Comparative Analysis Among Near-Operational Evapotranspiration Products for the Nile Basin Based on Earth Observations

December 8th 2014
R.W. Hofste

Delft University of Technology

National University of Singapore

Comparative Analysis Among Near-Operational Evapotranspiration Products for the Nile Basin Based on Earth Observations First Steps Towards an Ensemble Product

Rutger Willem Hofste
December 2014
1355260

Master thesis in partial fulfilment of the requirements for the double degree of:
Double Master of Science
in Hydraulic Engineering and Water Resources Management
at the Delft University of Technology and the National University of Singapore

Graduation Committee

Prof. dr. Wim Bastiaanssen (W.G.M.)	Delft University of Technology
Prof. dr.Massimo Menenti (M.)	Delft University of Technology
Dr. Yoshihide Wada (Y.)	University of Utrecht
Prof. Alan Ziegler (A.D.)	National University Singapore
Dr.Ir. Miriam Coenders (A.M.J.)	Delft University of Technology

TU Delft Delft University of Technology

NUS National University of Singapore

Comparative Analysis Among Near-Operational Evapotranspiration Products for the Nile Basin Based on Earth Observations; First Steps Towards an Ensemble ET Product

R.W. Hofste[a,*]

[a]*Delft University of Technology, Department of Water Management, Stevinweg 1, Delft, The Netherlands*

Abstract

Following the success of open access rainfall products based on earth observation data, similar global products are now under development for actual evapotranspiration. In this research, seven prototype evapotranspiration products based on MODIS and geostationary satellites were compared for the Nile basin. They include the special MOD16 version of the Nile basin, SEBS, SSEBSop, ALEXI, GLEAM, LSA-SAF and CMRSET. Most datasets are not yet released, and are in their testing phase before being disseminated to a wide audience. These remotely sensed ET products are fundamentally different in their parameterizations, ranging from reference ET based fractions products to two-layer turbulent transfer schemes for sensible and latent heat. Furthermore the used spectral radiances vary widely from near-infrared and shortwave infrared reflectance to thermal infrared and microwaves emissions. This study covers the time span 2005 -2010 for the Nile basin. The independent validation datasets are based on flux towers and augmented with water balances of catchments and subbasins of the Nile system. Significant differences between the different evapotranspiration products are observed both on subbasin and field scale. A comparison based on land-use class shows that for some land-use classes the estimation of the varies products correspond whereas for other land-use classes the differences between the products is substantial. The different input data and different parameterization of the ET products, together with the large observed differences in ET values for certain land-classes and areas, calls for an ensemble product. A set of ET products based on different level of physics and spectral data is the necessary basis for creating an ensemble ET product. The validation reveals that using an ensemble product instead of a single remote sensing evapotranspiration product yields more reliable and consistent results when compared on a subbasin and basin level. Three ensemble products have been proposed and evaluated in this study: A mean product, an ensemble product based on land-class selection and an ensemble product based on outlier statistics. The ensemble mean product based on land-use class selection is a logical solution because it recognizes that not all ET products have an equally good performance for various land surface conditions. Although all proposed ensemble products show better results when compared to subbasin water-balances, more research is needed to further formulate the best ensemble ET product.

1. Introduction

Water stress increases due to climate change, economic development and population growth in many river basins worldwide. The hydrological community is working hard on getting better data on the hydrological processes and quantify the amount of utilized and utilizable water resources for mitigating adverse impacts of water stress. Hydrological models are the classical tool for getting a consistent picture of water stocks, flows and fluxes across river basins, especially when basins are ungauged (Hrachowitz et al., 2013). While the vast majority of the hydrological models is calibrated on the basis of river flows (Wood et al., 2011) (Schuol et al., 2008) (van Dijk et al., 2014), there is gradual realization within the hydrological and water resources community that hydrological model calibration on the basis of remotely sensed ET data will improve the simulation of spatio-temporal variability of key hydrological process (Schuurmans et al., 2003)(Winsemius et al., 2008)(Cheema et al., 2014).

Evaporation, transpiration and interception (together they form evapotranspiation or ET) are the most important terms that constitute consumptive use (Perry, 2007). The remote sensing community has been active since the seventies to develop algorithms that can compute ET (Jackson (Jackson, 1977) (Menenti et al., 1989) (Soer, 1980) (SEGUIN and ITIER, 1983). Many fundamentally different remote sensing (R-S) products were developed since then, and most of them are calibrated to minimize the error with ground observations, notably flux tower observations and water balances. Bastiaanssen et al. (2014) reviewed the attainable accuracies of the accumulated ET values for minimum durations of a growing season up to one year. They concluded that at this longer time scale, most stud-

*Corresponding author
Email address: R.W.Hofste@student.tudelft.nl (R.W. Hofste)

ies are able to obtain an accuracy of 95% with a standard deviation of 5 %. It is only more recently that some algorithms are applied in an operational context. The ET model developed by D'Urso et al. (1992) is for instance applied operationally to irrigated stone fruit farmers in Southern Italy. The METRIC model (Allen et al., 2007) is applied to support the compliance of water rights and groundwater recharge in the Snake river plain in Idaho (USA). The SEBAL model (Bastiaanssen et al., 1994) is implemented for monitoring the water productivity of grape crops in the Western Cape. Explicit spatial information on water depletion and net withdrawal processes in complex river basins can benefit from a standard framework. The water accounting+ framework can provide explicit quantitative information on the hydrological status of complex river basins (Karimi et al., 2013).Water accounting+ requires ET data to be available for all river basins in the world, and in a standard data format with sufficient spatial detail and preferably at a daily basis (Karimi et al., 2013). Such data demand can only be fulfilled with operational and global scale 1 km ET products. The ET data from remote sensing has to be reproducible, verifiable and reliable, if they are included in water resources assessment reporting. The bandwidth of uncertainty should be identified as to properly inform the decision makers on the consumption of water, and the withdrawals of surface and groundwater (Hoekstra and Chapagain, 2007). There is a strong need for global ET products, similar to the remotely sensed rainfall products (Serrat-Capdevila et al., 2014) (Hessels, 2015). The CHIRPS algorithm is a good example of an ensemble rainfall product that is making the best use of other existing estimates of spatial rainfall (Funk et al., 2014). For soil moisture, also some results are published that suggest an ensemble product to be superior (Hain et al., 2011) (Parinussa et al., 2014). For ET, such an ensemble product does not yet exist, mainly because there are insufficient ET products available to base an ensemble mean procedure upon.

Cloud cover is a large obstacle in order to estimate land surface spectral reflectance and emittance. The land surface temperature is the most sensitive to the absorption of atmospheric gasses. The advantage of geostationary satellite observations is that half hourly land surface temperatures can be measured. This increases the chances for getting cloud free products. The MODIS products create a cloud free product on the basis of 8 and 16 day thermal infrared time series. The practical implication of missing data, is that gaps in the ET fluxes need to be filled. Most of the global products focus on monthly accumulated ET values, and they thus all have to make an assessment for cloud covered days. This in fact is the weak point of operational products, and it is interesting to observe that most publications provide little background information on how this issue is mitigated.

The comparison of different ET products based on satellite observations is not new. (Trambauer et al., 2013) compared several R-S ET products including the original MOD16 and GLEAM products. (Trambauer et al., 2013) concluded that the choice of meteorological input data has a significant effect on the output. More recently, (Hu et al., 2014) compared and validated MOD16 and LSA SAF products over Europe. They concluded that at local scale, LSA SAF is closer to eddy covariance observations of 15 flux towers than the standard MOD16 product, and that both products perform worse for areas with a shortage of moisture. Yilmaz et al. (2014) compared ALEXI results with the standard MOD16 product and also with the prognostic Noah land surface model. They concluded that ALEXI is better to capture hydrological process not related to rainfall, such as the ET from irrigation systems. They made another interesting observation: " The lack of reliable and openly distributed in situ data for the basin is particularly problematic given the long and continuing history of trans-boundary tensions associated with the use of Nile waters, as well as the fact that the basin is home to 238 million people who depend on the water of the basin in one way or another". The LandFlux-EVAL project evaluated and compared ET datasets as part of the LandFlux initiative of the Global Energy and Water Cycle Experiment (GEWEX), see Mueller et al. (2013). Their study compares spatial ET estimates from remote sensing algorithms, land surface models and hydrological models. While this can be regarded as a blend of the best feasible characteristics of products, it is no longer an independent dataset that could be used for the calibration of hydrological and land surface models. Also for the purpose of water accounting, it is preferred to have access to independently gathered datasets.

The current paper is one of the first studies that compared products purely based on remote sensing. Although LSA SAF and GLEAM are considered as hybrid models in which the soil water balances is having an essential contribution in the calculation of the surface energy fluxes, they are included in the analysis because their performance is very sensitive to the remote sensing input data. Their role and weight in the paper is minor, and the majority of the analysis is achieved with pure remote sensing models. The main objective of this paper is to assess whether the performance of several quasi-operational[1] and global scale ET products is similar. Another objective of this paper is to investigate whether an integrated ET product would be preferred for embedding into global water accounting procedures. Different ensemble products are proposed and validated, and the errors are estimated.

2. Theoretical framework

Several low resolution[2] remote sensing evapotranspiration (R-S ET) products exist. An ET R-S output dataset

[1]Quasi operational products provide an output within a reasonable amount of time after the observation. Depending on the satellite overpass and computation time this is normally within a month.

[2]Approximately 1 km resolution

is called a product when the data is produced under all circumstances including periods with consistent cloud cover. Another requirement for being a product is that the data should be reported with clearly defined time intervals, and preferably on a daily basis. The formulation of the algorithm should be described in the international literature, and preferably a manual with step by step explanations should be available. Often, the basic observations are based on spectral radiance products from the MODIS radiometer aboard the Terra and Aqua satellites. MSG data and observations from other geostationary satellites such as GOES and GSM are employed in addition.

ET products matching the aforementioned criteria will be evaluated for this research, including MOD16 Cleugh et al. (2007) Mu et al. (2011), MOD16NBI (Mu, 2013), SSEBop Senay et al. (2007) Velpuri et al. (2013) , ALEXI Kustas and Norman (1997) Anderson et al. (2007)), GLEAM (Miralles et al. (2011)), SEBS (Su (2002) ; Jia et al. (2003)), LSA-SAF (van den Hurk et al. (2000); Ghilain et al. (2011)) and CMRSET (Guerschman et al. (2009)). GLEAM uses in addition soil moisture data from microwave radiometers that implies that the data is provided with a very coarse resolution of 0.25° only. Yet, its global coverage and explicit recognition of interception makes the product attractive for the application in large river basins. The original MOD16 product is rejected since it did not provide data for bare areas and deserts, being a common land use class in the Nile basin. A special version of MOD16 developed for the Nile Basin Initiative was made available for the specific purpose of comparison with other models. The ET data produced by the ALEXI version7 became available during the execution of this research, and has been used for the analysis. The ET data created by the current version of LSA-SAF is only available from 2011 onward, and the data will therefore not be discussed in detail, because the main period of the comparison and validation study relates to the period 2005 to 2010. The importance of LSA-SAF for Africa is however recognized, and we therefore will make some preliminary analysis for the sake of being complete. This period is chosen because it has the highest overlap among all available ET datasets. The different ET products are described briefly below providing the reader with an understanding of some fundamental differences of the parameterization. An overview of certain key characteristics can be found in Table 1.

2.1. ALEXI, mainly based on land surface temperature

The basis of the Atmosphere Land Exchange Inverse (ALEXI) ET product is an operational two source surface energy balance model for soil and vegetation. Observations of the time rate of change in radiometric surface temperature during the morning hours using the geostationary GOES satellite is used as primary input, along with the temperature lapse rate of the atmospheric boundary layer during the same period. The ALEXI is applied at two times during the morning boundary layer growth phase (1.5 hours before and 5.5 hours after sunrise), using radiometric temperature data obtained from a geostationary platform like GOES. Energy closure over this interval is provided by a simple slab model of the atmospheric boundary layer, relating the rise in air temperature in the mixed layer to the time-integrated influx of sensible heat from the land surface. Given a diagnosis of air temperature at the interface between the surface and atmospheric boundary layer (ABL) model components, the vertical temperature gradient and sensible heat flux in the surface layer can be estimated. Since ALEXI is based on atmospheric model output data, the algorithm is designed to limit the use of routine weather station data (Anderson et al., 2011) . Latent heat from the canopy is estimated under the first-guess hypothesis of unstressed conditions using the Priestley and Taylor formulation. This will yield into an assessment of the canopy surface temperature Tc. Sensible heat from the soil and canopy are constrained by estimates of soil and canopy temperature, Ts and Tc, retrieved in the solution of the energy balance equation systems using:

$$T_{rad}(\theta) \sim\ T_c \cdot f(\theta) + T_s \cdot (1 - f(\theta))$$

In which: θ is the viewing angle of the sensor in [rad], $T_{rad}(\theta)$ the composite directional radiometric temperature in [K], $f(\theta)$ is the apparent vegetation cover fraction in [-], T_c is the directional radiometric canopy temperature in [K] and T_s is the directional radiometric surface temperature in [K](Anderson et al., 2007).

From this composite directional radiometric temperature and air temperature profile is extrapolated using a temperature lapse rate. A lapse rate is a change in temperature as a function of a change in elevation. For this, the atmospheric boundary layer method is used , as mentioned before. While the lapse rate provide essential information of the total sensible heat from a composite landscape, the pixel-to-pixel variability of surface temperature rise determines the spatial distribution of the sensible heat flux. ALEXI is thus a typical sensible heat flux model. Latent heat fluxes of soil and canopy s at the sensor overpass time are upscaled to daytime evapotranspiration (ET) assuming a self-preservation of the evaporative fraction ($\lambda E/Rn - G_0$) during daytime hours Brutsaert and Sugita (1992). A correction multiplicative coefficient of 1.1 for evaporative fraction is introduced to account for a 10% systematic underestimation of daytime average ETvalues. ALEXI version 7 with a spatial resolution of approximately 3 km based on MSG data is used in this study (Cammalleri et al., 2014).

2.2. CMRSET, mainly based on residual moisture index

The CSIRO MODIS Reflectance-based Scaling ET (CMRSET) ET product uses several spectral radiances from the MODIS radiometer (Guerschman et al., 2009). The product's goal is to generate output without the need for ancillary land use data. The basis for the product is the Priestley-Taylor

Table 1: R-S ET products overview. MBT stand for Microwave Brightness Temperature.

	ALEXI	CMRSET	GLEAM	MOD16-NBI	SEBS	SSEBop
General	LST	Residual moisture index	MBT	LAI	LST	LST
Spatial scale	0.027degree	0.05degree	0.25 degree	0.083degree	0.05degree	0.009degree
Temporal scale	Daily	Monthly	Daily	8-day	Monthly	Monthly
Available from	2007	2000	2007	2005	2000	2001
Available until	2012	2013	2010	2010	2013	2012
Standard projection	GCS_WGS_1984	GCS_WGS_1984	GCS_WGS_1984	GCS_WGS_1984		GCS_WGS_1984
Data structure	Binary/Geotiff	Matlab MAT	NetCDF	Geotiff	Matlab MAT	Geotiff
Open-access	Request	Request	Request	Request	Request	Request
Manual	0	1	0	1	0	0
Original Calibration area	USA	Australia	Global	Global	Arizona	
Used calibration area	MENEA east	uncalibrated	Global	Nile basin	Global	Africa
Primary input						
RS LAI/EVI	1	1	1	1	1	0
RS albedo	1	1	0	1	1	1
RS LST	1	0	0	1	1	1
RS soil moisture	0	0	1	0	0	0
RS solar radiation	1	0	1	1	1	0
RS land use	1	0	0	1	0	1
Primary output						
Soil evaporation	1	1	1	1	Fraction approach	Fraction approach
Vegetation transpiration	1	1	1	1	Fraction approach	Fraction approach
Interception	0	1	1	1	Fraction approach	Fraction approach
Open water	0	1	1	1	1	1
Snow and ice	0	0	1	0	0	0

(P-T) equation. The P-T potential ET is then scaled using the Enhanced Vegetation Index (EVI) data and Global Vegetation Moisture Index (GVMI) obtained from MODIS:

$$\text{EVI} = G \cdot \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_1 \cdot \rho_{red} - C_2 \cdot \rho_{blue} + L}$$

$$\text{GVMI} = \frac{(\rho_{NIR} + 0.1) - (\rho_{swir2} + 0.02)}{(\rho_{NIR} + 0.1) + (\rho_{SWIR2} + 0.02)}$$

In which: EVI is the enhanced vegetation in [-], GVMI is the Global Vegetation Moisture Index [-], ρ_{red}, ρ_{NIR}, ρ_{blue} and ρ_{SWIR2} and reflectances in red (645 nm), near-infrared (860 nm), blue (469 nm) and shortwave infrared (1640 nm) respectively and correspond to MODIS bands 1, 2, 3 and 6.In the EVI formula, G, C_1, C_2 and L are parameters that account for aerosol scattering and absorption and their values are 2.5, 6, 7.5 and 1 (Huete et al., 2002).

The GVMI allows separation between bare soils and open water when EVI is low. When EVI is high, the GVMI provides information about vegetation water content. The product has been developed, calibrated and tested on the Australian continent. The Residual Moisture Index (RMI) was calculated for each pixel as the vertical distance between its corresponding GVMI and a baseline.

$$\text{RMI} = max(0, \text{GVMI} - (k_{rmi} \cdot \text{EVI} + C_{rmi}))$$

In which: k_{rmi} is the slope and C_{rmi} is the intercept between the linear $\text{GVMI} = f(\text{EVI})$ function.

Furthermore, the crop coefficient kc is made a function of RMI to reduce the value of potential ET according to the ET P-T value to an actual value of ET:

$$K_c = K_{c-max} \cdot (1 - exp(-a \cdot \text{EVI}_r^\gamma - b \cdot \text{RMI}^\beta)$$

In which: K_{c-max} is the maximum value for the crop factor [-] and EVI_r^γ is the rescaled EVI value.

2.3. GLEAM, mainly based on microwave brightness temperatures

The Global Land surface Evaporation: the Amsterdam Methodology (GLEAM) has a spatial resolution of 0.25. The product distinguishes between 3 sources of ET namely bare soil, short vegetation and vegetation with tall canopy. Snow and ice are modeled separately and lakes and open water are not included in the product. The GLEAM model is based on the Priestley-Taylor (P-T) equation augmented with Soil Moisture (AMSR-E), vegetation density (LPRM) and estimation of rainfall interception fraction (CMORPH and TRMM). The product uses four modules: the first module calculates the interception for the vegetation with tall canopy class. The second module distributes incoming precipitation over the root zone using a water budget method. A water balance is run with todays precipitation and yesterdays ET fluxes, including a deep percolation flux that varies with moisture content in the root zone. Based on the second module and vegetation and soil parameters, a third module calculates the total stress separately for tall canopies, short vegetation and bare soil. Soil physical data is required for this model. The fourth module combines all aforementioned parameters with the P-T equation and calculates the total ET:

$$\text{ET} = S \cdot ET_{P-T} + 0.93I$$

*In which:*ET is the actual ET in [mm/day] S is the total stress factor in [-],ET_{P-T} is the potential ET based on Priestley and Taylor equation in [mm/day] and I is interception in [mm/day]

The factor 0.93 originates from Gash and Stewart (1977) that reveals that a small fraction of the energy required for interception evaporation is taken from the net available energy, and that the majority of the interception is an additive term in the ET process. The global product with 0.25 resolution is used for this study (Miralles et al., 2011).

2.4. MOD16-NBI, mainly based on leaf area index

The launch of the MODerate-resolution Imaging Spectroradiometer (MODIS) instrument on the AQUA and TERRA satellites enabled scientists to get access to a multi-spectral radiance dataset of the land surface covering the world. One of the official approved NASA products in association with MODIS data is MOD16 (Mu et al., 2013) (Mu et al., 2009). While MOD16 was launched not so long ago, it is the longest duration operational ET product that can be accessed via http://www.ntsg.umt.edu/project/mod16 . MOD16 uses the land use dependent Leaf Area Index functions based on the Enhanced Vegetation Index (EVI), augmented with meteorological reanalysis data. The product distinguishes between three types of evaporation: soil evaporation, wet canopy evaporation (interception) and plant transpiration. The transpiration product is based on the surface conductance model of (Nemani and Running, 1989) in combination with the Penman - Monteith (P-M) equation (Allen et al., 1998). For a vegetation canopy, the canopy resistance $r_{s,canopy}$ in $[s/m]$ is parameterized by:

$$r_{s,canopy} = \frac{1}{c_L} m(T_{min}) m(\text{VPD}) \text{LAI}$$

In which: $r_{s,canopy}$ is the Canopy resistance in $[s/m]$. c_L is the stomatal conductance per unit leaf area when all environmental parameters are at optimum levels for full stomatal opening in $[mmol/(m^2 \cdot s)]$. $m(T_{min})$ is an air temperature multiplier in [-]. $m(\text{VPD})$ a Vapour Pressure Deficit multiplier in [-] and LAI the Leaf Area Index in $[m^2/m^2]$

For the surface resistance and net incoming radiation, required for the determination of ET over deserts and water bodies, MODIS-based data on land use, LAI/fPAR and albedo were used. The fPAR is the fraction of Absorbed Photosynthetically Active Radiation. The actual soil evaporation λE_{soil} $[\mathrm{W/m^2}]$ is calculated using potential soil evaporation $\lambda E_{soil,pot}$ $[\mathrm{W/m^2}]$ and soil moisture constraint, assuming that soil moisture is in equilibrium with the adjacent atmospheric moisture as:

$$\lambda E_{soil} = \lambda E_{wet,soil} + \lambda E_{soil,pot}(\frac{RH}{100})^{\mathrm{VDP}/250}$$

In which: λE_{soil} is the soil evaporation in $[\mathrm{W/m^2}]$, $\lambda E_{soil,pot}$ is the potential evaporation $[\mathrm{W/m^2}]$, $\lambda E_{wet,soil}$ is the wet soil evaporation in $[\mathrm{W/m^2}]$, RH is the relative humidity in [%] and VPD is the vapour pressure deficit in [pa].

Because the original MOD16 algorithm excludes open water bodies, urban areas and deserts which are omnipresent in the Nile basin, a special version of MOD16 has been tailor made for the Nile Basin which includes these missing land use classes.

2.5. SEBS,mainly based on land surface temperature

The archive of the Surface Energy Balance System (SEBS) has been recently opened for applications in hydrological and climatological studies through the online SEBS user group. SEBS is also an example of a sensible heat flux model that is based on local minima and maxima of the sensible heat flux (Su, 2002) following the SEBI concept proposed by Menenti and Coudhury (1993). These hypothetical values are computed from extreme cases of latent heat flux, assuming zero ET when the surface resistance is infinite (rs = ∞ [s/m]) and free or maximum ET when the surface resistance is zero (rs = 0 [s/m]). This will result in an upper limit $(T_0 - T_a)_u$ for every pixel, as well as a lower limit $(T_0 - T_a)_l$. The equations for the upper and lower limit are displayed below:

$$(T_0 - T_a)_u = r_{rh,u} \cdot \frac{R_n - G_0}{\rho_a \cdot C_p}$$

$$(T_0 - T_a)_l = \frac{r_{ah,l}\frac{R_n - G_0}{\rho_a \cdot C_p} - \frac{e_0 - e_a}{\gamma}}{1 + \frac{\Delta}{\gamma}}$$

In which: R_n is the net radiation in $[\mathrm{W/m^2}]$, G_0 is the soil heat flux in $[\mathrm{W/m^2 day}]$, $r_{ah,l}$ is the aerodynamic resistance to heat transfer in $[\mathrm{s/m}]$, G_0 is the ground heat flux in $[\mathrm{W/m^2 day}]$, ρ_a is the density of air in $[\mathrm{kg/m^3}]$, C_p is the specific hear in water in $[\mathrm{kJ/kgK}]$, e_0 is the saturation vapour pressure in [Pa], e_a is the actual vapour pressure in [Pa], γ is the psychromatic constant in $[\mathrm{kPa/K}$ and Δ is the rate of change of saturation vapour pressure with temperature [Pa/K].

The radiometric surface temperature observations T_{rad} [K] of MODIS are used to compute also a true $(T_0 - T_a)$ [K] value that lies per definition in between these upper and lower limits, and this is used to estimate the relative ET (Λ_r) [-] as:

$$\Lambda_r = \frac{\lambda E}{\lambda E_p} = 1 - \frac{\frac{T_0 - T_a}{r_{ah}} - \frac{(T_0 - T_a)l}{r_{ah,l}}}{\frac{(T_0 - T_a)u}{r_{ah,u}} - \frac{(T_0 - T_a)l}{r_{ah,l}}}$$

In which: The parameters have been described in the previous equation

The SEBS model computes the latent and sensible heat fluxes according to turbulent flow transport at the land surface including advanced parameterizations of the aerodynamic parameter based on an inversed Stanton number (Su, 2002). The air temperature ought to be taken from either the Planetary Boundary Model using Bulk Atmospheric Similarity (BAS) theory of Sinnhuber et al. (2014)) or the Monin-Obukhov similarity (MOS) hypothesis for surface layer scaling if the air temperature is taken at a reference height within the near-surface layer. This study uses a 0.05 ° spatial resolution dataset.

2.6. SSEBop, mainly based on land surface temperature

The United States Geological Survey (USGS) developed an ET product based on the Simplified Surface Energy Balance Operational (SSEBop) algorithm (Senay et al., 2011). The idea behind the Simplified Surface Energy Balance Operational (SSEBop) algorithm is to integrate reference evapotranspiration or ET_0 with LST data to account for soil moisture induced evaporative stress. The reference evapotranspiration ET_0 is determined using the P-M equation (Allen et al., 1998) using the Earth Resource Observation System (EROS) and meteorological data from the National Oceanic and Atmospheric Administration (NOAA). The conversion from ET0 to a potential value for ET is accomplished using a crop coefficient kc:

$$\mathrm{ET}a = \mathrm{ET}_f \cdot k_c \mathrm{ET}o$$

In which: ET_0 is the grass reference ET for the location in $[\mathrm{mm/day}]$; k_c is a coefficient that scales the ET0 into the level of a maximum ET experienced by an aerodynamically rougher i.e. more resistance crop such as alfalfa in [-] and $\mathrm{ET}f$ is an ET fraction approach in [-]. The fraction approach is using the prevailing LST in combination with a hot and cold pixel selection is then used to reduce the potential ET (ET_0) to actual ET ($\mathrm{ET}a$). Following Bastiaanssen et al. (1994) and Allen et al. (2007), the ET fraction will determine a linear combination of hot and cold pixel properties.

$$ET_f = \frac{Th - Ts}{Th - Tc} = \frac{Th - Ts}{dt}$$

In which: ET_f is the ET fraction in [-], Ts is the satellite-observed land surface temperature of the pixel whose $\mathrm{ET}f$ is being evaluated on a given image date in [K], Th is the estimated temperature at the idealized reference hot and dry condition of the pixel for the same time period in [K],

Tc is the estimated temperature at the idealized cold and wet reference point in[K]. The difference between Th and Tc is simply the dT in [K].

In some desert areas when the Ts becomes more than Th, a negative ET_f value will be set to 0 and thus assigning 0 for $\mathrm{ET}a$.In this case, dT is pre-defined for the study location as explained in Savoca et al. (2013) and Senay et al. (2013) using the formulation below. It is calculated under clear-sky assumption and does not change from year to year, but is unique for each day and location.

$$dT = \frac{R_n r_{ah}}{\rho_a C_p}$$

In which: C_p is the specific heat of air at constant pressure or approximately 1.013 $[\mathrm{kJ/kgK}]$, Rn is clear-sky net radiation in $[\mathrm{MJ/(m^2d)}]$, r_{ah} is the aerodynamic resistance to heat flow from a hypothetical bare and dry surface in $[\mathrm{s/m}]$(Senay et al., 2013); ρ_a is the density of air in $[\mathrm{kg/m^3}]$ The Tc data is based upon the 8-daily MODIS thermal product. The USGS is publishing ET anomalies and is investigating the accuracies of ET in various locations, before this dataset will be disseminated to the public.

3. Materials and methods

3.1. Study area

The water of the Nile ranks among the highest disputed waters in the world. Two significant changes in the Nile river basin call for quantitative water assessment tools such as Water Accounting+. First there is the political upheaval of 2011, toppling the Egyptian president Mubarak. The resulting Egyptian discord gave the other riparian countries an impetus to further challenge the current water distribution. Second there is the socio-economic development of upstream Nile basin countries, notably Ethiopia which is rapidly developing her hydro power and irrigation potential. The historical distribution of the waters of the Nile, recorded in colonial style agreements is thus increasingly prone to alternation. In order to maximize the potential of the Nile and substantiate the debates about the distribution of the water, water assessment tools or models should be used. One of the determining parameters in these water assessment tools is evapotranspiration (ET). Some consultancy reports already include remote sensing ET products and even give advice on water resource management based on models with the remote sensing ET data as input. The remote sensing ET data should however be handled with care since intercomparison shows large differences between the remote sensing ET products

This research focuses on the Nile basin for three reasons. First, the riparian countries have a strong incentive to pay attention the Nile's hydrology for socio-economic reasons, making ET estimates relevant. Second, the area of the Nile basin is large enough to compare multiple remote sensing ET products on different scales ranging from basin to field scales. Third, the Nile basin consists of multiple climate zones, elevations and land-classes ranging from humid tropical forests to barren deserts. The downside of the choice for the Nile basin is the limited availability of ground observations. Since water is a political issue, information is not always shared. The last reason is at the same time a call for remote sensing observations.

The climate of the Nile basin ranges from arid (BWh) at the Nile Delta to Tropical wet (AW) near lake Victoria in the south. Part of the Ethiopian highlands have a temperate climate (C) (Peel, 2007). The Nile basin is shared by 10 countries. The basin boundaries of the Nile have in this study been taken from the Nile Basin Initiative (NBI) [3]. The Nile consists of two main branches: the Blue Nile and the White Nile. The Blue Nile originates from Lake Tana in the Ethiopian highlands. The White Nile is a continuation of the outflow from Lake Victoria. The Kagera river that originates from Burundi is one of the major tributaries to Lake Victoria. The Blue Nile accounts for approximately 2/3 of the discharge at the confluence near Khartoum in Sudan (FAO, 2011). Although the White Nile has a substantial flow, its magnitude is strongly reduced after it passes through the marshlands of the Sudd, Bhar El Ghazal, Marsha and Sobat (Mohamed and Savenije, 2014). The total ET of the Nile Basin is approximately 98% of the total rainfall hence the runoff coefficient is not more than 2% on a multi-annual basis Bastiaanssen et al. (2014) Karimi and Bastiaanssen (2014). This implies that the knowledge on the distribution of ET is very important, and that a cap of ET has been reached already. The non-consumed water is needed for leaching of salts and dilution of contaminated water. Further increases in ET is only feasible if elsewhere the ET is reduced. Increasing benefits and livelihoods from consumptive use is the challenge of the Nile basin.

3.2. Materials

The GlobCover 2008 regional dataset provides a global land cover classification system (ESA, 2008). The high number of different land classes is however less useful for the interpretation of the current spatial ET layers, and the land classes were reclassified using a majority resampling method [4]. This will make it simpler to evaluate and interpret the various ET products.

Except for SEBS and LSA SAF that have dedicated data platforms to pull the data from, all data have been collected from the original developers. The Nile Basin Initiative have asked the original developers of the MOD16

[3]The Nile basin initiative or NBI is an organization to promote better use of the waters of the Nile especially through trans-boundary cooperation and information sharing

[4]A coarser pixel that is resampled uing majority resampling takes the value of the majority of the smaller pixels. This method is useful for datasets with classes, such as land-use.

Table 2: Available ground observations, WB is water balance, EC is Eddy covariance tower and P is precipitation

Name	Type	Majority Surface type	Area $[km^2]$	Elevation $[m]$	P used	Start	End
Demokeya	EC	Mosaic Forest	n/a	529	n/a	01/2005	12/2009
Jinja	EC	Mosaic Vegetation	n/a	1181	n/a	10/2003	12/2003
Eburu	EC	Woodlands	n/a	2628	n/a	10/1998	03/1999
Ndabibi	EC	Grassland	n/a	2094	n/a	10/1998	03/1999
Blue Nile	WB	rainfed Crops	176920	1863	CHIRPS	01/2002	12/2012
Nile Delta	WB	Irrigated area	15031	6	n/a	01/1989	12/1989
Nasser	EC	Water bodies	5027	179	n/a	01/1995	12/2004
Gelgel	WB	Rainfed Crops	1656	2322	Multiple	01/2005	12/2006
Gummera	WB	Rainfed Crops	1279	2278	Multiple	01/2005	12/2006
Qarun	WB	Open water	228	-42	n/a	01/1989	12/1989

ET product to prepare a special version for the Nile with more attention to bare soil, water bodies and the ET of irrigated land. This special dataset was made available to the first and second author for scientific purposes (Mu, 2013). The time period of MOD16NBI elapses from 2005 to 2010. SSEBop has been developed for the period 2000 to 2012 covering the entire world with 1 km spatial resolution. The time step is monthly. The data is under embargo released for the sake of validation and comparison. The original ALEXI data produces by USDA in association with NOAA comes daily, and it has a 3 km pixel size due to the MSG data being used. This dataset was produced as part of an ongoing study towards water management in the Middle East and North Africa, and was for this reason by chance available. While there are plans to cover ALEXI globally, this has not been accomplished yet. Time series for the Middle East and Africa covers a period of 2007 to 2012 although strange LSA-SAF artefact are present in the 2007 data. Therefore the ALEXI product is used from 2008 onward. The CMRSET dataset was prepared by the CSIRO in collaboration with the Australian National University, and their global ET product has a spatial resolution of approximately 5 km. While they are now working on preparing a global product at 500 m pixel size, the Nile basin study was resampled to 3 km pixels.

Although the density of ground observations in the Nile basin is relatively low and not enough to draw firm conclusions, comparing the several ET products with ground observations provides useful first insight in the performance of the different ET products nevertheless. Especially the availability of independent ET data at different spatial scales is attractive to appraise the accuracy of the ET products. Most R-S ET products were calibrated and validated in other parts of the world where they have been developed, and it is interesting to compare them for the Nile basin. Most R-S ET products are available from 2005 until 2010 (see table 1) and hence coinciding ground observations had to be collected for the same period.

Three flux tower observations are available within the Nile basin and two additional towers are located just outside the Nile basin. The locations of the flux towers are depicted in figure 2 and additional information can be found in table 2. Four flux tower observations were used in this research. The Demokeya and Jinja flux towers are both part of the FluxNET program. The Jinja observations are only available for three months whereas for the Demokeya site, more data is available (Ardö et al., 2008). The ET measurments over lake Nasser are composites of multiple stations and averaged over multiple years (Elsawwaf and Willems, 2012). The results from the Eburu and Ndabibi and Jinja flux sites (Farah et al., 2004) have been omitted in this paper.

Besides flux towers, several water balances were used. The water balance of these basins has been thoroughly investigated before by Tekleab et al. (2011) and van Griensven et al. (2012). These water balances could be employed to test the data of the global R-S ET products. The area integrated rainfall remains a factor of uncertainty. Also inter-basin transfer among subbasins may take place, both via the mean stream as well as through ungauged smaller streams and overland flow. These hydrological flows have impact on the accuracy of the bulk ET of catchments and river basins because usually the ET is computed as the rest term of the catchment water balance. Nevertheless, the R-S ET products are compared in several subbasins decreasing in size. The largest basin is the Ethiopian Blue Nile subbasin with an area of 176920 km^2 which is over four times the area of the Netherlands. The discharge at the Ethiopian - Sudanese border was available for this study, and the Sudanese component of the Blue Nile basin is thus excluded from the analysis. For this water balance comparison, precipitation data from CHIRPS was used. The majority land-use class along with some other additional information can be found in table 2. The water-balance for the Ethiopian Blue Nile subbasin is based on multi-annual averages.Mohamed and Savenije (2014) Following the conclusions of Hessels (2015), the rainfall product CHIRPS has been selected for making a first order estimate of ET by considering the difference between rainfall and outflow from the Blue Nile basin. Duan (2014) estimated the water balance of the Gelgel and Gummera basins using observations of Tekleab et al. (2011). For the Gelgel and Gummera basins multiple analysis were executed including TRMMM, a calibrated version of TRMM, CHIRPS

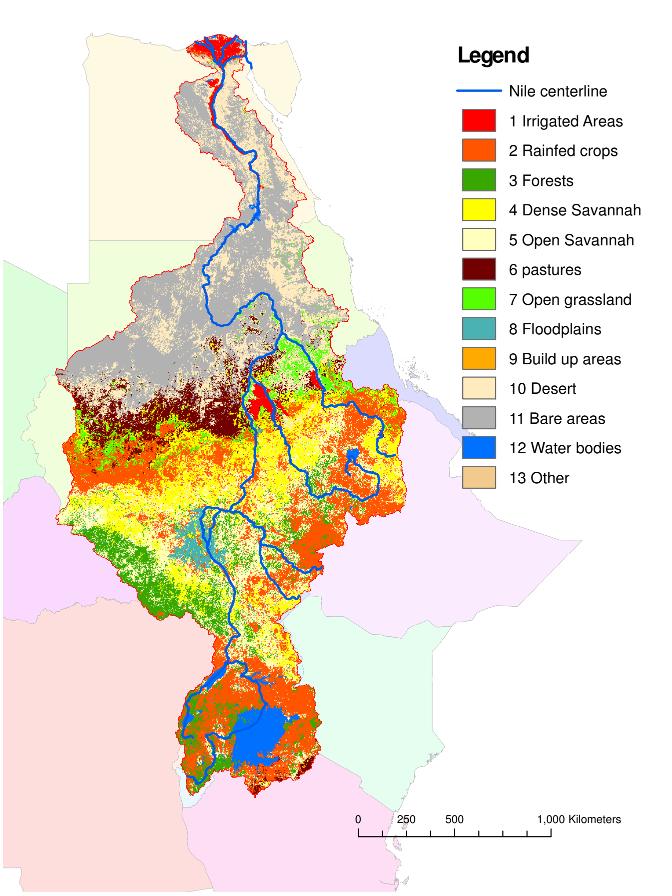

Figure 1: Land-use classes at 0.027° resolution. Boundary box Nile basin approximately 32°N 40°E 4°S 23°W

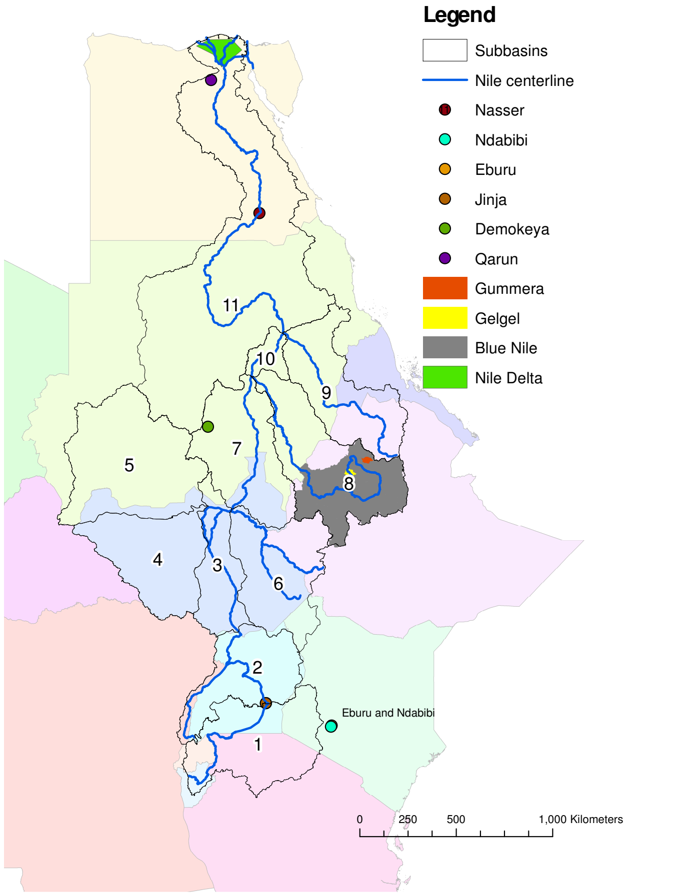

Figure 2: Ground measurement locations and subbasins. Boundary box Nile basin approximately 32°N 40°E 4°S 23°W

and the mean of 13 precipitation products. Eventually the comparison with the calibrated TRMM is used in this research.

There are two important data sources for ET in the Nile Basin that are worth exploring. They are the FAO (2011) and a more recent dataset of Bastiaanssen et al. (2014) that both calibrated ET from discharge observations and rainfall. The water balance is based on long term discharge data collected by Sutcliffe and Parks for various subbasins augmented with precipitation data by CRU 1960-1990 (Sutcliffe and Parks, 1999). The subbasins are depicted in figure 3 and are different from the subbasins used by the Nile Basin Initiative (NBI). The errors in the CRU precipitation product can easily exceed 10% Hessels (2015), so taking care is required. Nevertheless, the FAO-Nile program estimated the actual ET using a calibrated water balance model for the period 2005 to 2010 to be 1991 $[\mathrm{km}^3/\mathrm{y}]$ or 628 $[\mathrm{mm}/\mathrm{y}]$, if an area of $3,170,419$ $[\mathrm{km}^2]$ is considered (FAO, 2011). Bastiaanssen et al. (2014) calibrated the SSEBop model and reported a volumetric ET of 1987 $[\mathrm{km}^3/\mathrm{year}]$ or 616 $[\mathrm{mm}/\mathrm{y}]$ for an area of $3,229,039$ $[\mathrm{km}^2]$, covering the period 2005 to 2010. These results are very similar, and so does the spatial distribution. The dataset of FAO Nile will be used to independently verify some first ensemble mean ET products..

Two additional ET observations are used in this study to increase the understanding of irrigated areas and the downstream hydrological events.

3.3. Methods

Three standard spatial resolutions are used in this research for comparing different ET products: 0.009°, 0.027° and 0.27°. The original resolution of the Globcover land cover data is 0.009° in both latitudinal direction and longitudinal direction. This equals approximately 1kmx1km near the equator. Only the MOD16 and SSEBop products provide data on such high a resolution. Therefore the products have been resampled to 0.027° resolution. The coarser resolution covers exactly nine cells of the finer resolution and hence normal averaging was used. The land-use class dataset has been resampled using a majority resampling method as described before with the land-use class resampling. The SEBS and CRMSET products are interpolated from 0.05° to 0.027° using nearest neighbor resampling, introducing an error in the resampled products. In order to mitigate this error a third and final resolution is used to compare the products at 0.27° resolution. This resolution allows the comparison of GLEAM, SEBS and CMRSET without resampling to a finer resolution. A resolution of 0.27 ° is chosen because it is close to 0.25° (GLEAM) and a multitude of 0.027 °, allowing resampling using the average of the 100 cells at 0.027 resolution.

All R-S ET products are compared on a 0.027° and 0.27° resolution in a WGS84 geographical coordinate system. The products are compared once a month and the standard units are mm/month. If necessary the products are reprojected and re-sampled using the nearest neighbour re-sampling technique to a resolution which is close to a multitude of 0.027.

The MOD16NBI R-S ET product for example, was resampled to 0.009° (0.027° divided by 3) and then to 0.027° using normal averaging. For the SEBS and CMRSET products the original resolution of 0.05° is lower than the preferred 0.027° resolution and hence nearest neighbour resampling was used. An error is introduced when resampling to a higher resolution occurs. The comparison at 0.027° resolution can be used to evaluate this error.

Even when measured with advanced devices such as eddy covariance techniques, ground observations of ET exhibit large error margins of up to 20% (Wilson et al., 2002). It is also possible to compare R-S ET products with ground observations on a pixel basis. The footprint of a flux tower varies with wind wind velocities and wind direction and is different than the 3x3 km pixel size. Although the footprint of the flux towers are different from the pixels, the values can still be used. Visual inspection using satellite imagery shows that the surroundings are approximately the same as the flux tower site. The surrounding of the Jinja flux tower site are however not homogeneous.

4. Results

4.1. Comparison with ground observations

First, the basin-wide mean annual ET values for the Nile basin have been compared. The values appear to be rather variable, ranging from as low as 450 up to 1000 $[\mathrm{mm/y}]$. Figure 3 and table 3 depict the multi-annual average accumulated ET. The precipitation has been added as a reference. Because rainfall data has its own intrinsic uncertainty, the minimum and maximum annual precipitation product is selected from the suite of TRMM, GPCP, CRU, ECMWF, ARC, CHIRPS, RFE, PERSIANN, CMORPH, GSMaP and Tamsat products. Figure 3 shows that the SEBS model exhibits the highest ET values, and that the values even exceed the maximum rainfall product for every year during the period 2005-2010, including the very wet year 2008. Despite the complex turbulence physics, SEBS appears to be over-estimating ET. The data for LSA SAF demonstrated the opposite results for the years 2011 to 2013. This is not shown because this period lies outside the 2005-2010 time span for which most datasets are available. The ET values of LSA SAF appear to be systematically lower than the data of the other ET products. One obvious reason is that the HTESSEL model is less suitable for irrigated crops and flood plains, hence all land areas exposed to blue water withdrawals are not properly captured by LSA SAF. The basin-wide average ET for a longer period lies around 620 $[\mathrm{mm/y}]$. For this reason, the model results of SEBS and LSA SAF will not be discussed further. The results of ALEXI, SSEBop and CMRSET are near to the longer term average ET of 620 $[\mathrm{mm/y}]$)see Table 3. They will be further inspected together with MOD16NBI and GLEAM.

When compared to flux tower observations of Demokeya, it is clear that some R-S ET products show relatively good correspondence for example ALEXI7 with $r^2 = 0.83$ and $NRMSE = 0.14$ while other products do not capture differences between the different years.

The result for lake Nasser shows a remarkable good fit for MOD16NBI. The relative importance of water-balances located within the Nile basin is higher for products that concern only the Nile than for say products calibrated on a continental scale. ALEXI doesn't provide a value for open water ET and the SSEBop and CMRSET tend to underestimate the total ET value for lake Nasser. Because lake Nasser is located next to a desert and the waterlevel and hence area varies throughout the year, it is possible that pixels identified as open water are disturbed by the aforementioned effects.

Also Lake Qarun with ET = 1785 [mm/y] (Ramadan, 1989). The ET for a mixture of crops in the Nile Delta is 1180 [mm/yr] (El Quosy and El Guindy, 1989). At a presence of 10% build up areas with an ET of 100 [mm/y], the ET for the composite Nile Delta should be 1072 [mm/y]. These results are depicted in figure 4.

The difference between R-S ET products and the ground observations often reduces with increasing size of the area of interest. 4 demonstrates that the ET products show large differences on varying spatial aggregation level. Savannah and open water evaporation from a fresh water (Lake Nasser) and salt water (Lake Qarun) body are included in the dataset for fields scale. ALEXI performed well for the savannah of Sudan. The Nile Delta is used to represent a contiguous area of typical Mediterranean crops. Because ALEXI has the best agreements with the known ET volumes of the entire Nile basin, the subbasins and the flux towers, it is fair to conclude that ALEXI is among the more accurate operational ET products that currently exists. There is however not a single preferred global ET product, i.e. a model that is always superior. Each ET product shows its own behaviour with spatial scale. This is related to the type of single land use class considered: the accuracy is not identical for every hydroecosystem because every system has unique radiation, vegetation, aerodynamic and moisture characteristics. Certain spectral observations will have better capabilities than

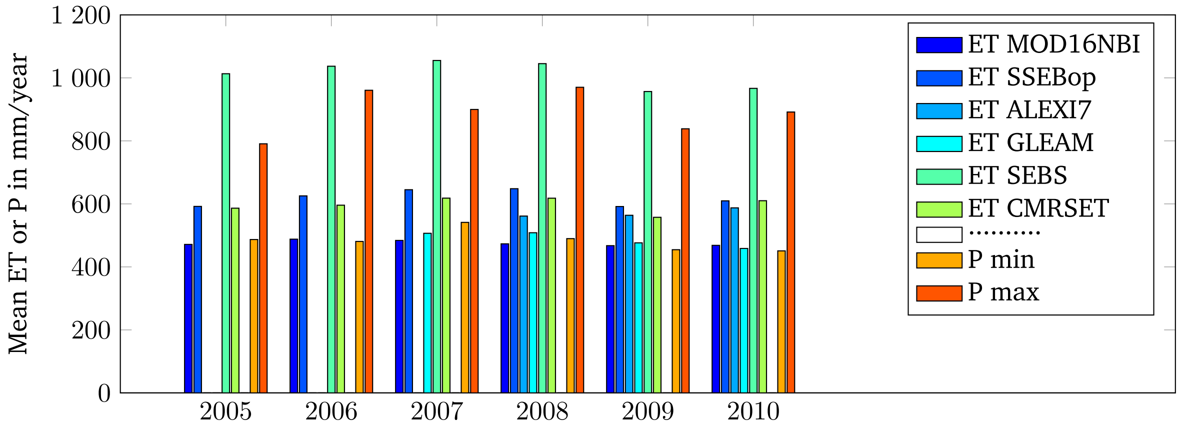

Figure 3: yearly ET and P for several products

Table 3: Mean annual ET and P values for the entire Nile basin in mm/year

	MOD16NBI	SSEBop	ALEXI	GLEAM	SEBS	CMRSET	P min	Pmax
2005	471	592	n/a	n/a	1013	586	487	791
2006	488	625	n/a	n/a	1037	596	481	961
2007	484	645	n/a	507	1055	618	541	900
2008	473	648	561	508	1045	618	490	970
2009	467	592	564	476	957	557	455	838
2010	468	610	587	458	967	610	451	892

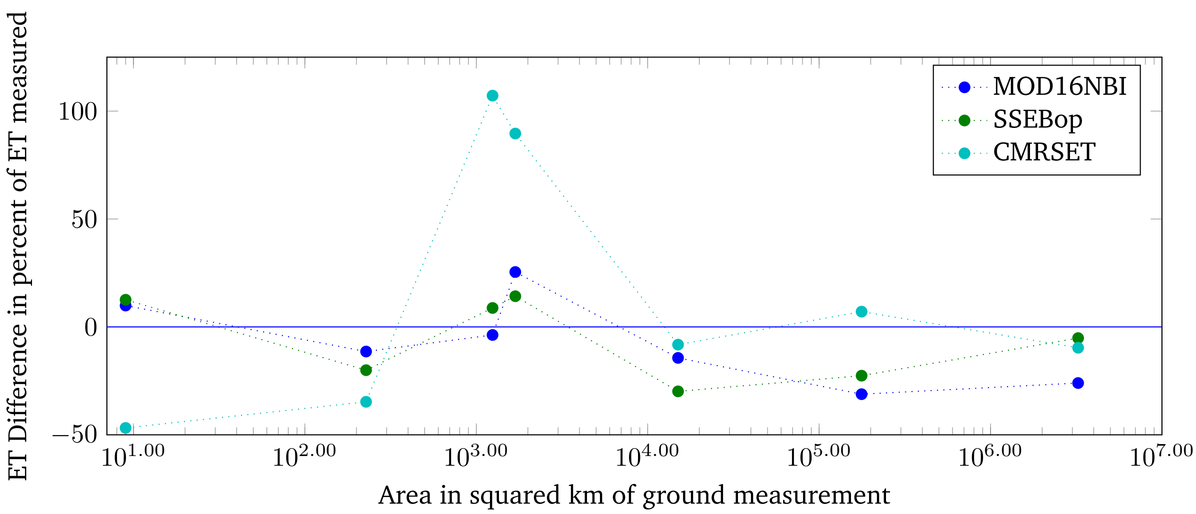

Figure 4: ET error as fraction of measured ET for different ground observations on different spatial scales. From left to right: flux tower Demokeya, Lake Garun basin, Gummera basin, Gelgel basin, Nile Delta, Ethiopian Blue Nile basin and entire Nile basin.

others to interpret these conditions and compute the impact on ET with the equations provided above. The dominant land use in a subbasin is an important factor on the performance, although mountains affect the amount of shortwave radiation observed. The ultimate goal is to fully utilize the benefits of the different paramterization methods and sensors and hence an ensemble ET product is examined.

4.2. Comparison among ET products

The difference between the monthly R-S ET products and the mean of all four products is shown in figure 6. The figure shows values for 2008 but other years show a similar pattern. Overall differences up to 50% of the total annual ET can be witnessed in certain areas. The most obvious differences are the high ET values of the SSEBop products in the flood plains of South Sudan and the high values of CMRSET in the Ethiopian highlands to the east. The MOD16NBI ET product show less extreme values for particular areas as does ALEXI. The ALEXI product is somewhat higher than the average of the four products for irrigated areas in Gezira and the Nile delta. The cause of the difference between the different R-S ET products is examined by comparing the ET difference maps with the land-use classes, Budyko parameter[5], the digital elevation map, the mean precipitation and the standard deviation between multiple R-S precipitation products. Monthly scatter plots for 2005-2010 for all possible combinations of the four products have been used together with monthly ET difference maps and the parameterization to get an understanding of the underlying reasons for the differences in ET estimates. Based on this comparison the correlation between the differences in ET estimation and land-use class is larger than between ET estimation and Budyko parameter, ET estimation and elevation, ET estimation and average precipitation or ET estimation and standard deviation between R-S precipitation products. The mean precipitation and standard deviation per pixel and per month of six R-S precipitation products is based on TRMM, ARC, CHIRPS, PERSIANN, CMORPH and GSMaP.

The products are also compared mutually using a pixel-to-pixel variability. The statistics of this inter-comparison are depicted in table 2. Monthly values for the available years have been compared with the other ET product. The correlation (r) between ALEXI7 and MOD16NBI of 0.87 is the highest, whereas the correlation (r) between CMRSET and SSEBop of 0.70 is the lowest. Looking at the root mean squared error (RMSE), the RMSE between MOD16NBI and ALEXI is the lowest with 25.12 [mm/month]. The largest RMSE is 35.46 mm/month between CRMSET and SSEBop. The interim conclusion is that ALEXI and MOD16NBI show more pixel scale similarities than for any other ET product combination. This practically implies that the spatial distribution of ET data within subbasins is comparable.

The normalized RMSE in this case is obtained by dividing the RMSE with the highest monthly mean of one of the two dataset. The NRMSE ranges between 0.07 [-] and 0.12 [-] for the same product pairs as mentioned before. Scatter plots of the inter-product comparison show however that often one product shows no ET (0 mm/month) while other products show a value significantly above 0 mm/month.

At the 0.027° resolution it is possible to compare the different products per land-use class. The monthly ET products can be accumulated to obtain yearly values. The box plots show the different annual values per product. For land-use class 1 Irrigated areas, MOD16NBI shows the lowest values and only a small inter-annual variation. Another relatively large difference is the annual ET of the SSEBop product versus the other three products for land-use class Open Savannah. This is mostly a result of a larger ET value provided by the SSEBop product during September and October. ET estimates for land-use class Pastures is almost the same for all four products and constant between multiple years. For land-use class Open Grassland, the CMRSET R-S ET product yields somewhat higher values. Looking at the annual difference maps in figure 5, the CMRSET exhibits two areas of higher ET estimates than the other products. Land-use class Open Grassland is primarily located Northwest of the Ethiopian Highlands. Another land-use class that shows significant differences both between multiple years and between different R-S ET products is Floodplains. The majority of the pixels of this land-use class is located in the Sudd wetlands in South Sudan. Flood plains physically act as open water surfaces during a significantly long period. A monthly changing land-use class dataset would have increased the reliability of this comparison study. The ALEXI product doesn't provide values for open water altogether. Hence, the average value for floodplains is provided by pixels during times when the floodplains are dry and thus an underestimation is expected. As can be seen from figure 7, not ALEXI but CMRSET shows the lowest multi-annual value for this land-use class. The difference for land-use class Build Up Areas is also quite large however this land-use class accounts for only 0.13% of the total amount of cells within the Nile Basin. This is not the case for land-use class Deserts and Bare Areas. A known caveat of the CMRSET product is the overestimation of dark rock outcrops. The Malha wells (15.217°N, 26.369°E) in Sudan is an example of such a dark rock outcrop were CMRSET ET estimates are approximately 400 [mm/year] above the mean of the four products.

5. Ensemble ET products

5.1. Different averaging procedures

Based on analysis of the parameterization, comparison with the ground observations and the inter-comparison,

[5]Net Radiation / Precipitation

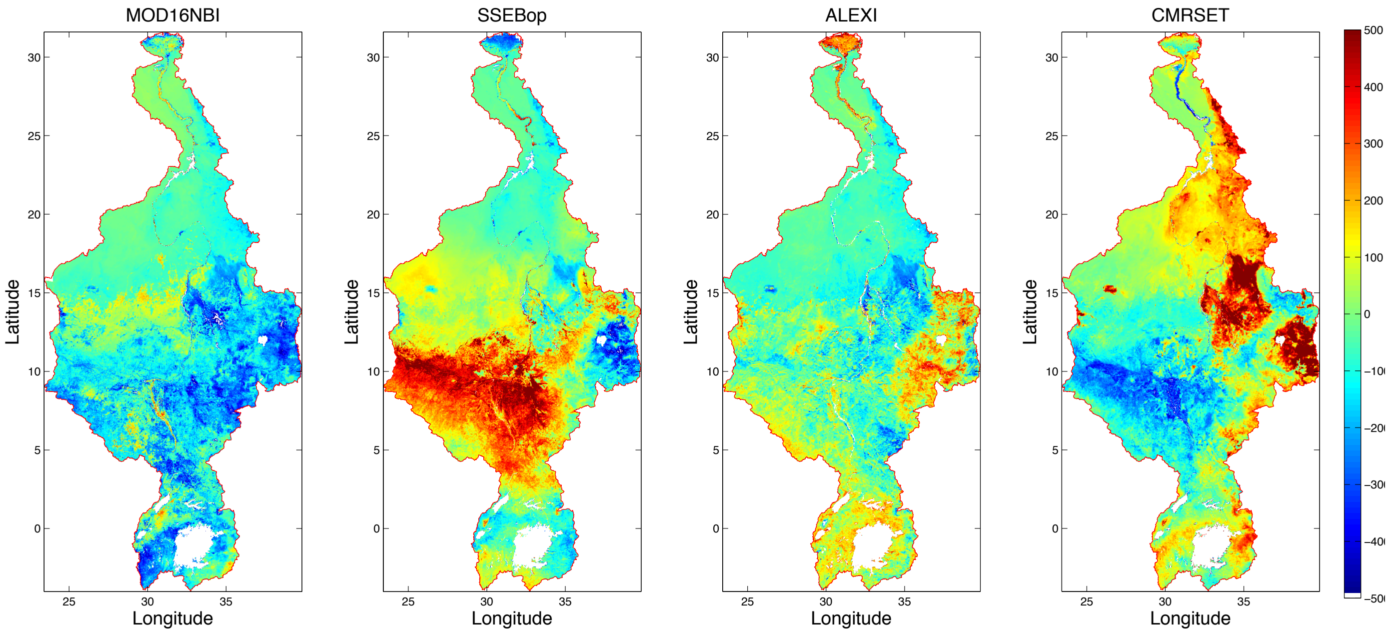

Figure 5: Difference between individual products and mean of all four products at 0.027 degree resolution for 2008, other years show similar patterns

Table 4: Statistical parameters for a multi-annual monthly comparison per pixel

		MOD16NBI	SSEBop	ALEXI	CMRSET
MOD16NBI	r	1.00			
	rmse	0.00			
	nrmse	0.00			
SSEBop	r	0.82	1.00		
	rmse	30.34	0.00		
	nrmse	0.08	0.00		
ALEXI	r	0.87	0.83	1.00	
	rmse	25.12	28.74	0.00	
	nrmse	0.07	0.11	0.00	
CMRSET	r	0.78	0.70	0.83	1.00
	rmse	27.97	35.46	26.54	0.00
	nrmse	0.08	0.12	0.09	0.00
Average	r	0.83	0.78	0.84	0.77
	rmse	27.81	31.51	26.80	29.99
	nrmse	0.08	0.11	0.09	0.10

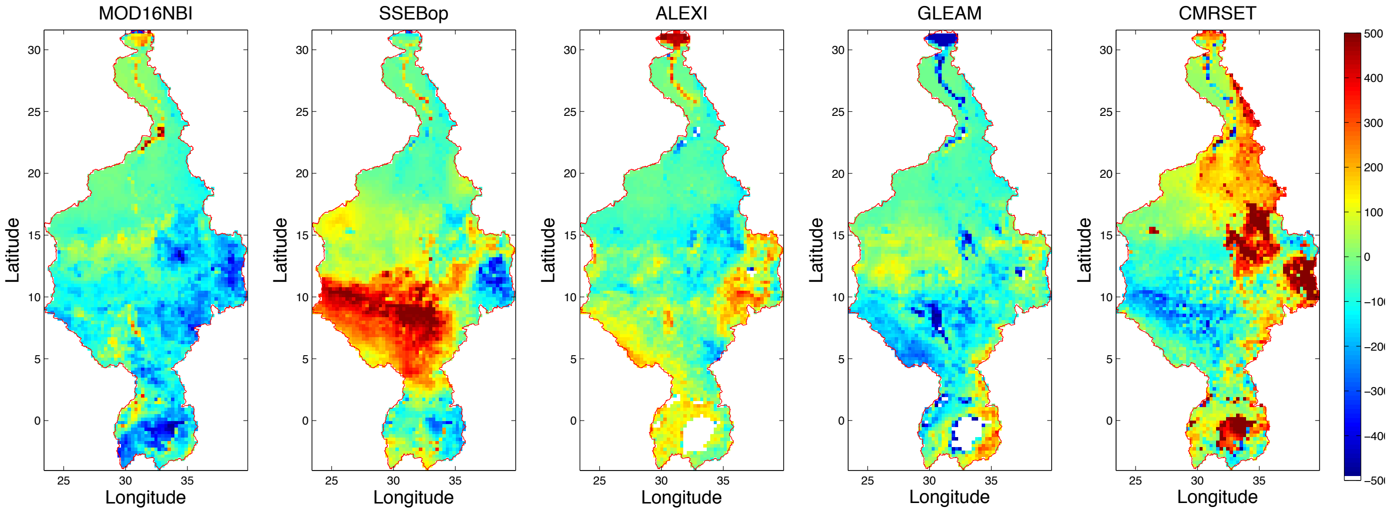

Figure 6: Difference between individual products and mean of all five products at 0.27 degree resolution for 2008, other years show similar patterns

each product has some pros and cons. The products face a trade-off between simplicity, accuracy on basin and long term scale and the urge to follow specific ET patterns on a small spatial and temporal scale. Since no standard parameterization exist and most models are calibrated and validated on global or continental scales, large differences occur when compared at basin and subbasin level. When ET products are use in other srtudies, for example to estimate discharge, the choice of ET product can have a huge effect on the conclusions. The proposed standard ensemble product includes the average ET value on a monthly interval with standard deviation between the different products and the minimum-maximum bandwidth. There are however multiple options to combine the available R-S ET products into a standardized new product. In this paper three possibilities to create an ensemble product are explored: Normal mean, expert knowledge and an ensemble product based on outlier elimination.

The first and most straight-forward method is to take the mean of all four products without any imposed restrictions. The downside of this approach is that known or unknown problems with some of the models identified (CMRSET: dark outcrops; ALEXI: water bodies; SSEBop: high values for savannah in september; MOD16NBI: irrigated and flooded areas) are overlooked and will be falsely included in the dataset. The advantage is that no arbitrary choices or biased views will affect the averaging procedure. The normal mean ensemble product is based on MOD16NBI, SSEBop, ALEXI7 and CMRSET on a monthly timescale and at 0.027° resolution.

Another coarser version with 0.27° resolution of the normal mean ensemble product includes the extra R-S ET product: GLEAM.

Based on the findings of the combination of ET products and land use classes, a preliminary expert knowledge table is created (table 5). The MOD16NBI R-S ET product is rejected in areas where water sources other than precipitation is the main water supply. Although the SSEBop R-S ET product is higher than the other products for land-use class Open Savannah, it is unclear whether this is incorrect. More ground observations for this land-use class are required, especially during September and October. As a result, all land-use classes of SSEBop are included. The ALEXI7 product doesn't provide values for land-use class Open Water. Furthermore land-use class Deserts is not included in the final product. The CMRSET R-S ET product is not used for land-use classes Irrigated Areas, Rainfed Crops and Floodplains.

Another straightforward method to improve the ensemble product is by using outlier analysis. The method uses the same 0.027° and 0.27° resolution monthly input datasets. For each cell a maximum coefficient of variation is used to eliminate the outliers. In this case the threshold value for CV is arbitrarily set to 0.5 [-]. Since only four datasets are used for the 0.027° resolution and five for 0.27° it is not always possible to determine outliers. The minimum number of R-S ET products used per cell is two. When more R-S ET products will be added to the ensemble products, the performance of this method is likely to increase. This research only shows this method as a proof of concept without optimization of the maximum coefficient of variation threshold. The reason why no optimiza-

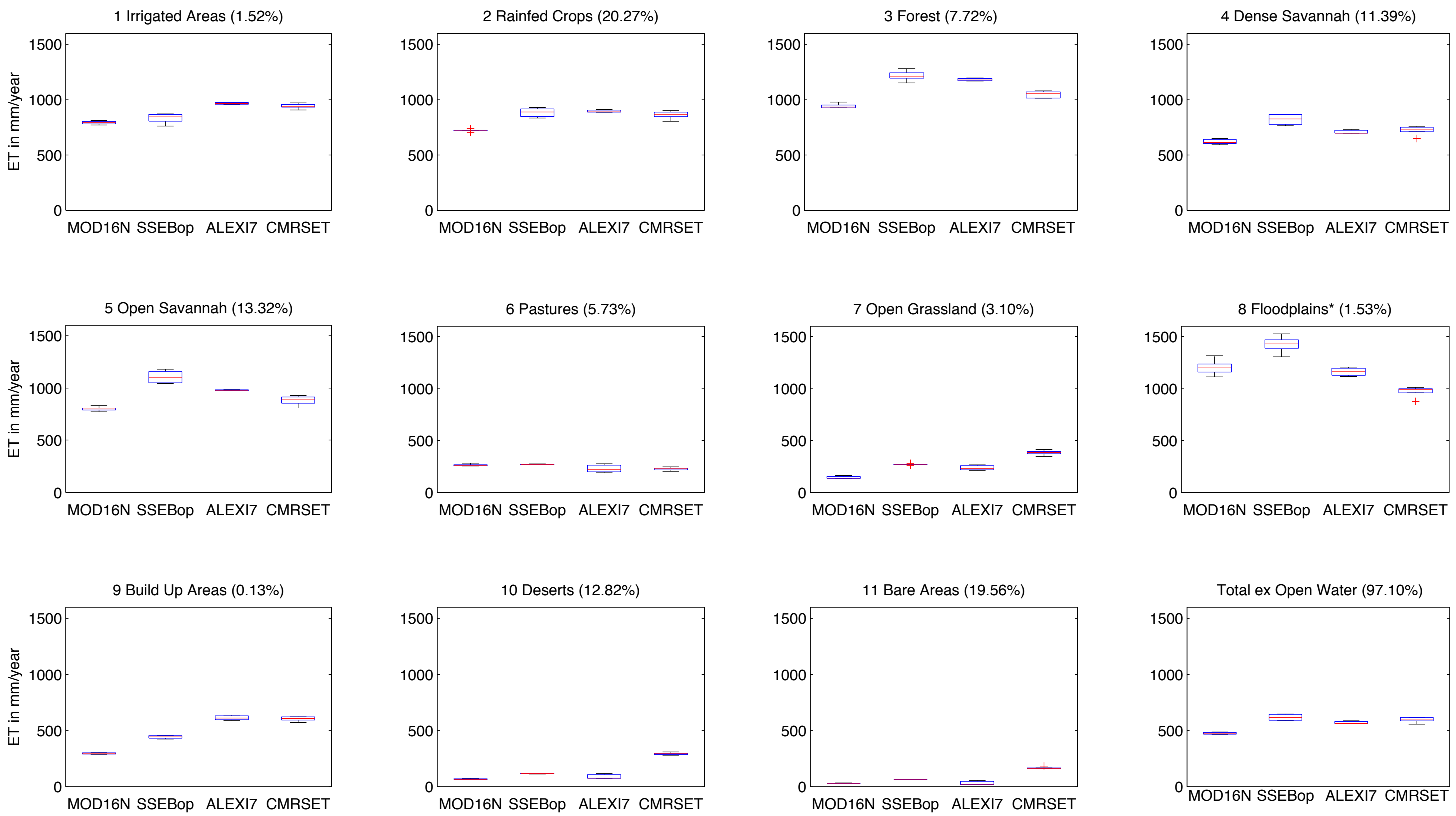

Figure 7: Difference between individual products and mean of all five products at 0.27 degree resolution for 2008, other years show similar patterns

Table 5: Land-use classes used per product to create the expert knowledge ensemble product

Landclass	MOD16NBI	SSEBop	ALEXI7	CMRSET
1 Irrigated Areas	0	1	0	0
2 Rainfed Crops	1	1	1	0
3 Forest	0	1	1	1
4 Dense Savannah	1	1	1	1
5 Open Savannah	1	1	1	1
6 Pastures	1	1	1	1
7 Open Grassland	1	1	1	1
8 Floodplains	0	1	1	0
9 Build Up Areas	1	1	1	1
10 Desert	1	1	0	1
11 Bare Areas	1	1	1	1
12 Water Bodies	1	1	0	1

tion of the threshold is done is twofold: first the number of available datasets for this method is too small and second the available data for validation may include errors up to 20%. The number of rejected R-S ET products per cell varies but a clear pattern is visible. In the dry areas in the north, more R-S ET products are rejected whereas in the humid south, often no products are rejected. The mean number of products on which a cell is based is 2.83.

The cumulative ET for the water-balance, the different R-S ET products and the three ensemble products is depicted in 9 and table 6. Overall the correspondence between the ensemble products and the water balance is much better than the correspondence between the individual R-S ET products and the water-balance. Only for basin 5 an individual product performs better than one of the ensembles. For a subbasin and basin level it is therefore better to use an ensemble product rather than an individual product. Figure 9 shows the difference between the ensemble products and the water balance for 2008.

5.2. Validation of ensemble ET products

Overall the performance of the normal mean ensemble is good compared to the other ensemble products i.e. the products based on expert knowledge (expert) and outlier analysis (CV). Four subbasins exhibit mean ensemble ET estimates lower than the FAO-Nile reference base, and seven exhibit mean ET estimates being higher. There is no systematic trend in differences for the upstream and downstream ends of the basin. The positive difference is the largest in subbasin 4 and 9. In subbasin 4 Bahr el Ghazal, the SSEBop ET estimates for land-use class 5 cause the overestimation whereas for subbasin 9, Atbarah the CMRSET ET product cause an overestimation. The overestimation of CMRSET in the eastern part of the Nile basin cannot be attributed to a single land-use class because it occurs in deserts and open grasslands. The negative difference between the normal mean ensemble and the FAO-ET is most distinct in subbasins 3,Bhar el Ghazal - Sudd, 8 Blue Nile and 10, Main Nile downstream of Khartoum. Only the SSEBop product provides an ET estimate higher that the FAO-ET. Note that the FAO-ET data shows an under-estimation of 21.9% in subbasin 3. In subbasin 8, MOD16NBI provides ET estimates being substantially lower than FAO-ET. There is no single land-use class accountable for causing this lower estimates although some irrigated areas lie within subbasin 8. Subbasin 10 is the smallest subbasin causing larger relative errors.

The ensemble product based on expert knowledge performs better than any individual product for most of the basins as can be seen from figure 8 8 and table 6. The performance of the expert knowledge and outlier analysis ensemble are worse than for the mean ensemble. Compared with the mean ensemble product, the expert knowledge ensemble product performs better in only four subbasins. The absolute error increases in seven subbasins notably in subbasin four and five. This is due to the elimination of MOD16NBI forests and to a lesser extent the irrigated areas. Except for subbasin six, the sign of the error is the same as for the normal mean ensemble. Furthermore no systematic upstream, downstream error occurs

The ensemble product based on a threshold value of 0.5 for the coefficient of variation performs better than a standard mean in 6 out of 11 subbasins. In subbasin 6, 7, 9 and 11 the performance is much better compared to the normal mean and expert knowledge ensemble products. In subbasin 2, 3, 5, 8 and 10 however the performance of this approach is the worst. More products are rejected in the downstream subbasins. On average, less than 0.5 product is rejected in all upstream subbasins [6]. Overall performance increases more downstream with subbasin 10 being the big exception.

6. Conclusions and way forward

Several institutions and universities are currently developing global near-operational R-S ET products. These products have different levels of complexity and utilize different spectral radiances, although often data from the

[6] basinnumber less than 9

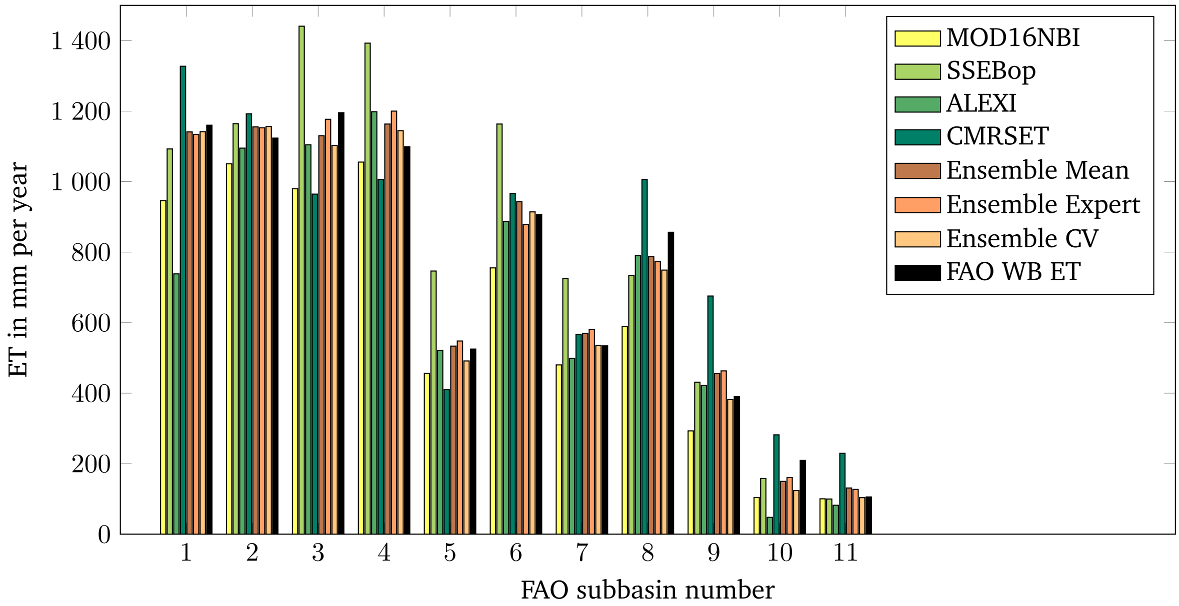

Figure 8: Total ET in mm/year for individual R-S ET products and ensemble products compared to FAO water-balances

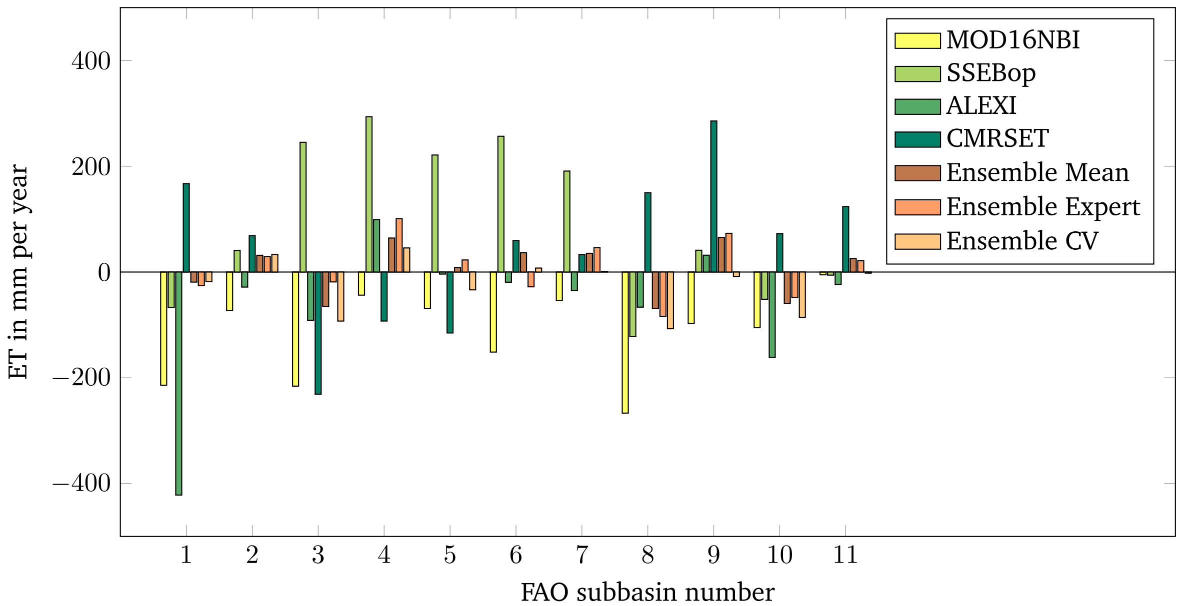

Figure 9: ET Difference in mm/year for individual R-S ET products and ensemble products compared to FAO water-balances

Table 6: Comparison of multiple ET ensemble products with the NBI subbasin water-balance and the FAO-subbasin. Not that these subbasins differ and are not the same as in figure 8

No.	Subbasin	NBI-Nile [mm/y]	Area [km^2]	FAO-Nile [mm/y]	Area [km^2]	Normal mean ensemble [mm/y]	Expert ensemble [mm/y]	CV ensemble [mm/y]
1	Main Nile d/s Atbara	107	877866	124	983375	131	127	104
2	Atbara	453	237044	397	231492	456	463	382
3	Main Nile d/s Khartoum	180	34523	211	35338	150	161	124
4	Blue Nile	737	308198	863	307262	787	773	749
5	White Nile	617	260943	554	237429	570	580	535
6	Bahr el Ghazal & el Arab	13	606428	749	717069	848	874	818
7	Pibor-Akabo-Sobat	1012	246779	907	230369	943	879	914
8	Bahr el Jebel	1144	136400	1196	80433	1130	1177	1103
9	Kyoga-Albert	1047	197253	1124	156839	1155	1153	1157
10	Lake Victoria basin	1018	264985	1160	249433	1141	1134	1142
	Total basin	616	3170419	628	3229039	619	620	590

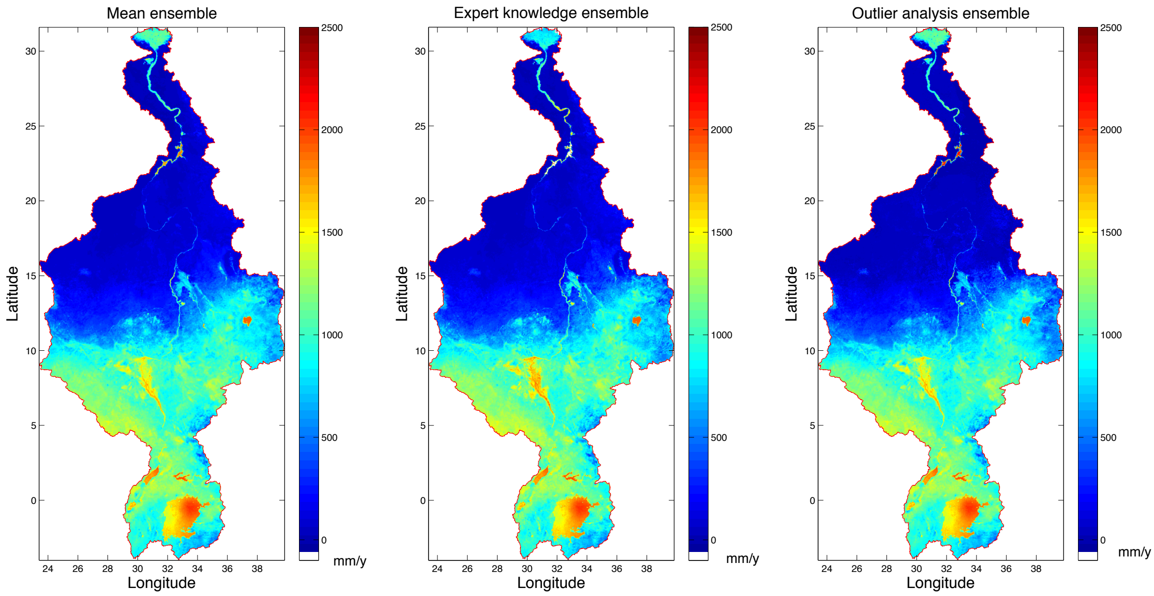

Figure 10: Annual ET values in mm/year at 0.027 degree resolution for the three ensemble products

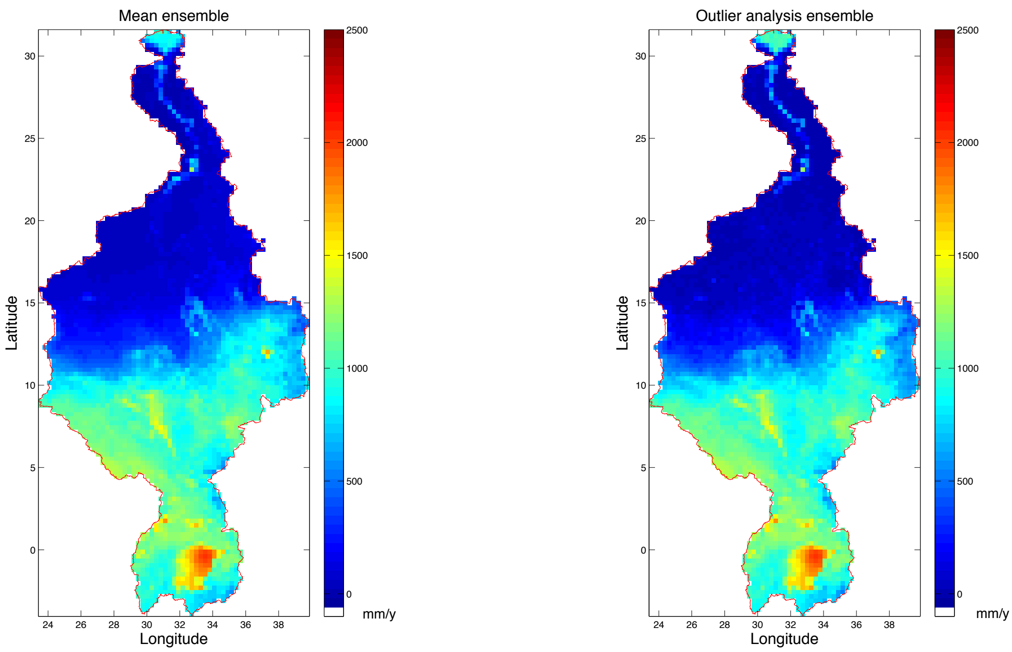

Figure 11: Annual ET values in mm/year at 0.27 degree resolution for the mean and outlier analysis ensemble products

MODIS radiometer is employed. This paper is one of the first attempts to compare the performance of more than 5 models that are entirely based on remote sensing techniques, thus uncertainty from soil information and simulation of water balances is excluded. The ET dataset investigated is unique and it is the first time that these prototype products are compared, although comparisons between ALEXI vs. MOD16, LSA SAF vs. MOD16 and CMRSET vs. GLEAM have been executed before. An independent remote sensing based ET dataset is suitable for calibration of hydrological models (King et al., 2011) (Gao and Hrachowitz, 2014) and examine consumptive use in water scarce river basins (van Eekelen et al., 2015).

Despite that the quality of the ground observations is not ideal, it can be concluded that SEBS systematically overestimates ET fluxes, most probably to the classical problems of coupling an instantaneous surface radiation temperature with an independently obtained instantaneous air temperature (Hall and Huemmrich, 1992). The original MOD16 (not the NBI version) and LSA SAF systematically underestimates the ET fluxes, which can be explained by the poor parameterization to detect moist land surface in both models (e.g. irrigated land, wetlands, inundation areas). The bulk surface conductance in MOD16 does not describe wet surfaces and LSA SAF does not have an adequate solution for irrigation and inundations. The ET results of GLEAM also tend to be at the lower side, but they can be accepted for the majority of the land surfaces. Most models perform well for pastures and rainfed crops.

ALEXI, SSEBop, CMRSET and MOD16NBI appear to have realistic ET values at larger aggregated areas, although certain subbasin scale results were disappointing. ALEXI and MOD16NBI have a high correlation (r=0.87) and low RMSE (25.12 [mm/month]) mutually, which implies that their pixel values and spatial structure have the largest similarities of all ET products. Since ALEXI revealed also a good agreement with the flux towers (except for water, for which it doesn't provide values), it is the preferred model. This can most probably be explained by the balanced model concept between a thorough theoretical background taken from TSEB and the practical solution by measuring thermal lapse rates and morning warming processes at pixel basis. This does not mean that ALEXI is also a single preferred model, because it showed lower values for pastures and grassland relative to notably the SSEBop product. Whether these higher values of SSEBop are correct is an impetus for further resarch.

Since every model has its advantages and disadvantages, it is preferred to develop a remotely sensed ensemble ET product for the global scale. Preferably this should be worked out for 1 km or lower ET products since this allows applications in agriculture. While the analysis suggest that this can be best done by means of developing a key between robust ET products and land use classes as in 5. The results of a mean ensemble and outlier analysis ensemble currently outweigh the expert judgements however future inprovements in both outlier analysis and expert knowledge determination can alter this. Especially variation in the number and seasonal variation of land-use classes should be examined. The ensemble mean products have a lower deviation with the FAO-ET values than the individual ET products, and this suggest that ensemble averaging is the way forward. The performance of the three approaches towards an ensemble product is remarkably close. Although the ensemble product based on outlier analysis significantly exhibits the best performance in six subbasins, it is the worst in the other five subbasins. Using the mean ensemble product is relatively safe: it is the intermediate performing approach in 7 out of 11 subbasins. The results of the ensemble product based on outlier elimination look promising. The number of available products however is a big limitation. This implies that it is worth undertaking more efforts in investigating the best way to average the ET products more on the basis of hydro-ecosystems. Experiences from ensemble rainfall products such as CHIRPS should be carefully studied and eventually applied to develop a superior ET ensemble product.

Allen, R., Tasumi, M., and Morse, A. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)Applications. *Journal of Irrigation ...*, (August):395–406.

Allen, R. G., Perreira, L., Raes, D., and Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *FAO Irrigation and drainage paper*, 300(56).

Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., González-Dugo, M. P., Cammalleri, C., D'Urso, G., Pimstein, A., and Gao, F. (2011). Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrology and Earth System Sciences*, 15(1):223–239.

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Otkin, J. a., and Kustas, W. P. (2007). A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *Journal of Geophysical Research*, 112(D10):D10117.

Ardö, J., Mölder, M., El-Tahir, B. A., and Elkhidir, H. A. M. (2008). Seasonal variation of carbon fluxes in a sparse savanna in semi arid Sudan. *Carbon balance and management*, 3(1):7.

Bastiaanssen, W., Hoekman, D., and Roebeling, R. (1994). A methodology for the assessment of surface resistance and soil water storage variability at mesoscale based on remote sensing measurements. *IAHS press*, 02:63.

Bastiaanssen, W., Karimi, P., Rebelo, L.-M., Duan, Z., Senay, G., Muttuwatte, L., and Smakhtin, V. (2014). Earth Observation Based Assessment of the Water Production and Water Consumption of Nile Basin Agro-Ecosystems. *Remote Sensing*, 6(11):10306–10334.

Brutsaert, W. and Sugita, M. (1992). Application of self-preservation in the diurnal evolution of the surface energy budget to determine daily evaporation. *Journal of Geophysical Research*, 97(D17):18377.

Cammalleri, C., Anderson, M., Gao, F., Hain, C., and Kustas, W. (2014). Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion. *Agricultural and Forest Meteorology*, 186:1–11.

Cheema, M. J. M., Immerzeel, W. W., and Bastiaanssen, W. G. M. (2014). Spatial quantification of groundwater abstraction in the Irrigated Indus Basin. *Groundwater*, 52(1):25–36.

Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W. (2007). Regional evaporation estimates from flux tower and MODIS satellite data. *Remote Sensing of Environment*, 106(3):285–304.

Duan, Z. (2014). *Estimating Water Balance Components of Lakes and Reservoirs Using various Open Access Satellite Databases*. PhD thesis.

D'Urso, G., Querner, E., and Morabito, J. (1992). Integration of hydrological simulation models with remotely sensed data: an application to irrigation management. *Advances in Planning, Design and Management of Irrigation Systems as Related to Sustainable Land Use*, 2:463–472.

El Quosy, D. and El Guindy, S. (1989). Water and Salt Balance of the Nile Delta, A new Approach. *Land drainage in Egypt*, pages 243–253.

Elsawwaf, M. and Willems, P. (2012). Analysis of the climate variability on Lake Nasser evaporation based on the Bowen ratio energy budget method. *Journal of environmental biology / Academy of Environmental Biology, India*, 33(2 Suppl):475–85.

ESA (2008). GlobCover Land Cover v2 2008 database. European Space Agency, European Space Agency GlobCover Project, led by MEDIAS-France.

FAO (2011). Information Products for Nile Basin Water Resources Management Synthesis Report. Technical report, FAO.

Farah, H., Bastiaanssen, W., and Feddes, R. (2004). Evaluation of the temporal variability of the evaporative fraction in a tropical watershed. *International Journal of Applied Earth Observation and Geoinformation*, 5(2):129–140.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Rowland, J., Romero, B., Husak, G., Michaelsen, J., and Verdin, P. (2014). A Quasi-Global Precipitation Time Series for Drought Monitoring Data Series 832, 4p. Technical report.

Gao, H. and Hrachowitz, M. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical . . .* , pages 1–8.

Ghilain, N., Arboleda, A., and Gellens-Meulenberghs, F. (2011). Evapotranspiration modelling at large scale using near-real time MSG SEVIRI derived data. *Hydrology and Earth System Sciences*, 15(3):771–786.

Guerschman, J. P., Van Dijk, A. I., Mattersdorf, G., Beringer, J., Hutley, L. B., Leuning, R., Pipunic, R. C., and Sherman, B. S. (2009). Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia. *Journal of Hydrology*, 369(1-2):107–119.

Hain, C. R., Crow, W. T., Mecikalski, J. R., Anderson, M. C., and Holmes, T. (2011). An intercomparison of available soil moisture estimates from thermal infrared and passive microwave remote sensing and land surface modeling. *Journal of Geophysical Research*, 116(D15):D15107.

Hall, F. and Huemmrich, K. (1992). Satellite remote sensing of surface energy balance: Success, failures, and unresolved issues in FIFE. *Journal of . . .* , 97.

Hessels, T. D. U. o. T. (2015). Working title: Remote Sensing Precipitation products in the Nile Basin.

Hoekstra, A. Y. and Chapagain, A. K. (2007). The water footprints of Morocco and the Netherlands: Global water use as a result of domestic consumption of agricultural commodities. *Ecological Economics*, 64(1):143–151.

Hrachowitz, M., Savenije, H., Blöschl, G., McDonnell, J., Sivapalan, M., Pomeroy, J., Arheimer, B., Blume, T., Clark, M., Ehret, U., Fenicia, F., Freer, J., Gelfan, A., Gupta, H., Hughes, D., Hut, R., Montanari, A., Pande, S., Tetzlaff, D., Troch, P., Uhlenbrook, S., Wagener, T., Winsemius, H., Woods, R., Zehe, E., and Cudennec, C. (2013). A decade of Predictions in Ungauged Basins (PUB)a review. *Hydrological Sciences Journal*, 58(6):1198–1255.

Hu, G., Jia, L., and Menenti, M. (2014). Comparison of MOD16 and LSA-SAF MSG evapotranspiration products over Europe for 2011. *Remote Sensing of Environment*, 16.

Huete, a., Didan, K., Miura, T., Rodriguez, E., Gao, X., and Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1-2):195–213.

Jackson, R. (1977). Wheat canopy temperature: a practical tool for evaluating water requirements. *Water Resources Research*, 13(3).

Jia, L., Su, Z., van den Hurk, B., Menenti, M., Moene, A., a.R De Bruin, H., Yrisarry, J. B., Ibanez, M., and Cuesta, A. (2003). Estimation of sensible heat flux using the Surface Energy Balance System (SEBS) and ATSR measurements. *Physics and Chemistry of the Earth, Parts A/B/C*, 28(1-3):75–88.

Karimi, P. and Bastiaanssen, W. (2014). Spatial evapotranspiration, rainfall and land use data in water accounting Part 1: Review of the accuracy of the remote sensing data P. *Hydrology and Earth System Sciences Discussions*, 11(1):1–29.

Karimi, P., Bastiaanssen, W. G. M., and Molden, D. (2013). Water Accounting Plus (WA+) a water accounting procedure for complex river basins based on satellite measurements. *Hydrology and Earth System Sciences*, 17(7):2459–2472.

King, E. A., Niel, T. G. V., Dijk, A. I. J. M. V., Wang, Z., Paget, M. J., Raupach, T., Guerschman, J., Haverd, V., Mcvicar, T. R., Miltenburg, I., Raupach, M. R., and Zhang, Y. (2011). Actual Evapotranspiration Estimates for Australia Inter-comparison and Evaluation. Technical Report April.

Kustas, W. and Norman, J. (1997). A twosource approach for estimating turbulent fluxes using multiple angle thermal infrared observations. *Water Resources Research*, 33(6):1495–1508.

Menenti, M. and Coudhury, B. (1993). Parameterization of land surface evaporation by means of location dependent potential evaporation and surface temperature range. *proceedings of an international symposium held at Yokohama, Japan*, 1:561–568.

Menenti, M., Visser, T., Morabito, J., and Drovandi, A. (1989). Appraisal of irrigation performance with satellite data and geo-referenced information. *Irrigation Theory and Practice*, 1:785–801.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. a. M., Gash, J. H., Meesters, a. G. C. a., and Dolman, a. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, 15(2):453–469.

Mohamed, Y. and Savenije, H. H. G. (2014). Impact of climate variability on the hydrology of the Sudd wetland: signals derived from long term (19002000) water balance computations. *Wetlands Ecology and Management*, 22(2):191–198.

Mu, Q., Jones, L. a., Kimball, J. S., McDonald, K. C., and Running, S. W. (2009). Satellite assessment of land surface evapotranspiration for the pan-Arctic domain. *Water Resources Research*, 45(9):n/a–n/a.

Mu, Q., Zhao, M., and Running, S. W. (2013). MODIS Global Terrestrial Evapotranspiration (ET) Product (NASA MOD16A2/A3) Algorithm Theoretical Basis Document Collection 5. Technical report, Nasa.

Mu, Q. U. o. M. (2013). MODIS 1-km 2 Terrestrial Evapotranspiration (ET) Product for the Nile Basin Algorithm Theoretical Basis Document. Technical report.

Mu, Q. U. o. M., Zhao, M. U. o. M., and Running, S. W. U. o. M. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, 115(8):1781–1800.

Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. a., Dolman, a. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan, F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y., and Seneviratne, S. I. (2013). Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis. *Hydrology and Earth System Sciences*, 17(10):3707–3720.

Nemani, R. R. and Running, S. W. (1989). Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. *Journal of Applied meteorology*, 28(4):276–284.

Parinussa, R. M., Yilmaz, M. T., Anderson, M. C., Hain, C. R., and de Jeu, R. a. M. (2014). An intercomparison of remotely sensed soil moisture products at various spatial scales over the Iberian Peninsula. *Hydrological Processes*, 28(18):4865–4876.

Peel, M. (2007). Updated world map of the Köppen-Geiger climate classification. *. . . and earth system . . .* , 11(5):1633–1644.

Perry, C. (2007). Efficient irrigation; inefficient communication; flawed recommendations. *Irrigation and Drainage*, 56(4):367–378.

Ramadan (1989). Lake Qarun. *unknown*.

Savoca, M., Senay, G., and Maupin, M. (2013). Actual Evapotranspiration Modeling Using the Operational Simplified Surface Energy Balance (SSEBop) Approach Scientific. Technical report.

Schuol, J., Abbaspour, K. C., Yang, H., Srinivasan, R., and Zehnder, A. J. B. (2008). Modeling blue and green water availability in Africa. *Water Resources Research*, 44(7):n/a–n/a.

Schuurmans, J., Troch, P., Veldhuizen, A., Bastiaanssen, W., and Bierkens, M. (2003). Assimilation of remotely sensed latent heat flux in a distributed hydrological model. *Advances in Water Resources*, 26(2):151–159.

SEGUIN, B. and ITIER, B. (1983). Using midday surface temperature to estimate daily evaporation from satellite thermal IR data. *International Journal of Remote Sensing*, 4(2):371–383.

Senay, G., Budde, M., and Verdin, J. (2011). Enhancing the Simplified Surface Energy Balance (SSEB) approach for estimating landscape ET: Validation with the METRIC model. *Agricultural Water Management*, 98(4):606–618.

Senay, G. B., Bohms, S., Singh, R. K., Gowda, P. H., Velpuri, N. M., Alemu, H., and 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.

Senay, G. B., Budde, M., Verdin, J. P., and Melesse, A. M. (2007). A Coupled Remote Sensing and Simplified Surface Energy Balance Approach to Estimate Actual Evapotranspiration from Irrigated Fields. *Sensors*, 7(6):979–1000.

Serrat-Capdevila, A., Valdes, J. B., and Stakhiv, E. Z. (2014). Water Management Applications for Satellite Precipitation Products: Synthesis and Recommendations. *JAWRA Journal of the American Water Resources Association*, 50(2):509–525.

Sinnhuber, M., Funke, B., von Clarmann, T., Lopez-Puertas, M., and Stiller, G. P. (2014). Development of a 10year (20012010) 0.1 dataset of land-surface energy balance for mainland China. *Atmospheric Chemistry and Physics Discussions*, 14(1):1–29.

Soer, G. (1980). Estimation of regional evapotranspiration and soil moisture conditions using remotely sensed crop surface temperatures. *Remote Sensing of Environment*, 9(1):27–45.

Su, Z. (2002). The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences*, 6(1):85–100.

Sutcliffe, J. and Parks, Y. (1999). *The Hydrology of the Nile*, volume 5. IAHS, 5 edition.

Tekleab, S., Uhlenbrook, S., Mohamed, Y., Savenije, H. H. G., Temesgen, M., and Wenninger, J. (2011). Water balance modeling of Upper Blue Nile catchments using a top-down approach. *Hydrology and Earth System Sciences*, 15(7):2179–2193.

Trambauer, P., Dutra, E., Maskey, S., Werner, M., Pappenberger, F., van Beek, L. P. H., and Uhlenbrook, S. (2013). *Comparison of different evaporation estimates over the African continent*, volume 10.

van den Hurk, B. J. J. M., Viterbo, P., Beljaars, A. C. M., and Betts, A. K. (2000). *Offline validation of the ERA40 surface scheme*. European Centre for Medium-Range Weather Forecasts.

van Dijk, a. I. J. M., Renzullo, L. J., Wada, Y., and Tregoning, P. (2014). A global water cycle reanalysis (20032012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble. *Hydrology and Earth System Sciences*, 18(8):2955–2973.

van Eekelen, M., Bastiaanssen, W., Jarmain, C., Jackson, B., Ferreira, F., van der Zaag, P., Saraiva Okello, a., Bosch, J., Dye, P., Bastidas-Obando, E., Dost, R., and Luxemburg, W. (2015). A novel approach to estimate direct and indirect water withdrawals from satellite measurements: A case study from the Incomati basin. *Agriculture, Ecosystems & Environment*, 200:126–142.

van Griensven, a., Ndomba, P., Yalew, S., and Kilonzo, F. (2012). Critical review of SWAT applications in the upper Nile basin countries. *Hydrology and Earth System Sciences*, 16(9):3371–3381.

Velpuri, N., Senay, G., Singh, R., Bohms, S., and Verdin, J. (2013). A comprehensive evaluation of two MODIS evapotranspiration products over the conterminous United States: Using point and gridded FLUXNET and water balance ET. *Remote Sensing of Environment*, 139:35–49.

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., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., and Verma, S. (2002). Energy balance closure at FLUXNET sites. *Agricultural and Forest Meteorology*, 113(1-4):223–243.

Winsemius, H. C., Savenije, H. H. G., and Bastiaanssen, W. G. M. (2008). Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins? *Hydrology and Earth System Sciences Discussions*, 5(4):2293–2318.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P. (2011). Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resources Research*, 47(5):W05301.

Yilmaz, M. T., Anderson, M. C., Zaitchik, B., Hain, C. R., Crow, W. T., Ozdogan, M., Chun, J. A., and Evans, J. (2014). Comparison of prognostic and diagnostic surface flux modeling approaches over the Nile River basin. *Water Resources Research*, 50(1):386–408.