Establishing and Operating a Regional Network for Field Measurement and Sensing of Crop Water Consumption (Evapotranspiration)

NENA Regional ET-Network

Output 9: Comparison Between Field Data and Remote Sensing Data Analyzed

Comparing Evapotranspiration Retrieved Through Various Remote Sensing based Models with Ground Measured Data

Submitted by
International Center for Agricultural Research in the Dry Areas (ICARDA

to
Food and Agriculture Organization of the United Nations (UN-FAO)
23 December 2021

Authors, Contributors and Citation:
Biradar, C., Krishna, G., Govind, A., Nangia, V., Gamal, R., Zitouna, R., Jomaa, I., Mazahrih, N., Mosad, A., El Meknassi, E., Mansouri, A., Haddad, M., Steduto, P., Swelam, A., Zaib, N., (2022). Comparing Actual Evapotranspiration Retrieved Through Various Remote Sensing based Models with Ground Measured Data, Technical Report. International Center for Agricultural Research in the Dry Areas (ICARDA), Cairo, Egypt. ICARDA, 2 Port Said St, Victoria Square, Ismail El-Shaaer Building, Maadi, Cairo, Egypt

DISCLAIMER

Table of Contents

Comparing Evapotranspiration Retrieved through Various Remote Sensing based Models with Ground measured data

Summary: Evapotranspiration (ET) measurement on a real-time basis at a large scale with high temporal frequency is important for agricultural water use and water resource planning. Identifying the differences between potential ET and crop ET at fine spatial resolutions is essential for management of agriculture and water resources. It could contribute to the water deficit assessment at the individual farm level. This study evaluates the potential of remote sensing-based models at five different locations in the Near East and North African (NENA) region. Ground-based ET was determined for these locations during different crop growing seasons. Remote sensing-based ET estimates generated through three remote sensing-based models (WaPOR, METRIC, and SEBAL) were compared with ground-based estimates. This comparative analysis evaluated the remote sensing-based modelling approaches and identified the model that generated the closest estimates to the ground estimates. Results showed that SEBAL and METRIC performed well except few locations. SEBAL showed highest index of agreement (d) (0.49 to 0.91) between observed and estimated ET data followed by METRIC (0.40 to 0.66) and WaPOR (0.31 to 0.61) for all seasons. Overall, the model's performance parameters showed that SEBAL had the smallest nRMSE and uncertainty and the largest Ratio of Performance to Deviation (RPD) and (d) compared to the other two models. RPD has been used as the standard way to report the quality of a model. Therefore, the results of this study reveal that SEBAL performed well for study locations followed by METRIC and WaPOR. Another finding for the study locations is that the ground ET determination technique based on soil moisture depletion (SMD) was less correlated with RS-based ET than other techniques.

Keywords: Evapotranspiration, ET, RS ET, WaPOR, SEBAL, METRIC, Field Data,

Comparing Evapotranspiration Retrieved through Various Remote Sensing based Models with Ground-based Data

1. Introduction

ET refers to the evapotranspiration of a crop assessed under diverse conditions of water availability, including situations where the ET rate might have been limited but soil water, thus exhibiting values below its maximum potential. Quantification of water consumption at a large scale, specifically for irrigated areas is quite vital for water resources planning, judicious water use, and water regulation. For arid and semiarid Mediterranean environments, evapotranspiration (ET) is one of the most critical mechanisms of the energy budget and land-surface water. Ecosystem complexity can be sensed effectively by ET variability. There are many ground-based ET determination techniques like eddy covariance (Aguilos et al., 2019; Baldocchi et al., 2001), scintillometry (Hemakumara et al., 2003), soil water balance (Cuenca et al.,1997; Eastham et al., 1988), the Bowen ratio (Angus and Watts 1984), soil moisture depletion, surface renewal (Paw et al., 1995) and soil-weighing lysimeters (Edwards, 1986; Jensen et al., 1990). These techniques provide ET approximations of varying degrees of accuracy and precision. To scale up ET to a very large scale, we need a more comprehensive approach aligned with such methods. To this end, remote sensing could be a great tool because of its wide area coverage and synoptic view, particularly if offering high temporal and fine spatial resolution data availability. Spatial evapotranspiration (ET) estimates are essential components to get an idea about soil moisture. In countries where field size is quite small, high spatial resolution ET can play a crucial role for water advisories to the farmers. The energy conservation principle can be applied to the surface energy balance for retrieval of ET. Several models provide an approximation of evapotranspiration as a residual term of the land surface energy balance, including SEBAL, METRIC, ETLOOK-WaPOR, etc. (Ynag et al., 2013; Wagle et al., 2017; Senay et al., 2013; Allen et al., 2007; Bastiaanssen et al., 1998). Besides this there is one model (ETLook) that uses soil moisture estimates from the passive microwave sensor AMSRE (Bastiaanssen et al., 2012; Pelgrum et al. 2012). This study provides a good intercomparison between popularly used Remote Sensing based ET models and ground estimated ET. In this study, METRIC and WaPOR based ET (ET) has been acquired from an already developed platform by Allen et al. and FAO, respectively, whereas we have developed the SEBAL model for actual ET estimation by utilizing the power of the Google cloud computing platform Google Earth Engine. This study also correlates multiple RS-based ET estimates with ground-based ET estimates. This study has a main objective of evaluating the potential of RS-based models by identifying the best remote sensing (RS) based ET estimation approach for the NENA region by comparing RS-based ET observation with ground-based ET determinations at multiple locations, multiple seasons, and different crops. The SEBAL model-based RS-ET is high spatial and temporal resolution ET and could be useful for determining crop water consumption for small holder farming system.

2. Material and Methods

2.1 Field Data Locations

In Egypt, the ET study location is situated at Sakha governorate farms in the Nile River plain. The ET study location in Jordan is in Dyar Ala District. Mountains cover this area from both sides. In Lebanon, the ET study site is in Tal Amara. Mountains cover this area too from both sides. In Morocco, the ET study site is located in Birched, Casablanca. This area has mountainous terrain in a nearby location. The ET facility is located at Gendouba in Tunisia. Spatial location of the study sites has been displayed in *Figure 1*. *Table 2* shows the lat long positions and ground-based ET estimation method for each site.

Figure 1: Locations where ET data for this study was obtained. Points and triangles depict exact ground observation locations.

Table 1: Countries, organizations and the sites and the on-going methodologies in Season 2

Country	Field Station & Institution	Latitude/Longitude	ET_a Method /Crop in Season-2
Egypt	Sakha, Nile Delta (ARC)	31°15’52.00’N 30°46’06.00’’E	- Surface Energy Balance (using Eddy Covariance for sensible heat flux - H) - Surface Energy balance (CORDOVA-ET)
Jordan	Dyar Ala, Jordan Valley (NARC)	32°19'00.00'N 35°34'43.00''E	- Weighing lysimeter - Surface Energy balance (CORDOVA-ET)

Lebanon	Tal Amara (LARI)	33°51'51.00'N 35°59'05.00''E	- Soil moisture depletion - Surface Energy balance (CORDOVA-ET)
Morocco	Birched, Casablanca (private farm)	33°34'12.00'N 7°37'13.00''E	- Soil moisture depletion - Eddy Covariance for latent heat flux (λE) - Surface Energy balance (CORDOVA-ET)
Tunisia	Gendouba (INGC-INRGREF)	36°32'47.83'N 9°00'50.00''E	- Eddy covariance for latent heat flux (λE) - Surface Energy Balance (using Eddy Covariance for sensible heat flux - H) - Surface Energy balance (CORDOVA-ET)

2.2 Description of the field sites

There are several methods to determine ET. Typically, only one is considered direct: (i) the water budget, which includes the weighing lysimeters and the soil water balance methods. The water vapor transfer methods, which includes eddy covariance and Bowen ratio are less direct in that they sense properties of the air well above the crop and soil surfaces where evapotranspiration occurs. The surface energy balance can also be used to estimate ET and is considered an indirect method because it depends on estimates of net radiation, sensible heat, and soil heat fluxes so that ET can be estimated as the residual of the energy balance. The four methods that the NENA-ETNet utilized for the field ET determinations in the second season (Summer 2020) are: 1) Eddy covariance; 2) Weighing Lysimeters; 3) Soil water balance, and 4) Surface Energy Balance. Each station in all selected five countries has more than one ET determination facility. There were several issues associated with the repair and maintenance of the equipment. The complete knowledge on the functionality of the instruments was of paramount importance in helping to address the repair and maintenance issues during the first season and to keep the equipment in good shape to have a full crop cycle measurement in the following seasons and to ensure quality data among all countries. These sites also have the CORDOVA-ET Stations deployed. Briefly, the CORDOVA-ET station contains the sensors needed to calculate the energy balance components of the crop surface under consideration. The field measurement in the NENA-ETNet Network began in the winter season 2019/2020 using the existing ET facilities. This report is targeted to elaborate on the obtained results from the second crop season in the five countries of the Network.

According to the protocol established for data acquisition and reporting, the country coordinators gathered the four seasons of field data at three levels of detail, L1, L2 and L3. The L1 and L2 depict half-hourly data. L1 and L2 are similar with the difference that L2 Data is gap-filled using standard methodologies (for those gaps that are less than three hours long during the day). L3 is the highest level of data processing and is available at a daily time step. This is the data used for inter-comparison between different instruments (e.g., Eddy covariance and CORDOVA-ET) and between other estimates (e.g., ET_EC from eddy covariance and ET_Lysi from weighing lysimeter). This L3 product should be used for calibration of remote sensing algorithms and testing of various products generated based on those algorithms. It can also be used to calibrate and test crop simulation models such as APSIM, CropSyst or AquaCrop with the whole set of valuable data and the metadata reported from each site for each crop and season. L2 datasets may be used to calibrate and test crop models that operate at sub-daily time steps.

There is a rigorous peer-review process, and iterative improvement of the initially submitted L1, L2 and L3 data before the Network arrives at a finalized set of datasets for a given season. The submitted data are heavily scrutinized regarding data gaps, magnitudes, trends and reporting style accuracy. The spurious data acquired due to methodological and sensor errors are kept as-is so that it helps us to analyze the accuracy and efficacy of the methods of data acquisition (e.g., ET values obtained by a given process or system). The finalized data arearchived for data analysis and synthesis.

2.3 Ground-based Evapotranspiration (ET) Estimation

Evapotranspiration was determined for every country using the best estimation technique for that particular area. For each location, observed ETa data was retrieved for three seasons from December 2019 to May 2021. *Table* 2 shows the month-wise exact span of the seasons. Season 3 and 4 had continuous and more data points (*Figure* 2 and *Figure 3*) than seasons 1 and 2 for ground-based ET values.

Table 2: Month-wise classification of multiple crop seasons for all five study locations with name of crop for that season

Country	Season-1	Season-2	Season-3	Season-4
Egypt	Winter Wheat *(Dec 1,2019 – May 13, 2020)*	Summer Maize *(July 20, 2020 – Oct 21, 2020)*	Winter Wheat *(Nov. 25, 2020 – April 30,2021)*	Summer Rice *(June 16, 2021 – Oct. 20, 2021)*
Jordan	Winter Wheat *(Dec 25, 2019 – May 5, 2020)*	Summer Maize *(July 15, 2020 – Oct. 20, 2020)*	Fodder Vetch *(Jan. 13, 2021 – April 30, 2021)*	Maize *(June 14, 2021 – Sep. 9, 2021)*
Lebanon	Wheat *(Dec 7, 2019 – July 2, 2020)*	Potato- Fallow *(March 1, 2020 – July 31, 2020)*	Faba bean *(Dec. 3, 2020 – May 6, 2021)*	Maize *(June 17, 2021 – Oct. 8, 2021)*
Morocco	Maize *(Feb 23,2020 – July 3, 2020*	Beetroot *(Aug. 27, 2020 – Nov. 11, 2020)*	Durum Wheat *(Jan 11, 2021 – May 31, 2021)*	NA
Tunisia	Wheat *(Dec 3, 2019 – June 23, 2020)*	Maize *(July 19,2020 – Nov 3, 2020)*	Faba bean *(Dec. 19, 2020 – May 26, 2021)*	Sorghum *(Aug. 9, 2021 – Nov. 2, 2021)*

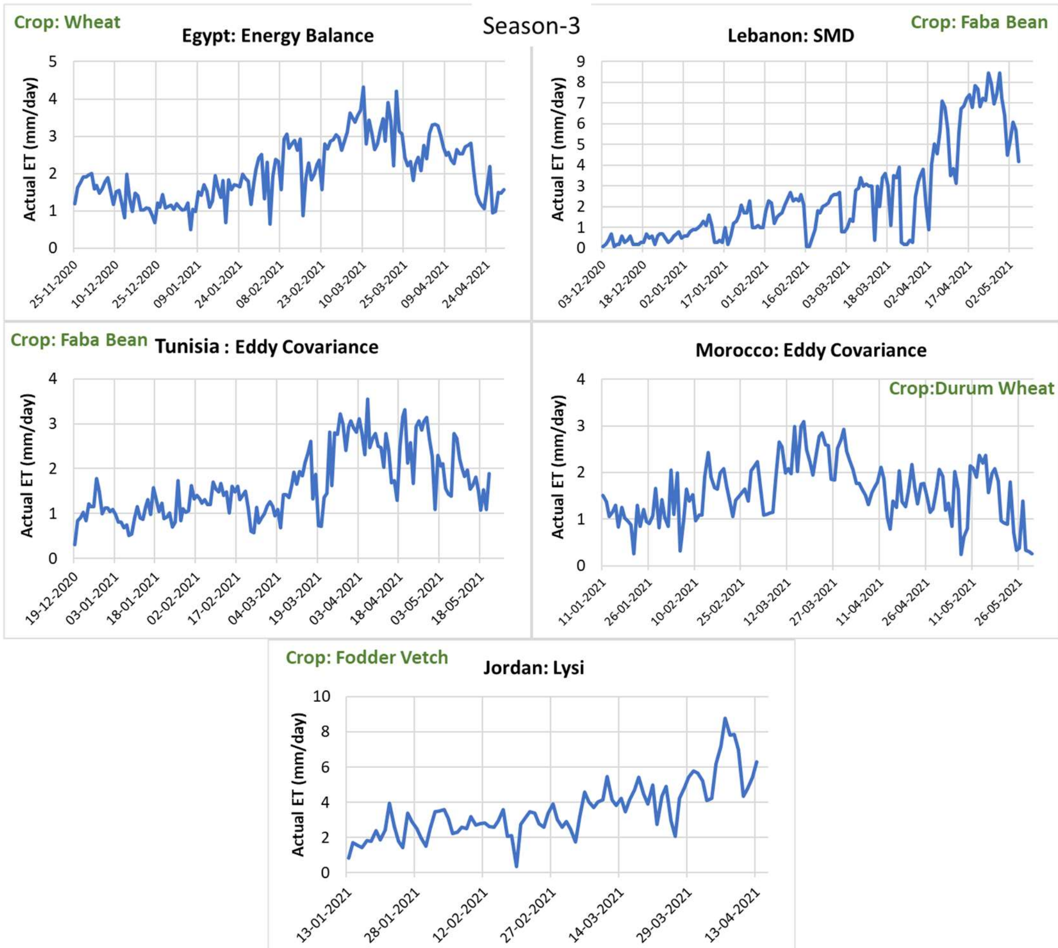

Figure 2: Ground-based ET plots for season 3.

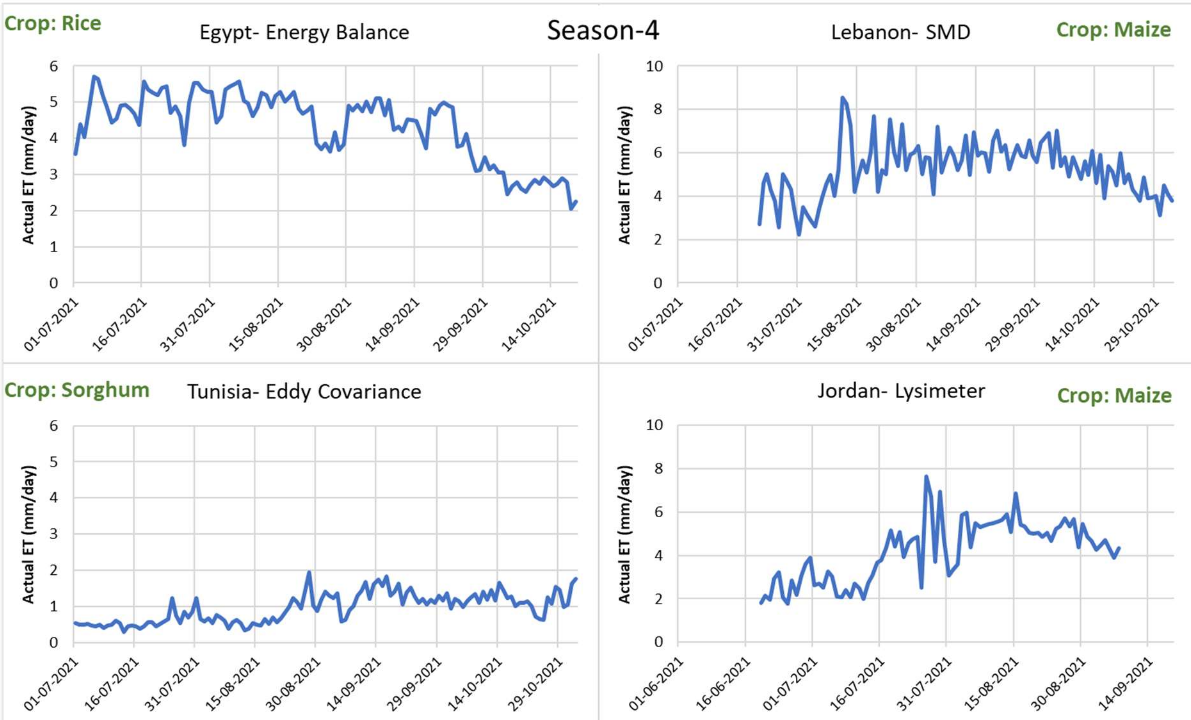

Figure 3: Ground-based ET plots for season 4.

Based on the recommendations of the agrometeorologist and ground-based teams, the ET estimation method for each country was decided. In Egypt, the energy balance method was the chosen observed ET estimation method, whereas, for Morocco and Tunisia, the Eddy Covariance method was chosen. In Jordan, Lysimeter-derived estimates of the ETa were chosen, and for Lebanon, the Soil Moisture Deficit (SMD) technique was chosen for ET observations.

3. Remote Sensing ET Models' Description

3.1 Water Productivity Open-access Portal (WaPOR)

The FAO developed WaPOR portal provides ET data for continental to national levels (*WaPOR, FAO*). Sub-national level ET data is available for very few experimental locations at 30 meters spatial resolution. The continental ET product is at 250 meters spatial resolution, and the national level estimate is at 100 meters. At present, the portal offers ET for dekadal, annual, and monthly frequencies for Africa and Near East regions. For our comparative analytics, we used national-level products at a monthly frequency. This was the best dataset for study countries on WaPOR (available for download) as sub-national level products don't cover our study area sites.

As per the WaPOR portal's description, from January 2020 onwards, all the base input layers (NDVI, albedo and fAPAR) for 100 m products were derived from the Copernicus Sentinel-2 satellite data. Before this until December 2019, Proba-V satellite data were used for the same input data.

The WaPOR ET estimate is based on the ET-Look model described by Bastiaanssen et al. (2012). Here, the ET is the sum of the soil evaporation (E), canopy transpiration (T), and evaporation from rainfall intercepted by leaves (I). The monthly total is obtained by taking the ET in mm/day, multiplying by the number of days in a decade, and summing the decades of each month. The broad approach for ET-Look (WaPOR) ETa computation is given in *Figure 4*.

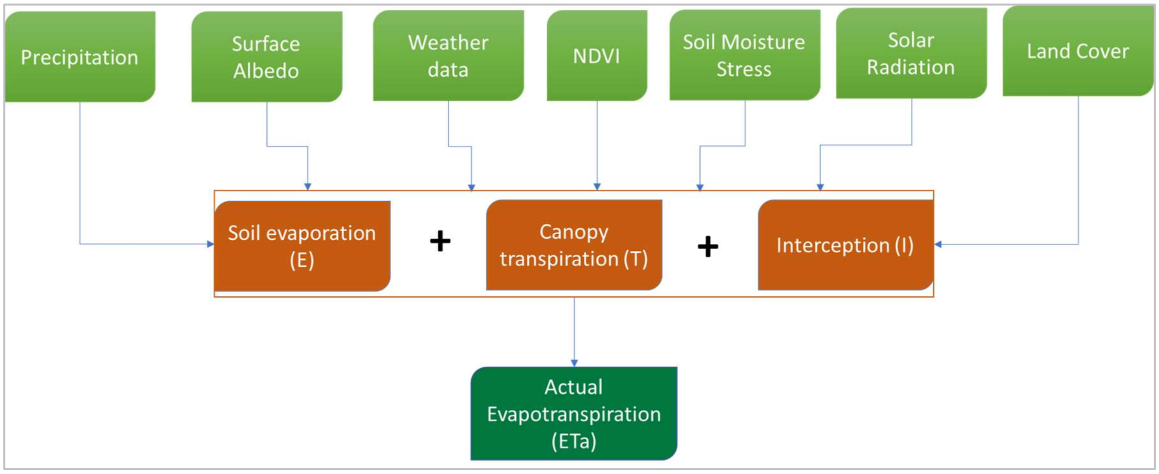

Figure 4: Brief indicative Methodology for ET estimation using ET-Look model, ET data available on Water Productivity Open-access Portal (WaPOR)

3.2 Mapping EvapoTranspiration at high Resolution with Internalized Calibration (METRIC)

The METRIC model was developed based on SEBAL (Surface Energy Balance Algorithm for Land) algorithm, which was proposed by Bastiaanssen et al. in 1998. The development of METRIC was started in 2000. Broad methodological steps are given in *Figure 5*. The SEBAL model was developed to model ET with minimum ground-based measurements. METRIC uses the alfalfa reference for calibration because of the near-maximum ET represented by alfalfa. The primary difference between SEBAL and METRIC is that the Reference ET is used to calibrate sensible heat flux (H) and near surface temperature difference (dT) in METRIC but not in SEBAL (Allen & Kilic et al., 2021).

In this study, we downloaded the METRIC based ET using the EEFLUX Google Earth Engine (GEE) driven app. This platform uses Landsat 7 and Landsat 8 satellite data (30-meter spatial resolution) and its thermal bands (120-meter spatial resolution for Landsat 7 and 100-meter for Landsat 8). Due to the use of both 7 and 8 series, imagery is available at eight-day intervals. The Landsat-7 Data has stripped gaps. The EEFLUX GEE app allows users to download the ET imagery for the area of interest. Then, using a GIS overlay, the pixel value of the point of interest can be extracted. The resolution of the ET product is 30 meters.

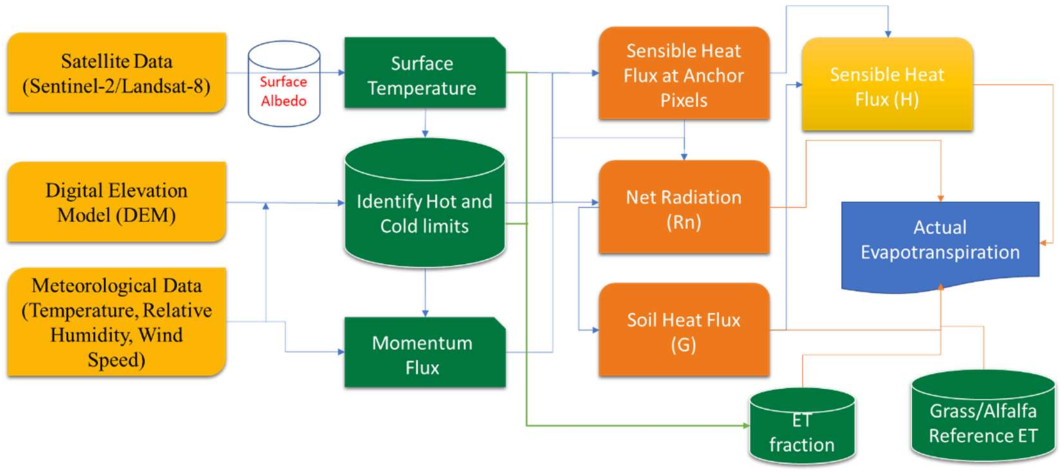

Figure 5: Broad steps for SEBAL and METRIC models

3.3 Surface Energy Balance Algorithm for Land (SEBAL)

A brief description of the SEBAL model can be expressed as given below-

$$LE = R_n - G - H \qquad \text{EQUATION 1}$$

where LE is the latent energy consumed by ET, Rn is net radiation (sum of all incoming and outgoing short-wave and long-wave radiation at the surface), G is sensible heat flux into and out of the ground, and H is sensible heat flux transmitted to the air. Energy absorbed into the canopy and photosynthesis is generally less than a few % and is ignored (Allen et al., 2011).

The model was scripted in GEE (*Figure 6*, *Figure 7*) by utilizing Sentinel-2 satellite data of 10-meter spatial resolution and MODIS Land Surface Temperature (LST) data of 1 km resolution. The ET estimate produced by this model has a 10-meter spatial resolution and 5 days temporal resolution. Note that the spatial and temporal resolutions of the three methods used were different, specifically 10 m, 30 m and 100 m spatial resolution at 5 days, 8 days and monthly frequency for SEBAL, METRIC and WaPOR respectively.

MODIS LST has a daily frequency and sentinel 2 series satellite data has a 5-day temporal resolution. Most of the time, we got only 2-4 (out of at least 6) sentinel-2 imageries due to cloud cover. This frequent availability of satellite data generates more probability of getting cloud free scenes compared to less frequent data. Land Surface Temperature (LST) at higher spatial resolution was obtained by downscaling MODIS LST through the disaggregation method. The approach suggested by Sánchez et al,2020 was tried to be implemented for the LST disaggregation. Although disaggregated to a lower resolution, there might be a variation in LST because of coarse pixel size. IT may impact ET values if there are a lot of variations in the LULC temperature due to different land use features. But We are mostly concerned with agricultural areas. That is a homogenous area, therefore, we didn't observe a higher degree of variation in LST values. With the help of MODIS LST and Landsat 7-8 combined LST, we intend to develop a 5-day interval interpolated LST product that we would utilize in our SEBAL Model to improve its efficiency. We are trying to generate a 10-meter spatial resolution LST with 5 days frequency by interpolating the Landsat-7 and 8 LST. This may improve the efficiency of our SEBAL model.

This high spatial and temporal resolution based and SEBAL model derived RS-ET can be used for crop water consumption mapping at smallholder farmers' level.

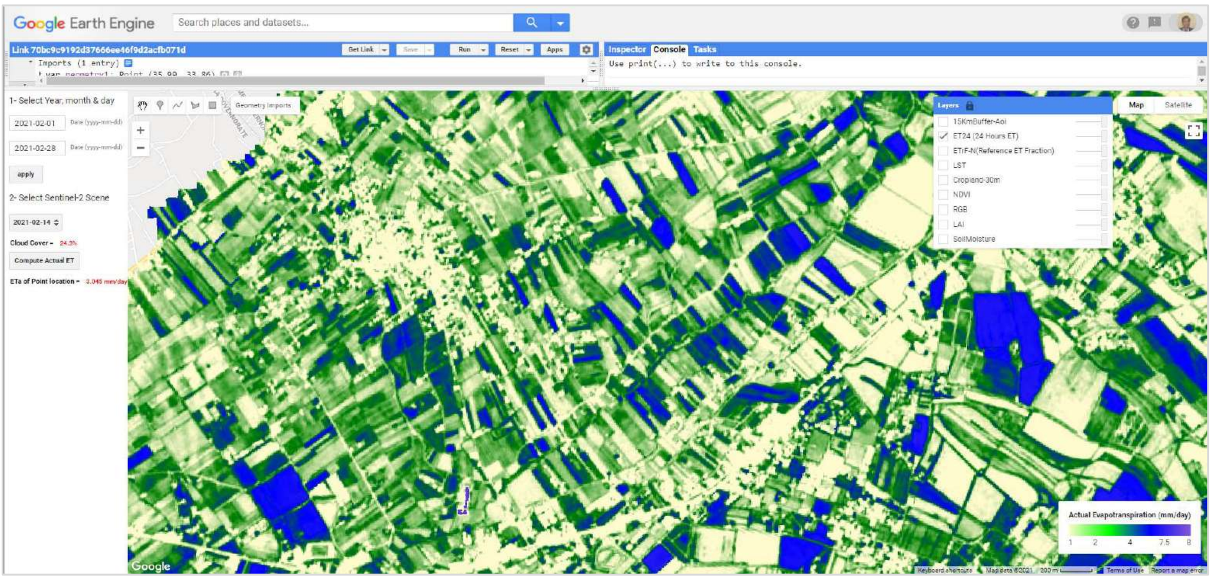

Figure 6: Screenshot of SEBAL GEE model

4. Remote Sensing based ET Data extraction and arrangement

The ET from the SEBAL model was extracted using the location point's pixel value extraction method. We also tried to extract SEBAL ET using a 3- by 3-pixel window and a 9- by 9-pixel window. But little variation in ET was observed for the Egypt location. Therefore, the point extraction method was utilized throughout the analysis. Similarly, the location point's pixel value extraction was applied to obtain the METRIC ET values. The WaPOR system enables a user to insert a latitude and longitude value of the point of interest, and then a time series can be generated by providing date ranges. Overall, all the RS-based ET were extracted using the location point's pixel value extraction method.

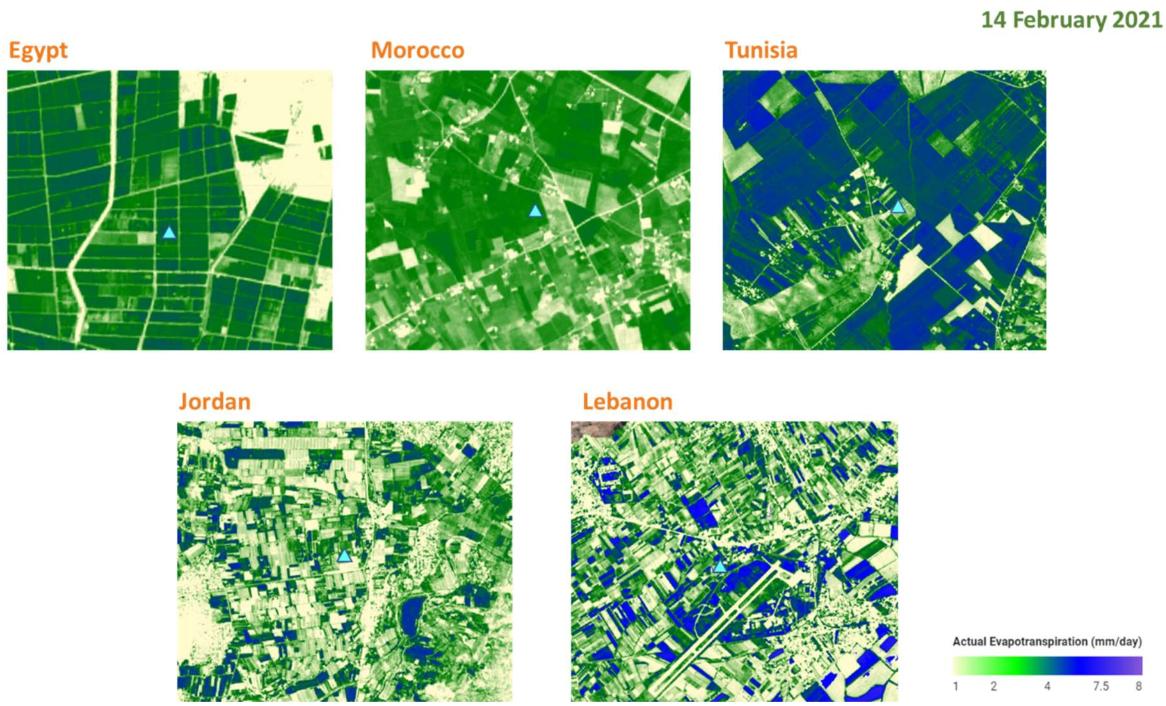

Figure 7: SEBAL GEE based evapotranspiration of study locations during peak vegetative stage of different crops during 2021. The triangle (cyan coloured) depicts the Ground ET observation site

4.1 Model's Performance Indicator

The regression modelling generates coefficient of determination (r^2), which is not an exhaustive method to evaluate the model's efficiency. To evaluate the model's performance, we computed normalized Root Mean Squared Error (nRMSE), uncertainty/standard error, Ratio of Performance to Deviation (RPD) and index of agreement (d). The index of agreement (d) proposed by Willmott (1981) is considered as a standardized measure of the degree of model prediction error which varies between 0 and 1 (AgriMetSoft (2019)). The value of d near 1 shows the model's perfect agreement between observed and estimated values. Smaller values of RMSE & uncertainty and larger values of RPD and d reflect a model's high efficiency. Mathematically, these parameters can be explained as given below-

$$normalizedRMSE\ (nRMSE) = \sqrt{\frac{\sum_{i=1}^{N}(O_i - P_i)}{N}} \Big/ \bar{O} \qquad \text{EQUATION 2}$$

$$Uncertainity\ = \frac{Std\ Dev}{\sqrt{N}} \qquad \text{EQUATION 3}$$

$$Uncertainity\ \% \ = \frac{Std\ Dev}{\sqrt{N}} \Big/ max \times 100 \qquad \text{EQUATION 4}$$

$$RPD\ = \frac{Std\ Dev}{RMSE} \qquad \text{EQUATION 5}$$

$$\boldsymbol{Index\ of\ Agreement\ (d)} = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(|P_i - \bar{O}| + O_i - \bar{O})^2} \quad \text{EQUATION 6}$$

Where O = Observed values, P = predicted/estimated value, N = number of data points.

5. Comparison of Remote Sensing based ET models with ground-based estimates

To determine the efficacy and best suitability of remote sensing (RS) based models, a detailed comparative analysis was carried out between each RS based model mentioned in this paper and observed ET from ground stations.

As mentioned in the model description section, the ET obtained from WaPOR was of monthly frequency. The ground-based ET (GB-ET henceforth) was of daily frequency for season-3 and roughly fortnightly for seasons 1 and 2. Here, to compare GB-ET with WaPOR ET, we used two approaches. In the first approach, we converted daily data to monthly data by summing all month days and comparing it to WaPOR ET. In another approach, to see the response and variability, we converted the monthly WaPOR ET data into daily data by dividing it by the number of days in a particular month and then comparing those daily ET values with GB-ET.

5.1 Remote Sensing based ET Models' intercomparison with Ground based ET for mutual satellite overpasses

The scatter plot depicting all locations in a plot for each model separately for each season and corresponding models' performance metrics were generated to identify the best modelling technique for ET estimation using remote sensing. Season wise plots, models' performance metrics and description has been given below. Here, this is to note that ground-based and RS-based ET values were selected for the dates of satellite pass only. As mentioned earlier, the RS-ET obtained from FAO's WaPOR is of monthly frequency. SEBAL based RS-ET was based on sentinel-2 satellite data. The frequency of the satellite image was 5 days and ET could be obtained at 5-day frequency except for the cloudy days. WaPOR monthly RS-ET data was scaled to the daily RS-ET using the daily reference ET of the respective locations. For daily WaPOR based ET generation, the monthly value of WaPOR based RS-ET was divided by the reference ET for that month. This output value was used for scaling the monthly ET to daily ET by multiplying it to the daily reference ET. The reference ET was calculated by site-specific weather parameters. For comparative plotting, the WaPOR derived daily ET values were obtained synchronized with the SEBAL ET dates because SEBAL ET dates are more compared to METRIC ET dates. This was done just to compare daily approximation of WaPOR, SEBAL and METRIC derived daily RS-ET (for the dates of satellite overpass only) with corresponding GB-ET. RS-ET values were compared with the ground-based ET values determined on the same day of satellite passing. METRIC based ET was also synchronized with SEBAL based ET dates, but a smaller number of values were obtained as METRIC is based on 8 days Landsat series satellite data and cloud cover during rainy season map widen the gap between two images. For Jordan, the least number of ET values can be obtained because here cloud-free satellite data were available for a smaller number of days compared to other study locations.

5.1.1 Observations for Season-1

For Season-1, scatterplots (*Figure 8*) and models' performance metrics (*Table 3*) were generated for all five study locations employing three RS-based ET estimation models. For Egypt's location, the largest R^2 (0.71) and lowest nRMSE (19%) were observed by using the SEBAL model. In fact, the SEBAL model generated the lowest nRMSE and better index of agreement (d) for each location. The overall nRMSE was the lowest for SEBAL compared to METRIC and WaPOR generated outputs. The overall RPD values were the best for SEBAL based ET followed by WaPOR and METRIC. The d values were largest for SEBAL followed by METRIC and WaPOR. The largest d value (near 1) denotes better agreement between ground-based and estimated data. The ET regression modelling for Lebanon showed the largest underprediction using all RS-based models.

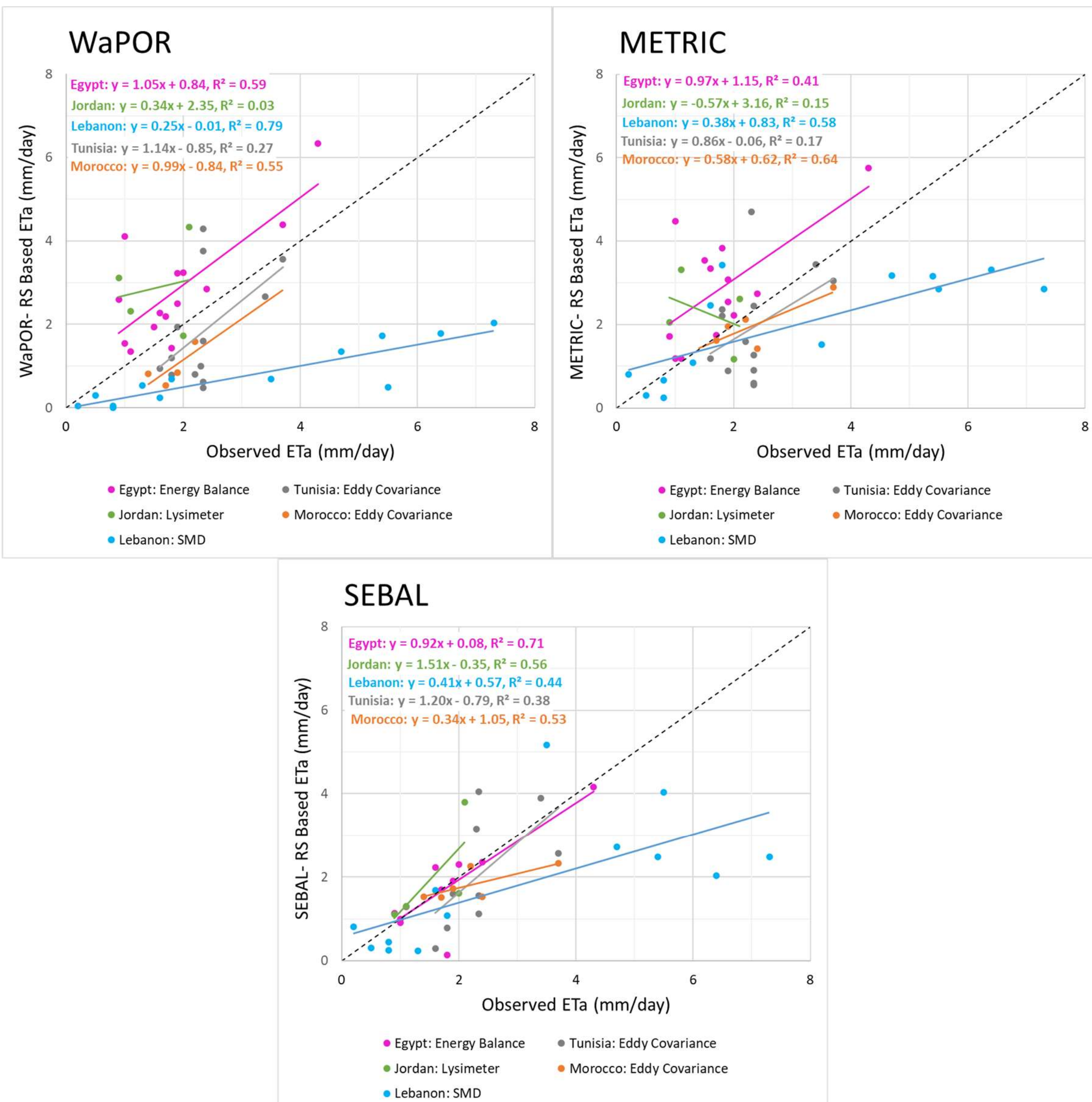

Figure 8: Model-wise intercomparison of RS-ET with observed ET for study locations for season 1. The dotted line represents the identity/reference line

This observation infers that soil moisture depletion (SMD) based ET estimation showed lowest correlation with the RS-Derived ET values as most of the RS based ET values were below the reference line. Both WaPOR and METRIC overpredicted the ET values whereas most SEBAL based ET values were observed to be near the reference line. Each RS-based model underpredicted the ET values for Morocco and Tunisia locations. ET values for Jordan generated poor R^2 values using WaPOR and METRIC (0.03 and 0.15 respectively) compared to SEBAL (0.56).

Table 3: Models' Performance metrics for Remote Sensing based ET Model's intercomparison analysis for season 1. Cells coloured in green depicts the best model for the season based on nRMSE, d and RPD-

				Season-1
Location	**nRMSE (%)**	**Uncertainty (%)**	**RPD**	**Index of Agreement**
		WaPOR		
EGYPT	44	8.5	1.1	0.71
Tunisia	51	9.9	1.1	0.51
Lebanon	88	2.7	0.3	0.41
Jordan	58	14.1	0.7	0.46
Morocco	40	6.1	0.9	0.30
Overall	56	8.2	1.1	0.37
		METRIC		
EGYPT	60	8.6	0.8	0.56
Tunisia	50	9.3	1.1	0.46
Lebanon	63	13.7	0.6	0.69
Jordan	110	8.8	0.5	0.27
Morocco	32	2.4	1.2	0.78
Overall	71	8.6	1.0	0.57
		SEBAL		
EGYPT	19	6.7	0.8	0.91
Tunisia	41	12.3	0.9	0.65
Lebanon	71	5.9	0.9	0.66
Jordan	27	5.8	1.6	0.69
Morocco	19	5.6	1.4	0.62
Overall	35	7.3	1.5	0.70

5.1.2 Observations for Season-2

For Season-2, both SEBAL and METRIC performed well in terms of small nRMSE and higher values of the d and RPD (*Table 4*). The largest d (0.88 for Jordan) was generated by the WaPOR model followed by SEBAL

(0.87 for Egypt) and METRIC (0.84 for Jordan). Largest d was accompanied by smallest nRMSE and highest RPD values show better performance of the model for these locations. For the Lebanon location, SMD-based GB-ET was negatively correlated with RS-ET and generated negative d (*Figure 9*). Lower RS-ET efficiency was observed for the Morocco location too. SEBAL based RS-ET was the best correlation among all locations and all models for season-2. For Jordan, WaPOR followed by METRIC and SEBAL performed well as d is above 0.8 except for SEBAL (0.48). The correlation between GB-ET and RS-ET was the least for this season for the Morocco location. For Morocco location, ET values were observed of lower range (between 0 to 4) and the correlation between GB-ET and RS-ET was unexpectedly poor.

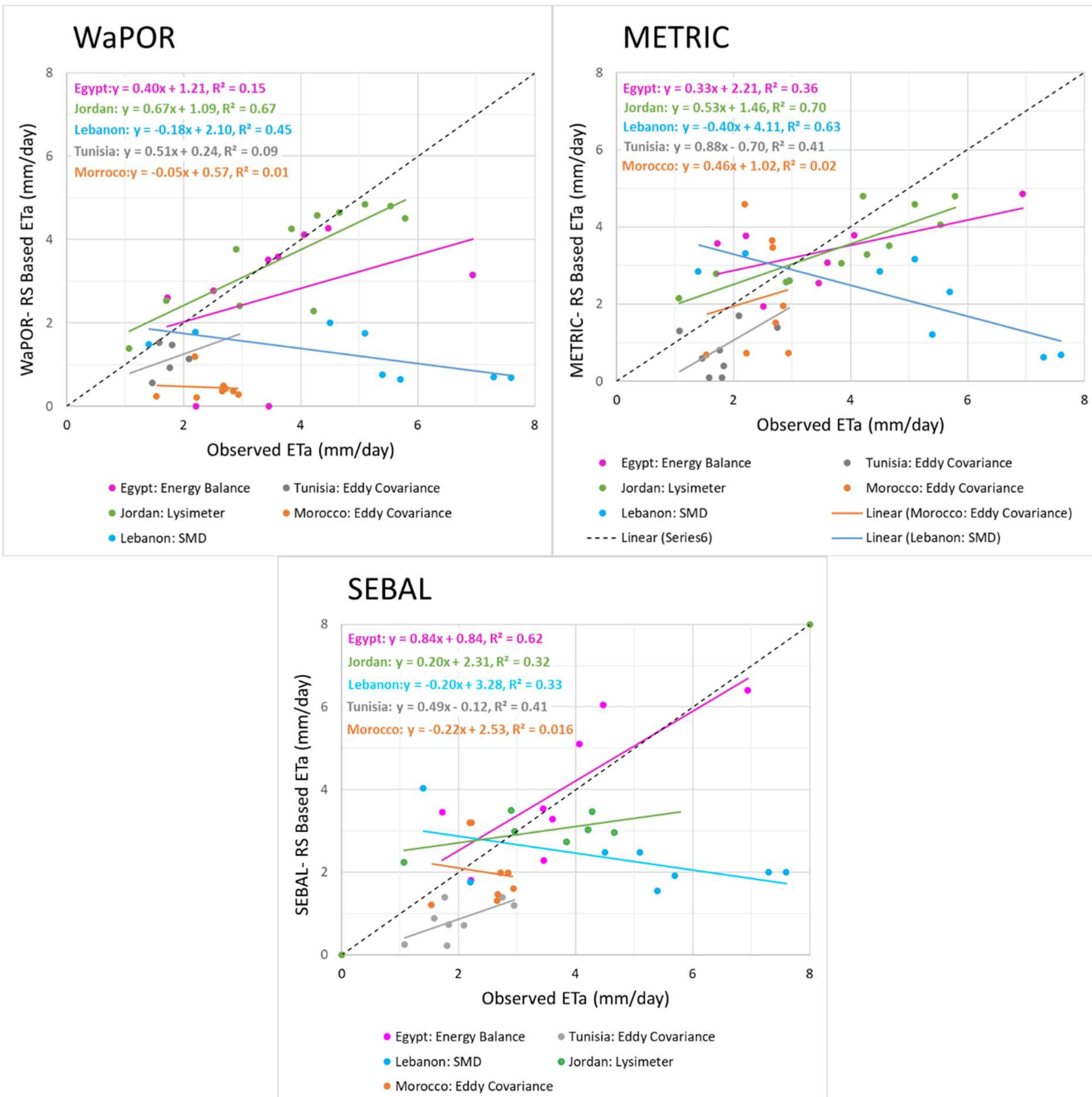

Figure 9: Model-wise intercomparison of RS-ET with observed ET for study locations for season 2. The dotted line represents the identity/reference line

ET observations were less in number for season-1 and season-2 compared to season-3 and season-4. For Lebanon and Jordan locations, lower ET values can be attributed to heterogeneous land cover features.

Table 4: Models' Performance metrics for Remote Sensing based ET Model's intercomparison analysis for season 2. Cells coloured in green depicts the best model for the season based on nRMSE, d and RPD-

				Season-2
Location	**nRMSE (%)**	**Uncertainty%**	**RPD**	**Index of Agreement**
		WaPOR		
EGYPT	71	7.7	0.9	0.49
Tunisia	84	6.0	0.2	0.16
Lebanon	90	2.7	0.1	-0.18
Jordan	22	6.5	1.8	0.88
Morocco	85	3.8	0.2	-0.01
Overall	70	3.2	0.8	0.38
		METRIC		
EGYPT	38	5.2	0.7	0.67
Tunisia	57	9.4	0.7	0.48
Lebanon	83	5.2	0.3	-0.71
Jordan	24	5.0	1.7	0.84
Morocco	60	18.5	0.3	0.14
Overall	52	3.2	1.0	0.40
		SEBAL		
EGYPT	26	8.6	1.6	0.87
Tunisia	63	5.6	0.4	0.30
Lebanon	75	3.6	0.2	-0.35
Jordan	28	2.8	1.4	0.48
Morocco	42	9.6	0.5	0.21
Overall	47	3.2	1.1	0.49

5.1.3 Observations for Season-3

For Season-3, the GB-ET data was better compared to previous seasons. In this season, SEBAL outperformed the other two RS-ET models (*Figure 10*). The SEBAL was followed by METRIC and WaPOR. Overall, a larger coefficient of determination values was observed. RS-ET generated by SEBAL was the nearest to the reference line. This itself tells that SEBAL outperformed other two models. SEBAL showed the larger d for Lebanon followed by Egypt. The SEBAL and METRIC estimated RS-ET satisfactorily for all locations but the highest accuracy was observed for Jordan followed by Egypt as smaller nRMSE & uncertainty values and higher RPD values were observed (*Table 5*). Unexpectedly SEBAL performed poorly in the prediction of RS-ET for Tunisia where both d and RPD were smaller (0.32 and 0.6) and comparatively higher nRMSE (50%). For Lebanon, both WaPOR and METRIC underpredicted the RS-ET. Underprediction of RS-ET for Lebanon was larger for WaPOR compared to METRIC followed by SEBAL. This underprediction can be attributed to the coarse spatial resolution of the WaPOR model. WaPOR predicted RS-ET value for Tunisia with higher accuracy compared to METRIC and SEBAL. For Jordan, the largest r^2 and RPD values were observed throughout the seasons. Though d was comparatively less, the nRMSE was exceptionally low (10% only). That can also be seen in the SEBAL plot of *Figure 10* where most of the ET points are located near the identity line.

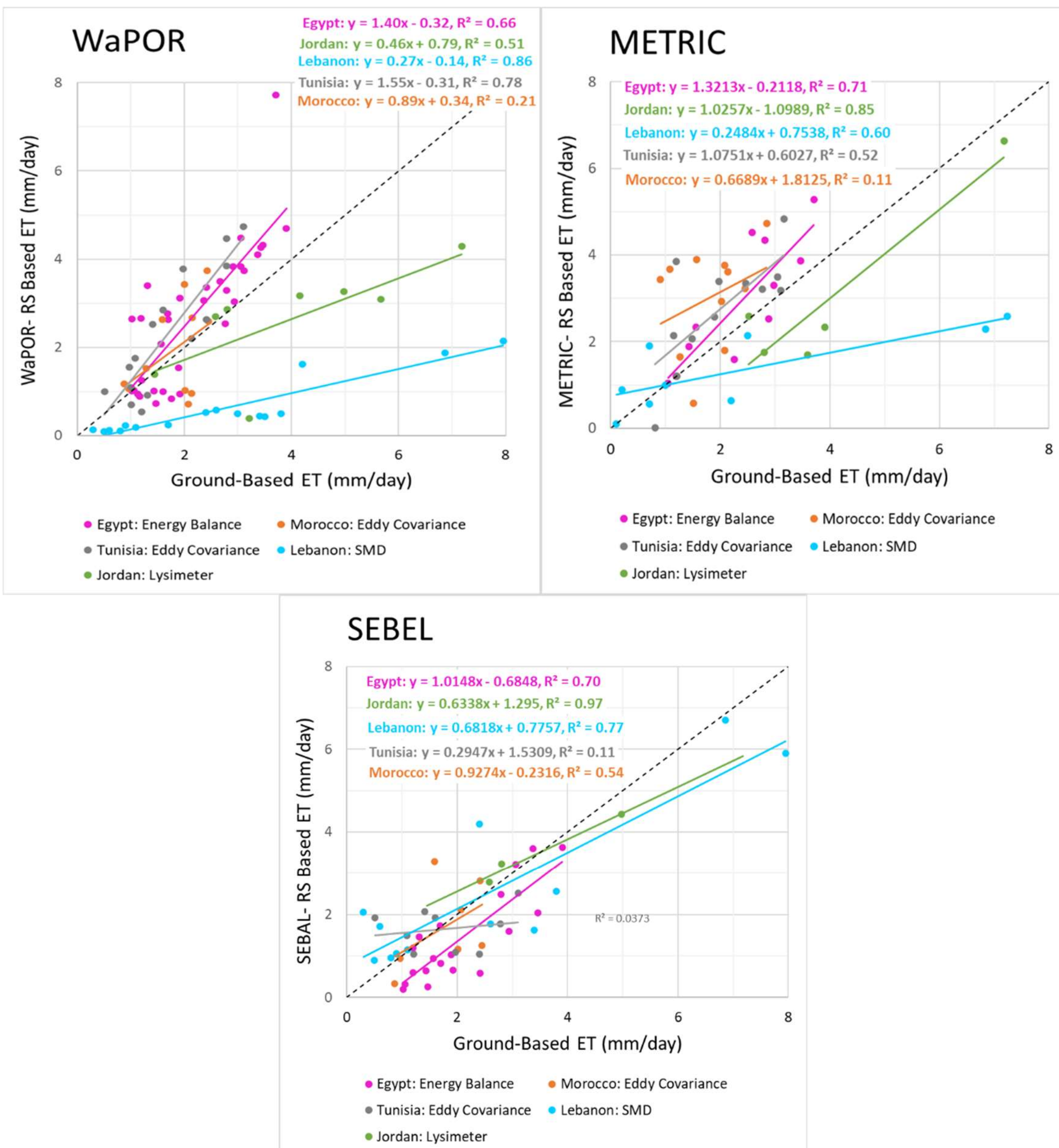

Figure 10: Model-wise intercomparison of RS-ET with observed ET for study locations for season 3. The dotted line represents the identity/reference line

Table 5: Models' Performance metrics for Remote Sensing based ET Model's intercomparison analysis for season 3. Cells coloured in green depicts the best model for the season based on nRMSE, d and RPD-

				Season-3
Location	**nRMSE (%)**	**Uncertainty (%)**	**RPD**	**Index of Agreement**
		WaPOR		
EGYPT	40	3.5	1.8	0.78
Tunisia	58	7.5	1.4	0.79
Lebanon	97	8.1	0.3	0.42
Jordan	46	10.0	1.0	0.63
Morocco	51	4.4	0.6	0.53
Overall	60	6.7	1.0	0.61
		METRIC		
EGYPT	41	8.3	1.5	0.82
Morocco	91	7.7	0.7	0.25
Tunisia	56	7.7	1.1	0.69
Lebanon	108	11.5	0.3	0.56
Jordan	31	13.9	1.7	0.88
Overall	66	3.1	0.9	0.65
		SEBAL		
EGYPT	40	7.0	1.3	0.81
Morocco	44	12.3	1.4	0.59
Tunisia	50	7.3	0.6	0.32
Lebanon	44	8.5	1.6	0.94
Jordan	10	11.1	2.0	0.59
Overall	41	3.1	1.5	0.89

5.1.4 Observations for Season-4

For Season 4, we received data from four locations except Morocco. The data from Morocco location was not obtained because farmer not plantted any crop and they decied to move the station. Our SEBAL model is based on ERA-5 Land hourly climate data. This data is available till September 29, 2021 (last checked on 12th December 2021). Therefore, we didn't extract October month's RS-ET values for all locations by all models. During this season larger ET values were observed for the Egypt location whereas smaller ET values were seen for Tunisia location (*Figure 11*). Satellite data for the Tunisia location was infrequent compared to other locations. WaPOR performed well for the Egypt location followed by SEBAL and METRIC. nRMSE was quite small for RS-ET estimated by all models. WaPOR model performed poorly for all other locations as both RPD and d were quite low and nRMSE was high (*Table 6*). For Lebanon and

Jordan, this poor performance can be attributed to the coarse spatial resolution of the WaPOR based RS-ET. METRIC worked comparatively well for Jordan and Egypt locations with some underprediction of RS-ET values. The poor correlation was observed for Lebanon by METRIC as all the points were underpredicted and were observed far from the identity line. The SEBAL based RS-ET was comparatively better. High underprediction of RS-ET values was observed for Lebanon location whereas some overprediction of RS-ET values for Egypt and Jordan was observed. For Tunisia the performance of SEBAL was good and most of the points were observed near identity line and this is corroborated by a higher value of d (0.87) throughout this season.

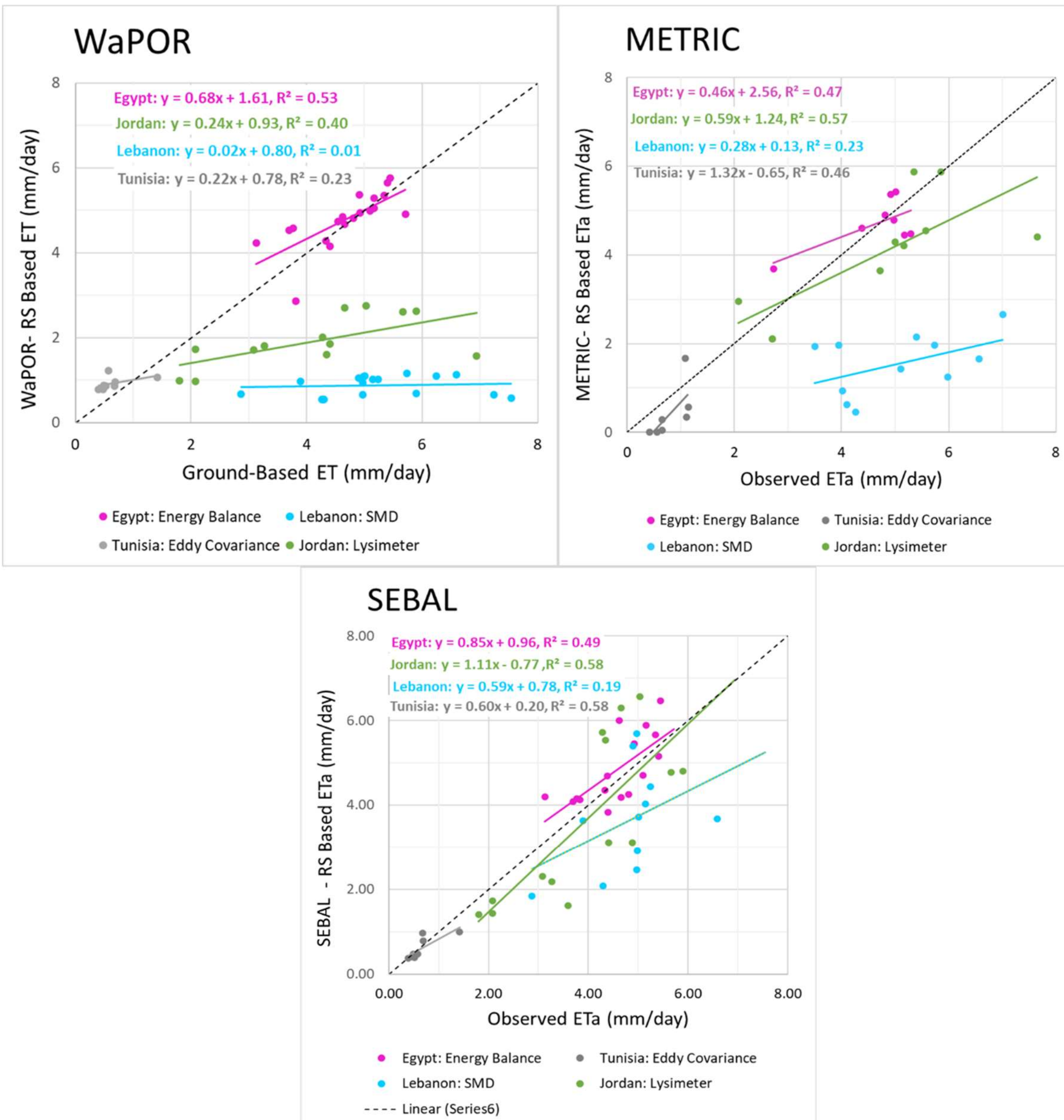

Figure 11: Model-wise intercomparison of RS-ET with observed ET for study locations for season 4. The dotted line represents the identity/reference line

Table 6: Models' Performance metrics for Remote Sensing based ET Model's intercomparison analysis for season 4. Cells coloured in green depicts the best model for the season based on nRMSE, d and RPD-

				Season-4
Location	**nRMSE (%)**	**Uncertainty%**	**RPD**	**Index of Agreement**
		WaPOR		
EGYPT	11	2.58	1.30	0.73
Tunisia	60	4.40	0.40	0.36
Lebanon	86	4.79	0.05	0.00
Jordan	61	6.09	0.24	0.26
Overall	54	4.14	0.90	0.31
		METRIC		
EGYPT	12	7.78	0.98	0.78
Tunisia	70	9.45	1.04	0.56
Lebanon	72	7.55	0.19	0.10
Jordan	27	7.17	1.36	0.80
Overall	45	5.36	1.25	0.66
		SEBAL		
EGYPT	16	3.20	1.14	0.78
Tunisia	20	8.43	2.02	0.87
Lebanon	30	6.63	0.52	0.41
Jordan	28	7.83	1.13	0.82
Overall	23	4.12	1.53	0.91

5.1.5 Overall Observations

Based on all observations from scatterplots and models' performance metrics, SEBAL based RS-ET estimation was found to be the best method for this study. SEBAL based RS-ET values were more accurate compared to the other two models. This fact can also be statistically proved as lower nRMSE and higher d values were observed for SEBAL based RS-ET. SEBAL failed to predict accurately for Tunisia during season 3. The SMD method-based GB-ET was found to be least correlated with RS-ET during all seasons. For all seasons, SEBAL followed METRIC and WaPOR was the trend of performance of RS-based ET models. WaPOR based RS-ET accuracy was higher for Egypt location except for season 2. METRIC worked well for all seasons except for seasons 2 and 3 for Morocco. METRIC-based RS-ET was underpredicted for Jordan and Lebanon. The poor correlation reported in Lebanon for all seasons is related to the overprediction of smaller ET values and the underprediction of larger ET values. This can also be attributed to the 30 m resolution of METRIC ET, which might have contributed to the underprediction of ETa values for Lebanon and Jordan, where non-agricultural land-uses are located near the agricultural land of the ground-based ET estimation locations.

Most aligned regression lines to the identity line, low RMSE & uncertainty and comparatively higher RPD values indicate that SEBEL can be considered the best prediction model for RS-based ETa estimation followed by METRIC and WaPOR. Here, this is worth mentioning that SEBEL has produced ETa at fine (10 m) spatial resolution compared to METRIC (30 m) and WaPOR (100m). This finer resolution ETa could be beneficial for smallholder farming to pave the way for clever water use. This study also reflects that fine resolution ETa can be considerably useful in areas where homogeneity is absent, or non-agriculture features are present. Fine resolution RS-based ET could solve the problem of mixed pixel, hence could provide more accurate RS-ET values for small farms.

6. Limitations of the study

The ground-based ET observations were determined using different methods (i.e., EB, EC, SMD, and Lysi) for diverse locations (one method for each location). Therefore, there are likely underlying variations in the values of the ET obtained. Contrary to this, each RS-based ET model has a universally applicable method for all locations. Therefore, some errors may creep in after comparing ground-based ET (obtained from different methods) for the different areas with RS-based ET data.

METRIC model is based on Landsat series data and has 8 days frequency whereas the SEBAL model is based on Sentinel-2 data with the 5-day frequency. During comparative analysis, there might be some difference due to different input satellite imagery. This is one of the limitations. We have scaled WaPOR based monthly ET values to daily ET values based on daily reference ET. This could introduce some error.

7. Concluding remarks and Future scope

This research work investigates the use of remote sensing-based evapotranspiration modelling at a broader scale for assessing ET in smallholder farming. Smallholder farming-based RS-ET estimation would be more feasible and accurate after its robust optimization and time to time validation at different geographies. The results reflect that the most reliable remote sensing-based method for ET modelling in this study was SEBAL (smallest RMSE and uncertainty values and larger RPD values), followed by METRIC and WaPOR. Most of the ETa predictions by SEBAL, moderately underpredicted ground ET measurements. On the contrary, METRIC had moderate over predictions of measured ET. Probably due to coarse spatial resolution, the WaPOR model exhibited the largest RMSE, higher uncertainty and smallest RPD values.

There is a need to establish at least one common ground-based ET estimation technique across the locations for better comparison with RS model output. Another aspect can be the use of the same RS data for each remote sensing-based model. Inverse calibrated ET methods have great strength, especially for reasons where less data is available, like in the NENA region for site-specific advisories. The SEBAL model presented here could be upgraded by developing five-day interval land surface temperature data using interpolation of 8-16 days interval Landsat LST. This may bring improvements in the model. Besides this, to interpolate monthly or weekly RS based ET values to daily ET values, a separate analysis could be conducted by utilizing a suitable Mateo input. More modelling techniques like Operational Simplified Surface Energy Balance (SSEBop) could also be included for comparison with ground-based estimates.

8. References

AgriMetSoft (2019). Online Calculators. Available on: https://agrimetsoft.com/calculators/Index%20of%20Agreement

Aguilos, M., Stahl, C., Burban, B., Hérault, B., Courtois, E., Coste, S., Wagner, F., Ziegler, C., Takagi, K., & Bonal, D. (2018). Interannual and seasonal variations in ecosystem transpiration and water use efficiency in a tropical rainforest. *Forests*, *10*(1), 14. https://doi.org/10.3390/f10010014

Allen R, Kilic A (2021) METRIC -- Mapping EvapoTranspiration at high Resolution with Internalized Calibration in FAO water scarcity initiative Webinar Series 'Remote Sensing Determination of Evapotranspiration, Algorithms, strengths and weaknesses, uncertainty and best fit-for-purpose'Allen R, Morton C, Kamble B, Kilic A, Huntington J, Thau D, et al. (2015) EEFlux: A Landsat-based evapotranspiration mapping tool on the Google Earth Engine. 2015 ASABE / IA Irrigation symposium: emerging technologies for sustainable irrigation—a tribute to the career of terry howell, Sr Conference Proceedings. St. Joseph, MI: ASABE; pp. 1–11.

Allen, R. G., Tasumi, M., & Trezza, R. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (metric)—model. *Journal of Irrigation and Drainage Engineering*, *133*(4), 380–394. https://doi.org/10.1061/(asce)0733-9437(2007)133:4(380)

Angus, D. E., & Watts, P. J. (1984). Evapotranspiration — how good is the Bowen Ratio Method? *Agricultural Water Management*, *8*(1-3), 133–150. https://doi.org/10.1016/0378-3774(84)90050-7

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., … Wofsy, S. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*(11), 2415–2434.

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. *Journal of Hydrology*, *212-213*, 198–212. https://doi.org/10.1016/s0022-1694(98)00253-4

Bastiaanssen, W. G., Cheema, M. J., Immerzeel, W. W., Miltenburg, I. J., & Pelgrum, H. (2012). Surface Energy Balance and actual evapotranspiration of the transboundary Indus Basin estimated from satellite measurements and the ETLOOK model. *Water Resources Research*, *48*(11). https://doi.org/10.1029/2011wr010482

Cuenca, R. H., Stangel, D. E., & Kelly, S. F. (1997). Soil Water Balance in a boreal forest. *Journal of Geophysical Research: Atmospheres*, *102*(D24), 29355–29365. https://doi.org/10.1029/97jd02312

Eastham, J., Rose, C. W., Cameron, D. M., Rance, S. J., & Talsma, T. (1988). The effect of tree spacing on evaporation from an agroforestry experiment. *Agricultural and Forest Meteorology*, *42*(4), 355–368. https://doi.org/10.1016/0168-1923(88)90043-3

Edwards, W. R. (1986). Precision weighing lysimetry for trees, using a simplified tared-balance design. *Tree Physiology*, *1*(2), 127–144. https://doi.org/10.1093/treephys/1.2.127

Hemakumara, H. M., Chandrapala, L., & Moene, A. F. (2003). Evapotranspiration fluxes over mixed vegetation areas measured from large aperture scintillometer. *Agricultural Water Management*, *58*(2), 109–122. https://doi.org/10.1016/s0378-3774(02)00131-2

Jensen, M.E.; Burman, R.D.; Allen, R.G. (1990) Evapotranspiration and Irrigation Water Requirements; ASCE Manuals and Reports on Engineering Practice No. 70; ASCE: Reston, VA, USA.

Kyaw Tha Paw U, Qiu, J., Su, H.-B., Watanabe, T., & Brunet, Y. (1995). Surface Renewal Analysis: A new method to obtain scalar fluxes. *Agricultural and Forest Meteorology*, *74*(1-2), 119–137. https://doi.org/10.1016/0168-1923(94)02182-j

Montazar, A., Krueger, R., Corwin, D., Pourreza, A., Little, C., Rios, S., & Snyder, R. L. (2020). Determination of actual evapotranspiration and crop coefficients of California date palms using the residual of energy balance approach. *Water*, *12*(8), 2253. https://doi.org/10.3390/w12082253

Pelgrum, H.; Miltenburg, I.J.; Cheema, M.J.M.; Klaasse, A.; Bastiaanssen, W.G.M. (2012) ET Look: A novel continental evapotranspiration algorithm. IAHS-AISH Publ., 352, 120–123.

Sánchez, J. M., Galve, J. M., González-Piqueras, J., López-Urrea, R., Niclòs, R., & Calera, A. (2020). Monitoring 10-m LST from the Combination MODIS/Sentinel-2, Validation in a High Contrast Semi-Arid Agroecosystem. *Remote Sensing, 12*(9), 1453. doi:10.3390/rs12091453

Senay, G. B., Bohms, S., Singh, R. K., Gowda, P. H., Velpuri, N. M., Alemu, H., & Verdin, J. P. (2013). Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach. *JAWRA Journal of the American Water Resources Association*, *49*(3), 577–591. https://doi.org/10.1111/jawr.12057

Wagle, P., Bhattarai, N., Gowda, P. H., & Kakani, V. G. (2017). Performance of five surface energy balance models for estimating daily evapotranspiration in high biomass sorghum. *ISPRS Journal of Photogrammetry and Remote Sensing*, *128*, 192–203. https://doi.org/10.1016/j.isprsjprs.2017.03.022

WaPOR, FAO. FAO Water Productivity. (n.d.). Retrieved November 19, 2021, from https://wapor.apps.fao.org/home/WAPOR_2/1.

Wegehenkel, M.; Gerke, H.H. (2015) Water table effects on measured and simulated fluxes in weighing lysimeters for differently- textured soils. J. Hydrol. Hydromech., 63, 82–92, doi:10.1515/johh-2015-0004.

Willmott, C. J. (1981). On the validation of Models. *Physical Geography*, *2*(2), 184–194. https://doi.org/10.1080/02723646.1981.10642213

Yang, Y., Scott, R. L., & Shang, S. (2013). Modeling evapotranspiration and its partitioning over a semiarid shrub ecosystem from satellite imagery: A multiple validation. *Journal of Applied Remote Sensing*, *7*(1), 073495. https://doi.org/10.1117/1.jrs.7.073495