https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

Can we trust remote sensing ET products over Africa?

Imeshi Weerasinghe[1], Ann van Griensven[1,2], Wim Bastiaanssen[2], Marloes Mul[2], and Li Jia[3,4]

[1]Vrije Universiteit Brussels, Brussels, Belgium
[2]IHE Delft Institute for Water Education, Delft, Netherlands
[3]Chinese Academy of Sciences, Beijing, China
[4]Joint Center for Global Change Studies, Beijing, China

Correspondence: Imeshi WEERASINGHE (Imeshi.Nadishka.Weerasinghe@vub.be)

Abstract. Evapotranspiration (ET) is one of the most important components in the water cycle. However, there are relatively few direct measurements of ET (using flux towers), whereas various disciplines ranging from hydrology to agricultural and climate sciences, require information on the spatial and temporal distribution of ET at regional and global scale. Due to limited data availability, attention has turned toward satellite based products to fill observational gaps. Various remote sensing data products have been developed, providing a large range of ET estimations. Across Africa only a limited number of flux towers are available which are insufficient for systematic evaluation of remotely sensed (RS) derived ET products. Thus we propose a methodology for evaluating RS derived ET data at the basin scale using a general water balance (WB) approach, where ET is equal to precipitation minus discharge for long-term annual averages. Firstly, RS ET products are compared with WB inferred ET for basins without long-term trends present. The RS products are then assessed according to spatial characteristics through analysing two land cover elements across Africa, irrigated areas and water bodies. A cluster analysis is also conducted to identify similarities between individual ET products. Finally, the RS products are evaluated against the Budyko equation. The results show that CMRSET, SSEBop and WaPOR rank highest in terms of estimation of long-term annual average mean ET across basins with low biases. Along with ETMonitor, the same three products rank highest in spatial distribution of ET patterns across Africa. GLEAM and MOD16 consistently rank the lowest in most criteria evaluation. Many of the products analysed in this study can be trusted depending on the study under question, keeping in mind some of these products have large biases in magnitude estimation. However our recommendation would be the three highest ranked products being CMRSET, SSEBop and WaPOR.

1 Introduction

Evapotranspiration (ET) or the water vapor flux is an important component in the water cycle and is widely studied due to its implications in hydrology to agricultural and climate sciences (Trambauer et al., 2014). Growing attention has been given to estimating ET fluxes at regional and global scales for a wide variety of reasons, for example, understanding the partitioning

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

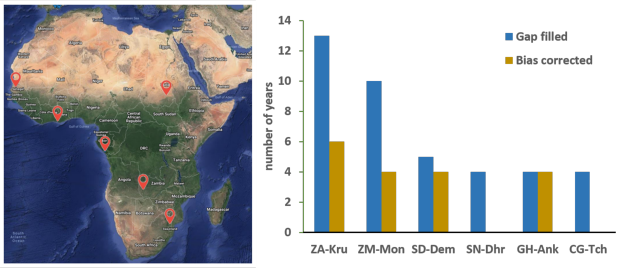

Figure 1. (left) distribution of flux towers with LE data across Africa. (right) Number of years of available data at the six flux tower sites across Africa for both gap filled and bias corrected LE

of energy and water at the earth's surface and their feedbacks; how the different external drivers of ET vary regionally and; understanding the impacts of potential changes on the hydrological cycle under a changing climate, to name a few (Teuling et al., 2009; Vinukollu et al., 2011a; Mu et al., 2011). However, the estimation of ET at large scales has always been a difficult task due to direct measurement of ET being possible only at point locations, for example using flux towers (Trambauer et al., 2014). Flux tower data can be openly accessed through FLUXNET, a global network of micrometeorological flux measurement sites that measure the exchange of CO2, water vapor and energy between the biosphere and the atmosphere (Baldocchi et al., 2001). From the latest FLUXNET 2015 dataset, there are only six eddy covariance sites in Africa, from which latent heat (LE) measurements can be obtained, which can be converted to ET. Figure 1 shows the distribution and data availability of the sites. Gap filled LE data using the Marginal Distribution Sampling (MDS) technique is available at these locations however, a general lack of energy balance closure is found at many sites (Wilson et al., 2002). For this reason LE can also be obtained with a correction factor applied for energy balance closure and thus, reduces the number of data points and sites available for use. Due to the limited data availability of observed point data for the entirety of the African continent a method of evaluating ET estimations using data other than point measurements is required.

Recent advances in satellite based products provide promising data to fill these observational gaps (Alkema et al., 2011; Miralles et al., 2016). ET cannot be directly measured by satellite based measurements, but can be derived from physical variables that can be observed from space, such as latent heat and surface heat using the surface energy balance. In addition, due to passing frequencies and cloud interference, interpolations in time are required. In this respect remote sensing derived ET cannot be interpreted as direct satellite observations but as model outputs based on satellite forcing data (Miralles et al., 2016). Satellite observations often give useful information on the spatial variability, however the products tend to suffer from a large bias. Therefore, large-scale estimations of ET are most commonly products of remote sensing based models, hydrological models and land-surface models (Trambauer et al., 2014). More recently, remote sensing ET products have been developed using Machine Learning (ML) approaches such as Model Tree Ensemble (MTE) or Artificial Neural Network (ANN) combined with observed flux tower data or model outputs used as training sets (Tramontana et al., 2016; Jiménez et al., 2011; Jung et al., 2017; Alemohammad et al., 2017).

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

With this large range of approaches to estimate ET, large differences are observed among the products and therefore, validation is required. Since it is difficult to validate ET estimates using observed data, an alternate method of inferring ET for a river basin is used. Assuming the change in water storage (soil moisture, lakes, deltas) is negligible at the river basin scale, ET becomes equal to precipitation (P) minus discharge (Q) (Miralles et al., 2016, 2011; Vinukollu et al., 2011b). Using this general Water Balance (WB), it is possible to gain understanding of the magnitude of ET within a given basin and hence to estimate biases in ET estimation by the different ET products. Unfortunately, the period of observation for measured discharge for certain basins is limited or do not overlap with RS derived estimations of ET. For this reason long-term annual averages for time series without trends are used.

This study focuses on a methodology for evaluating RS derived ET products from discharge observations and observation based precipitation to derive ET using a WB approach, at the continental scale over Africa using long-term averages for non-overlapping time periods. A trend analysis is conducted in order to justify the use of different time periods. Spatial variability is analysed using specific land cover elements that tend to have a higher or lower ET such as water bodies and irrigated areas.

2 Data

The remote sensing derived ET products being evaluated in this study include WaPOR, GLEAM, MOD16, SSEBop, WECANN, FLUXNET-MTE, ETMonitor and CMRSET. All data are projected and gridded on a $0.0022\ ^\circ \times 0.0022^\circ$ geographic grid and averaged at yearly temporal resolution. Table 1 summarizes the characteristics of the remote sensing products being used.

2.1 Remotely sensed ET products

2.1.1 GLEAM

The Global Land Evaporation Amsterdam Model (GLEAM) is a physically based model that estimates terrestrial evapotranspiration using satellite observations (Miralles et al., 2011, 2016). It consists of three different calculation schemes, namely, (1) rainfall interception driven by rainfall and vegetation observations; (2) potential evaporation calculated using the Priestley and Taylor (P-T) equation (Priestley and Taylor, 1972) and driven by satellite observations; and (3) a stress factor attenuating potential evaporation based on a semi-empirical relationship between microwave vegetation and optical depth (VOD) observations and root zone soil moisture estimates (Alemohammad et al., 2017). GLEAM ET estimates are provided at daily temporal resolution from 1980-2013 and $0.25\ ^\circ \times 0.25^\circ$ spatial resolution.

2.1.2 WaPOR

The Food and Agriculture Organisation's data portal to monitor **Wa**ter **P**roductivity through **O**pen access of **R**emotely sensed derived data (WaPOR) offers products related to water productivity (WP) derived mainly from freely available remote sensing satellite data (FAO, 2018). Actual evapotranspiration estimates are the sum of the soil evaporation (E) and canopy transpiration

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 1. Characteristics of remotely sensed ET products

Product	Temporal Coverage	Spatial Coverage	Temporal Resolution	Spatial Resolution	Estimation Approach	Input Data Source	Reference
CMRSET	2000-2013	Global	8-daily	$0.0022^\circ \times 0.0022^\circ$	P-T Equation, relationship between EVI and GVMI	MODIS	(Guerschman et al., 2009)
ETMonitor	2008-2013	Global	daily	$0.005^\circ \times 0.005^\circ$	P-M, Gash model, Shuttleworth-Wallace	MODIS	Zheng et al. (2016)
GLEAM	1980-2015	Global	Daily	$0.25^\circ \times 0.25^\circ$	P-T Equation, soil stress factor	AMSR-E, LPRM, CMORPH, TRMM	Martens et al. (2017); Miralles et al. (2011)
MOD16A3	2000-2014	Global	Monthly	$0.0083^\circ \times 0.0083^\circ$	P-M Equation, surface conductance model	MODIS	Mu et al. (2011, 2007)
MTE	1982-2012	Global	Monthly	$0.5^\circ \times 0.5^\circ$	MTE approach, training using in-situ observations, flux tower data	Eddy Covariance, in-situ	Jung et al. (2011)
SSEBop	2003-2017	Global	Monthly	$0.0096^\circ \times 0.0096^\circ$	P-M Equation, ET fractions from T_s estimates	MODIS	Senay et al. (2007, 2011a, b)
WaPOR	2009-2017	Africa	Dekadal	$0.0022^\circ \times 0.0022^\circ$	P-M Equation, calculates E, T and I separately	MODIS, GEOS-5/MERRA	FAO (2018)
WECANN	2007-2015	Global	Monthly	$1^\circ \times 1^\circ$	ANN approach, training using observations and model based LE	GOME-2	Alemohammad et al. (2017)

(T). Calculation of E and T are based on the ETLook model described in Bastiaanssen et al. (2012) using the Penman-Montieth (P-M) equation (Montieth, 1965) adapted for remote sensing input and solved separately for E and T. For T the coupling with

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019


Hydrology and Earth System Sciences
Discussions

the soil is made via the root zone soil moisture content whereas for the E the coupling is made via the soil moisture content of the topsoil (FAO, 2018). The actual evapotranspiration and Interception (ETIa) maps used in this study are provided at a spatial resolution of $0.0022° \times 0.0022°$ and at either dekaadal or annual temporal resolution for the period 2009-present.

2.1.3 MOD16

MODIS Global Evapotranspiration Project (MOD16) estimates terrestrial evapotranspiration by using satellite remote sensing data. Terrestrial ET includes evaporation from wet and moist soil, rain water intercepted by the canopy and transpiration through stomata from plant leaves and stems. ET datasets are calculated using (Mu et al., 2011) improved algorithm from the initial developed algorithm in (Mu et al., 2007) and is based on the P-M equation. Improvements include; evaporation from wet soil; nighttime ET; simplified calculation of vegetative fraction cover; adding soil heat flux; improving estimates of stomatal conductance, aerodynamic resistance and boundary layer resistance and separating dry and wet canopy surfaces (Mu et al., 2011). MOD16 ET estimates are provided at a spatial resolution of $0.0083° \times 0.0083°$ and at either 8-daily, monthly or annual temporal resolution for the period 2000-2014.

2.1.4 SSEBop

The operational Simplified Surface Energy Balance (SSEBop) model estimates ET as a function of the land surface temperature (T_s) from remotely sensed data and reference ET (ETo) from global weather datasets using the Simplified Surface Energy Balance (SSEB) method developed by Senay et al. (2007, 2011b, a). The Surface Energy Balance (SEB) is first solved for each pixel for a reference crop condition using the standard P-M equation and is adjusted according to T_s through an ET fraction approach, which accounts for the spatial variability of water availability and vegetation health in the landscape (Savoca et al., 2013). SSEBop uses pre-defined, seasonally dynamic boundary conditions that are unique to each pixel for "hot/dry" and "cold/wet" reference points defined in Bastiaanssen et al. (2014) and Allen et al. (2007). SSEBop ET estimates are provided at a spatial resolution of $0.0096° \times 0.0096°$ and at either monthly or annual temporal resolution for the period 2001-2017.

2.1.5 WECANN

The Water, Energy and Carbon Cycle with Artificial Neural Networks (WECANN) retrieves monthly estimates of Latent Heat Flux (LE) using the Artificial Neural Network (ANN) approach. The LE estimates, converted to ET in this study using a coefficient, uses remotely sensed solar-induced fluorescence (SIF) estimates along with remotely sensed estimates of precipitation, temperature, soil moisture, snow cover and net radiation as inputs. Different observations and/or model-based estimates of LE are used to produce the training dataset using a Bayesian perspective (Alemohammad et al., 2017). WECANN LE estimates are provided at a spatial resolution of $1° \times 1°$ and at monthly temporal resolution for the period 2007-2015.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019


Hydrology and Earth System Sciences
Discussions

2.1.6 FLUXNET-MTE

The FLUXNET Model Tree Ensemble (FLUXNET-MTE) provides global fluxes of LE, converted to ET in this study, derived from empirical upscaling of eddy covariance measurements from the FLUXNET global network (Baldocchi et al., 2001). The MTE method uses an ensemble learning algorithm by training the MTEs for LE using site-level explanatory variables and fluxes and then applying these established MTEs using gridded datasets of the same explanatory variables (Jung et al., 2011). MTE LE estimates cover a period from 1982-2012 at a spatial resolution of $0.5° \times 0.5°$ and at a monthly temporal resolution.

2.1.7 ETMonitor

ETMonitor is a process based model using mainly satellite observations to estimate ET at the global scale. In order to calculate ET, different modules for different land cover classes are used, including soil evaporation and plant transpiration for soil-plant systems based on the Shuttleworth-Wallace (Shuttleworth and Wallace, 1985) model, an analytical module for rainfall interception loss by vegetation canopies, a water evaporation module for water bodies based on the P-M equation and a sublimation module for snow/ice surfaces (Zheng et al., 2016). The ET estimates are available globally, covering a period from 2008-2012 at a spatial resolution of $0.0096° \times 0.0096°$ and at daily temporal resolution.

2.1.8 CMRSET

CMRSET provides estimates of ET based on surface reflectances from MODIS-Terra and interpolated climate data. The algorithm uses Enhanced Vegetation Indices (EVI) through its relationship with Leaf Area Index (LAI) and Global Vegetation Moisture Indices (GVMI) which provides information on vegetation water content and allows the separation of surface water and bare soil to scale derived P-T potential evapotranspiration (Guerschman et al., 2009). CMRSET ET estimates are available at a spatial resolution of $0.0022° \times 0.0022°$ and an 8-day temporal resolution for the period 2000-2013.

2.1.9 Multi-Product Mean

The Multi-Product Mean (MPM) is obtained by calculating the mean of the eight aforementioned RS ET products used in this study. The product has a spatial resolution of $0.0022° \times 0.0022°$.

2.2 Precipitation data

2.2.1 EWEMBI

EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected (EWEMBI) is a global precipitation product. Data sources of EWEMBI are ERA-Interim reanalysis data (ERAI), WATCH forcing data methodology applied to ERA-Interim reanalysis data (WFDEI), eartH2Observe forcing data (E2OBS) and NASA/GEWEX Surface Radiation Budget data (SRB) (Dee et al., 2011; Weedon et al., 2014; Stackhouse et al., 2011). The dataset covers the entire globe at a spatial resolution of $0.5° \times 0.5°$ and daily temporal resolution from 1979 to 2013.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019


Hydrology and Earth System Sciences
Discussions

2.2.2 CHIRPS

Climate Hazards group Infrared Precipitation with Stations (CHIRPS) dataset uses the Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis version 7 (TMPA 3B42 v7) to calibrate global Cold Cloud Duration (CCD) rainfall estimates as well as a 'smart' interpolation of gauge data from Meteorological Organization's Global Telecommunication System (GTS) (Funk et al., 2015). The product is available for $50°$S to $50°$N and all latitudes at a spatial resolution of $0.05° \times 0.05°$ at daily, pentadal and monthly temporal resolution covering the period from 1981-2017.

2.2.3 MSWEP

The Multi-Source Weighted Ensemble Precipitation (MSWEP) is a gridded precipitation product based on gauge (WorldClim, GHCND, GSOD, GPCC, and others), satellite (CMORPH, GridSat, GSMaP, and TMPA 3B42RT) and reanalysis (ERA-Interim and JRA-55) data (Beck et al., 2019). The dataset provides precipitation estimates globally at a spatial resolution of $0.1° \times 0.1°$ and temporal resolution of 3-hourly covering the period 1979-2017.

2.3 Discharge data

Discharge data was obtained from the Global Runoff Data Centre (GRDC) for the majority of basins and from the Vrije Universiteit Brussels (VUB) Department of Hydrology and Hydraulic Engineering (HYDR) for the Nile and Blue Nile basins. All data was initially obtained at either daily or monthly temporal resolution and aggregated to monthly and yearly averages.

2.4 Reference potential evapotranspiration data

The datasets used for reference potential evapotranspiration (PET) was developed by Deltares (Sperna Weiland et al., 2015). The datasets are derived from the WFDEI dataset with a resolution of $0.5° \times 0.5°$ and downscaled based on a high resolution Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission at 90m resolution (Sperna Weiland et al., 2015). Three datasets were used for PET based on the Hargreaves, P-M and P-T approaches respectively. The global PET datasets have a spatial resolution of $0.083° \times 0.083°$ and daily temporal resolution covering the period 1979-2012.

3 Methodology

A methodology to evaluate RS derived ET estimations is presented next:

1. Preprocessing and data analyses
2. Comparison using WB inferred ET estimates
3. Performance with characteristics land cover elements
4. Evaluation using the Budyko curve

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

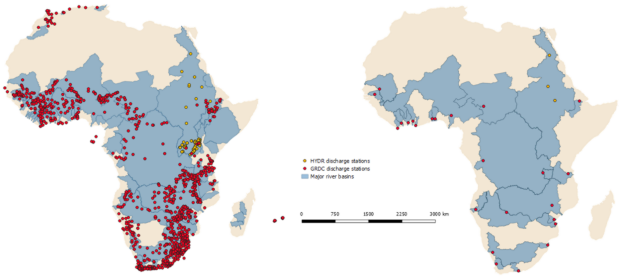

Figure 2. (left) All major basins in Africa and all discharge stations; (right) Major basins in Africa with available discharge data at outlet

3.1 Preprocessing and data analyses

Due to the limited availability of direct observations of ET across Africa, we infer ET estimates at the river basin level using the WB approach. The long-term WB assumes a negligible change in storage (discussed further in Section 5) and therefore the total inflow (P) is equal to the total outflow (ET and Q) and therefore ET is equal to P minus Q, according to the following equation:

$$ET = P - Q \tag{1}$$

For all the major river basins in Africa, discharge data from GRDC and other sources were analysed at their outlets based on data availability and quality. As seen in Fig. 2, from fifty four major basins in Africa we found twenty seven basins with sufficient and quality discharge data at the outlet. The spread of only these twenty seven basins covers the majority of the African continent. Since direct observations of precipitation from gauges were not used, three different precipitation products, as described above, are used for comparison. Using these available discharge data and precipitation data from EWEMBI, CHIRPS and MSWEP, the annual average ET was estimated across each of the twenty seven basins using equation 1 and the long-term average was calculated.

As mentioned previously, using the general WB to infer average ET across different basins poses the problem of limited to no overlapping time periods between the data sources. Thus, we investigated whether or not annual trends can be detected from the inferred ET. If the data show no major trends across the different basins then it can be justified to evaluate the ET estimations using long-term averages from different time periods (discussed in Section 5).

The Mann-Kendall (MK) (Mann, 1945; Kendall, 1948) test was used to identify whether a monotonic upward or downward trend is present in the inferred ET estimates. The MK test is non-parametric (distribution free) and best used as an exploratory analysis to identify where changes are significant or of large magnitude (Matzke et al., 2014) and should only be used where seasonal trends are not present. Considering annual averages are used in this study, the MK test was deemed appropriate. A python function mk_test developed by Matzke et al. (2014) was used to conduct the MK test. The function returns the following outputs:

- trend: the type of trend (increasing, decreasing or no trend)

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019


Hydrology and Earth System Sciences
Discussions

- h: hypothesis testing, returns True (if trend is present) or False (if trend is absent)
- p: p value of the significance test (low value ≤ 0.05 for true and high value >0.05 for false)
- z: normalised test statistic

After conducting the test, if trends were present, these basins were discounted from the analyses.

Lastly, a cluster analysis was performed, using the method followed by Wartenburger et al. (2018) on the RS products and the MPM to investigate the overall level of similarity between the individual products in terms of spatial variability. The long-term average map for each product and the MPM were used whereby the pairwise Euclidean distance between each dataset for each pixel was calculated and evaluated. Each of the maps used were resampled to $0.0096° \times 0.0096°$ for computation efficiency.

3.2 Comparison using WB inferred ET estimates

In order to conduct comparisons of ET estimations, all RS derived products were projected to WGS 84, EPSG:4326 on a $0.0022\ ° \times 0.0022°$ grid, the highest spatial resolution of the products being analysed. The nearest neighbors interpolation method was used for any resampling required from course to high resolution. The estimations were then combined to give a single map of the long-term annual average ET across Africa. The time periods averaged for each product can be found in Table 1. These maps were then clipped for each of the basins being analysed and the basin mean long-term annual average ET recorded. From these results the correlation, average difference and weighted average difference with the estimated WB ET using all three precipitation products was calculated. The ranking of the RS ET estimations for the correlation, average and weighted average difference was based on the mean performance against the three WB ET estimates derived using EWEMBI, CHIRPS and MSWEP precipitation data.

3.3 Performance with characteristics land cover elements

Two types of land cover elements were evaluated in this study. A map with areas equipped for irrigation actually irrigated by FAO and Rheinische Friedrich-Wilhelms-University (Siebert et al., 2013) and a map of water bodies was obtained from the Global Reservoir and Dam (GRanD) database (Lehner et al., 2011) were used to evaluate how well the ET products identified spatial characteristics. Two steps were used, firstly the maps were evaluated visually using the same colour scale. Secondly, since for water bodies the ET should be more or less equal to the PET, the long-term annual average mean ET estimates across water bodies by the products were compared with the long-term annual average mean PET estimates across water bodies by calculating the difference between them. For irrigated areas, average crop coefficients (kc=ET/PET) for maize, wheat and sugarcane estimated by FAO were used as a reference. Thus, the long-term annual average mean ET estimates across irrigated areas were divided by the long-term annual average mean PET estimates across irrigated areas to find the average crop coefficient (kc) across irrigated areas. The difference between the reference kc from FAO and estimated kc using RS ET estimates and PET derived using Hargreaves, P-M and P-T were then found. Ranking was based firstly on visual inspection of the distribution of irrigated areas and water bodies between the products and secondly on the smallest average difference of mean ET of irrigated areas and water bodies with the specified reference conditions.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

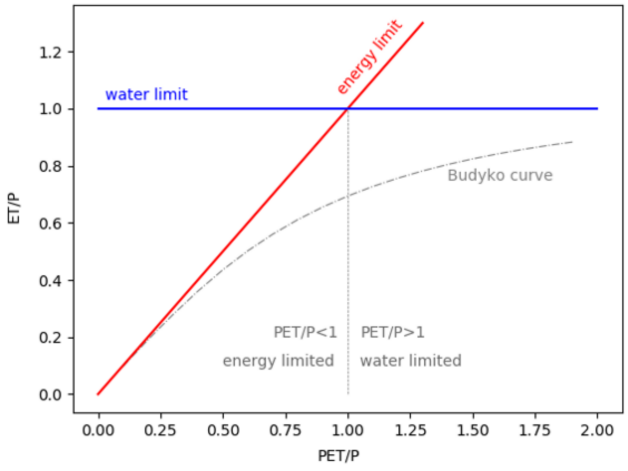

Figure 3. Budyko curve showing the energy limit and water limit

3.4 Evaluation using the Budyko curve

The Budyko equation partitions precipitation into streamflow and ET by describing the relationship between mean annual ET and long-term average water and energy balance at catchment scales (Sposito, 2017) as seen in Fig. 3. Budyko (1974) developed this approach for the physics of catchment ET by postulating on the phase transformation of green water to vapor and thus that ET reflects not only the partitioning of water but also radiant energy at the vadoze zone and atmosphere interface (Sposito, 2017; Gerrits et al., 2009) following equation 2.

$$\left[\frac{PET}{P} tanh\left(\frac{1}{\frac{PET}{P}}\right)\left(1 - exp^{-\frac{PET}{P}}\right)\right]^{0.5} \quad (2)$$

Since the Budyko curve provides a reference condition for the water balance assuming it correctly describes the partitioning of P into Q, then we can use this information to see how well our products perform in estimating ET. For each of the basins under study, we calculated ET/P and PET/P and plotted these against the Budyko curve. We derived long-term annual average basin mean PET estimates for Hargreaves, P-M and P-T approaches. We also used P from EWEMBI, CHIRPS and MSWEP separately to compare the results. The ranking of the RS products using the Budyko evaluation is based on the smallest difference with the Budyko ET estimations for the average ET across the basins for the three PET approaches, Hargreaves, P-M and P-T.

4 Results

In this section we present the obtained results for the different methodology stages.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

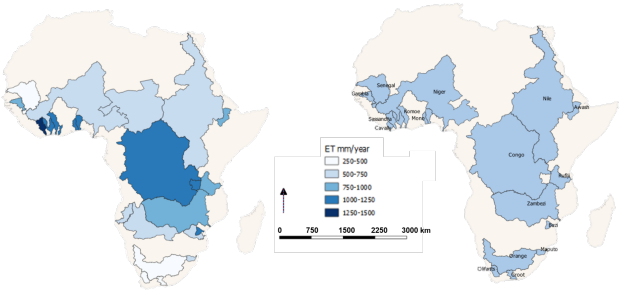

Figure 4. (right) ET estimation for 28 major basins in Africa using P-Q (left) Final basins being analysed after trend analyses

4.1 Preprocessing and data analyses

Figure 4 (left) shows the annual average ET estimates for the twenty seven basins with available discharge and precipitation data. The spread of the ET across the basins seems to be consistent with the climate, where basins in the semi-arid to arid northern and southern parts of Africa show lower ET than the more centrally located basins known to be more tropical.

For the MK test to be accurate a minimum of ten data points should be used. However, of the twenty seven basins being tested, eleven basins did not have sufficient data points for an accurate analyses invalidating these results. Table 2 shows results from conducting a MK test for monotonic trends in the ET estimates inferred from the WB approach for the remaining twenty seven basins across Africa with available discharge data. ET estimates for twenty two of the basins show no trends, while three basins show trends, Cunene and Okavango increasing trends and the Nile a decreasing trend. Two basins, Rufiji and Tana, did not have any overlapping precipitation and discharge data to calculate ET for analyses. For the basins with fewer than ten data points, the MK test was conducted on the collected precipitation and discharge data used to calculate ET. From the eleven basins analysed, five basins, the Blue Nile, Lake Chad, Save, Tana and Void, show a increasing or decreasing trend in either the precipitation or discharge as seen in Table 2. Thus from the MK trend analyses conducted on ET, P and Q estimated, seven basins showed a trend in at least one of the three variables and thus were eliminated from the study. Figure 4 (right) shows the final twenty basins being analysed after elimination based on a monotonic trend being present. The number and spread of the final basins being analysed still gives a good coverage across the African continent.

Two groupings or clusters are observed when looking at the similarity between individual products and the MPM (Fig. 5). We see one cluster formed with three products, CMRSET, SSEBop and WaPOR, with SSEBop and WaPOR being slightly more similar than with CMRSET. And a second cluster with the remaining products and the MPM showing the most similar products being WECANN with MPM and GLEAM with MTE.

4.2 Comparison using WB inferred ET estimates

Figure 6 shows the correlation of the long-term annual average basin mean ET estimates and the different RS ET products, the MPM and the WB inferred ET using the different precipitation products across the twenty basins. For all products the

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 2. Mann-Kendall test results for all basins on evapotranspiration, precpitation (EWEMBI) and discharge

Basin	Variable	Data Availability	Trend	hypothesis	p-value	z-value	no. of samples
Awash	ET	1990-2004	no trend	false	0.2496	-1.1514	14
Bandama	ET	1979-1996	no trend	false	0.7619	-0.3030	18
		not enough data points					4
Blue Nile	P	1979-2016	no trend	false	0.6875	-0.4023	38
	Q	1900-1982	decreasing	true	0.0009	-3.3271	83
		not enough data points					5
Buzi	P	1979-2016	no trend	false	0.4210	-0.8046	38
	Q	1957-1983	no trend	false	1.0	0.0	23
Cavally	ET	1979-1996	no trend	false	0.54449	-0.6060	18
Congo	ET	1979-2010	no trend	false	0.0830	-1.7336	31
Cunene	ET	1980-2015	increasing	true	0.0003	3.5823	36
		not enough data points					5
Gambia	P	1979-2016	no trend	false	0.2579	1.1315	38
	Q	1979, 1981-82, 1984,1988	no trend	false	0.8065	0.2449	5
Groot	ET	1979-2015	no trend	false	0.1697	1.3733	37
Kamoe	ET	1979-1996	no trend	false	0.3633	-0.9091	18
		not enough data points					4
Lake Chad	P	1979-2016	increasing	true	0.0194	2.3384	38
	Q	1983-1986	no trend	false	0.3081	-1.0190	4
		not enough data points					5
Maputo	P	1979-2016	no trend	false	0.3393	-0.9555	38
	Q	1953-1983	no trend	false	0.1261	-1.5297	31
Mono	ET	1979-2007	no trend	false	0.5115	-0.6565	29
Niger	ET	1979-2006	no trend	false	0.6214	0.4939	28
		not enough data points					6
Nile	P	1979-2016	no trend	false	0.2909	1.0560	38
	Q	1912-1984	no trend	false	0.0693	1.8164	56
Okavango	ET	1979-2014	increasing	true	0.0127	2.4926	36
Olifant	ET	1979-2014	no trend	false	0.9457	0.0681	36
Orange	ET	1979-2016	no trend	false	0.6691	0.4274	38
Queme	ET	1979-80, 1982-84, 1990-2005, 2007	no trend	false	0.3377	0.9587	22
			no overlapping data				
Rufiji	P	1979-2016	no trend	False	0.6508	-0.4526	38
	Q	1954-1978	no trend	False	0.9741	-0.0324	20
Sassandra	ET	1979-1996	no trend	false	0.8796	0.1515	18
		not enough data points					3
Save	P	1979-2016	no trend	False	0.8801	0.1509	38
	Q	1968-1981	increasing	True	0.0118	2.5183	14
Senegal	ET	1979-1989	no trend	false	0.2129	1.2456	11
			no overlapping data				
Tana	P	1979-2016	decreasing	True	0.0006	-3.4447	38
	Q	1975-1978	no trend	False	0.7341	-0.3397	4
		not enough data points					8
Upper Blue Nile	P	1979-2016	no trend	False	0.6875	-0.4023	38
	Q	1961-1983	no trend	False	0.1339	-1.4988	26
		not enough data points					3
Void	P	1979-2016	no trend	False	0.1251	-1.5338	38
	Q	1979-1981	increasing	True	0.0483	1.9748	7
Zambezi	ET	1979-1990	no trend	false	0.5371	0.6172	12

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

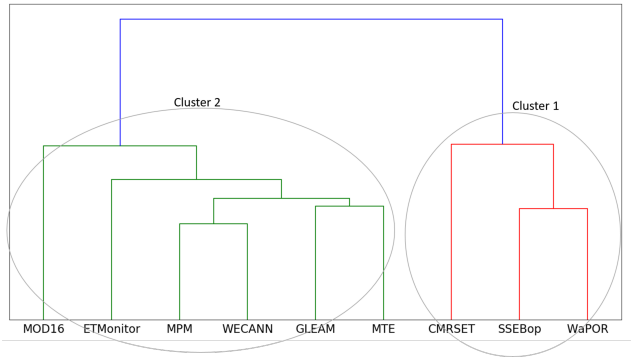

Figure 5. Dendogram after performing a cluster analysis showing the overall level of similarity between the RS products and MPM

correlation between the WB inferred ET using EWEMBI, CHIRPS and MSWEP precipitation is relatively high ranging from 88 to 95 percent. MTE, WaPOR and WECANN estimates show the highest correlations while SSEBop shows the lowest correlations with WB estimations.

Figures 7 and 8 shows the average percentage difference and weighted average percentage difference in the long-term annual average basin mean ET estimates between the RS products and the WB ET respectively. The trends and percentage differences across the products are very similar in both figures. WaPOR, SSEBop and CRMSET show the smallest biases, maximum of 6 and 9 percent, compared with the WB ET estimates derived using the EWEMBI and CHIRPS precipitation respectively. While SSEBop, the MPM and MTE show the smallest biases, maximum 10 percent, compared with WB ET estimates derived using MSWEP precipitation. GLEAM and ETMonitor show the largest biases, maximum 28, 32 and 22 percent, compared with WB estimates using EWEMBI, CHIRPS and MSWEP precipitation respectively when looking at average percentage difference. While GLEAM and MOD16 show the largest biases, maximum 36, 36 and 27 percent, compared with WB estimates using EWEMBI, CHIRPS and MSWEP precipitation products respectively when looking at the weighted average percentage difference. The mean difference ranges between 13-319 mm/year from a total average of 849 mm and 8-394 mm/year from a total weighted average of 1152 mm.

Table 3 shows the ranking of the RS products based on the mean WB ET derived using the three precipitation products for each of the calculated statistics. Considering the correlation is relatively high for all the products, we see that the higher ranked products are WaPOR, SSEBop and CMRSET, while GLEAM and ETMonitor are ranked as the lowest.

4.3 Performance with characteristics land cover elements

Figure 9 shows a section of the Nile basin where large irrigation occurs from the Nile Delta in Egypt all the way down to the Gezira scheme in Sudan. This area was selected as it was easiest to view the differences between products on how well they performed in showing the spatial distribution of ET since the ET is relatively higher in these areas than surrounding areas. Most of the products are able to capture the spatial distribution of irrigation patterns in this area with some products performing better than others except for GLEAM. Even the courser products, WECANN and MTE can also slightly capture higher ET in these

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

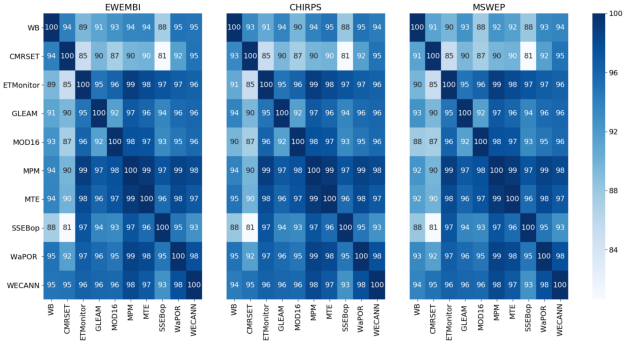

Figure 6. Correlation between long-term mean WB inferred ET and RS derived ET across basins using three different precipitation products (EWEMBI (left), CHIRPS (middle) and MSWEP (right))

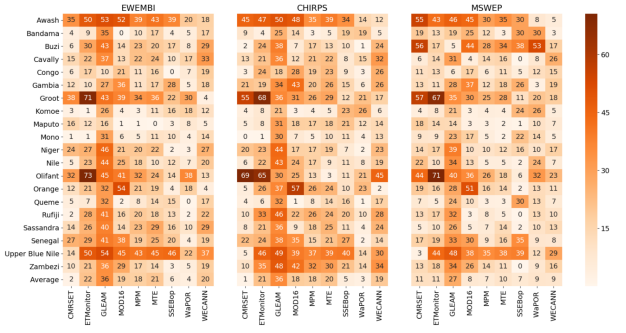

Figure 7. Percentage difference between long-term mean WB inferred ET and RS derived ET across basins using three different precipitation products (EWEMBI (left), CHIRPS (middle) and MSWEP (right))

larger irrigation areas. As expected the higher resolution products, WaPOR, CMRSET, SSEBop and ETMonitor capture the spatial patterns of ET across these areas very well. From visual inspection we ranked the performance of each of the products in capturing the spatial distribution of ET in irrigated areas as seen in Table 3. We see that WaPOR and CMRSET rank the highest while GLEAM and WECANN rank the lowest.

Figure 10 also compares the products at capturing spatial characteristics, this time using water bodies. We zoomed into Lake Victoria to clearly be able to identify and visualise the spatial patterns of ET across a large lake. The majority of products here do not estimate ET across water bodies. Only CMRSET, ETMonitor, SSEBop and WaPOR estimate ET across Lake Victoria. We can see that CMRSET and ETMonitor show higher ET across the lake than SSEBop and WaPOR which show better characterisation of ET across water bodies, thus these products were ranked higher than the other two. The ranking based on visual inspection and magnitude can be found in Table 3. For all products that did not estimate ET across water bodies the ranking was set to 9.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

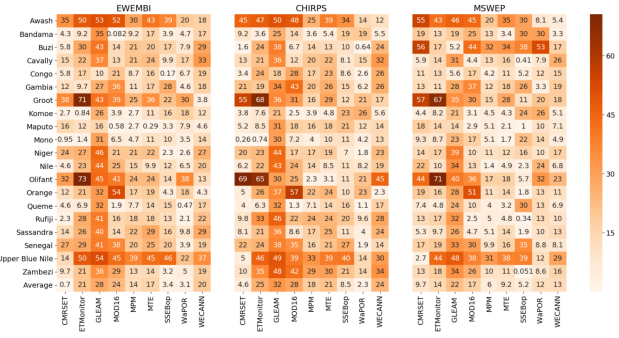

Figure 8. Weighted average (based on area) percentage difference between long-term mean WB inferred ET and RS derived ET across basins using three different precipitation products (EWEMBI (left), CHIRPS (middle) and MSWEP (right))

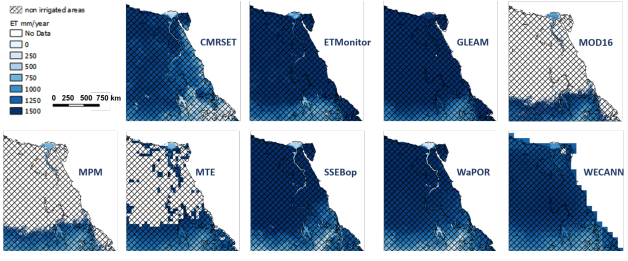

Figure 9. Comparison of RS products in representing irrigated areas. Zoomed to part of the Nile basin.

Figure 11 shows the difference between the reference crop coefficients of maize, wheat and sugarcane with the estimates long-term annual average mean crop coefficient across irrigated areas in Africa. It is clear that all products show underestimations in irrigated areas when compared with the reference crop coefficients. The shapefile used for defining the irrigated areas shows very small areas that are smaller than the highest resolution pixels from our products. Thus some of these irrigated areas are calculating ET but are not being accounted as irrigated areas within our products which may account for the underestimation. We see that the three products that have the smallest difference with the reference crop coefficients are consistently CMRSET, WaPOR and SSEBop. Figure 12 shows the difference between the the long-term annual average mean ET across water bodies from the RS ET estimates and PET estimates using the Hargreaves, P-M and P-T approaches. Only four products are presented ETMonitor, CMRSET, WaPOR and SSEBop while all other products do not calculate ET across water bodies however, some water bodies are included in those products due to the resolution of the data. Therefore, only CMRSET, ETMonitor, SSEBop and WaPOR have estimations of ET across water bodies. All four products tend to estimate ET across water bodies relatively well with small differences with the PET estimates across water bodies. All products underestimate ET across water bodies except for CMRSET which overestimates ET.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

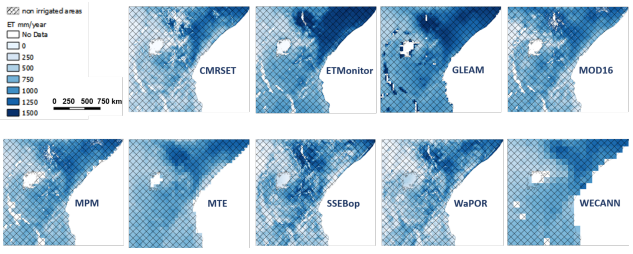

Figure 10. Comparison of RS products in representing water bodies. Zoomed to part of the Nile basin.

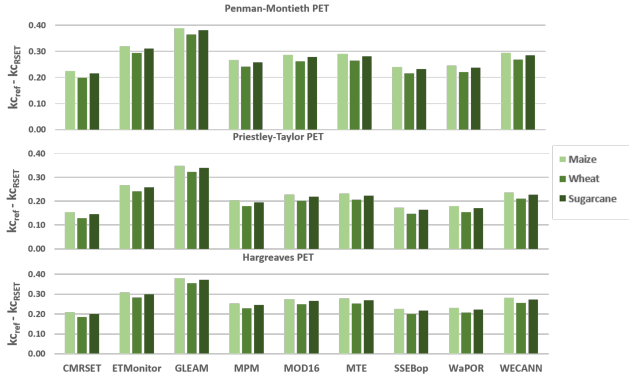

Figure 11. Average difference across long-term ET and PET estimates using (top) P-M (middle) P-T and (bottom) Hargreaves approaches for irrigated areas

Table 3 shows the ranking of the different RS products based on the mean ET across irrigated areas and water bodies compared with PET estimates. We see that the highest ranked products for irrigated areas are CMRSET and SSEBop and the lowest are GLEAM and ETMonitor, while for water bodies the highest ranked products are ETMonitor and CMRSET.

4.4 Evaluation using the Budyko curve

The results of the Budyko analysis are shown in Figs. 13, 14 and 15 which shows estimations of long-term annual average basin mean ET using the WB and RS products plotted against PET/P estimates calculated using EWEMBI, CHIRPS and MSWEP precipitation respectively. In each figure long-term annual average basin mean PET was calculated using Hargreaves, P-M and P-T approaches. We see that for all three precipitation products the different ET estimations across the basins follow the same trends with small differences in values. However, we see that the water limit and energy limit are exceeded by some ET models in figures 13, 14 and 15, thus when using EWEMBI, CHIRPS or MSWEP precipitation respectively. Exceeding the energy limit implies water is being lost through the groundwater system for example and exceeding the water limit suggests there is an additional input of water beyond precipitation. SSEBop, WECANN and CMRSET exceed the water limit in a more basins relative to other products, however their ET estimations are not necessarily further from ET estimations using the Budyko

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

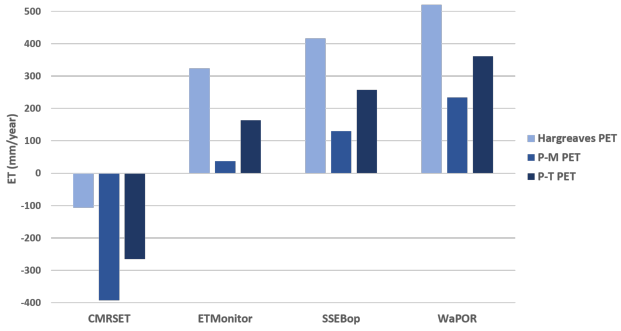

Figure 12. Average difference across long-term ET and PET estimates using (top) P-M (middle) P-T and (bottom) Hargreaves approaches for water bodies

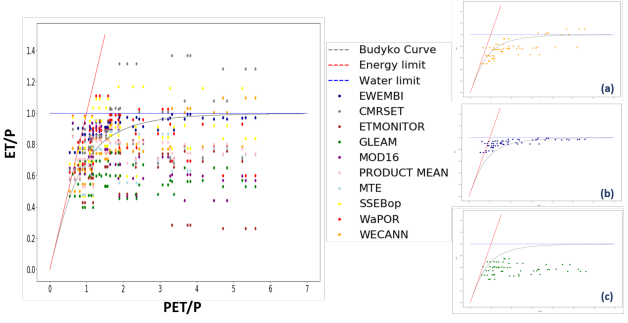

Figure 13. Evaluation of EWEMBI WB and RS derived ET estimates using the Budyko curve with PET estimates from Hargreaves, PM and PT approaches. Figure (a) WECANN ET estimations (smallest difference with Budyko curve), Fig. (b) WB ET estimations and Fig. (c) GLEAM ET estimations (largest difference with Budyko curve) plotted on the Budyko curve.

approach as given by equation 2. Figure (b) in Figs. 13, 14 and 15 plots the WB ET estimations on the Budyko curve and suggests these ET estimations across the different basins follows the curve relatively well, especially when using EWEMBI precipitation. Figures (a) and (c) of Figs. 13, 14 and 15 show the highest and lowest performing RS product respectively, in terms of being closest to the Budyko curve. Table 3 shows the ranking of the different RS products based on the Budyko evaluation. We see that the highest ranked products are WECANN and MTE and the lowest ranked products are GLEAM and MOD16.

5 Discussion

We make two assumptions in this paper regarding the methodology applied for evaluating RS derived ET estimates. The first assumption is that if no trends are present in long-term annual average mean WB inferred ET estimates across a basin, then long-term annual average mean ET estimates across basins can be compared with different time periods. This is true if long-

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

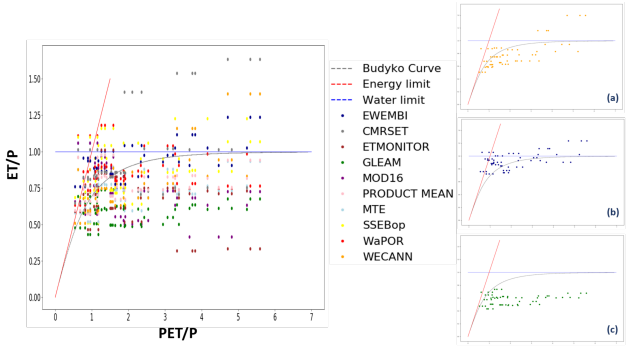

Figure 14. Evaluation of CHIRPS WB and RS derived ET estimates using the Budyko curve with PET estimates from Hargreaves, PM and PT approaches. Figure (a) WECANN ET estimations (smallest difference with Budyko curve), Fig. (b) WB ET estimations and Fig. (c) GLEAM ET estimations (largest difference with Budyko curve) plotted on the Budyko curve.

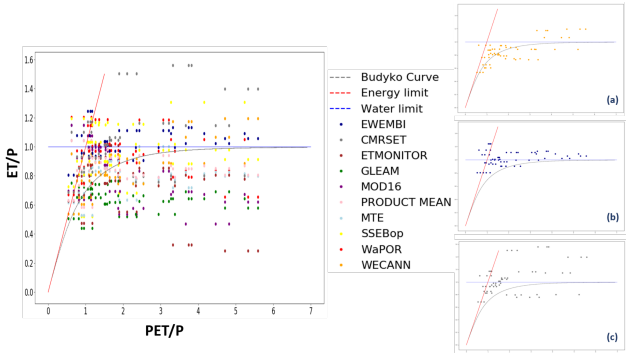

Figure 15. Evaluation of MSWEP WB and RS derived ET estimates using the Budyko curve with PET estimates from Hargreaves, PM and PT approaches. Figure (a) WECANN ET estimations (smallest difference with Budyko curve), Fig. (b) WB ET estimations and Fig. (c) CMRSET ET estimations (largest difference with Budyko curve) plotted on the Budyko curve.

term trends in global ET are not visibly present. However, Jung et al. (2010) claim that there have been declining trends in global ET estimates in the recent past along with the last major El Niño event in 1998 with largest regional contributions to the declining trend in Australia and Southern Africa. The exact opposite effect is reported by Zhang et al. (2016) which claims significant increases in global land ET trends especially in Australia and Southern Africa. Other studies also focus on investigating trends in long-term ET and do not come to a consensus as to the cause or direction of the trend (Miralles et al., 2014; Douville et al., 2013; Jung et al., 2010; Zhang et al., 2016).

For this first assumption to hold, we must also address the possibility that regardless of whether there are no trends present, the mean ET from one period may be different from another period due to precipitation variability. In this case we analysed four basins for which the calculated WB ET estimations covered the different periods of all RS ET products being evaluated. For each of the different RS product periods and for each of the four basins the corresponding mean WB ET was found. This

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

Table 3. Ranking of the RS products based on the different evaluation steps of the proposed methodology

	CMRSET	ETMonitor	GLEAM	MOD16	MPM	MTE	SSEBop	WaPOR	WECANN
Correlation	4	7	4	7	4	1	9	1	1
Average	1	8	9	5	4	7	3	2	6
Weighted Average	1	7	9	8	4	5	2	3	6
Visual inspection									
Irrigated Areas	1	4	9	6	5	8	3	2	7
Water Bodies	1	1	9	9	9	9	3	3	9
Difference ET and Reference Condition									
Irrigated Areas	1	8	9	4	5	5	2	3	7
Water Bodies	2	1	9	9	9	9	3	4	9
Budyko	6	7	9	8	4	2	5	3	1
Final ranking (with visual)	1	4	9	8	5	6	3	2	6
Final ranking (without visual)	1	7	9	8	5	4	3	2	5

was then subtracted from the calculated WB ET long-term mean. From Table 4 we see that the percentage differences in mean for the different periods ranges from 0 to a maximum of 7.4 percent of basin ET. Thus, in this study our assumption holds that if no significant trend can be found in annual long-term ET estimates then different time periods can be used due to lack of overlapping data.

The second assumption is that the water balance can be simplified to equation 1 where for annual long-term average estimates the change in storage is negligible. Many studies make this assumption for long-term averages and basin scale averages (Du et al., 2016; Taniguchi et al., 2003; Wang and Alimohammadi, 2012; Carter, 2001; Budyko, 1974). However a recent study by Rodell et al. (2018) quantified trends in terrestrial water storage using the Gravity Recovery and Climate Experiment (GRACE) for the period 2002- 2016. The largest annual trend found in this study is 20 mm per year and for the African continent can be found across sections of the Congo, Zambezi, Okavango, Cunene, Save and Rufiji basins. Of these basins Okavango, Cunene and Save are not used in this study and thus are not affected. Assuming a contribution of the largest trend in storage for the other basins this represents a maximum of 2.3 percent of the long-term annual average mean basin ET. Therefore we assumed negligible change in storage for our calculations.

The comparison between the RS products were carried out at the highest spatial resolution of the different products which is $0.0022° \times 0.0022°$. As we are resampling from coarse resolution to higher resolution the nearest neighbor method employed for completing the resampling is sufficient as the magnitude and spatial characteristics will not be altered or lost (Porwal and Katiyar, 2014; Gurjar and Padmanabhan, 2005). It must also be kept in mind that the initial spatial resolution and the temporal period under comparison are not the same for each product and this may effect the ranking that we are considering. Also many of the coarser resolution products do not estimate ET across water bodies and this may therefore explain the large biases in

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

Table 4. Differences in mean WB ET estimations for varying RS product periods

					Period					
	Total	CMRSET	ETMonitor	GLEAM	MOD16	MTE	SSEBop	WaPOR	WECANN	Average
Congo										
	1979-2010	2000-2010	2008-2010	1980-2010	2000-2010	1982-2010	2003-2010	2009-2010	2007-2010	
ET mm/year	1186	1203	1159	1196	1203	1194	1194	1168	1193	1189
Difference mm/year		17	27	10	17	8	8	18	7	14
% difference		1.4	2.3	0.8	1.4	0.7	0.7	1.5	0.6	1.2
Groot										
	1979-2015	2000-2013	2008-2013	1980-2015	2000-2014	1982-2012	2003-2015	2009-2015	2007-2015	
ET mm/year	373	390	381	377	390	371	387	396	392	386
Difference mm/year		17	8	4	17	2	14	23	19	13
% difference		4.4	2.1	1.1	4.4	0.5	3.6	5.8	4.9	3.4
Olifant										
	1979-2014	2000-2013	2008-2013	1980-2014	2000-2014	1982-2012	2003-2014	2009-2014	2007-2014	
ET mm/year	278	279	293	284	278	286	272	275	296	283
Difference mm/year		1	15	6	0	8	6	3	18	7
% difference		0.4	5.1	2.1	0.0	2.8	2.2	1.1	6.1	2.5
Orange										
	1979-2015	2000-2013	2008-2013	1980-2015	2000-2014	1982-2012	2003-2015	2009-2015	2007-2015	
ET mm/year	349	377	376	356	374	362	351	350	351	362
Difference mm/year		28	27	7	25	13	2	1	2	13
% difference		7.4	7.2	2.0	6.7	3.6	0.6	0.3	0.6	3.6

certain products when comparing ET estimations with the WB ET estimations. WaPOR, ETMonitor and WECANN have less than 10 years in total coverage in order to calculate their long-term annual average.

We used the assumption that where there is ample water ET equals PET (McMahon et al., 2013) and thus applied this assumption for evaluating our ET products for irrigated areas and water bodies. The assumption seems to hold quite well for some of the products when evaluating water bodies. We also used the assumption that for irrigated areas ET/PET should equal the crop coefficient. However all products seem to be underestimating the long-term annual average ET over irrigated areas when compared with crop coefficients. One of the largest factors in this is that we take the entire irrigated area mean to calculate one crop coefficient, where we know there are many different crops being grown and thus should account for many different crop coefficients that is not feasible in this study. Also areas much smaller than the highest resolution product are being taken into consideration in the irrigated areas and thus ET in these areas are not being calculated as irrigated area pixels. This can effect the calculation of ET and in most cases would be underestimated. Therefore the resolution of the product in this case, is also a factor for this underestimation.

Evaluation of the spatial characteristics is completed using two steps, the comparison of land cover elements with PET estimates and visual interpretation. There are two issues involved in this spatial comparison. Firstly, the evaluation is taking

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

place based on products originating with different resolutions. Thus, the view that higher resolution products will outperform the coarser resolution products, which is generally the case. However, we can also see that coarser resolution products, namely WECANN and MTE, outperform the higher resolution product GLEAM in these spatial characteristics and thus, this is not always the case. Also, the spatial resolution of the ET estimates used may also be a critical element in determining which product is of use for the user. Secondly, the visual interpretation can be viewed as quite arbitrary and subjective to the evaluator's eye. This again is the case, however by using land cover elements that are large and easy to visualize, such as irrigated areas and water bodies, the relative subjectivity can be reduced.

Looking at the overall level of similarity between the products in Fig. 5 we can see that for the cluster between CRMSET, SSEBop and WaPOR all products use MODIS as an input. SSEBop and WaPOR both use the P-M method for the calculation of ET, while CMRSET uses the P-T method. ETMonitor and MOD16 also use MODIS as an input with MODIS using the P-M method for ET calculation and ETMonitor using both Shuttleworth-Wallace and the P-M method, however both are found in the second cluster. The remaining products within the second cluster use different inputs and different ET estimation methods. Thus, no patterns can be inferred through the cluster analysis by looking at the input or ET calculation method. What is clear is that the first cluster contains the products which overall rank the best in terms of ET estimation based on the proposed methodology.

The overall ranking for each product was based on the average ranking of the different comparative elements. An overall ranking was performed including the visual inspection of the land cover elements, however was also performed without including the visual inspection due to this being rather subjective based on the analyst. This does not affect the ranking of the top three or the lowest two ranked products but changes the order of the products ranked in the middle. CMRSET, WaPOR and SSEBop are consistently ranked 1, 2 and 3, respectively. The lowest ranked products in both cases are GLEAM and MOD16. MPM is consistently ranked 5 in both cases. MTE and WECANN rank higher without visual inspection from positions 6 to 4 and 6 to 5, respectively. ETMonitor's ranking position changes the most ranking lower without visual inspection going from position 4 to 7.

6 Conclusions

This study focuses on the question of whether or not we can trust remote sensing ET products over Africa. By trying to overcome the problem of the lack of data for validation and evaluation purposes the methodology proposed can identify which products perform well in terms of biases, magnitudes and spatial characteristics. Using observations of discharge and observation based precipitation products to infer long-term annual average mean ET estimates at the basin scale and overcoming the lack of overlapping data for comparison by using different time periods for calculation of our long-term annual averages, RS derived ET estimations were evaluated. Based on the different elements being analysed CMRSET, WaPOR and SSEBop capture the magnitude of ET showing small biases in the long-term annual average mean ET across basins. The same products also capture the spatial distribution of the ET patterns well along with ETMonitor. WECANN performs well in both the correlation and Budyko analysis. The high correlation statistics indicate a good spatial distribution of WECANN ET magnitudes but the

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

product seems to show bias in ET estimations. This is contradictory with WECANN ranking high in the Budyko analysis which indicates small differences with ET estimates using the Budyko curve. GLEAM and MOD16 are consistently ranked low in both spatial pattern analysis and in terms of ET magnitude estimation. Therefore, if we answer our question of whether to trust remote sensing estimates of ET across Africa, the answer is not black and white. Yes, in general we can trust some products at least based on the products under evaluation in this study. CMRSET, WaPOR and SSEBop show low biases in estimations and a good spatial distribution of ET patterns. Each of these products have relatively high resolutions and both CMRSET and SSEBop are global products. Depending on the study under question, other products can also be used, however the bias in magnitudes need to be kept in mind. From this analysis at the African scale, there are better products to use than GLEAM and MOD16 which do not perform well in many of the evaluated criteria.

Author contributions. IW and AVG conceived and designed the alternate methodology for evaluation of large scale RS ET products. IW performed the required data analysis using scripts written by IW. IW and AVG prepared the structure of the manuscript. IW wrote the initial draft of the paper. AVG and WB supervised the research and contributed to improving the manuscript prior to submission. MM contributed to improving the manuscript prior to submission. LJ made availabile ETMonitor data that is not openly accessible.

"Data Availability" - data used in this analysis that is openly accessible can be accessed when requested by emailing the first author.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. This research is supported by the ERA4CS CIREG and TORUS projects. Most spatial data layers can be accessed through the public domain, however we would still like to thanks the teams providing the CMRSET, GLEAM, MOD16, MTE, SSEBop, WaPOR, WECANN, CHIRPS, MSWEP, EWEMBI and PET datasets. We thank the team at Global Runoff Data Centre (GRDC), 56068 Koblenz, Germany for providing the discharge data. We thank Wim Thiery and Steven Eisenreich for their advice on improvements to the scientific content.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Hydrology and Earth System Sciences
Discussions

References

Alemohammad, S. H., Fang, B., Konings, A. G., Aires, F., Green, J. K., Kolassa, J., Miralles, D., Prigent, C., and Gentine, P.: Water, Energy, and Carbon with Artificial Neural Networks (WECANN): a statistically based estimate of global surface turbulent fluxes and gross primary productivity using solar-induced fluorescence, Biogeosciences, 14, 4101–4124, https://doi.org/10.5194/bg-14-4101-2017, https://doi.org/10.5194/bg-14-4101-2017, 2017.

Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., and Pelletier, F.: Estimating the Total Fertility Rate from Multiple Imperfect Data Sources and Assessing its Uncertainty, Center for Statistics and the Social Sciences, University of Washington, 50, 1–19, https://doi.org/10.1029/2011RG000373.1.INTRODUCTION, http://www.csss.washington.edu/Papers/wp89.pdf, 2011.

Allen, R. G., Tasumi, M., and Trezza, R.: Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC) Model, Journal of irrigation and drainage engineering, 133, 380–394, https://doi.org/10.1061/ASCE0733-94372007133:4380, https://ascelibrary.org/doi/pdf/10.1061/{%}28ASCE{%}290733-9437{%}282007{%}29133{%}3A4{%}28380{%}29, 2007.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Run-ning, 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., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wil-son, K., and Wofsy, S.: 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, 2415–2434, 2001.

Bastiaanssen, W. G., Karimi, P., Rebelo, L. M., Duan, Z., Senay, G., Muthuwatte, L., and Smakhtin, V.: Earth observation based assessment of the water production and water consumption of Nile basin agro-ecosystems, Remote Sensing, 6, 10 306–10 334, https://doi.org/10.3390/rs61110306, 2014.

Bastiaanssen, W. G. M., Cheema, M. J. M., Immerzeel, W. W., Miltenburg, I. J., and Pelgrum, H.: Surface energy balance and actual evapotranspiration of the transboundary Indus Basin estimated from satellite measurements and the ETLook model, Water Resource Research, 48, https://doi.org/10.1029/2011WR010482, http://nsidc.org/data/ae{_}land3.html, 2012.

Beck, H., Yang, L., Pan, M., Wood, E. F., and William, L.: MSWEP V2 global 3-hourly 0.1° precipitation: methodology and quantitative appraisal, Bulletin of the American Meteorological Society, early rele, https://doi.org/10.1175/BAMS-D-17-0138.1, 2019.

Budyko, M. I.: Climate and life, Academic Press, New York, english edn., 1974.

Carter, J. M.: Hydrologic Budgets Methods for Estimating Basin Yield and Recharge, Tech. rep., 2001.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., M. A. Balmased, G. B., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimb, L., Vitart, F., and J. N. Thépautaan: The ERA-Interim reanalysis: configuration and performance ofthe data assimilation system., Royal Meteorology Society, 137, 553–597, 2011.

Douville, H., Ribes, A., Decharme, B., Alkama, R., and Sheffield, J.: Anthropogenic influence on multidecadal changes in reconstructed global evapotranspiration, Nature Climate Change, 3, 2013.

Du, C., Sun, F., Yu, J., Liu, X., and Chen, Y.: New interpretation of the role of water balance in an extended Budyko hypothesis in arid regions, Hydrology and Earth System Sciences, 20, 393–409, https://doi.org/10.5194/hess-20-393-2016, 2016.

FAO: WaPOR Database Methodology:Level 1 Data using remote sensing in support of solutions to reduce agricultural water productivity gaps, Tech. rep., FAO, Rome, http://www.fao.org/3/I7315EN/i7315en.pdf, 2018.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes, Scientific Data, 2, 1–21, https://doi.org/10.1038/sdata.2015.66, 2015.

Gerrits, A. M., Savenije, H. H., Veling, E. J., and Pfister, L.: Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model, Water Resources Research, 45, 1–15, https://doi.org/10.1029/2008WR007308, 2009.

Guerschman, J. P., Van Dijk, A. I., Mattersdorf, G., Beringer, J., Hutley, L. B., Leuning, R., Pipunic, R. C., and Sherman, B. S.: Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia, Journal of Hydrology, 369, 107–119, https://doi.org/10.1016/j.jhydrol.2009.02.013, http://dx.doi.org/10.1016/j.jhydrol.2009.02.013, 2009.

Gurjar, S. B. and Padmanabhan, N.: Study of various resampling techniques for high-resolution remote sensing imagery, Journal of the Indian Society of Remote Sensing, 33, 113–120, https://doi.org/10.1007/BF02989999, 2005.

Jiménez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M. F., Wood, E. F., Rossow, W. B., Balsamo, G., Betts, A. K., Dirmeyer, P. A., Fisher, J. B., Jung, M., Kanamitsu, M., Reichle, R. H., Reichstein, M., Rodell, M., Sheffield, J., Tu, K., and Wang, K.: Global intercomparison of 12 land surface heat flux estimates, Journal of Geophysical Research Atmospheres, 116, 1–27, https://doi.org/10.1029/2010JD014545, 2011.

Jung, H. C., Getirana, A., Policelli, F., McNally, A., Arsenault, K. R., Kumar, S., Tadesse, T., and Peters-Lidard, C. D.: Upper Blue Nile basin water budget from a multi-model perspective, Journal of Hydrology, 555, 535–546, https://doi.org/10.1016/j.jhydrol.2017.10.040, https://doi.org/10.1016/j.jhydrol.2017.10.040, 2017.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., De Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, Nature, 467, 951–954, https://doi.org/10.1038/nature09396, 2010.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, Journal of Geophysical Research: Biogeosciences, 116, 1–16, https://doi.org/10.1029/2010JG001566, 2011.

Kendall, M.: Rank correlation methods, Griffin, Oxford, England, 1948.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., and Frenken, K.: Global Reservoir and Dam (GRanD) database, Tech. Rep. March, https://doi.org/10.1128/JCM.01792-10, http://www.gwsp.org/products/grand-database.html, 2011.

Mann, H.: Non-parametric tests against trend, Econometrica, 13, 163–171, 1945.

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev, 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, www.geosci-model-dev.net/10/1903/2017/, 2017.

Matzke, B. D., Wilson, J. E., Newburn, L., Dowson, S. T., Hathaway, J. E., Sego, L. H., Bramer, L. M., and Pulsipher, B. A.: Mann-Kendall Test For Monotonic Trend, Tech. rep., Pacific Northwest National Laboratory, Richland, Washington, 2014.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., and McVicar, T. R.: Estimating actual , potential , reference crop and pan evaporation using standard meteorological data : a pragmatic synthesis, Hydrol. Earth Syst. Sci, 17, 1331–1363, https://doi.org/10.5194/hess-17-1331-2013, 2013.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Hydrology and Earth System Sciences Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci, 15, 453–469, https://doi.org/10.5194/hess-15-453-2011, www.hydrol-earth-syst-sci.net/15/453/2011/, 2011.

Miralles, D. G., Van Den Berg, M. J., Gash, J. H., Parinussa, R. M., De Jeu, R. A. M., Beck, H. E., Holmes, T. R. H., Jiménez, C., Verhoest, N. E. C., Dorigo, W. A., Teuling, A. J., and Johannes Dolman, A.: El Niño-La Niña cycle and recent trends in continental evaporation, Nature Climate Change, 4, 122–126, https://doi.org/10.1038/nclimate2068, http://dx.doi.org/10.1038/nclimate2068, 2014.

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., Mccabe, M. F., Hirschi, M., Martens, B., Dolman, A. J., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project-Part 2: Evaluation of global terrestrial evaporation data sets, Hydrol. Earth Syst. Sci, 20, 823–842, https://doi.org/10.5194/hess-20-823-2016, www.hydrol-earth-syst-sci.net/20/823/2016/, 2016.

Montieth, L. J.: State and movement of water in living organisms, in: 19th Symposium of Evaporation and the Environment, Cambridge University Press, Swansea, London, 1965.

Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W.: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sensing of Environment, 111, 519–536, https://doi.org/10.1016/j.rse.2006.07.007, www.elsevier.com/locate/rse, 2007.

Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorith, Remote Sensing of Environment, 115, 1781–1800, https://doi.org/10.1016/j.rse.2007.04.015, http://www.taiga.net/wolfcreek/Proceed-, 2011.

Porwal, S. and Katiyar, S. K.: Performance evaluation of various resampling techniques on IRS imagery, 2014 7th International Conference on Contemporary Computing, IC3 2014, pp. 489–494, https://doi.org/10.1109/IC3.2014.6897222, 2014.

Priestley, B. and Taylor, R.: On the assessment of surface heat flux and exporation using large-scale parameters, Monthly Weather Review, 100, 81–92, 1972.

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., and Lo, M.-H.: Emerging trends in global freshwater availability, Nature, https://doi.org/10.1038/s41586-018-0123-1, https://doi.org/10.1038/s41586-018-0123-1, 2018.

Savoca, M. E., Senay, G. B., Maupin, M. A., Kenny, J. F., and Perry, C. A.: Actual Evapotranspiration Modeling Using the Operational Simplified Surface Energy Balance (SSEBop) Approach, Tech. rep., U.S. Geological Survey Scientific Investigations Report 2013-5126, http://pubs.usgs.gov/sir/2013/5126., 2013.

Senay, G., Budde, M., and Verdin, J.: Enhancing the Simplified Surface Energy Balance (SSEB) approach for estimating landscape ET-Validation with the METRIC model, Agricultural Water Management, 98, 606–618, 2011a.

Senay, G., Leake, S., Nagler, P., Artan, G., Dickinson, J., Cordova, J., and Glenn, E.: No Title, Hydrological Processes, 25, 4037–4049, 2011b.

Senay, G. B., Budde, M., Verdin, J. P., and Assefa M. Melesse: A Coupled Remote Sensing and Simplified Surface Energy Balance Approach to Estimate Actual Evapotranspiration from Irrigated Fields, Sensors, 7, 979–1000, 2007.

Shuttleworth, W. and Wallace, J.: Evaporation from sparse crops – an energy combination theory., Royal Meteorology Society, 111, 839–855, 1985.

Siebert, S., Henrich, V., Frenken, K., and Burke, J.: Global Map of Irrigation Areas version 5, 2013.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Sperna Weiland, F., Lopez, P., van Dijk, A., and Schellekens, J.: Global high-resolution reference potential evaporation, 21st International Congress on Modelling and Simulation, pp. 2548–2554, 2015.

Sposito, G.: Understanding the budyko equation, Water (Switzerland), 9, 1–14, https://doi.org/10.3390/w9040236, 2017.

Stackhouse, P. W., Gupta, S. K., Cox, S. J., Mikovitz, C., Zhang, T., and Hinkelman, L. M.: The NASA/GEWEX surface radiation budget release 3.0: 24.5-year dataset, GEWEX News, 21, 10–12, 2011.

Taniguchi, M., Burnett, W., Cable, J., and Turnerd, J.: Assessment methodologies for submarine groundwater discharge, Land and Marine Hydrogeology, pp. 1–23, 2003.

Teuling, A. J., Hirschi, M., Ohmura, A., Wild, M., Reichstein, M., Ciais, P., Buchmann, N., Ammann, C., Montagnani, L., Richardson, A. D., Wohlfahrt, G., and Seneviratne, S. I.: A regional perspective on trends in continental evaporation, Geophysical Research Letters, 36, 1–5, https://doi.org/10.1029/2008GL036584, 2009.

Trambauer, P., Dutra, E., Maskey, S., Werner, M., Pappenberger, F., Van Beek, L. P. H., and Uhlenbrook, S.: Comparison of different evaporation estimates over the African continent, Hydrol. Earth Syst. Sci, 18, 193–212, https://doi.org/10.5194/hess-18-193-2014, www.hydrol-earth-syst-sci.net/18/193/2014/, 2014.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, Biogeosciences, 13, 4291–4313, https://doi.org/10.5194/bg-13-4291-2016, www.biogeosciences.net/13/4291/2016/, 2016.

Vinukollu, R. K., Meynadier, R., Sheffield, J., and Wood, E. F.: Multi-model, multi-sensor estimates of global evapotranspiration: climatology, uncertainties and trends, Hydrological Processes, 25, 3993–4010, https://doi.org/10.1002/hyp.8393, 2011a.

Vinukollu, R. K., Wood, E. F., Ferguson, C. R., and Fisher, J. B.: Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches, https://doi.org/10.1016/j.rse.2010.11.006, http://www.wmo.ch/pages/, 2011b.

Wang, D. and Alimohammadi, N.: Responses of annual runoff, evaporation, and storage change to climate variability at the watershed scale, Water Resources Research, 48, https://doi.org/10.1029/2011WR011444, 2012.

Wartenburger, R., Seneviratne, S. I., Hirschi, M., Chang, J., Ciais, P., Deryng, D., Elliott, J., Folberth, C., Gosling, S. N., Gudmundsson, L., Henrot, A.-J., Hickler, T., Ito, A., Khabarov, N., Kim, H., Leng, G., Liu, J., Liu, X., Masaki, Y., Morfopoulos, C., Muller, C., Schmied, H., Nishina, K., Orth, R., Pokhrel, Y., Pugh, T., Satoh, Y., Schaphoff, S., Schmid, E., Sheffield, J., Stacke, T., Steinkamp, J., Tang, Q., Thiery, W., Wada, Y., Wang, X., Weedon, G., Yang, H., and Zhou, T.: Evapotranspiration simulations in ISIMIP2a-Evaluation of spatio-temporal characteristics with a comprehensive ensemble of independent datasets, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aac4bb, https://doi.org/10.1088/1748-9326/aac4bb, 2018.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH ForcingData methodology applied to ERA-Interim reanalysis data, Water Resource Research, 50, 7505–7514, https://doi.org/doi:10.1002/2014WR015638, 2014.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B. E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., and Verma, S.: Energy balance closure at FLUXNET sites, Tech. rep., https://nature.berkeley.edu/ahg/pubs/Energy.pdf, 2002.

https://doi.org/10.5194/hess-2019-233
Preprint. Discussion started: 14 June 2019

Zhang, Y., Peña-Arancibia, J. L., Mcvicar, T. R., Chiew, F. H. S., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang, Y., Liu, Y. Y., Miralles, D. G., and Pan, M.: Multi-decadal trends in global terrestrial evapotranspiration and its components OPEN, Scientific Reports, 6, https://doi.org/10.1038/srep19124, www.nature.com/scientificreports/, 2016.

Zheng, C., Jia, L., Hu, G., Lu, J., Wang, K., and Li, Z.: Global evapotranspiration derived by ETMonitor model based on earth observations, International Geoscience and Remote Sensing Symposium (IGARSS), 2016-Novem, 222–225, https://doi.org/10.1109/IGARSS.2016.7729049, 2016.