Automated mapping of evapotranspiration from irrigated crops using global climate grids and MODIS data: Comparison of energy- and vegetation-based methods with observation towers

Trent W. Biggs

Michael Marshall

Alex Messina

**Abstract**

Automated methods to map evapotranspiration (ET) at regional and global scales often use vegetation indices to estimate ET, and have been validated mostly over rainfed areas. Here, two automated methods for estimating ET, one vegetation-based (MOD16) and one temperature-based (surface energy balance algorithm, SEBAL) are compared over irrigated areas using only global climate grids (GMAO-MERRA) and remote sensing data (MODIS) as input, validated with eddy flux correlation towers over several irrigated crops (maize, rice, cotton) in the Central Valley of California. Automated selection of calibration pixels in SEBAL was based on quantiles of the cumulative distribution of land surface temperature (TR) and normalized difference vegetation index (NDVI). For maize, seasonal MOD16 ET had lower relative mean bias (rMB 0-10%) compared to SEBAL ET (rMB 14-17%) but MOD16 significantly underestimated ET from rice (rMB -50%) and cotton fields (-78%), with the largest error in the early stages of the growing season when NDVI is low but land surface wetness index (LSWI) is high. SEBAL also underestimated ET from rice (rMB -24%) and cotton (-28%) but with lower rMB than MOD16. Error in SEBAL evaporative fraction (Λ=ET/Rn), which controls for errors in Rn, was lower for rice and cotton but higher for maize due to underestimation of Rn in SEBAL. Field-measured crop height, rather than fractional vegetation cover (FVC) or aboveground wet biomass (AWB), explained most of the variation in error among crops in different growth stages, though the error also correlated with the evaporative fraction. Overall, MOD16 performed poorly over short crops with high Λ (cotton, rice, H<1m) but outperformed SEBAL over tall crops with lower Λ (maize, H>2m). Contrary to expectation, MOD16 error did not correlate with FVC, suggesting that the error in MOD16 ET was not only due to underestimation of soil evaporation in early growth stages but also to underestimation of transpiration in mid and late parts of the growing season. SEBAL ET was sensitive to the size of the domain, the quantile used to select calibration pixels, and the parameters of the model relating NDVI to surface roughness length (z0m). We conclude that automated vegetation-based methods like MOD16 may underestimate ET from short irrigated crops with high Λ, particularly early in the growing season, while SEBAL may have difficulty in areas where the NDVI-z0m relationship varies with crop type. The global product MOD16 requires modification before being used in estimation of ET from irrigated croplands, and automated temperature-based methods like SEBAL can provide accurate estimates over irrigated crops using only global climate grids and remote sensing data. Regional and global products and assessments of crop water use could incorporate different methods depending on the land cover and crop type.

**Introduction**

Regional and global estimates of evapotranspiration (ET) are important for understanding water and energy balances, and operational estimates of ET at large geographic scales are useful for water resources management, irrigation timing, and drought impact assessment. Satellite-based methods have been developed that can estimate ET at ~1km spatial resolution with regional and global coverage. The methods include empirical relationships between ET and vegetation indices (Glenn et al., 2010), process-based models driven primarily by vegetation indices (Fisher et al., 2008; Mu et al., 2011b), and models based on surface energy balances (SEB), including one-source (Allen et al., 2007; Bastiaanssen et al., 1998; Su, 2002), and two-source models (Anderson et al., 2013), both of which use thermal imagery to estimate radiometric land surface temperature (TR) and the sensible and latent heat flux. SEB models depend primarily on TR to estimate land surface fluxes for a given amount of net radiation, so here we call them temperature-based methods.

Automated and operational global ET products based on satellite data and global climate grids are available at 1 km spatial and 8-day or monthly temporal resolution, including MOD16 (Mu et al., 2011b) and the PT-JPL model (Fisher et al., 2008). Both of these models are based primarily on vegetation indices, which are used to estimate parameters in either the Penman-Monteith equation (Mu et al., 2011b) or the Priestley-Taylor equation (Fisher et al., 2008). The meteorological data needed, including air temperature, relative humidity, and net radiation, are derived from global climate grids (e.g. Global Modeling and Assimilation Office, GMAO), and no ground-level meteorological data is needed. Both methods have been evaluated primarily over rainfed ecosystems. Less well quantified is their performance in irrigated agricultural land, which may have different relationships between vegetation indices and ET. In particular, the soil evaporation component in both MOD16 and PT-JPL is based on relative humidity (h). In irrigated areas in semi-arid environments, soil evaporation may be high over a particular field when vegetation cover is low at the beginning of the cropping season, despite low h in the corresponding global climate grid cell. Temperature-based models should perform well under such conditions, since they are less dependent on a close correlation between ET and a vegetation index.

Temperature-based methods may be cumbersome to apply on seasonal time scales, since they can require manual calibration to pixels in the imagery. Automated methods can be used to select calibration pixels in SEBAL and the closely related METRIC model, with similar performance to manual calibration (Morton et al., 2013). For regional and global application, global gridded climate datasets would replace the need for ground-based meteorological measurements, as has been done for MOD16. There has not been a systematic comparison of automated temperature-based methods with the most widely-used global ET products that use vegetation-based models (MOD16 and PT-JPL), particularly over irrigated areas. Such a comparison would also test for calibration uncertainty in SEBAL due to pixel selection, model parameters, and the domain over which the model is implemented (Long et al., 2011). Model domain can be important if there are regional variations in the relationship between the radiometric land surface temperature (TR) and the sensible heat flux.

The interpretation of seasonal patterns in ET and model errors is enhanced by incorporating other spectral information about vegetation and land surface wetness. Crop growth stages, including periods of inundation and transplanting, can be detected using a combination of vegetation indices and other indices that are sensitive to both soil and vegetation water content, such as the land surface water index (LSWI) (Xiao et al., 2005). Such indices, combined with ancillary data on crop biomass, fractional vegetation cover, and soil moisture, can be used to identify conditions under which different ET mapping methods perform well or need further improvement.

The objectives of this paper are to implement an automated temperature-based model of ET (SEBAL) using only global climate grids and remote sensing (MODIS) data as input, and to compare SEBAL ET with ET from both ET measured at eddy flux correlation towers and the global MOD16 dataset for different irrigated crops. The primary research questions are:

1. How do weekly and seasonal estimates of ET from MOD16 and SEBAL compare in irrigated areas?

2. Do the differences in ET estimated using MOD16 and SEBAL relate to growth stage of the crop and/or land surface wetness?

3. What are the implications for automated regional and global monitoring of ET from irrigated croplands?

**Study area**

The study area includes 6 eddy flux correlation towers in the Central Valley of California, United States (*Figure 1*, Table 1). Crops at the tower sites included rice (N=3), maize (N=2), and cotton (N=1). Data was available for different time periods from 2009 to 2013, with the most availability in 2011 and 2012 (Table 1). One rice site (Us.Twt) had hourly water level recordings, and soil moisture was available at the maize sites.

All 6 towers are bounded by a study area boundary that is 2.5 degrees longitude by 4 degrees latitude (*Figure 1*). The modeling domain was further restricted to elevations that covered the range of the tower elevations (-20 m to +100 m), which reduced the impact of adiabatic lapse rates on land surface temperature and ET. The study area can be divided into northern and southern sections defined by the locations of the towers (*Figure 1*). The southern section was further divided into a smaller section to compare the effect of different domain sizes on SEBAL ET (see section xx below). These various domains were used to test the impact of domain size on SEBAL ET. The vegetation around each of the towers is uniform (Appendix Table xx?), making it unnecessary to calculate the footprint.

The cropping season begins in April-May and ends in September. Data on irrigation application dates and water level were available for some sites (Table 1).

**Methods**

*Flux tower data: Energy balance closure*

Two techniques were used to measure LE at the towers, either singly or in combination: surface renewal (SR) (French et al., 2012) and eddy-covariance (EC) (Dabberdt et al., 1993; Twine et al., 2000). SR calculates LE as the residual of the energy balance measured at the tower (LE=Rn-G-H). SR uses information in high frequency temperature data to model the vertical H flux. SR may be applied in both unstable and stable conditions, and does not require measurements of wind speed (French et al., 2012). EC measures all components of the energy balance separately using vertical velocity fluctuations and concentrations (Dabberdt et al., 1993).

In the EC method, the sum of LE and H may not equal available energy (Rn-G) if the key assumptions are not met, such as zero net vertical advective flux (Twine et al., 2000). Closure is achieved by adjusting values of LE and/or H so that LE+H equals Rn-G. Energy balance closure can be achieved using either the latent energy closure method or the energy-balance-Bowen-ratio (EBBR) method (Twine et al., 2000). The latent energy closure method or “residual-LE closure” method assumes that all missing energy is latent heat. The Bowen ratio method assumes that the ratio of sensible heat (H) to latent heat (LE) is constant between the measured values and the actual, corrected values.

Both SR and EC were performed at three towers, EC alone at four towers, and SR alone at two towers (Table 1). For towers with EC, the latent energy closure method was used to close the energy balance. Where both SR and EC were performed, the difference between mean LE was very small (<1%), suggesting that LE values from SR and EC are comparable.

*Surface energy balance algorithm*

The surface energy balance algorithm (SEBAL) calculates an evaporative fraction at the moment of satellite overpass as the residual in the surface energy balance:

|  |  |
| --- | --- |
|  | ( 1 ) |

where Rn is net radiation, G is the soil heat flux, and H is the sensible heat flux, all at the instant of satellite overpass. G is a function of vegetation cover (NDVI) and albedo. Rn is calculated as:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  | | --- | --- | --- | --- | | |  |  | | --- | --- | |  | ( 2 ) | |  | |  |  | |  |

## where *α* is broadband blue-sky albedo (dimensionless), *SW↓* is incoming shortwave radiation (W/m2), *LW↓* is downwelling longwave radiation and *LW↑* is upwelling longwave radiation (W/m2) at the surface. Instantaneous H is calculated as:

|  |  |
| --- | --- |
|  | ( 3) |

where *ρair*is the density of air (kg/m3), *Cp* is the specific heat capacity of air (J/kg/K), a and b are empirical coefficients determined for each image, TR is the radiometric surface temperature (also called the land surface temperature (LST)) from the daily product MOD11A1 v5, and Rah is the aerodynamic resistance to turbulent heat transport from z1 to z2.

Determining values for the parameters a, b and Rah in ( 3) requires identifying pixels where H and TR are known. One pixel is selected that is “wet”, where H is assumed to be 0 and λE is assumed equal to Rn-G, and another pixel is selected that is “dry”, where H is assumed to be Rn-G and λE is assumed to be 0. Initial values of a and b are based on the Rn and TR at the wet and dry pixels. H is then calculated again using ( 3), this time accounting for unstable atmospheric conditions using the Monin-Obukhov (MO) equations (Allen et al., 2011; Bastiaanssen et al., 2002, 1998). The values of a, b, and Rah are then solved iteratively by updating the values of each until the result converges on H=0 for the wet pixel and H=Rn-G for the dry pixel.

A stability-corrected value of Rah is calculated as:

|  |  |
| --- | --- |
|  | ( 4 ) |

Where z1 and z2 are the heights of the evaporating surface and the height of the wind speed measurements, u\* is the friction velocity, k is a constant (0.41) and is the Monin-Obukhov stability correction for heat transport for stable conditions.

u\* is estimated as: (Eq 10.8 Morse):

|  |  |
| --- | --- |
|  | ( 5 ) |

where U200 is wind speed at the blending height (200m), estimated from observed wind speed at some height (2-3m) and an assumed logarithmic wind speed profile. The roughness length for momentum transport (z0m) is assumed to be a fixed fraction of crop height (0.123h), and is either derived from a land use map or modelled as a function of a vegetation index (Allen et al., 2007; Bastiaanssen, 2000; Morse et al., 2000):

|  |  |
| --- | --- |
|  | ( 6 ) |

where c and d are coefficients calibrated to a given region. Since land use maps showing detailed crop growth stages and type are often not available for regional applications, we used ( 6 ) to estimate z0m. We calibrated the values of c and d to crop height data collected in the field. Since several pairs of c and d values were possible and yielded different errors for different crops, we ran SEBAL with four different parameter sets of c and d, some of which were calibrated to short crops like rice, and others to tall crops like maize in different growth stages.

ET over a 24-hour period (ET24) is then calculated as ΛRn24, as G is assumed zero on a daily time step. This assumes that Λ is constant over a 24-hour period. This may not be the case in some situations (Van Niel et al., 2011), but is a common assumption in applications of SEBAL.

*GMAO-MERRA data*

## Following the MOD16 algorithm, *SW↓*, *LW↓* and *LW↑* for SEBAL were taken from the Global Modeling and Assimilation Office (GMAO) Modern-Era Retrospective Analysis for Research and Applications for Land (MERRA-Land, dataset Shortname MAT1NXRAD, <http://disc.sci.gsfc.nasa.gov/daac-bin/DataHoldings.pl>, access date Sept 11, 2013) (Reichle et al., 2011). The data includes hourly average *SW↓*, *LW↓* and *LW↑* for each day of the study period, with spatial resolution of 0.5 latitude by 0.67 degrees longitude. *SW↓*, *LW↓* and *LW↑* were extracted for the time of satellite overpass (10:30am local standard time). Daily values of *SW↓*, *LW↓* and *LW↑* were taken as the daily average of the hourly values. α was taken from the MODIS albedo product (MCD43A).

## Also following MOD16, wind speed at 2 m above the displacement height at the time of overpass (10:30am local time) was taken from the GMAO-MERRA dataset (MAT1NXSLV). All gridded MERRA data were downscaled to the resolution of the MODIS data (1 km) using bilinear interpolation on a 2x2 grid using the resample algorithm in the statistical package R. Similar downscaling methods are used in MOD16 (Mu et al., 2011a).

*Automation of SEBAL*

Wet and dry pixels are often chosen manually given user knowledge of the area. In order to produce a time series of daily ET estimates to calculate seasonal ET, we use an automated procedure based on quantiles (q) of the observed TR and NDVI data to select the wet and dry pixels. For the wet pixel, a subset is generated where TR is lower than a given percentile (e.g. q=0.005, q=0.01), and the final wet pixel is the one with the highest NDVI within that subset. Use of NDVI assumes that vegetated surfaces better represent wet pixel conditions than water bodies. For the dry pixel, a subset is generated where TR is higher than a given percentile (e.g 0.005, 0.01), and the final dry pixel is has the lowest NDVI in the subset. The impact of different quantiles for selecting wet and dry pixels was tested by using q=0.005 and q=0.01. An additional run with a very low percentile (q=10-6) was used to select the pixels with the minimum and maximum TR as the wet and dry pixels.

Several previous implementations of SEBAL have excluded different land cover types from the pixel selection and model implementation, including urban areas (Conrad et al., 2007) and water bodies (Morse et al., 2000). In our application of SEBAL, land cover data (MCD12Q1) was used to remove both urban areas and water.

The domain of implementation of SEBAL can also have important impacts on the resulting ET values. If a domain includes a large area, Rn-G and temperature may vary with latitude, causing underestimation of ET in locations with high Rn-G, since the relationship between TR and dT is assumed constant over the image. In order to test the impact of study area size, we first implemented SEBAL over the area that includes all towers and elevations between -20 and +100 m (*Figure 1*). Then we performed SEBAL over just the northern and southern sections separately. An additional subset of the southern section that covers 1 degree latitude by 1 degree longitude was created, hereafter referred to as the “small southern” domain, due in part to an observed discontinuity in TR values at 37º N.

*Comparison of SEBAL and MOD16 with tower data*

ET estimated by both SEBAL and MOD16 was compared with ET estimated at the towers. We used both a 1x1 pixel window and a 3x3 pixel window. A 3x3 pixel window has been used in several other studies, including the primary validation study for MOD16 (Mu et al., 2011b) to compensate for any issues with georeferencing of imagery, location of the tower near boundaries between two or more pixels, or heterogeneity of the land surface near the tower.

Errors in ET can be due to either errors in Rn or errors in Λ. In order to remove the impact of errors in Rn, we also compared Λ from SEBAL and Kc from MOD16 with Λ and/or Kc from the towers. Since hourly data were not available at most towers, the Kc for MOD16 and Λ for SEBAL were calculated using 24-hour averages.

*Vegetation and wetness indices*

Vegetation and wetness indices were used to interpret the differences between MOD16 and SEBAL estimates of ET over the growing season. The enhanced vegetation index (EVI) from 16-day MOD13A2 were used as a proxy for vegetation growth. EVI was used instead of NDVI because NDVI saturates over dense vegetation. The Land Surface Water Index (LSWI) was used as a proxy for moisture in both vegetation and soil (Xiao et al., 2005):

|  |  |
| --- | --- |
|  | ( 7 ) |

where ρ is reflectance and nir1 and swir2 indicate near infrared (841-876 nm, MODIS band 2) and shortwave infrared (2105-2155 nm, MODIS band 7). The net difference water index (NDWI) (X Xiao et al., 2002) substitutes ρnir2 (1230-1250 nm, MODIS band 5) for ρswir2. LSWI and NDWI correlate with NDVI, EVI, and vegetation cover, but LSWI and NDWI are also high over wet surfaces, such as rice paddy during the flooding and transplanting stages (Xiangming Xiao et al., 2002).

*Field data*

Crop height, aboveground wet biomass (ABW) and fractional crop cover were measured in the field for three dates corresponding to sprouting, flowering, and grain- or bud-filling stages.

\*\* More detail here from Michael? \*\*

*Irrigation practices*

Flooding and planting of rice fields occur in May. The rice site with water level data (US-Twt) is flooded for approximately one month prior to planting, and then drained and planted in late April (2011) or mid-May (2012). During planting, the field is drained. The field is flooded again within a week of planting. Flooded conditions are maintained until mid-(2011) or end-September (2012).

Maize is planted in April, but not irrigated until July. One of the maize sites (StaW) was kept saturated at the surface throughout the growing season, while the second maize site (StaD) had a water table below the ground surface (more here ????). Cotton fields are planted in June and irrigated several times during the growing season.

**Results**

*Energy balance closure*

At the one tower where energy balance was not closed (US.Twt), the slope of the OLR line between Rn-G and H+LE was 0.58 with an intercept of 33.2 W m-2 and an R2 of 0.76. Compared to OLR statistics from other FLUXNET sites (Wilson et al., 2002), the slope (0.58) is within the lower ranges of other sites (0.55-0.99, mean 0.79), the intercept (33.2) is in the upper ranges of other sites (-32.9 to 32.6, mean 3.7), and the R2 is within the range (0.64-0.96) but lower than the average (0.86). The EBR was 1.01 for 2010, 0.82 in 2011, and 0.85 in 2012 for a three-year mean of 0.89, which is slightly higher than the mean of other FLUXNET sites (0.84). The OLR slope for the daily means where Rn-G >0 was 0.64 with an intercept of 25.5 W m-2 and an R2 of 0.74. A few outlier values of Rn-G (>300 W m-2) influenced the OLR slope; removal of those outliers improved the slope to 0.73 and an intercept of 17.8 W m-2, R2 of 0.81. Overall, OLR indicates that the data from Us-Twt are within the ranges reported by other FLUXNET sites and is acceptable for analysis.

*Net radiation and reference ET*

Rn estimated by GMAO-MERRA was lower than Rn from the towers for all towers and years except Us.Twt (*Figure 2*, Table 2). The relative mean bias ranged from -0.30 at Wil to +0.04 at Us.Twt in 2011 (Table 2). The relative RMSE (rRMSE) ranged between 0.23 and 0.55. Errors were highest in spring and early summer; excluding May and June decreased the rRMSE at Us.Twt from 0.26-0.55 to 0.18.

Reference ET (ETo) from MOD16 was very close to ETo estimated at the flux towers (Table 3). The maximum error in seasonal ETo was 15%, with an average bias of 5% and mean absolute error of 7%.

*Roughness length functions*

There was no optimal pair of values of c and d from the NDVI-z0m relationship; rice and maize had very different heights and z0m for a given NDVI value (*Figure 3*). In order to test the effect of using different parameter sets in (6), we fit c and d to different crops and growth stages. Set 1 was fit to rice, set 2 to sprouting maize, set 3 to senescent maize, and set 4 to a combination of senescent rice, sprouting and flowering cotton, and sprouting and flowering maize (*Figure 3*). Formal parameter estimation minimizing least squares yielded equations that did not fit targeted features in the NDVI-z0m relationship, in particular maize, so we used manual parameter adjustment and visual inspection to provide a range of NDVI-z0m relationships.

*Automated SEBAL: Pixel selection thresholds and domain effects*

Pixels selected with the automated technique occurred at the extremes of the NDVI-TR space. In the beginning of the growing season, the wet pixel sometimes had low NDVI, likely due to low vegetation cover but high soil moisture and high ET. In the middle of the growing season (xxx July, Jday 2012, 181 layer 10), the wet pixel typically had high NDVI, indicating that the wet pixel was of a vegetated surface. The automation of SEBAL resulted in high temporal variability in ET, particularly where clouds covered some fraction of the scene. Taking the 8-day moving average or seasonal averages smoothed out this variability and suggests that long time series of ET are feasible using automated pixel selection.

?? Plots of NDVI vs LST and pixels?

*SEBAL compared to tower data: Λ and ET*

ET and Λ from SEBAL were higher than observed at the towers for the maize sites, but lower than observed at the towers for rice and cotton sites for all model runs (Figure 4, Table 4, Table 5). Error in Λ was smaller than error in ET for the rice and cotton sites but larger than error in ET for maize, reflecting the underestimation of Rn by the GMAO data.

The use of a two smaller domains reduced error in ET and Λ from SEBAL compared to using one large domain (*Figure 1*, Table 4, Table 5), with the largest improvements at the cotton site. Our objective is not to perform a systematic evaluation of the effect of domain size, but rather to test the impact of domain size on the comparison of SEBAL with tower and MOD16 ET at our particular tower sites. Implementation of SEBAL over the single large domain results in the selection of calibration pixels that do not accurately represent the relationship between TR and H over the whole domain. Future work could determine the critical maximum domain size that can be used for SEBAL with different accuracies.

The use of different percentiles (q) for selecting wet and dry pixels also affected ET and Λ values (Table 4, Table 5). Using the maximum TR in the image (q=10-5) increased ET estimates by 0-19% compared to ET estimated using q=0.01. There was no clear optimal value of q; the lowest errors in ET were obtained for q=0.01 for rice, but for q=10-5 for maize. A value q=0.005 kept all errors in mean seasonal ET less than or equal to 33%. The error in Λ was smaller (larger) than the error in ET at the rice and cotton (maize) sites due to the underestimation of Rn in the GMAO-MERRA data at most towers.

The parameters of the NDVI-z0m relationship affected the error in SEBAL ET and Λ, but there was no optimal value that minimized the error for all crops (Table 4, Table 5). Parameters that estimate low z0m values for a given NDVI, like those associated with z0m parameter set 1, resulted in the highest errors for tall crops (maize). The NDVI-z0m parameters that resulted in the lowest error were those that best fit a range of crop types (parameter set 4), though that parameters set had higher error in some crops (maize) for some values of q than other parameter sets.

The best combination of NDOM, q, c, and d was determined as that with the minimum crop-mean relative mean bias of SEBAL ET (Table 4). The best combination for Λ was slightly different from the best combination for ET (Table 5). For simplicity, we use the same parameter set (q=0.005, NDVI-z0m parameter set 4) for both ET and Λ in the error analysis and comparison between MOD16 and SEBAL.

*MOD16 compared to SEBAL and flux towers*

MOD16 ET was slightly more accurate than SEBAL over maize, but MOD16 ET was much lower than tower ET and less accurate than SEBAL over rice (mean rMB -51%) and cotton (rMB -78%) (Figure 4, Figure 5, Table 6, Table 7). For rice, the difference between MOD16 ET and tower ET was largest in the beginning of the growing season, which coincides with low NDVI but high LSWI (Figure 4, Figure 5) and corresponds to flooding and early growth stages in rice. During this initial stage, SEBAL had a lower error than MOD16, suggesting that MOD16 underestimates soil evaporation. However, error in MOD16 ET was also high during the middle of the growing season for rice and cotton, suggesting that MOD16 also underestimates transpiration during maximum vegetation coverage for rice and cotton.

The error in seasonal ET was a function of crop height (p<0.001, Figure 6). Error in MOD16 ET was greater than 50% for crops less than 75 cm tall. For rice, error in MOD16 decreased slightly with crop height. For tall crops (maize), error in SEBAL ET was similar to error in MOD16 ET except for the later growth stage (grain-bud filling), when SEBAL overestimated ET. FVC and AWB were not significantly correlated with the error in seasonal ET (p>0.1). Maize also had a low evaporative fraction (Λ) compared to rice and cotton, so the error in ET may also be due to the relative performance of SEBAL and MOD16 over crops with low Λ.

**Discussion**

*SEBAL automation and comparison with MOD16*

The temperature-based, automated SEBAL algorithm estimated seasonal ET to within 14-33% of ET observed at the towers for all crops. SEBAL ET estimates were a significant improvement from the vegetation-based MOD16 algorithm over rice and cotton, which underestimated seasonal ET by 34-78%, though MOD16 ET was slightly better than SEBAL ET over maize. The error in SEBAL evaporative fraction (Λ) was <10% for cotton and 2 of 3 rice sites, suggesting that SEBAL performs very well for short irrigated crops.

*Comparison with other studies of automated ET*

The error in seasonal mean ET from our automated SEBAL method is higher than that reported for automated METRIC implemented over alfalfa and pasture sites in western Nevada (Morton et al., 2013), where seasonal mean error was ~5% for all fields, with a range of -16% to +21%, compared to our overall error 22% for all towers with a range of -33% to +17%. The improved results for METRIC are likely due to a combination of the use of ground-level data in METRIC and the relative homogeneity of the crop type. The METRIC application included ground-based measurements of potential evapotranspiration and utilized ground-measured wind speed to determine friction velocities, compared to our implementation of SEBAL which used no ground-level data as input. The sites from Morton et al (2013) included irrigated alfalfa and pasture and non-irrigated pasture, where the parameters of the NDVI-z0m relationship may be expected to be more homogeneous over the landscape compared with the Central Valley. If only short crops are considered and the NDVI-z0m relationship fit to those crops (parameter set 1), the MODIS automated SEBAL also has a small mean error (18%), but the error over tall crops increases from 16% to 26%. The automated SEBAL results suggest that estimation of ET over irrigated sites, particularly those where the evaporative fraction is significantly less than one, is complicated by spatial variability in crop height and the NDVI-z0m relationship. Improvements in the estimation of crop height and z0m with remote sensing could significantly improve regional ET estimates in SEBAL.

The accuracy of remote sensing methods to estimate ET is summarized in several publications (Bastiaanssen et al., 2005; Kalma et al., 2008). (Teixeira et al., 2009) found that…

Michael, can you add a few sentences about the relative performance of temp and veg-based models (e.g. SEBAL vs JPL, MOD16)? We may be able to cut and paste from the chapter.

*Implications for global mapping of ET in irrigated areas*

Vegetation-based methods like MOD16 or the related PT-JPL model perform well in rainfed areas and in irrigated sites with low land surface wetness and/or high NDVI, which includes maize. MOD16 performed poorly compared with SEBAL over plots planted in rice, particularly at the beginning of the growing season when NDVI was low but land surface wetness was high. This period coincides with inundated conditions at the surface, when ET may be high but vegetation cover low. Since the soil evaporation component from both MOD16 and PT-JPL is low when relative humidity is low, both methods may systematically underestimate ET in irrigated fields with high soil evaporation in semi-humid environments like the Central Valley. Such conditions may also be found in other major irrigated areas of the world, including India and China. Improvements to the soil evaporation component of MOD16 are recommended. MOD16 authors have incorporated LST data to better estimate soil evaporation, and this new algorithm will be tested in irrigated areas (Mu, 2013). However, MOD16 ET was also low over rice and cotton for growth periods with high fractional vegetation cover, suggesting that MOD16 also underestimates transpiration.

**Conclusion**

Automated methods for estimating seasonal evapotranspiration (ET) have a wide range of applications. This study documents that 1) automated ET algorithms based on temperature (SEBAL) using only global climate grids and remote sensing data can estimate ET from irrigated crops to with errors ranging from -33 to +17%; 2) automated SEBAL is sensitive to the size of the domain, the parameters of the equation used to predict the surface roughness length (z0m) and the quantile used to select calibration pixels; 3) automated algorithms based on vegetation indices alone (MOD16) underestimate ET in irrigated crops by more than 50% for rice and cotton, but perform well for maize. The difference in performance by crop type may be due to differences in the surface roughness or to differences in the evaporative fraction by crop: MOD16 performed will over tall crops with a low evaporative fraction, while SEBAL performed well over short crops with a high evaporative fraction. Determination of the factors that affect the performance of each method, including separation of the effects of crop height and evaporative fraction, requires further research.

The initial hypothesis, that ET from vegetation-based methods would show the most error in the early stages of crop growth when soil evaporation is high, was partly validated, though MOD16 also underestimated ET from rice and cotton during full canopy cover, suggesting that it also underestimates transpiration. Further research is required to improve the estimate of crop height and to determine the range of crops over which each method is most valid. Future regional and global maps could use a map of crop type or height to determine which method should be used over a given pixel to produce the most accurate ET estimates at regional and global scales.

**References**

Allen, R., Irmak, A., Trezza, R., Hendrickx, J.M.H., Bastiaanssen, W., Kjaersgaard, J., 2011. Satellite-based ET estimation in agriculture using SEBAL and METRIC. Hydrol. Process. 25, 4011–4027. doi:10.1002/hyp.8408

Allen, R.G., Tasumi, M., Trezza, R., 2007. Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)—Model. J. Irrig. Drain. Eng. 133, 380–394. doi:10.1061/(ASCE)0733-9437(2007)133:4(380)

Anderson, M.C., Kustas, W.P., Hain, C.R., 2013. Mapping surface fluxes and moisture conditions from field to global scales using ALEXI/DisALEXI, in: Petropoulos, G.P. (Ed.), Remote Sensing of Energy Fluxes and Soil Moisture Content. Taylor and Francis, pp. 207–232.

Bastiaanssen, W.G.M., 2000. SEBAL-based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey. J. Hydrol. 229, 87–100. doi:http://dx.doi.org/10.1016/S0022-1694(99)00202-4

Bastiaanssen, W.G.M., Ahmad, M.D., Chemin, Y., 2002. Satellite surveillance of evaporative depletion across the Indus Basin. Water Resour. Res. 38, 1273, doi:10.1029/2001WR000386.

Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M., 1998. A remote sensing surface energy balance algorithm for land (SEBAL) 1. Formulation. J. Hydrol. 212-213, 198–212.

Bastiaanssen, W.G.M., Noordman, E.J.M., Pelgrum, H., Davids, G., Thoreson, B.P., Allen, R.G., 2005. SEBAL model with remotely sensed data to improve water-resources management under actual field conditions. J. Irrig. Drain. Eng. 131, 85–93.

Conrad, C., Dech, S., Hafeez, M., Lamers, J., Martius, C., Strunz, G., 2007. Mapping and assessing water use in a Central Asian irrigation system by utilizing MODIS remote sensing products. Irrig. Drain. Syst. 21, 197–218. doi:10.1007/s10795-007-9029-z

Dabberdt, W.F., Lenschow, D.H., Horst, T.W., Zimmerman, P.R., Oncley, S.P., Delany, A.C., 1993. Atmosphere-surface exchange measurements. Science (80-. ). 260, 1472–1481.

Fisher, J.B., Tu, K.P., Baldocchi, D.D., 2008. Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. Remote Sens. Environ. 112, 901–919.

French, A.N., Alfieri, J.G., Kustas, W.P., Prueger, J.H., Hipps, L.E., Chávez, J.L., Evett, S.R., Howell, T.A., Gowda, P.H., Hunsaker, D.J., Thorp, K.R., 2012. Estimation of surface energy fluxes using surface renewal and flux variance techniques over an advective irrigated agricultural site. Adv. Water Resour. 50, 91–105. doi:http://dx.doi.org/10.1016/j.advwatres.2012.07.007

Glenn, E.P., Nagler, P.L., Huete, A.R., 2010. Vegetation index methods for estimating evapotranspiration by remote sensing. Surv. Geophys. 31, 531–555.

Kalma, J.D., McVicar, T.R., McCabe, M.F., 2008. Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. Surv. Geophys. 29, 421–469.

Long, D., Singh, V.P., Li, Z.-L., 2011. How sensitive is SEBAL to changes in input variables, domain size and satellite sensor? J. Geophys. Res. 116, D21107. doi:10.1029/2011jd016542

Morse, A., Tasumi, M., Allen, R.G., Kramber, W.J., 2000. Application of the SEBAL methodology for estimating consumptive use of water and streamflow depletion in the Bear River Basin of Idaho through remote sensing.

Morton, C.G., Huntington, J.L., Pohll, G.M., Allen, R.G., McGwire, K.C., Bassett, S.D., 2013. Assessing Calibration Uncertainty and Automation for Estimating Evapotranspiration from Agricultural Areas Using METRIC. JAWRA J. Am. Water Resour. Assoc. 49, 549–562. doi:10.1111/jawr.12054

Mu, Q., 2013. MODIS 1-km2 Terrestrial Evapotranspiration (ET) Product for the Nile Basin Algorithm Theoretical Basis Document.

Mu, Q., Zhao, M., Running, S.W., 2011b. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sens. Environ. 115, 1781–1800.

Mu, Q., Zhao, M., Running, S.W., 2011a. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sens. Environ. 115, 1781–1800.

Reichle, R.H., Koster, R.D., De Lannoy, G.J.M., Forman, B.A., Liu, Q., Mahanama, S.P.P., Touré, A., 2011. Assessment and Enhancement of MERRA Land Surface Hydrology Estimates. J. Clim. 24, 6322–6338. doi:10.1175/JCLI-D-10-05033.1

Su, Z., 2002. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. Hydrol. Earth Syst. Sci. Discuss. 6, 85–100.

Teixeira, A.H. de C., Bastiaanssen, W.G.M., Ahmad, M.D., Bos, M.G., 2009. Reviewing SEBAL input parameters for assessing evapotranspiration and water productivity for the Low-Middle São Francisco River basin, Brazil: Part A: Calibration and validation. Agric. For. Meteorol. 149, 462–476.

Twine, T.E., Kustas, W.P., Norman, J.M., Cook, D.R., Houser, Pr., Meyers, T.P., Prueger, J.H., Starks, P.J., Wesely, M.L., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agric. For. Meteorol. 103, 279–300.

Van Niel, T.G., McVicar, T.R., Roderick, M.L., van Dijk, A.I.J.M., Renzullo, L.J., van Gorsel, E., 2011. Correcting for systematic error in satellite-derived latent heat flux due to assumptions in temporal scaling: Assessment from flux tower observations. J. Hydrol. 409, 140–148. doi:http://dx.doi.org/10.1016/j.jhydrol.2011.08.011

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., 2002. Energy balance closure at FLUXNET sites. Agric. For. Meteorol. 113, 223–243.

Xiao, X., Boles, S., Frolking, S., Salas, W., Moore, B., Li, C., He, L., Zhao, R., 2002. Landscape-scale characterization of cropland in China using Vegetation and Landsat TM images. Int. J. Remote Sens. 23, 3579–3594. doi:10.1080/01431160110106069

Xiao, X., Boles, S., Frolking, S., Salas, W., Moore Iii, B., Li, C., He, L., Zhao, R., 2002. Observation of flooding and rice transplanting of paddy rice fields at the site to landscape scales in China using VEGETATION sensor data. Int. J. Remote Sens. 23, 3009–3022.

Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., Salas, W., Moore Iii, B., 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. Remote Sens. Environ. 95, 480–492.

**Figure captions**

Figure 1. Map of study area, including flux tower sites and the boundaries of the study areas where SEBAL was applied. The white boundary indicates the area between -20 m and +100 m that was used as an additional mask in defining the model domain for SEBAL.

Figure 2. Plot of Rn from GMAO-MERRA vs Rn from the towers.

Figure 3. Plot of crop height versus NDVI at the tower sites with crop height data, and the NDVI-z0m relationships (6) used in the SEBAL model.

Figure 4. Time series of LSWI, EVI and TR (top plot in each panel), water table elevation (mid plot) and ET from SEBAL, MOD16, and towers (bottom plot in each panel) at the US-Twt rice site, 2011 and 2012. In the ET plots, SPR, FLW and GB indicate sprouting, flowering, and grain-bud filling.

Figure 5 Time series of LSWI, EVI and TR (top plot in each panel), water level, and ET from SEBAL, MOD16, and towers (bottom plot in each panel) at the US-Twt rice site in 2011 and 2012. In b), DR, PL, FL indicate drained, planted, and flooded. In c), SPR, FLW and GB indicate sprouting, flowering, and grain-bud filling.

Figure 6. Error in SEBAL and MOD16 versus field-measured crop height. GB indicates maize in the grain-bud filling growth stage.

Figure xx?. Map of season total ET SEBAL (left) MOD16 (right) at the field sites.

Figure xx? TR-NDVI plot showing hot and cold calibration pixels for different values of q.

**Tables**

Table 1. Characteristics of flux towers and data. EC indicates eddy covariance and SR indicates surface renewal.

Table 2. Error statistics for net radiation (Rn) of GMAO-MERRA compared to Rn at flux towers.

Table 3. Comparison of reference evapotranspiration (ETo) from MOD16 and towers at the towers.

Table 4. Sensitivity analysis of automated SEBAL ET to the number of domains (NDOM), quantile for selecting calibration pixels (q), and the parameters of the NDVI-z0m relationship. See Figure 3 for parameter values for each parameter set.

Table 5 Sensitivity analysis of automated SEBAL evaporative fraction (Λ) to the number of domains (NDOM), quantile for selecting calibration pixels (q), and the parameters of the NDVI-z0m relationship.

Table 6. Comparison of seasonal ET from automated SEBAL and MOD16 at each tower.

Table 7. Comparison of the evaporative fraction (Λ) from SEBAL and MOD16 at each tower.