev	Mid	Late	End
Wheat	Devel. Days	0	25	48	62	17	-
	k_{cb}	-	0.15	-	1.17	-	0.22
Cotton	Devel. Days	106	31	51	50	60	-
	k_{cb}	-	0.15	-	1.20	-	0.52

Table 8. Crop normalization parameters

Crop	Beta Dist.		NDVI Limits		k_{cb} Limits	
	α	β	Lower	Upper	Lower	Upper
Wheat	0.58	1.26	0.06	0.65	0.15	1.25
Cotton	0.42	0.83	0.10	0.74	0.15	1.25
Alfalfa	2.20	1.89	0.27	0.72	0.3	1.25