Article

Comparison of the Automatically Calibrated Google Evapotranspiration Application - EEFlux and the Manually Calibrated METRIC Application

Foad Foolad[1], Philip Blankenau[1,2], Ayse Kilic[1,2], Richard G. Allen[3], Justin L. Huntington[4], Tyler A. Erickson[5], Doruk Ozturk[2], Charles G. Morton[4], Samuel Ortega[2], Ian Ratcliffe[2], Trenton E. Franz[2], David Thau[5], Rebecca Moore[5], Noel Gorelick[5], Baburao Kamble[1,2], Peter Revelle[3], Ricardo Trezza[3], Wenguang Zhao[3] and Clarence W. Robison[3]

[1] Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln NE 68588, USA; foadfoolad@huskers.unl.edu (F. F.), pblankenau2@unl.edu (P. B.), akilic@unl.edu (A. K.), bkamble3@unl.edu (B. K.)

[2] School of Natural Resources, University of Nebraska-Lincoln; Lincoln NE 68583, USA; doruk@huskers.unl.edu (D. O.), sortegas88@hotmail.com (S. O.), iratcliffe2@unl.edu (I. R.), tfranz2@unl.edu (T. E. F.)

[3] University of Idaho, Kimberly Research and Extension Center, Kimberly, ID 83341, USA; rallen@uidaho.edu (R. G. A.), prevelle@gmail.com (P. R.), rtrezza@uidaho.edu (R. T.), wzhao@uidaho.edu (W. Z.), robison@uidaho.edu (C. W. R.)

[4] Desert Research Institute (DRI), Reno, NV 89512, USA; Justin.Huntington@dri.edu (J. L. H.), Charles.Morton@dri.edu (C. G. M.)

[5] Google Earth Engine and Google Earth Outreach, Google LLC, Mountain View, CA 94043, USA; tylerickson@gmail.com (T. A. E.), thau@google.com (D. T.), rmoore@google.com (R. M.), gorelick@google.com (N. G.)

* Correspondence: foadfoolad@huskers.unl.edu.; Tel.: +1-402-591-0817

Abstract: Reliable evapotranspiration (ET) estimation is a key factor for water resources planning, attaining sustainable water resources use, irrigation water management, and water regulation. During the past few decades, researchers have developed a variety of remote sensing techniques to estimate ET. The Earth Engine Evapotranspiration Flux (EEFlux) application uses Landsat imagery archives on the Google Earth Engine platform to calculate the daily evapotranspiration at the local field scale (30 m). Automatically calibrated for each Landsat image, the EEFlux application design is based on the widely vetted Mapping Evapotranspiration at high Resolution with Internalized Calibration (METRIC) model and produces ET estimation maps for any Landsat 5, 7 or 8 scene in a matter of seconds. In this research we evaluate the consistency and accuracy of EEFlux products that are produced when standard US and global assets are used. Processed METRIC products for 58 scenes distributed around the western and central United States were used as the baseline for comparison. The goal of this paper is to compare the results from EEFlux with the standard METRIC applications to illustrate the utility of the EEFlux products as they currently stand.

Given that EEFlux is derived from METRIC, differences are expected to occur due to differing calibration methods (automatic versus manual) and differing input datasets. The products compared include the fraction of reference ET (ETrF), actual ET (ETa), and surface energy balance components net radiation (Rn), ground heat flux (G), and sensible heat flux (H), as well as Ts, albedo and NDVI. The product comparisons show that the intermediate products of Ts, Albedo, and NDVI, and also Rn have similar values and behavior for both EEFlux and METRIC. Larger differences were found for H and G. Despite the more significant differences in H and G, results show that EEFlux is able to calculate ETrF and ETa values comparable to the values from trained expert METRIC users for agricultural areas. For non-agricultural areas such as semi-arid rangeland and forests, the automated EEFlux calibration algorithm needs to be improved in order to be able to reproduce ETrF and ETa that is similar to the manually calibrated METRIC products.

Keywords: Google Earth Engine, EEFlux, METRIC, Evapotranspiration, Landsat, Water Resources Management

---

1. Introduction

Reliable and accurate estimates of water consumption are essential for water rights management, water resources planning and water regulation, especially for agricultural fields that may have specifically attached water rights [1]. Over the past few decades, a variety remote sensing techniques have been used to quantify evapotranspiration (ET) at the field and larger scales over large range of agricultural and nonagricultural land uses [1–6]. Among the types of remote sensing of ET models, surface energy balance techniques are one of the more popular methods used. The Mapping Evapotranspiration at high Resolution with Internalized Calibration (METRIC) application [7,8] is one of the more widely used surface energy balance models in operational practice, and employs principles and techniques that originated with the Surface Energy Balance Algorithms for Land (SEBAL) [9].

The accuracy of METRIC ET has been evaluated using measured ET by Lysimeter, Bowen ratio and eddy covariance towers in a range of locations of the U.S. [10–16]. Because results of comparisons between METRIC ET and measured ET have been promising, and due to the physically-based employment of surface energy balance algorithms, METRIC is considered to be a well-established model that has been routinely applied as part of the water resources management operations in a number of states and federal agencies [17]. However, applying METRIC can often be time-consuming, since a well-trained expert is typically needed to calibrate and run the model. Calibration of METRIC is required for each Landsat scene and image date and entails the determination and assignment of extreme ranges in ET (high and low) to locations within an image. The step calibrates temperature-impacted components of the surface energy balance to reproduce the assigned ET range. Different users who might not be equally experienced can produce different results. To reduce the uncertainties associated with the calibration process, and to save time and money [15,18], designed automated calibration algorithms for the METRIC model to generate ET estimates comparable to ones manually produced from well-trained users. Comparison results have suggested that an automated calibration algorithm can estimate ET comparable to the ET estimated by trained users, and the variation within populations of ET produced with automated calibrations have mimicked the variation produced manually between different users [15].

Although the automated calibration of the METRIC application reduces some of the expertise requirements of ET production, users still have to accrue and assemble a variety of inputs including the satellite image, land cover map, digital elevation map, local weather data, and soils map, from a variety of sources and platforms. There can be a significant amount of pre-processing required for the different inputs before applying the algorithms. The input and data handling can be one of the most time consuming parts of the overall process. As a means to automate data assembling and handling and to speed the ET computation process, the Earth Engine Evapotranspiration Flux (EEFlux) application was designed and developed on the Google Earth Engine (GEE) platform based on the METRIC model [7]. EEFlux utilizes Landsat imagery archives stored on GEE, a cloud-based platform (see Allen et al., [10]). A web-based interface provides users with the ability to request ET estimation maps for any Landsat 5, 7 or 8 scene in a matter of seconds. EEFlux also provides rapid generation of intermediate product maps, such as surface temperature (T_s), normalized difference vegetation index (NDVI) and albedo maps for given Landsat scene that may be useful for other applications besides ET.

The goal of this paper is to compare the results from EEFlux with standard manually calibrated METRIC products to assess the utility and accuracy of EEFlux products as they currently stand. Though METRIC does not represent ground-truth, its standing in the scientific community is established, making it a reasonable benchmark for comparison. Further, given that EEFlux is derived from METRIC, it is useful to examine the differences between their products. Differences are expected due to the differing energy balance calibrations (automatic versus manual), versions of METRIC, geographic location and differing input datasets. Because of the continuing evolution of both METRIC and EEFlux, there are algorithmic differences beyond the energy balance calibrations, but these generally tend to have more minor impacts on the final ET products relative to calibration and input differences. Therefore, this paper does not seek to trace each algorithmic difference but touches on some of the significant known differences. The products compared include the fraction of reference ET (ET_rF), actual ET (ET_a), net radiation (R_n), ground heat flux (G), sensible heat flux (H), T_s, albedo and NDVI. Those products were gathered from 58 METRIC scenes in the western and central United States that were produced by trained individuals.

2. Materials and Methods

2.1 Study Area

A suite of images from different parts of the western and central U.S. were chosen to compare the performance of automatically calibrated EEFlux to manually calibrated METRIC, and locations within agricultural fields and non-agricultural land areas were examined. These areas were selected due to the importance of water in the areas and the significant impacts of water on the study areas' economies. In this comparison analysis, we used existing processed METRIC images that had been developed to identify or address particular water resources issues in key areas. Analyzing different regions of the U.S. provided a basis for examining regional differences in comparison statistics.

In total 58 Landsat image dates were evaluated in this study. Figure 1 shows the Landsat scene locations and study areas of the research. In central Nebraska, areas along the Platte River were the focus of study, where 15 Landsat images (Paths 29-30 and Rows 31-32), during summer 2002, were utilized. In western Wyoming, agricultural areas along the Green River were evaluated. That area falls into 2 Landsat rows on a single path (Path 37 and Rows 30-31). We utilized 9 Landsat images during summer 2011 for the comparison. Southern California was the third study area (Path 39 and Row 37). Due to its very dry climate, the California location had the highest frequency of cloudless images, so that we were able to evaluate 13 Landsat images from late January 2014 to early November 2014. A large irrigated area in southern Idaho comprised a fourth area containing 15 Landsat image dates from year 2016 (Path 40 and Row 30). That location represents a large irrigated region receiving irrigation water from the Snake River and from the Snake Plain Aquifer. The fifth location was comprised of agricultural areas in the Klamath basin of southern Oregon and northern California where we evaluated 6 Landsat images (Path 45 and Row 31), during the growing season of year 2004.

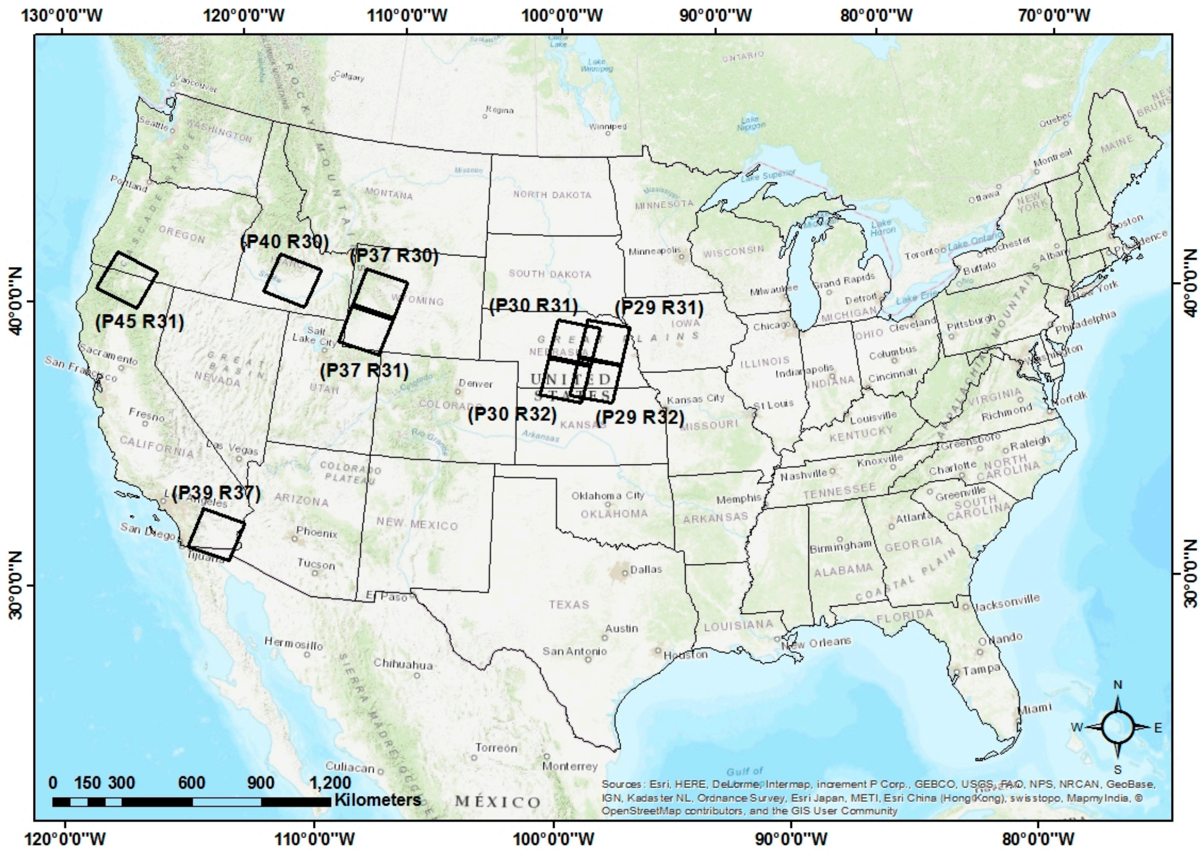

Figure 1. Locations of Landsat Scenes evaluated in this study.

2.2 Methods

Because the objective of this study was the comparison between the automatically calibrated EEFlux products to manually produced METRIC products, we discuss the primary differences between the two applications and refer the readers to primary documents that explain the details of the METRIC model (e.g., [1,7–9,17]). We note that the GEE-based EEFlux application is still being actively developed by the University of Nebraska-Lincoln (UNL), University of Idaho (UI) and Desert Research Institute (DRI). EEFlux production data from version 0.9.4 was used in this study.

In this section, we briefly explain the sampling methods we used and introduce the criteria used to compare EEFlux and METRIC products. We note that METRIC algorithms have been improved upon and evolved over time, with applications of METRIC in the study areas occurring over a number of different years (2002-2016), and using different versions of METRIC algorithms. The different versions of METRIC include differences in produced energy balance components that are generally minor, for example, in the calculation of ground heat flux and aerodynamic roughness.

2.2.1 Similarities and Differences between EEFlux and METRIC

EEFlux employs primary METRIC algorithms that conduct a full energy balance at the land surface and calculate latent heat energy (LE, W/m^2) on a pixel by pixel basis as a residual of the surface energy balance equation:

$$LE = R_n - G - H \quad (1)$$

where LE is heat energy used by water in its phase change from liquid to gas during the ET_a process, R_n is net radiation flux density (W/m^2); G is the ground heat flux density (W/m^2) representing sensible heat conducted into the ground; and H is the sensible heat flux density (W/m^2) convected into the air. LE is estimated at the exact time of the satellite overpass for each pixel. ET_a is then calculated by dividing LE by the latent heat of vaporization:

$$ET_{inst} = 3600 \frac{LE}{\lambda \rho_w} \quad (2)$$

where ET_{inst} is the instantaneous ET flux (mm h^{-1}); 3600 converts seconds to hours; ρ_w is the density of water (~1000 kg m^{-3}); and λ is the latent heat of vaporization (J kg^{-1}) that can be computed using T_s, which is the surface temperature (K):

$$\lambda = [2.501 - 0.00236(T_s - 273.15)] \times 10^6 \quad (3)$$

The ET_rF is calculated for each pixel as the ratio of the computed ET_{inst} from each pixel to the instantaneous tall crop reference evapotranspiration (ET_r):

$$ET_rF = \frac{ET_{inst}}{ET_r} \quad (4)$$

ET_rF is used as a vehicle for extrapolating ET from the instant of the overpass to the surrounding 24-hour period. Lastly, daily ET_a over the 24 hour period is calculated by multiplying ET_rF values for each individual pixel by the daily ET_r computed from local or gridded weather data, assuming consistency between ET_rF at overpass time and ET_rF for the 24-hour period [7]:

$$ET_a = ET_rF \times ET_r \quad (5)$$

Equivalency of instantaneous and 24-hour ETrF is applied to land uses that typically have an adequate water supply for full ET, including agriculture and wetland classes. For most other classes such as rangeland and forest, the well-known evaporative fraction, EF, [19] is used to extrapolate to the full day, where EF = $ET_{inst}/(Rn\text{-}G)_{inst}$. Both EEFlux and METRIC applications utilize hourly and daily ET_r computed for the tall reference crop of alfalfa to convert ET_rF to daily ET_a, where the tall alfalfa reference approximates maximum, energy-limited ET from a well-watered, extensive surface of vegetation. ET_r is computed using the ASCE Standardized Penman-Monteith method [20].

One of the primary differences between EEFlux and METRIC is in the use of sources of weather data in their calibration and calculations. METRIC generally uses ground-based hourly weather data from an agriculturally sited weather station to calculate ET_r for the solution of the surface energy balance equation during calibration and estimation of any background evaporation caused by recent precipitation events. EEFlux uses gridded hourly and daily weather data stored on Earth Engine. For locations processed in the US, EEFlux uses North American Land Data Assimilation System (NLDAS) (https://ldas.gsfc.nasa.gov/nldas/) [21] hourly weather data for calibration and GridMet gridded weather data [22] for determining background evaporation. In California, EEFlux uses spatial California Irrigation Management Information System (CIMIS) (https://cimis.water.ca.gov/) daily weather data, if available for the particular date, instead of GridMet. For locations outside of the conterminous United States, EEFlux uses the six-hourly CFSv2 operational analysis [23,24] and the Climate Forecast System Reanalysis (CFSR) (http://cfs.ncep.noaa.gov/cfsr/) [25] gridded weather data for all calculations.

The use of gridded weather data in EEFlux can explain, to some extent, differences between METRIC and EEFlux final products, including estimates for daily ET_a. This is discussed in more detail in the following sections. More detailed information on METRIC and EEFlux ET_r calculations is found elsewhere [10,26,27].

During calibration, METRIC and EEFlux solve the energy balance equation by applying an estimate for ET_a at low ET and high ET conditions and solving for $H = Rn - G - LE$. The low and high ET calibration end-points are referred to as hot and cold pixels. In METRIC, these end-points are searched for automatically or manually, and EEFlux, they are determined automatically. LE is computed by multiplying ET_r by the assumed fraction of ET_r at the calibration points (typically between 0 and 0.1 for the hot pixel and between 1 and 1.05 for the cold pixel). The estimate for instantaneous ET_r does not have a large effect on the ET_rF or ET_a values, since ET_rF is assigned to the end-point conditions. However, it does have an impact on the internally computed H, which is used to absorb and later correct for systematic biases in the other parameters, including Rn, G, albedo, aerodynamic roughness and ET_r [7].

A significant internal difference between EEFlux and METRIC is in the way they calculate G. Some versions of METRIC evaluated calculated G by the following equations depending on the pixel leaf area index (LAI) value:

$$\frac{G}{R_n} = 0.05 + 0.18e^{-0.521\,LAI} \qquad (\mathrm{LAI} \geq 0.5) \qquad (6a)$$

$$\frac{G}{R_n} = \frac{1.80(T_s - 273.15)}{R_n + 0.084} \qquad (\mathrm{LAI} < 0.5) \qquad (6b)$$

whereas later versions of METRIC calculated G as a function of sensible heat flux for LAI > 0.5 and equation 6b otherwise. Very recent versions of METRIC calculate G as a function of LAI only. The version of EEFlux evaluated calculated G as:

$$G = (0.1 + 0.17e^{-0.55\,LAI}) \times R_n \qquad (7)$$

LAI is estimated from surface-corrected NDVI. Due to the differences in calculation of G, the G products often do not match well between METRIC and EEFlux. These differences are carried into the calibration of H, as previously described, but are generally factored back out during calculation of ET_a due to the internal bias correction of METRIC and EEFlux. This is shown later in the results.

METRIC and EEFlux use similar methods for estimating aerodynamic roughness length for momentum transfer, z_{om}, used in calculating aerodynamic resistance in the calculation of H, sensible heat flow from the surface to the air. z_{om} is estimated as a function of estimated LAI for agricultural land classes and as fixed values for nonagricultural classes. METRIC and EEFlux apply a Perrier roughness function [28] for trees, where roughness is a convex function of amount of ground cover. Some versions of METRIC provide for local modification of land cover maps to specify orchard, vineyard and tall (corn) crops so that special estimation can be made for z_{om} as well as albedo and surface temperature to account for shadowing in deep canopies.

2.2.2 Sampling method and comparison criteria

For the comparisons, the highest percentage cloud-free images were selected for the five locations and, for the few images having minor cloud cover, a cloud mask was applied to avoid sampling from clouded areas. A minimum thermal threshold of 270 (K) was used to further screen sampling pixels to avoid thermal pixels lying near the edges of cloud masks or at the edge of gaps in Landsat 7 images caused by the Scan Line Corrector failure. Occasionally, thermal pixels in Landsat 7 images are contaminated by cubic convolution-averaged non-data values stemming from the original native thermal resolution of 60 m.

For the comparison, we randomly chose 1000 pixels from specified areas of interest in the Landsat scenes. These areas targeted primary agricultural areas and adjacent non-agricultural areas comprised of rangeland or forests. National Land Cover Database (NLCD) (https://www.mrlc.gov/) raster data were used to distinguish between agricultural and non-agricultural land covers during sampling. Pixels designated as 81 and 82 NLCD class numbers were used to represent agricultural areas. Non-agricultural pixels were sampled from among all pixels not labeled 81 or 82 in the area of interest. We used a 7×7 focal standard deviation on NDVI to avoid sampling from agricultural field edges, which usually contain mixed pixels, by selecting a pixel only when the standard deviation of the NDVI for those 49 pixels was less than 0.05. Pixels with negative values were removed from the sample selection.

Root Mean Square Error (RMSE) and Coefficient of Determination (R^2) were calculated for each set of data to compare EEFlux products with the same products from METRIC. In addition, slopes of EEFlux products vs. METRIC products with zero intercept were calculated to indicate when EEFlux underestimated or overestimated the products, on average, compared to METRIC. In this study, R^2 values higher than 0.8, RMSE values less than 15% of the average magnitude of each product, and slope values between 0.9 to 1.1 were conidered acceptable, in terms of expected error common to operationally produced spatial ET products [1,7,29–31].

3. Results

Five locations in the United States comprised of nine Landsat image scenes were used to compare the automatically calibrated EEFlux products to the manually calibrated METRIC products. Although the final and primary products of the applications are ET_rF and ET_a, we also compared intermediate products from the models including T_s, albedo, and NDVI, and the primary components of the energy balance: R_n, G, and H. EEFlux is a user-friendly web-based platform that enables users to download the intermediate products of T_s, albedo, and NDVI in addition to ET_rF and ET_a. Therefore, it is useful to confirm similarity with METRIC for those additional products.

We compared the intermediate and final products for each location and calculated R^2, RMSE, and slopes relative to the METRIC products. Figure 2 shows an example comparison for each product sampled from within agricultural fields in Path 29 Row 32 in central Nebraska for a Landsat 5 (2002/06/28) image. Additional graphs of the same format as Figure 2 are included for each location studied in the Supplemental Figures 1-8.

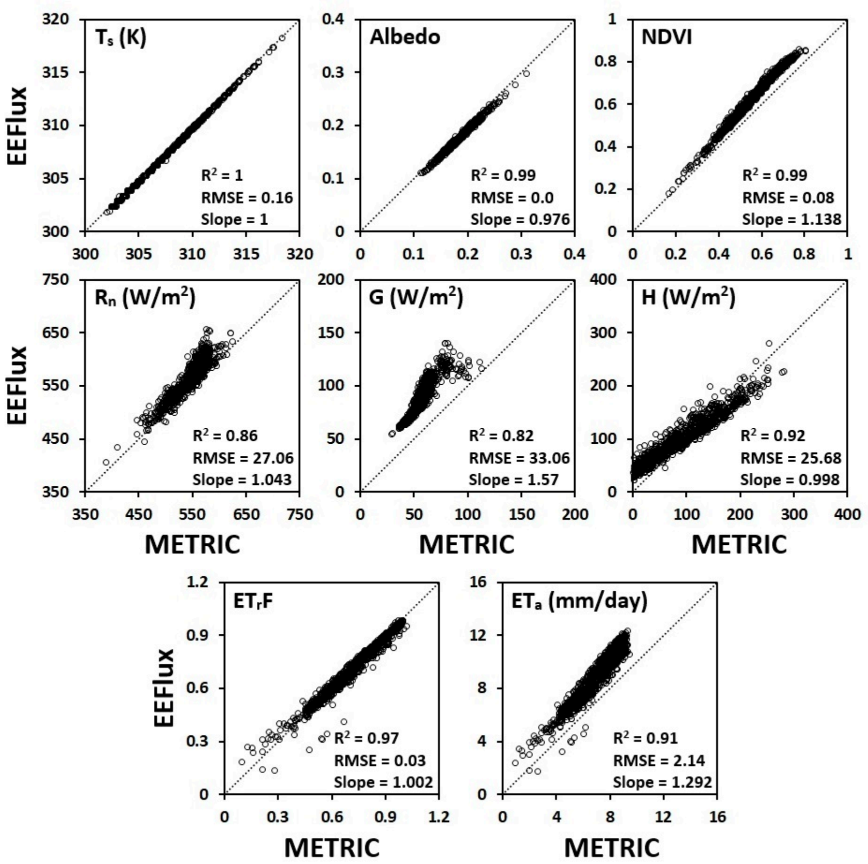

Figure 2. Comparison between various components of EEFlux and METRIC models for agricultural fields located in central Nebraska (Path 29 Row 32, Landsat 5, 2002/06/28).

The comparisons in Figure 2 indicate that the three intermediate products of T_s, Albedo, and NDVI have nearly identical values between EEFlux and METRIC. Their R^2 and slope values are nearly equal to 1 and they have very small RMSE values. The slope for NDVI is greater than 1 due to the particular METRIC version computing NDVI using top-of-atmosphere reflectance values rather than using surface reflectance values as is done in EEFlux. The R_n and H products are also similar between the two models, with R^2 and slope close to 1. Considering the magnitudes of the two products, RMSE values are relatively small. The EEFlux version evaluated uses a different equation to compute G, as compared to the METRIC version applied in Nebraska. Therefore, as expected, G values do not match well, with a positive offset in EEFlux estimates of about 20 W/m^2; However, the R^2 and RMSE values are still within the acceptable range. Moreover, due to the self-reducing bias reduction used internally in EEFlux and METRIC, the systematic bias in G largely cancels out during production of ET_rF [7].

The agreement found with the intermediate products and energy balance components are good indicators of strong correlation and similarity in algorithm performance between EEFlux and METRIC. ET_rF values from EEFlux and METRIC were very similar, with R^2 and slope close to 1 and RMSE value of 0.03. This indicates similarity in the energy balance calibration performed in EEFlux via the automated scheme and the manually-determined calibration in METRIC. For daily ET_a, however, EEFlux had a significant bias relative to METRIC, with RMSE exceeding 2 mm/d and slope of 1.3. The higher estimation of ET_a from EEFlux, given similarity in ET_rF, traces to the conversion of ET_rF to ET_a by multiplying by daily ET_r, which is derived from synoptic gridded weather data in EEFlux as compared to being derived from local measured point or gridded weather data collected from agricultural environments. The general aridity of synoptic weather data, with generally lower humidity content and higher air temperature than experienced under irrigated conditions, especially in semiarid and arid climates [32,33], causes overstatement of ET_r by the Penman-Monteith combination reference equation that presumes a well-watered surface and associated air temperature and humidity parameters [20]. This is discussed more in a later section.

3.1 Overall Summary of EEFlux vs METRIC comparisons

A summary of comparisons over all 58 images and five locations was compiled by combining all sampled data and calculating overall R^2, RMSE, and slope values. For individual image and location comparisons, the reader is referred to Supplemental Tables 1-6 that provide statistics for both agricultural and non-agricultural areas for each image date. Table 1 presents the overall R^2, RMSE, and slope values for all products for agricultural and non-agricultural areas. Intermediate products of T_s, Albedo, and NDVI were relatively similar between agricultural and non-agricultural classes, with R^2 and slope values close to 1 and with relatively small RMSE values. R_n estimates by EEFlux correlated well with those by METRIC, with an average R^2 value of 0.93 and slope of 1.02 for agricultural areas and average R^2 of 0.87 and slope of 1.02 for non-agricultural areas. Relative RMSE for R_n was less than 5%, on average, for R_n for both land covers. The other two energy balance components sampled (G and H) did not match as well between EEFlux and METRIC. The poor agreement for G is attributed to the previously noted differences between METRIC and EEFlux equations for G. Although the equations for G differed between EEFlux and the various METRIC versions, the average RMSE and slope indicate that EEFlux still calculated ET_rF and ET_a values that compared well to METRIC for agricultural areas, with R^2 values of 0.82 and 0.76 for ET_rF and ET_a, respectively. The relatively good agreement for ET_rF and the relatively poor agreement in H is partly explained by the systematic differences in estimates for G, which are embedded into the calibrated estimates for H, and that are then removed from the ET estimates during the ET production steps, due to the internal, systematic bias correction of METRIC and EEFlux. Differences in H are also traceable to the sources used to compute instantaneous ET_r as noted previously, where generally higher estimates in ET_r in EEFlux produce lower values for H during the surface energy balance calibration.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 July 2018 doi:10.20944/preprints201807.0040.v1

Because METRIC typically uses ground-based weather data for hourly and daily ET_r calculation, and EEFlux uses gridded weather data sets to derive ET_r, the calculated ET_r values used in computations can be different due to differences in origin of weather data and aridity biases common to the gridded weather data sets. While several of the METRIC applications applied only a single ET_r value for an entire Landsat image for both energy balance calibration and for interpolation to 24-hour periods, ET_r values used in EEFlux can vary across the image through the gridded weather data that has an approximately 12 km grid spacing for NLDAS-2 hourly data, for CONUS, and 4 km grid spacing for GRIDMET 24-hour data. In order to explore differences among ET_r values used in METRIC and EEFlux, we calculated averages of gridded ET_r values for each image date and associated ratios of those average values to the typically single scene-wide METRIC ET_r values. Table 1 summarizes average slopes of 24-hour EEFlux ET_r values to METRIC ET_r values. On average, over all five locations and the dates evaluated, the grid-based ET_r ran higher than ground-based calculated ET_r by ratios of 1.10 and 1.09 for agricultural and non-agricultural land uses, respectively. The approximately 10% higher ET_r estimation by the gridded data suggests that general ET applications with EEFlux can be biased 10% high solely due to the aridity bias of the gridded data sets [27,34]. This bias is the basis for ongoing studies and development of methods to identify and condition gridded data sets to remove aridity bias prior to calculation of reference ET, which represents near maximum ET in well-watered environments [32]. We further explored the ET_r biases for each individual date and location as described later in the discussion section.

Table 1. Average values for R^2, RMSE, and slope for EEFlux vs. METRIC, based on a comparison over all data (Ag sample size = 47838, Non-Ag sample size = 35110)

Product	Average R^2		Average Slope		Average RMSE	
	Ag	Non-Ag	Ag	Non-Ag	Ag	Non-Ag
T_s (K)	1.00	1.00	1.00	1.00	0.53	0.51
Albedo (0-1)	0.98	0.97	1.00	1.00	0.01	0.01
NDVI (0-1)	0.97	0.93	1.09	1.11	0.07	0.06
R_n (W/m^2)	0.93	0.87	1.02	1.02	26.8	31.6
G (W/m^2)	0.53	0.26	1.43	1.22	41.8	40.6
H (W/m^2)	0.47	0.37	1.03	0.94	69.0	71.5
ET_r (mm/day)	---	---	1.10	1.09	---	---
ET_rF (0-1.05)	0.82	0.45	0.94	0.64	0.13	0.21
ET_a (mm/day)	0.76	0.44	1.01	0.70	1.23	1.39

3.2 ET_rF and ET_a examples

For most applications, the primary products of EEFlux and METRIC that are of most interest are ET_rF and ET_a. Therefore, this results section focuses on those two products. Figure 3 illustrates ET_rF and ET_a correlations and behavior between EEFlux and METRIC over individual sample points for two locations (central Nebraska and southcentral Idaho) and two Landsat systems for agricultural areas. The top two rows of graphs show good EEFlux calibration and estimation relative to the METRIC calibration and estimation, producing relatively good R^2, RMSE, and slope values. The lower row of graphs illustrates a poorer calibration where EEFlux substantially underestimated ET_rF and ET_a especially in the lower end of the ET spectrum, as reflected in poor R^2, RMSE, and slope values. The poor agreement for the particular location and date indicate that the EEFlux automated calibration algorithms can fail under some conditions. As previously noted, those algorithms are under continued improvement by the UNL and UI developers. While the automated calibration of EEFlux is prone to producing poor calibrations under some circumstances, it should be noted that manually calibrated METRIC can also depart from the ground truth [35]. In the 2002/5/2 application shown in Figure 3, the METRIC application diagnosed a substantial impact of recent rain on elevating minimum ET_rF to no lower than 0.6 across the Landsat scene, even for bare soils. The EEFlux application, which used GRIDMET-based precipitation, did not diagnose that same evaporation residual, apparently due to low precipitation amounts present in the gridded data set, and EEFlux therefore projected minimum values for ET_rF of 0.0. This last illustration illustrates some of the challenges associated with what are sometimes labeled as ‘wet’ images, where atmospheric conditions are clear for processing, but the land surface is relatively wet from recent precipitation events.

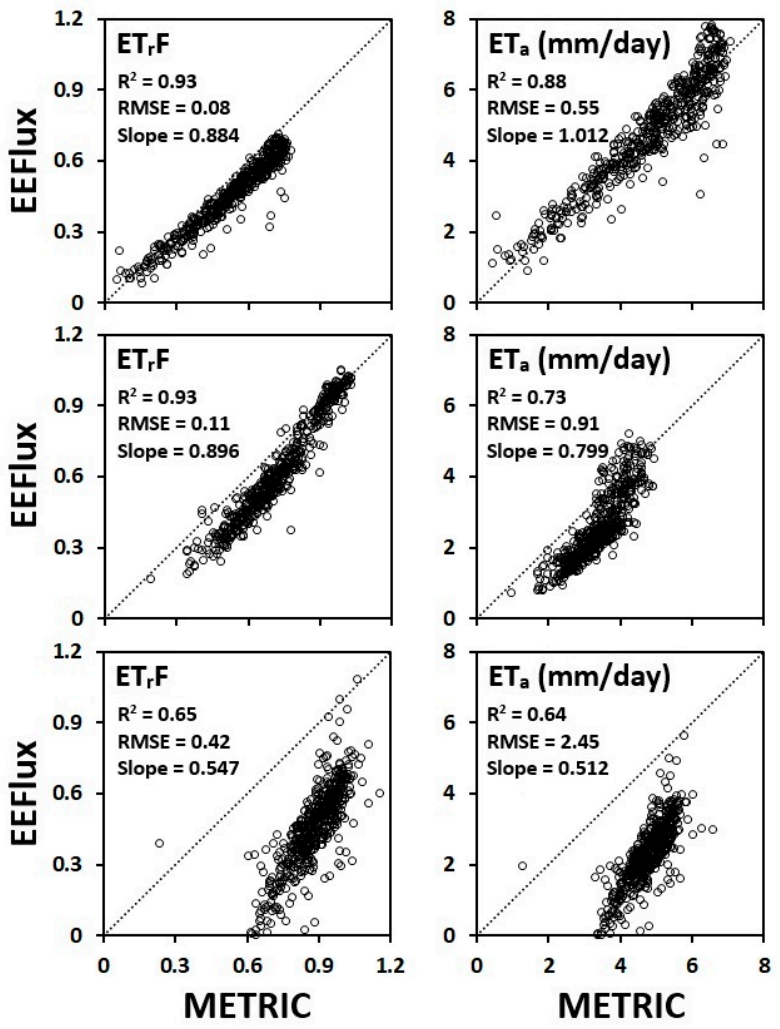

Figure 3. Examples of ET_rF and ET_a calibrations at agricultural fields in different locations. The upper two graphs: good calibration (P29 R31, Landsat 7, central Nebraska, 2002/9/8). The middle two graphs: relatively good calibration (P40 R30, Landsat 7, southcentral Idaho, 2016/9/27). The lower two graphs: poor calibration (P30 R31, Landsat 5, central Nebraska, 2002/5/2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 July 2018 doi:10.20944/preprints201807.0040.v1

In the following section, we explore the differences between EEFlux and METRIC by discussing average statistics determined for ET_rF and ET_a for each of five locations.

3.3 EEFlux ET_rF vs METRIC ET_rF for Individual Locations

Table 2 provides a statistical summary for ET_rF comparisons for each of the nine Landsat path and row locations evaluated that were located in five general USA locations. Statistics are provided for agricultural and non-agricultural land uses. Figure 4 illustrates average slope values for ET_rF for the different locations and Figure 5 presents average RMSE values for ET_rF. The supplemental Figure 9 provides similar plots showing average R^2 values for ET_rF. As shown in Table 2 and Figures 4 and 5, there was minor underestimation of ET_rF values by EEFlux, relative to METRIC, within agricultural land uses for some locations. However, the results were generally good, and EEFlux, on average, is judged to have produced reasonably accurate and useful ET_rF imagery, particularly in southern California, southern Oregon, the Green River area of Wyoming, and in southern Idaho, with average R^2 values higher than 0.84 and average slope values larger than 0.93, and where, in some of the areas, slopes were nearly 1.00. Moreover, the RMSE values in these areas were almost all less than 10% of the average magnitudes of ET_rF values (0-1.05). RMSE values of 10% are considered by Allen et al., [29] and Jensen and Allen [32] to be common to ET estimation and ET measurement. Within the agricultural fields in Nebraska, EEFlux performance was not as good or consistent as for the other locations. However, RMSE and R^2 values are still within our acceptable range, except for one scene area which had an ET_rF RMSE value of 0.28 and R^2 value of 0.69. This was previously illustrated in Figure 3 and is explained by the impact of recent rains, where EEFlux underestimated ET_rF for agricultural areas for several dates in central Nebraska.

R^2, slope and RMSE values in Table 2 and Figures 4 and 5 indicate that EEFlux ET_rF values did not match METRIC ET_rF values as strongly for non-agricultural land uses as they did for agricultural land uses. EEFlux tended to underestimate ET_rF for all non-agricultural land covers sampled and produced RMSE values that were higher than those for agricultural land uses within the same Landsat scene. Some of the differences are due to different means for estimating soil heat flux, for aerodynamic roughness of natural vegetation systems, and potentially due to impacts of the digital elevation model (DEM) used to estimate solar radiation and aerodynamic behavior in complex terrain that is characteristic of natural systems. Differences are also attributed to the weather data sources used in the application of the evaporative fraction (EF) function to nonagricultural land uses, where a ratio of ET_a to $R_n - G$ is used to transform ET_rF to 24-hour ET_rF values, rather than assuming that 24-hour

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 July 2018 doi:10.20944/preprints201807.0040.v1

ET_rF equals instantaneous ET_rF as is done for agricultural land uses [7]. The typically stronger ET_r from gridded weather data impacts this transformation. Causes of these differences, with location, continue to be investigated.

Table 2. Average values for R^2, slope and RMSE for ET_rF for each Landsat scene location evaluated. RMSE values are unitless.

Path	Row	Year	Processed Year	Ag ET_rF				Non-Ag ET_rF			
				n	R^2	Slope	RMSE	n	R^2	Slope	RMSE
29	31	2002	2014	2003	0.84	0.80	0.16	1063	0.83	0.63	0.26
29	32	2002	2014	2387	0.86	0.86	0.15	1309	0.32	0.42	0.30
30	31	2002	2014	3187	0.69	0.72	0.28	1910	0.19	0.40	0.42
30	32	2002	2014	3302	0.94	0.94	0.11	3906	0.50	0.55	0.28
37	30	2011	2013	4815	0.84	0.93	0.11	915	0.52	0.61	0.18
37	31	2011	2013	3608	0.89	1.05	0.10	1921	0.31	0.72	0.14
39	37	2014	2014	10152	0.86	1.00	0.13	6311	0.61	0.81	0.14
40	30	2016	2016	12164	0.89	0.95	0.10	12416	0.52	0.81	0.16
45	31	2004	2011	5765	0.89	0.98	0.10	5759	0.49	0.70	0.18

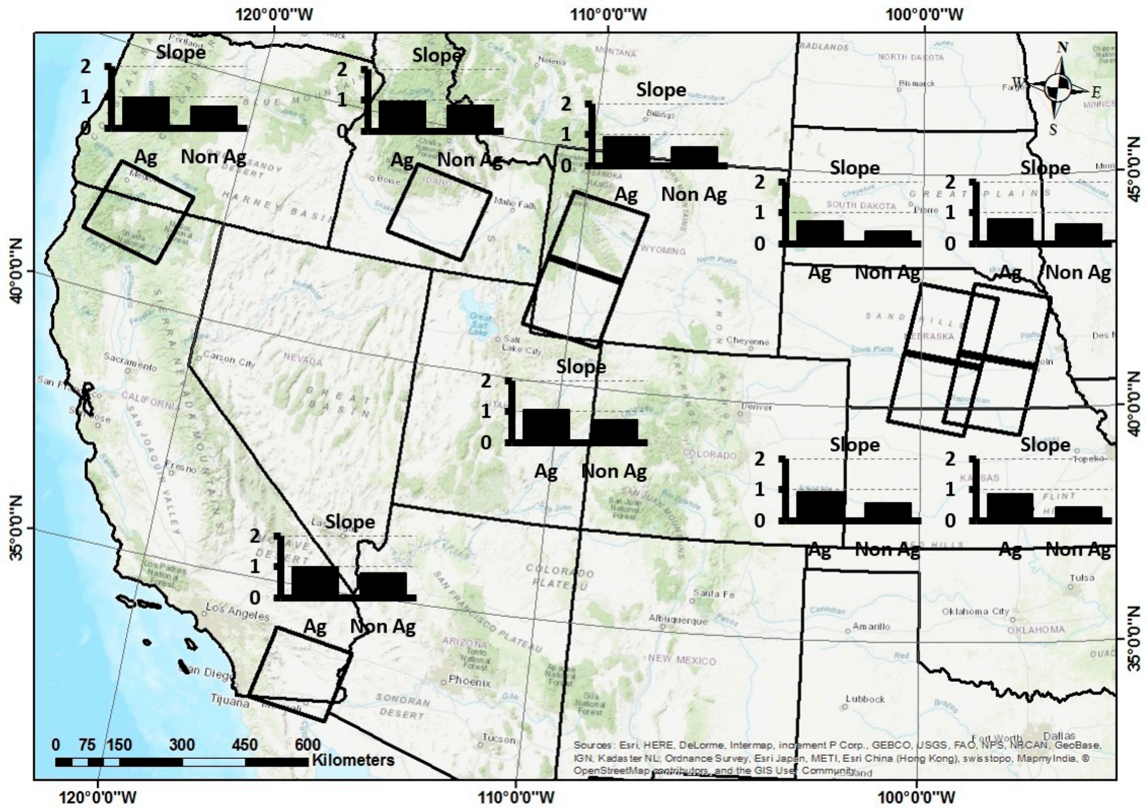

Figure 4. Average slope values for ET$_r$F for EEFlux vs. METRIC for different locations and scenes for agricultural and nonagricultural land uses.

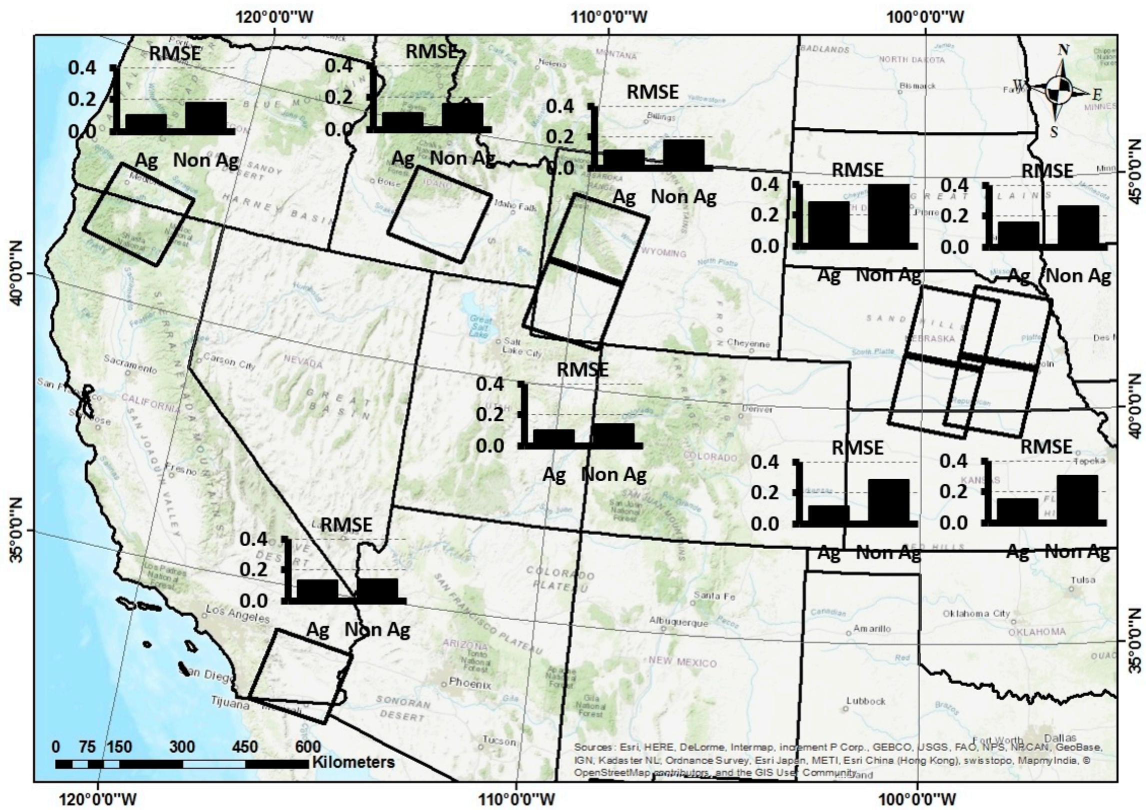

Figure 5. Average RMSE values for ET$_r$F for EEFlux vs. METRIC for different locations and scenes for agricultural and nonagricultural land uses.

3.4 EEFlux ET_a vs METRIC ET_a for Individual Locations

Table 3 provides a statistical summary for ET_a comparisons for the nine Landsat path and row locations evaluated, for both agricultural and non-agricultural land uses. Figures 6 and 7 show average slopes and RMSE values for ET_a. Supplemental Figure 10 provides similar plots for average R^2 values for ET_a. As shown in Table 3 and Figures 6 and 7, slope values increased over those for ET_rF for both agricultural and non-agricultural areas for most of the locations investigated. As discussed previously, that is largely a consequence of ET_r overestimation by use of the gridded weather data set [27,34]. R^2 and slope values were generally within the acceptable accuracy range for agricultural areas. R^2 values were mostly larger than 0.8 and RMSE values were generally in the range of 0.9 to 1.1 mm/d, except one location where it was 0.69 mm/d. Most R^2 values were less than 0.8 for non-agricultural land uses and RMSE values in all locations, except for southern California and southern Idaho, were larger for non-agricultural land uses as compared to agricultural lands. Slope values show that EEFlux tended to underestimate ET_a for non-agricultural land uses everywhere except for southern Idaho. In general, ET_a was substantially lower in non-agricultural land uses than in agricultural areas due to limits on ET imposed by precipitation amount. The agricultural areas sampled were generally all irrigated.

Table 3. Average values for R^2, slope and RMSE for 24-hour ET_a for each Landsat scene location evaluated. RMSE values have units of mm/d.

Path	Row	Year	Processed Year	Ag ET_a				Non-Ag ET_a			
				n	R^2	Slope	RMSE	n	R^2	Slope	RMSE
29	31	2002	2014	2003	0.84	0.92	0.93	1063	0.83	0.73	1.90
29	32	2002	2014	2387	0.87	1.11	1.76	1309	0.39	0.54	2.33
30	31	2002	2014	3187	0.50	0.69	1.89	1910	0.49	0.46	2.67
30	32	2002	2014	3302	0.86	0.91	0.92	3906	0.52	0.57	1.78
37	30	2011	2013	4815	0.83	0.91	1.11	915	0.58	0.54	1.58
37	31	2011	2013	3608	0.87	1.02	0.88	1921	0.34	0.62	1.13
39	37	2014	2014	10152	0.76	1.10	1.22	6311	0.51	0.96	0.97
40	30	2016	2016	12164	0.82	1.13	1.29	12416	0.53	1.05	1.15

45	31	2004	2011	5765	0.89	1.11	0.80	5759	0.54	0.82	0.86

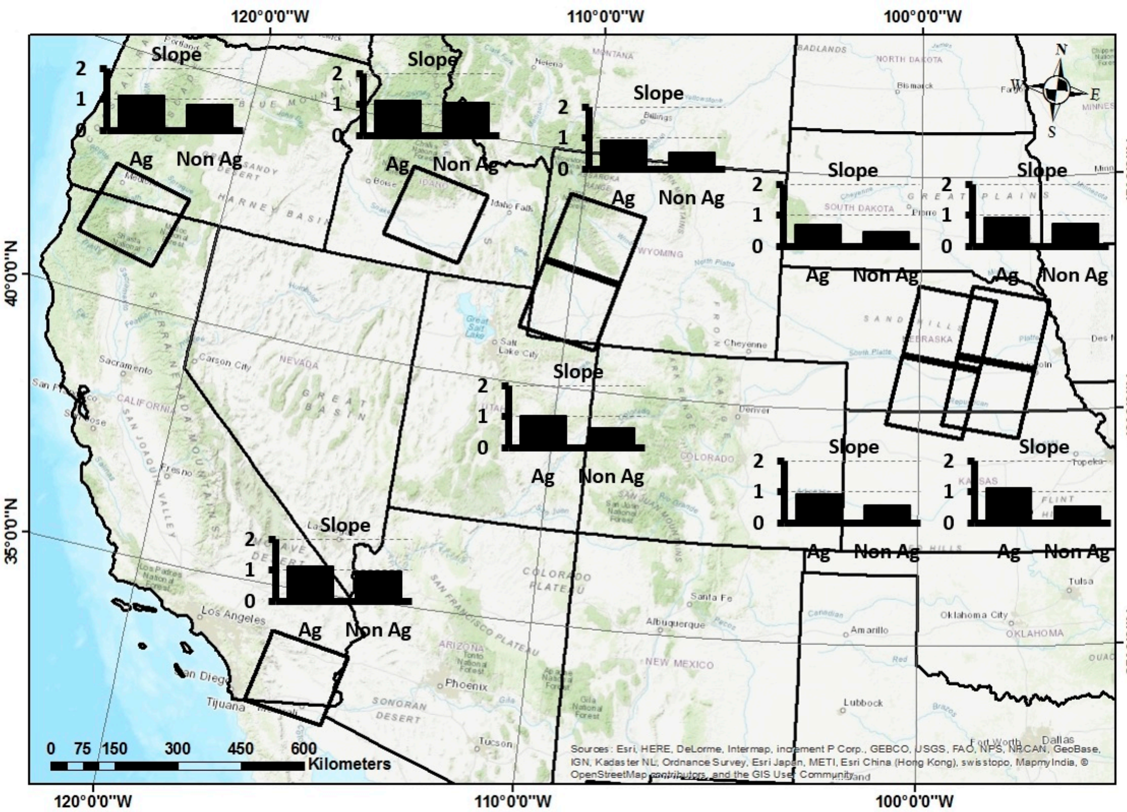

Figure 6. Average slope values for ET_a for EEFlux vs. METRIC for different locations and scenes for agricultural and nonagricultural land uses.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 July 2018 doi:10.20944/preprints201807.0040.v1

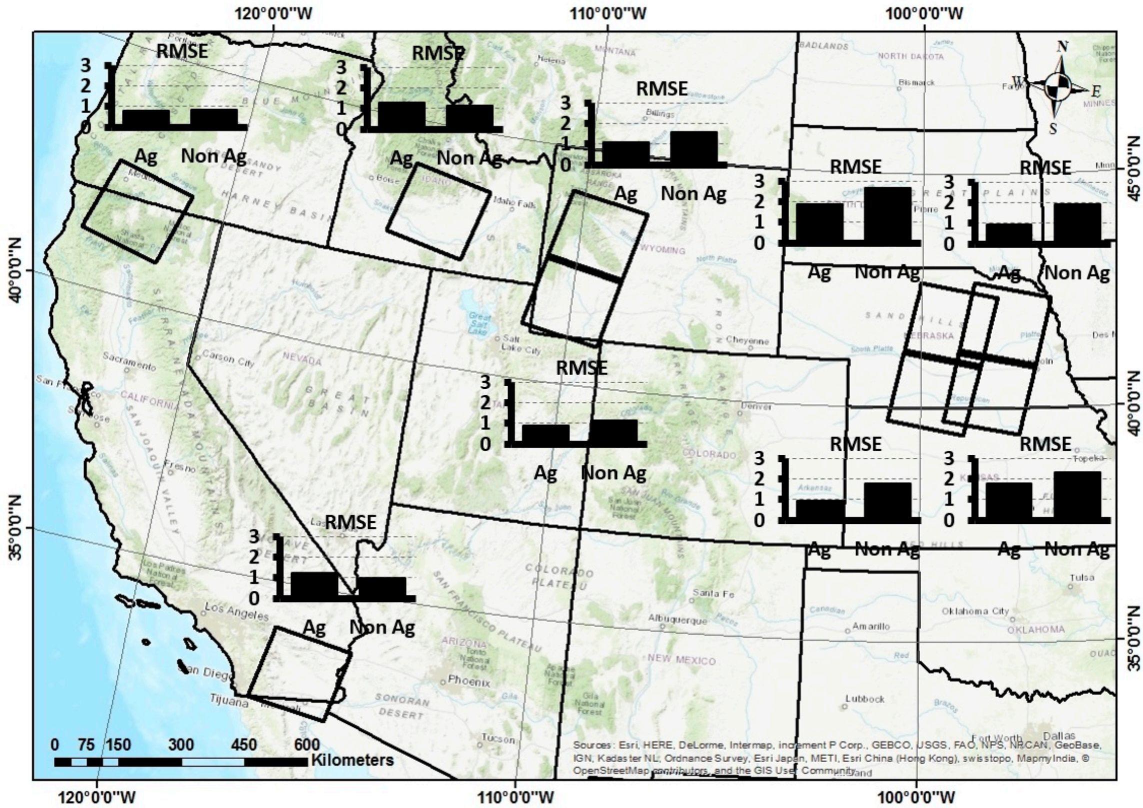

Figure 7. Average RMSE values (mm/d) for ET_a for EEFlux vs. METRIC for different locations and scenes for agricultural and nonagricultural land uses.

3.5 Time dependency of EEFlux performance

Because the study area in southern California had the broadest time series of processed images, we chose this location to explore the time dependency of EEFlux performance and to assess the impact of time of year on performances of the two processing systems. As described earlier we evaluated 13 processed Landsat 8 images for the southern California location. The first and last images evaluated were the 26th of January 2014 and the 10th of November 2014, respectively. Figure 8 shows R^2, slope, and RMSE values for ET_rF and ET_a for agricultural and non-agricultural land uses for different comparison dates. Generally, there was not any statistical correlation between the performance of EEFlux as compared to that of METRIC with time of year. While R^2 values for both ET_rF and ET_a were always higher for agricultural land uses as opposed to non-agricultural land uses, no trends through time were detected. The slope values were similar over time for both agricultural and non-agricultural land uses. However, slopes for non-agricultural ET_rF and ET_a do show a slight trend, decreasing from March through November. RMSE values for ET_rF, like R^2 and slope values did not follow any visible trend during 2014 in the agricultural land uses in southern California. However, as observed in the bottom plot of Figure 8, RMSE values for ET_a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 July 2018 doi:10.20944/preprints201807.0040.v1

increased for both land covers during summer time, indicating larger differences between EEFlux ET_a values and METRIC values during the primary growing season when ET_a was higher.

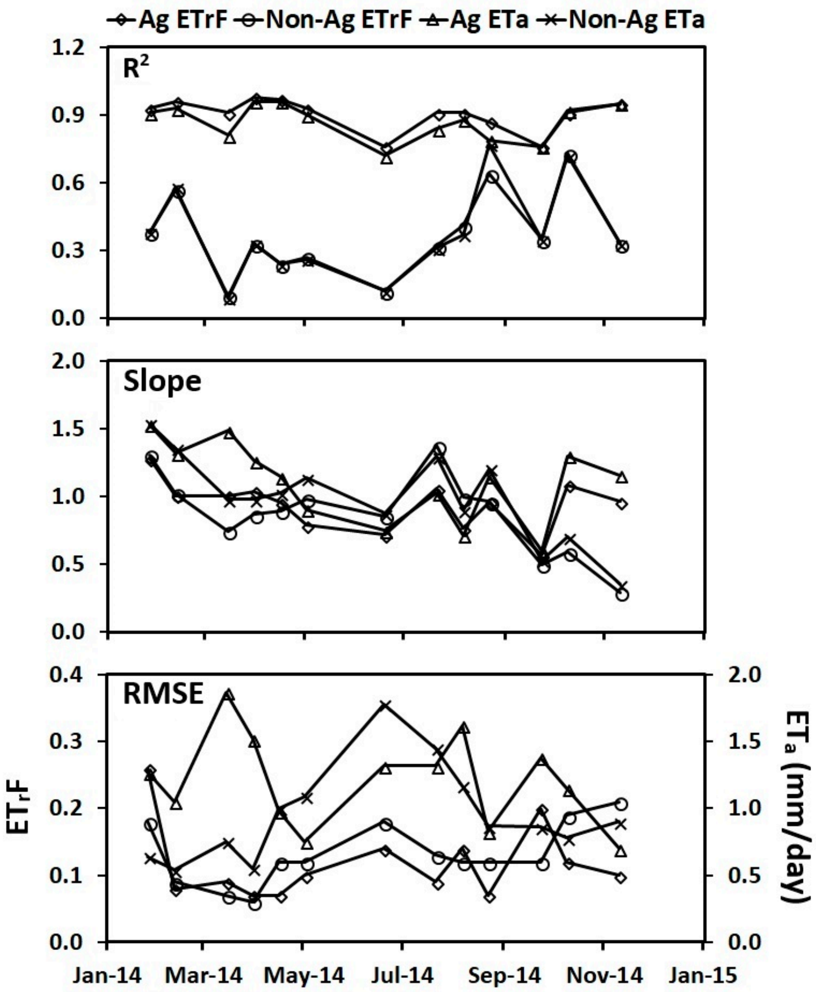

Figure 8. a) R^2, b) slope and c) RMSE values for ET_rF and ET_a products for EEFlux vs. METRIC for a series of comparison dates (Path 39 Row 37).

 doi:10.20944/preprints201807.0040.v1

4. Discussion

Based on the comparison results, we conclude that the implementation of EEFlux on GEE, including the automated internal calibration, has been relatively successful. EEFlux ET_rF and ET_a results matched those from manually applied METRIC applications for most of the agricultural areas evaluated. For some dates within central Nebraska, EEFlux performance was poorer than for the other locations for agricultural land uses. Some of the increased error is due to fewer Landsat images processed for that region due to extensive cloud clover. In one location we were able to evaluate only 3 Landsat image dates (Path 29 Row 31) and for the other three Worldwide Reference System (WRS) scene areas we evaluated 4 image dates; whereas we evaluated 13 Landsat Image dates in California and 15 image dates in Idaho. Having fewer image dates can result in more extreme means due to greater impacts of outliers and/or a smaller sample size. Other impacts, as noted, for central Nebraska is the tendency for more frequent and substantial rainfall during the growing season that increases the impact of background evaporation. This complicates the image calibration. In non-agricultural land uses, EEFlux did not match with METRIC as well as it did for agricultural land uses. This may be partially due to differences among G and H products and DEM sources used. As noted earlier, we evaluated EEFlux version 0.9.4 and, as EEFlux is still in progress, the automated calibration algorithms are expected to be improved in the future, which should result in even more accurate ET_rF and ET_a estimates.

5. Other Analyses

5.1 Source of Reference ET Estimation

Besides using ET_r for internal energy balance calibration and computation, EEFlux uses gridded weather data to extrapolate instantaneous daily ET_rF values to the 24-hour period, which is then multiplied by 24-hour ET_r to calculate daily ET_a values. Figure 9 shows ratios of gridded ET_r values versus the single ET_r values generally used in METRIC computations for each image date and location. As shown in Figure 9, for most dates and locations, the average gridded ET_r values used in EEFlux were higher than the associated single average gridded ET_r values used by METRIC, with variation within each location from about 0.9 to 1.3. As we discussed earlier, the average EEFlux-gridded ET_r was larger than the METRIC calculated, ground-based ET_r values by an average ratio of 1.10 and 1.09 for agricultural and non-agricultural land uses, respectively. The higher 24-hour ET_r estimation in EEFlux due to the gridded weather data source, leads to some degree of daily ET_a overestimation.

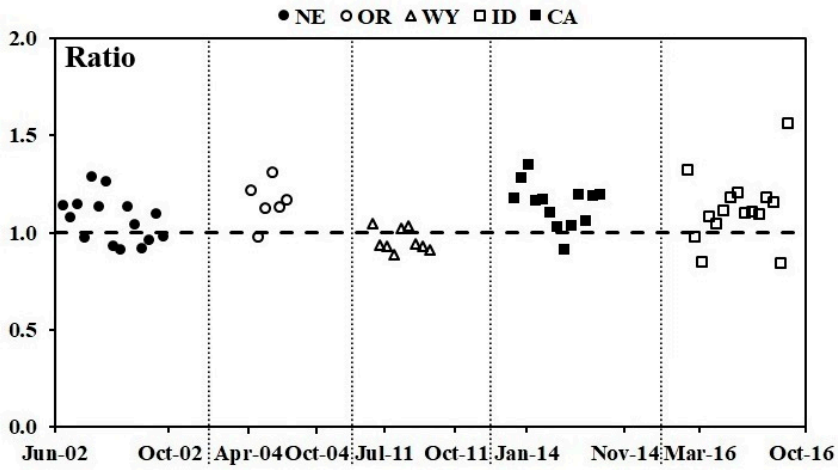

Figure 9. Ratios of calculated 24-hour ET_r used in EEFlux (based on gridded weather data) to that used in the METRIC model (calculated from ground-based weather station data) for five different Landsat scene locations and comparison days.

5.2 Impact of METRIC Calibration Style (User) on METRIC Estimation

Some of the differences noted between ET_rF and ET_a from EEFlux vs. METRIC could stem from the semi-subjective behavior for METRIC estimates that are traceable to the particular individual user and situation responsible for the METRIC application and calibration. To explore the impact of METRIC user, two different METRIC users with varying experience and expertise in ET image production applied similar METRIC algorithms independently during two different time periods, where they calibrated two image dates in central Nebraska (Path 29 Row 32) for year 2015. Figure 10 shows the results of comparisons for two processed Landsat 8 image dates for the agricultural land use. The top two comparisons belong to 18^{th} of July and the two in the bottom belong to 4^{th} of September. While R^2 of ET_rF and ET_a values are higher than 0.89 for both days, the RMSE and slope values are considered to be acceptable for only July 18^{th}, and is not in the acceptable range for September 4^{th}. The average R^2 of ET_rF and ET_a values for combination of all the data were 0.78 and 0.73, respectively. The combined slope values were 0.9 for ET_rF and 1.07 for ET_a values, which do fall within the acceptable ranges. Scatter in the comparisons is due to small differences in the METRIC version used or in internal parameter settings in METRIC such as corrections for low albedo in crops such as corn that have deep canopies [7]. Combined RMSE values were 0.14 for ET_rF and 0.98 mm/d for ET_a values. A comparison of these average R^2, slope and RMSE values with average values for EEFlux vs. METRIC summarized in Table.1, suggests that, for the locations evaluated, that the EEFlux automated calibration algorithm is generally able to estimate ET_rF and ET_a values for agricultural land uses that are comparable in accuracy and reproducibility to differences noted from METRIC when applied by different trained users. This finding is consistent with that of Medellín-Azuara et al., [14].

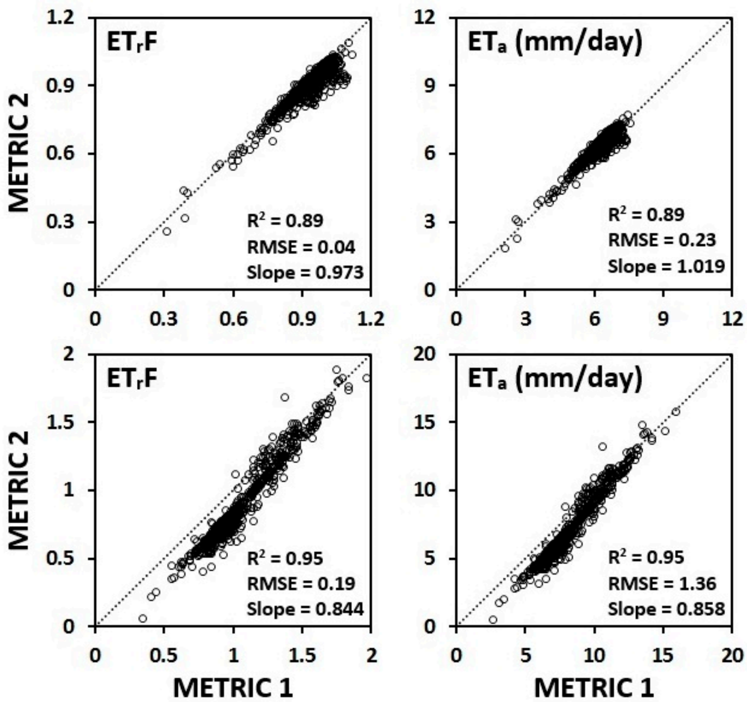

Figure 10. Comparison between METRIC products (ET$_r$F and ET$_a$) that were manually calibrated and produced by 2 different METRIC users. The top two comparisons are for 18th of July and the bottom two are for 4th of September.

5. Conclusions

The consistency and accuracy of ET products from the automatically calibrated GEE EEFlux application were evaluated by comparing EEFlux products to those from manually calibrated METRIC images for 58 Landsat images. Sets of Landsat images from five study locations distributed across central and western USA included both agricultural and non-agricultural land uses. The agricultural areas sampled were typically irrigated. The comparison results show that EEFlux is able to calculate ET$_r$F and ET$_a$ values in agricultural areas that are comparable to those produced by trained METRIC users and that are generally within accepted accuracy ranges. Differences between EEFlux and METRIC were larger for non-agricultural land uses showing room for improvement to the EEFlux algorithms. Differences noted could, in part, be the result of EEFlux struggling to account for background evaporation at the hot pixel calibration end point. Hot pixel bias in the hot pixel assigned ETrF tends to affect the non-agricultural pixels more than agricultural pixels because the non-agricultural pixels tend to have lower ET and are therefore more impacted by error or bias in the overall surface energy balance. Another likely reason for the poorer performance for non-agricultural land uses is a bias introduced during the application of EF to extrapolate instantaneous ET_rF to daily ET_rF, as discussed earlier. The EF relies on the instantaneous and 24-hour ET_r, R_n and G being accurate. We have established that both ET_r and G estimates deviate between METRIC and EEFlux, so we would expect to have different results in the non-agricultural areas. In fact, we should expect larger differences between METRIC and EEFlux in non-agricultural areas than in agricultural areas given that the instantaneous ET_rF used in the agricultural areas is robust in the face of biased G and instantaneous ET_r. While EEFlux is still a work in progress, it can be used to rapidly estimate ET_a for areas of interest. However, it is important to be aware of biases in 24-hour ET_a estimates due to aridity biases in the gridded weather data used by EEFlux. Results presented in this paper should provide a good overview of the general variability and error to be expected for ET_rF and ET_a estimates from EEFlux.

Acknowledgments

We acknowledge and greatly appreciate funding support and development support provided by Google Earth Engine, as well as very positive encouragement by Google Earth Engine to initiate the EEFlux development to promote global access to global ET estimates. EEFlux development was also supported by the National Aeronautics and Space Administration (NASA) and by the United States Geological Survey (USGS) by funding Landsat Science Team members Allen, Kilic and Huntington. The USGS and state and local water resources agencies, including Idaho Department of Water Resources, the Wyoming Office of the State Engineer, and the Central Platte Natural Resources District supported the development and production of the original METRIC ET products. Work was also supported by the Idaho and Nebraska Agricultural Experiment Stations.

Author Contributions: Conceptualization, F. F.; Data curation, F. F., P. B., S. O., and P. R.; Formal analysis, F. F.; Funding acquisition, A. K., and R. G. A.; Investigation, F. F., P. B., and R. G. A.; Methodology, F. F., P. B., A. K., R. G. A., J. L. H., D. O., C. G. M., T. E. F., B. K., and R. T.; Project administration, A. K., and R. G. A.; Resources, F. F., P. B., J. L. H., T. A. E., S. O., I. R., D. T., R. M., N. G., P. R, R. T., W. Z., and C. W. R; Software, P. B., T. A. E., D. O., C. G. M., S. O., I. R., D. T., R. M., N. G., and B. K.; Supervision, A. K., R. G. A., J. L. H., and T. A. E.; Visualization, F. F.; Writing – original draft, F. F., P. B., A. K., R. G. A., T. E. F., and P. R.; Writing – review & editing, F. F., P. B., A. K., R. G. A., T. E. F., and P. R.

Conflicts of Interest: The authors declare no conflict of interest.

References

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

2. Anderson, M. C.; Kustas, W. P.; Norman, J. M.; Hain, C. R.; Mecikalski, J. R.; Schultz, L.; González-Dugo, M. P.; Cammalleri, C.; d'Urso, G.; Pimstein, A.; Gao, F. Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrol Earth Syst Sci* **2011**, *15*, 223–239, doi:10.5194/hess-15-223-2011.

3. Bastiaanssen, W. G. M. Remote sensing in water resources management: the state of the art. *Remote Sens. Water Resour. Manag. State Art* **1998**.

4. Courault, D.; Seguin, B.; Olioso, A. Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches. *Irrig. Drain. Syst.* **2005**, *19*, 223–249, doi:10.1007/s10795-005-5186-0.

5. Kustas, W. P.; Norman, J. M. Use of remote sensing for evapotranspiration monitoring over land surfaces. *Hydrol. Sci. J.* **1996**, *41*, 495–516, doi:10.1080/02626669609491522.

6. Morton, F. I. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *J. Hydrol.* **1983**, *66*, 1–76, doi:10.1016/0022-1694(83)90177-4.

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

8. Allen, R. G.; Tasumi, M.; Trezza, R.; Kjaersgaard, J. METRIC: mapping evapotranspiration at high resolution–applications manual for Landsat satellite imagery. *Univ. Ida.* **2005**, 130.

9. Bastiaanssen, W. G. M.; Menenti, M.; Feddes, R. A.; Holtslag, A. A. M. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *J. Hydrol.* **1998**, *212*, 198–212, doi:10.1016/S0022-1694(98)00253-4.

10. Allen, R. G.; Morton, C.; Kamble, B.; Kilic, A.; Huntington, J.; Thau, D.; Gorelick, N.; Erickson, T.; Moore, R.; Trezza, R.; Ratcliffe, I.; Robison, C. EEFlux: A Landsat-based Evapotranspiration mapping tool on the Google Earth Engine. In; American Society of Agricultural and Biological Engineers, 2015; pp. 1–11.

11. Allen, R. G.; Tasumi, M.; Morse, A. T.; Trezza, R.; Wright, J. L.; Bastiaanssen, W. G. M.; Kramber, W.; Lorite, I.; Robison, C. W. Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)—Applications.

J. Irrig. Drain. Eng. **2007**, *133*, 395–406, doi:10.1061/(ASCE)0733-9437(2007)133:4(395).

12. Geli, H. M. E.; Neale, C. M. U.; Verdin, J. P. Estimating Crop Water Use with Remote Sensing: Development of Guidelines and Specifications – Part 2: Evapotranspiration Model Intercomparison. *US Geol. Surv. Sci. Investig. Rep. 2017* **2017** (under review).
13. Irmak, A.; Ratcliffe, I.; Ranade, P.; Hubbard, K. G.; Singh, R. K.; Kamble, B.; Kjaersgaard, J. ESTIMATION OF LAND SURFACE EVAPOTRANSPIRATION WITH A SATELLITE REMOTE SENSING PROCEDURE. *Gt. Plains Res.* **2011**, *21*, 73–88.
14. Medellín-Azuara, J.; Paw U, K. .; Jin, Y.; Lund, J. R. A Comparative Study for Estimating Crop Evapotranspiration in the Sacramento-San Joaquin Delta. **2018**.
15. Morton, C. G.; Huntington, J. L.; Pohll, G. M.; Allen, R. G.; McGwire, K. C.; Bassett, S. D. Assessing Calibration Uncertainty and Automation for Estimating Evapotranspiration from Agricultural Areas Using METRIC. *JAWRA J. Am. Water Resour. Assoc.* **2013**, *49*, 549–562, doi:10.1111/jawr.12054.
16. Tasumi, M.; Trezza, R.; Allen, R. G.; Wright, J. L. Operational aspects of satellite-based energy balance models for irrigated crops in the semi-arid U.S. *Irrig. Drain. Syst.* **2005**, *19*, 355–376, doi:10.1007/s10795-005-8138-9.
17. Irmak, A.; Allen, R. G.; Kjaersgaard, J.; Huntington, J.; Kamble, B.; Trezza, R.; Ratcliffe, I. Operational Remote Sensing of ET and Challenges. *Weed Technol.* **2012**, doi:10.5772/25174.
18. Allen, R. G.; Burnett, B.; Kramber, W.; Huntington, J.; Kjaersgaard, J.; Kilic, A.; Kelly, C.; Trezza, R. Automated Calibration of the METRIC-Landsat Evapotranspiration Process. *JAWRA J. Am. Water Resour. Assoc.* **2013**, *49*, 563–576, doi:10.1111/jawr.12056.
19. Crago, R. D. Conservation and variability of the evaporative fraction during the daytime. *J. Hydrol.* **1996**, *180*, 173–194, doi:10.1016/0022-1694(95)02903-6.
20. ASCE–EWRI. 2005. The ASCE standardized reference evapotranspiration equation. ASCE–EWRI Standardization of Reference Evapotranspiration Task Committee Rep., ASCE Reston, Va. 120 p.
21. Cosgrove, B. A.; Lohmann, D.; Mitchell, K. E.; Houser, P. R.; Wood, E. F.; Schaake, J. C.; Robock, A.; Marshall, C.; Sheffield, J.; Duan, Q.; Luo, L.; Higgins, R. W.; Pinker, R. T.; Tarpley, J. D.; Meng, J. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res. Atmospheres* **2003**, *108*, doi:10.1029/2002JD003118.

22. Abatzoglou, J. T. In *Development of gridded surface meteorological data for ecological applications and modelling*; 2013; Vol. 33.

23. Saha, S.; Moorthi, S.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Behringer, D.; Hou, Y.; Chuang, H.; Iredell, M.; Ek, M.; Meng, J.; Yang, R.; Mendez, M. P.; van den Dool, H.; Zhang, Q.; Wang, W.; Chen, M.; Becker, E. The NCEP Climate Forecast System Version 2. *J. Clim.* **2013**, *27*, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.

24. Yuan, X.; Wood, E. F.; Luo, L.; Pan, M. A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophys. Res. Lett.* **2011**, *38*, doi:10.1029/2011GL047792.

25. Saha, S.; Moorthi, S.; Pan, H.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Kistler, R.; Woollen, J.; Behringer, D.; Liu, H.; Stokes, D.; Grumbine, R.; Gayno, G.; Wang, J.; Hou, Y.; Chuang, H.; Juang, H. H.; Sela, J.; Iredell, M.; Treadon, R.; Kleist, D.; Van Delst, P.; Keyser, D.; Derber, J.; Ek, M.; Meng, J.; Wei, H.; Yang, R.; Lord, S.; van den Dool, H.; Kumar, A.; Wang, W.; Long, C.; Chelliah, M.; Xue, Y.; Huang, B.; Schemm, J.; Ebisuzaki, W.; Lin, R.; Xie, P.; Chen, M.; Zhou, S.; Higgins, W.; Zou, C.; Liu, Q.; Chen, Y.; Han, Y.; Cucurull, L.; Reynolds, R. W.; Rutledge, G.; Goldberg, M. The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.* **2010**, *91*, 1015–1058, doi:10.1175/2010BAMS3001.1.

26. Allen, R. G.; Walter, I. A.; Elliott, R. L.; Howell, T. A.; Itenfisu, D.; Jensen, M. E.; Snyder, R. L. *The ASCE Standardized Reference Evapotranspiration Equation*; ASCE Publications, 2005; ISBN 978-0-7844-7563-8.

27. Blankenau, P. Bias and Other Error in Gridded Weather Data Sets and Their Impacts on Estimating Reference Evapotranspiration. *Civ. Eng. Theses Diss. Stud. Res.* **2017**.

28. Perrier, A. *Land surface processes: vegetation. In land surface processes in atmospheric general circulation models, Eagelson P (ed.). Cambridge University Press, Cambridge. 395–448.*; Cambridge University Press, Cambridge, 1982;

29. Allen, R. G.; Pereira, L. S.; Howell, T. A.; Jensen, M. E. Evapotranspiration information reporting: I. Factors governing measurement accuracy. *Agric. Water Manag.* **2011**, *98*, 899–920, doi:10.1016/j.agwat.2010.12.015.

30. Gonzalez-Dugo, M. P.; Neale, C. M. U.; Mateos, L.; Kustas, W. P.; Prueger, J. H.; Anderson, M. C.; Li, F. A comparison of operational remote sensing-based models for estimating crop evapotranspiration. *Agric. For. Meteorol.* **2009**, *149*, 1843–1853, doi:10.1016/j.agrformet.2009.06.012.

31. Kalma, J. D.; McVicar, T. R.; McCabe, M. F. Estimating Land Surface Evaporation: A Review of Methods Using Remotely Sensed Surface Temperature Data. *Surv. Geophys.* **2008**, *29*, 421–469, doi:10.1007/s10712-008-9037-z.

 doi:10.20944/preprints201807.0040.v1

32. Jensen, M. E.; Allen, R. G. *Evaporation, Evapotranspiration, and Irrigation Water Requirements*; American Society of Civil Engineers., 2016;

33. Temesgen B.; Allen R. G.; Jensen D. T. Adjusting Temperature Parameters to Reflect Well-Watered Conditions. *J. Irrig. Drain. Eng.* **1999**, *125*, 26–33, doi:10.1061/(ASCE)0733-9437(1999)125:1(26).

34. Lewis, C. S.; Geli, H. M. E.; Neale, C. M. U. Comparison of the NLDAS Weather Forcing Model to Agrometeorological Measurements in the western United States. *J. Hydrol.* **2014**, *510*, 385–392, doi:10.1016/j.jhydrol.2013.12.040.

35. Anderson, M. C.; Kustas, W. P.; Alfieri, J. G.; Gao, F.; Hain, C.; Prueger, J. H.; Evett, S.; Colaizzi, P.; Howell, T.; Chávez, J. L. Mapping daily evapotranspiration at Landsat spatial scales during the BEAREX'08 field campaign. *Adv. Water Resour.* **2012**, *50*, 162–177, doi:10.1016/j.advwatres.2012.06.005.