Journal of Hydrology: Regional Studies 30 (2020) 100709

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Validation of seven global remotely sensed ET products across Thailand using water balance measurements and land use classifications

Nutchanart Sriwongsitanon[a,b,*], Thienchart Suwawong[a,b], Sansarith Thianpopirug[a,b], James Williams[c], Li Jia[d], Wim Bastiaanssen[e]

[a] Department of Water Resources Engineering, Faculty of Engineering, Kasetsart University, Bangkok, 10900, Thailand
[b] Remote Sensing Research Centre for Water Resources Management (SensWat), Faculty of Engineering, Kasetsart University, Bangkok, Thailand
[c] Civil and Environmental Engineering, The University of New South Wales, Sydney, Australia
[d] Institute of Remote Sensing and Digital Earth (RADI), Beijing, China
[e] IHE-Delft, Institute for Water Education, Westvest 7, 2611 AX, Delft, the Netherlands

ARTICLE INFO

Keywords:
Evapotranspiration
MOD16A2
Terrestrial water storage change
GRACE
Water balance
Land use
Thailand

ABSTRACT

Study region: Annual and monthly ET values from seven global remote sensing products (ALEXI, CMRSET, ETMonitor, GLEAM V3.3b, MOD16A2, SEBS V3 and SSEBop) were validated for 172 sub-basins in Thailand.

Study focus: This study describes a generalised validation procedure that uses rainfall (P), streamflow (Q) and storage change data (from the Gravity Recovery and Climate Experiment - $TWSC_{GRACE}$) and land use information. For each sub-basin, bulk ET was computed using the water balance framework and compared to estimates by ET products. Inverse water balance computations were applied to infer the storage change estimates from each product ($\Delta S = P - Q - ET_{RS}$), which were compared to $TWSC_{GRACE}$ to assess their monthly scale performances.

New hydrological insights for the region under study: All products performed very well on the annual basis (mean NSE > 0.96) and satisfactorily on the monthly scale (mean NSE > 0.65). Land use classifications from the Land Development Department were used to examine the ability of four candidates (CMRSET, MOD16A2, GLEAM V3.3b and ETMonitor) to provide ET estimates with correspondence to physical land use conditions. By also considering product resolutions and data accessibility, MOD16A2 was consensually shown to be the most promising product to be used for water resources management in Thailand. In addition to local applications, the outcomes emanate the potential for utilisation on the global scale which should be further investigated.

1. Introduction

Evapotranspiration (ET) is a major component of the hydrological cycle which generates several products and services that support society in terms of agronomy, economy, environment, micro-climates, industry and leisure (Bastiaanssen and Chandrapala, 2003). Globally, ET generally accounts for more than half of the total annual precipitation. Detailed information on ET is required to

* Corresponding author at: Department of Water Resources Engineering, Faculty of Engineering, Kasetsart University, 50 Paholyothin Rd., Ladyao, Jatujak, Bangkok, 10900, Thailand.
E-mail addresses: fengnns@ku.ac.th (N. Sriwongsitanon), thienchart.su@ku.th (T. Suwawong), sansarith@hotmail.com (S. Thianpopirug), james.kus41@gmail.com (J. Williams), liijia@yahoo.com (L. Jia), wim.bastiaanssen@gmail.com (W. Bastiaanssen).

https://doi.org/10.1016/j.ejrh.2020.100709
Received 18 December 2019; Received in revised form 22 June 2020; Accepted 22 June 2020

express these benefits in relation to consumptive use. ET data can be used for water accounting procedures (Karimi et al., 2013; Sriwongsitanon, 2018), to evaluate irrigation performance (Bastiaanssen et al., 1996), streamflow (Shilpakar et al., 2011), water productivity (Zwart et al., 2010) and withdrawals (van Eekelen et al., 2015), among others. ET is affected by land use, soil moisture availability and climate condition, which makes it highly variable across heterogeneous landscapes. Factors including leaf area index, root depths, albedo, surface roughness and transfer of energy and momentum all contribute to the variation in ET occurrence across different bioregions (Bonan et al., 1992). Dias et al. (2015) simulated mean annual ET using the Integrated Model of Land Surface Processes (INLAND) for small catchments in Brazil and found that mean annual ET was 39% lower in agricultural ecosystems than in natural ecosystems. Similarly, Costa and Foley (1997) demonstrated that forest clearing decreases annual ET by 12%.

To account for the lack of ET data from meteorological observations, numerical simulation models are often used to produce gridded ET maps. These are produced by either spatially distributed models based on raster grids (e.g. PCR-Globwb: Sutanudjaja et al., 2018 ; HBV: Winsemius et al., 2008) or semi-distributed models at which catchments are separated into units with similar hydrological responses (e.g. SWAT model: Srinivasan et al., 1998; SHE model: Abbott et al., 1986; Mike Basin; FlexTopo: Savenije, 2010). Several of these models have been applied for hydrological studies in Thailand (Akter and Babel, 2012; Yasin and Clemente, 2014). The general drawback of hydrological models is the specification of soil and vegetation parameters as well as human impacts that cannot be accurately simulated without intensive auxiliary local data.

An alternative methodology to assess the spatial variation of ET is the inclusion of earth observation data. While hundreds of ET algorithms have been developed since the 1970s, few are more generally applied. Some reviews of determining ET from earth observation data can be found in Courault et al. (2005); Allen et al. (2007); Kalma et al. (2008); Wang and Dickinson (2012) and Karimi and Bastiaanssen (2015). Nowadays, global-scale ET products are becoming increasingly available further to global precipitation products. Recent studies have demonstrated the assimilation of potential evapotranspiration to improve the representation of soil moisture, actual evapotranspiration, and streamflow (Rajib et al., 2018). With ongoing research, there are prospects for these new-era products to be beneficial for a wide range of hydrological applications. Typical ET products whose ET layers are open to access include MOD16A2 (Running et al., 2019), SEBS V3 (Su, 2002); and GLEAM V3.3b (Martens et al., 2017). ET products that are not yet facilitated by operational websites include CMRSET (Guerschman et al., 2009); SSEBop (Senay et al., 2013); ALEXI (Anderson et al., 1997) and ETMonitor (Hu and Jia, 2015). Comparison studies between various global scale ET products have been published earlier (e.g. McCabe et al., 2015; Michel et al., 2016; Wang-Erlandsson et al., 2016; Poortinga et al., 2017; Ha et al., 2017; da Motta Paca et al., 2019). Their general conclusions were that no ET products were deemed the most appropriate for heterogeneous landscapes and catchments. However, as noted by Lu et al. (2009), variations amongst ET products can lead to substantially differing estimations in annual runoff, water storage and irrigation requirements, necessitating the precautions required to select the most appropriate ET product to be adopted for water resources management.

A key shortcoming of ET products is their validation against flux tower data. These datasets are not very accurate, such that errors up to 20% are acceptable (Twine et al., 2000; Ramoelo et al., 2014). Advances in eddy covariance technology have led to more reliable results, but for older systems, caution is needed in using flux towers as ground truth. Instead, given the availability of rainfall and streamflow records in Thailand, we were able to devise a methodology to validate global ET products against the water balance, which requires sub-basin-scale rainfall, streamflow, and terrestrial water storage change data. This method has been adopted before in several studies such as Simons et al. (2016), who solved the water balance to evaluate ET at Son Tay gauging station in Vietnam (catchment area ~144,000 km2) which was compared to the estimates from MOD16A2, SSEBop, SEBS, ALEXI and CMRSET. Key limitations of this study pertain to the use of a single runoff station and the assumption that storage change over a ten-year period was negligible. Nevertheless, bulk ET data cannot sufficiently determine water consumption at specific locations within catchments, nor to describe the influence of land use on the ET regime. This demonstrates importance of capturing seasonal dynamics of ET, of which shall be explored in this study.

The objective of this study is to adopt the water balance framework to develop a validation procedure for seven global-scale ET products based on high-quality hydrological data from 1,743 gauged rainfall stations and 172 streamflow stations across the country, of which should help form a reliable basis. The accuracy of monthly-scale ET assessments is assessed by exploring terrestrial water storage changes (TWSC) throughout yearly cycles. As this cannot be easily physically obtained, the Gravity Recovery and Climate Experiment (GRACE) is used as a proxy for TWSC. Through the inversion of the water balance framework, the accuracy of each ET product can be evaluated with respect to the GRACE-based estimate. This also provides the opportunity to test the performance of GRACE at simulating monthly storage changes. Lastly, given the dependence of ET occurrences on land cover and soil moisture conditions, the ET products are examined for their ability to provide realistic ET estimates with accordance to its land use type to deduce the most suitable ET product for facilitating water management.

2. Study area and materials

2.1. Study area

The Kingdom of Thailand has a latitudinal extent of 5°37′N to 20°37′N and a longitudinal coverage of 97°22′E to 105°37′ E. The country has a total area of 514,050 km^2 and comprises of 25 main river basins and 254 sub-basins. As displayed in Fig. 1, the topography of Thailand comprises of the highest mountain ranges in the North and West, forming the border with Laos and Myanmar, respectively. This region is the source of several major river basins: including the Ping; Wang; Yom; Nan; Sakae-krang; Tha Chin and Pasak basins. These river basins flow south and converge to the Chao Phraya basin, which altogether has a total drainage area of 158,587 km^2, or roughly a third of Thailand's area. The convergence of these tributaries form a large floodplain in Central Thailand.

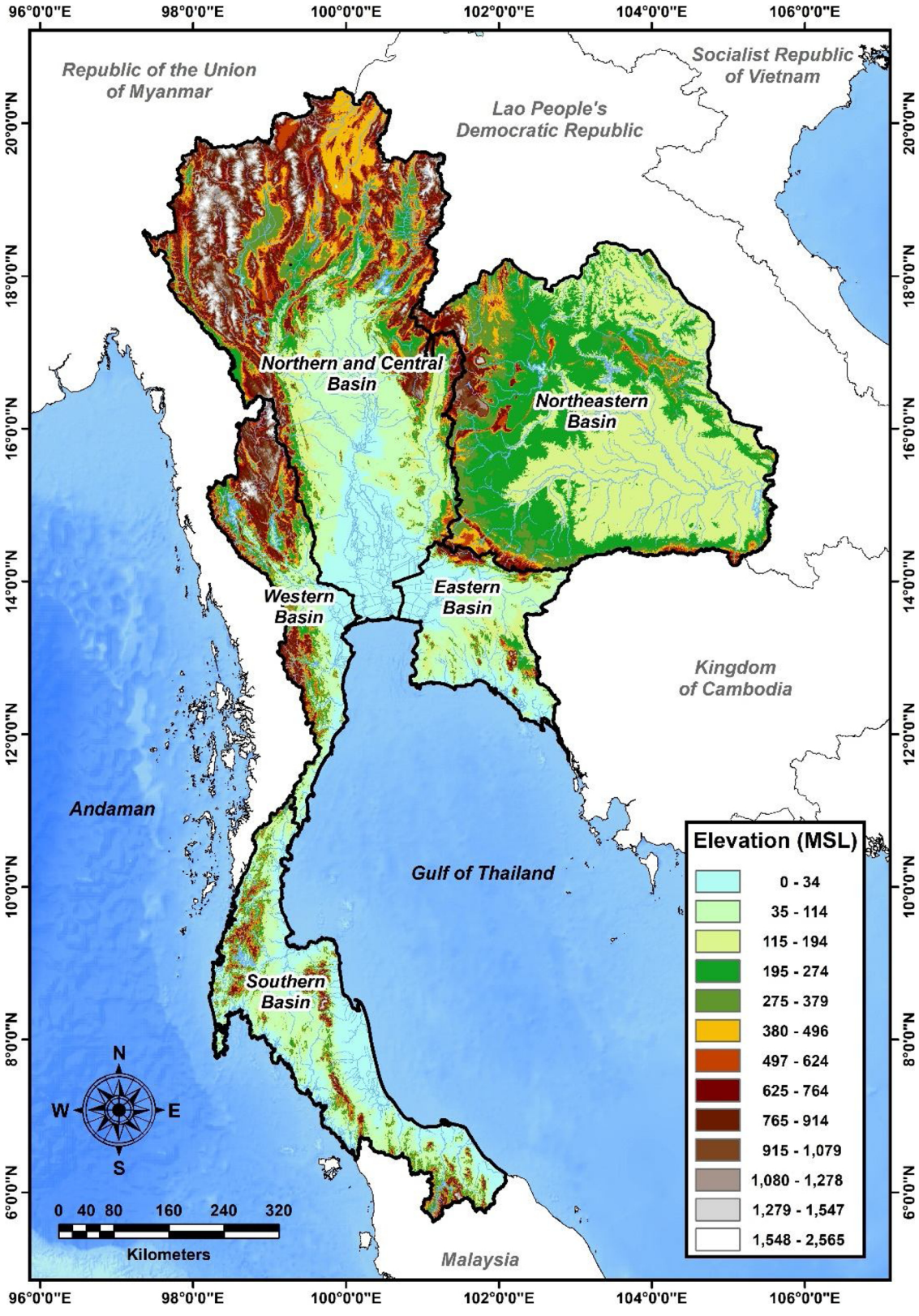


Fig. 1. The digital elevation of Thailand and the boundaries of the 5 regions.

The Eastern part of Thailand drains into the Gulf of Thailand. South Thailand consists of many mountainous rivers that flow directly to the sea. In this study, the country was divided into 5 regions: North and Central, Northeast, East, West and South.

The humid tropical climate of Thailand is influenced by the Southwestern and Northeastern monsoons. The Southwestern monsoon is responsible for bringing moisture to the entire country between May and October. The Northeastern monsoon brings dry weather to mainland Thailand between November and April. However, the monsoon provides Southern Thailand with extra rainfall as moisture is picked up in the Gulf of Thailand. The average annual rainfall depth between 2003 and 2013 of the country is approximately 1,426 mm/year, whereas for Southern Thailand alone, it is 2,132 mm/year. According to the Koppen climate classification, Thailand has a tropical savanna climate (Aw), with some coastal areas being classified as Equatorial climate (Af). The South essentially has a monsoon climate (Am).

2.2. Materials

2.2.1. Gauge rainfall data (P)

The Thailand Meteorological Department (TMD) and Royal Irrigation Department (RID) are responsible for measuring rainfall in Thailand. The TMD and RID individually operated 1150 and 582 gauged rainfall stations between 2003 and 2013, respectively. Other departments also operated 47 additional stations for their specific purposes. This brings the total nation-wide number of stations to 1779. Rainfall data from 2003 to 2013 were checked for their accuracy using the Double Mass Curve (DMC) method (Searcy and Hardison, 1960). Subsequently, 36 stations were eliminated, and daily gridded rainfall datasets with 1 km spatial resolution were generated for the remaining 1743 stations using the Inverse Distance Square (IDS) method (eg. De Silva et al., 2007). Daily rainfall was aggregated onto a monthly scale, whereafter zonal statistics were applied to estimate areal-averaged rainfall for the 172 sub-basins. The drainage area of each sub-basin was carefully determined using the SRTM-30 m model, which is a digital elevation model from the Shuttle Radar Topography Mission (SRTM) (https://www2.jpl.nasa.gov/srtm/).

2.2.2. Streamflow data (Q)

During the study period, daily streamflow data was measured at the outlet of 195 sub-basins by the RID. Similar to the method described in section 2.2.1, 23 sub-basins that contained incorrect streamflow data were eliminated, and monthly datasets were aggregated for the remaining 172 sub-basins to allow for water balance calculations to be performed. The catchment areas of these sub-basins vary from 25 to 108,566 km^2, most of which are located in Western, Northern and Eastern Thailand. The central floodplain region contains no streamflow stations as runoff is unattainable.

2.2.3. Terrestrial water storage change (TWSC)

To complete the computation of evapotranspiration (ET: see Eq. (1)) through the water balance framework, information on storage change (ΔS) is required in addition to rainfall (P) and streamflow (Q).

$$ET_{WB} = P - Q - \Delta S \quad (1)$$

The Gravity Recovery and Climate Experiment (GRACE) was adopted to represent monthly ΔS in terms of the terrestrial water storage change ($TWSC_{GRACE}$) (Wahr et al., 2004). The product already accounts for the distributed and unpredictable interaction between surface water, soil moisture and groundwater, complex reservoir operations and the gradual change of biomass weights. All in all, this alleviates the limited accessibility of physically obtaining ΔS across all sub-basins. Furthermore, the physical nature of monsoonal climate dictates the magnitude of ΔS is significant at monthly time scales as opposed to over the annual scale at which seasonal fluctuations cancel out. The storage change ($TWSC_{GRACE}$) over a given time period is computed by taking the difference of the monthly terrestrial water storage anomalies (TWSA: Rodell and Famiglietti, 1999) between time steps $t_{i+\Delta t}$ and t_i (Billah et al., 2015; see Eq. (2))

$$TWSC_{GRACE}(\Delta t) = TWSA(t_{i+\Delta t}) - TWSA(t_i) \quad (2)$$

Following the end of the GRACE mission in October 2017, the Gravity Recovery and Climate Experiment Follow-On (GRACE-FO) was launched in May 2018. This product provides spherical harmonic solutions of TWSA on one-degree grids (Swenson and Wahr, 2006). These TWSA datasets have been processed at (i) the Center for Space Research of the University of Texas (CSR), (ii) the Jet Propulsion Laboratory (JPL) and (iii) the GeoForschungsZentrum (GFZ). As suggested by Sakumura et al. (2014), the mean value from each dataset was determined to minimise uncertainties. The value of GRACE data has been demonstrated in many hydrological studies (e.g. Rodell and Famiglietti, 2002; Long et al., 2014; Jiang et al., 2014). In July 2018, GRACE-FO RL06 was released and it provides mascon solutions at 0.5-degree resolution. This gridded mascon product is processed by the JPL and is available from https://podaac-tools.jpl.nasa.gov/drive/files/allData/tellus/L3/mascon/RL06. Coastal resolution improvement (CRI) filters are applied to coastal mascons, of which provides better separation between land and ocean signals (Cooley and Landerer, 2019). Further, provided that the mascon-based solution provides greater resolution for smaller spatial regions, the mascon solution of GRACE-FO RL06 was preferred over the spherical harmonic-based solution.

2.2.4. Evapotranspiration (ET) products

Seven global ET products were chosen to be validated against the water balance and land use classifications. These include: the Atmosphere-Land Exchange Inverse model (ALEXI) developed by the United States Department of Agriculture (USDA); the CSIRO MODIS Reflectance-based Scaling ET model (CMRSET); the ETmonitor model from the Chinese Academy of Sciences; the Global Land

Table 1
Summary of the 7 global ET products.

ET product	Spectral measurements used	Energy balance	Directly downloadable	Spatial resolution		Original temporal resolution	Data Availability (Number of water years)
				Deg	km		
ALEXI	Red, NIR, TIR	2-layer Residual	No	0.050	5	Daily	2003-2013 (11)
CMRSET	Blue, Red, NIR, SWIR	–	No	0.050	5	Monthly	2003-2011 (9)
ETMonitor	Red, NIR, TIR, PMW, AMW	2 layer Penman-Monteith	No	0.010	1	Daily	2008-2012 (5)
GLEAM V3.3b	Red, NIR, PMW, AMW	2-layer Priestley & Taylor	Yes	0.250	25	Daily	2003-2013 (11)
MOD16A2	Red, NIR	3-layer Penman-Monteith	Yes	0.005	0.5	8 Days	2003-2013 (11)
SEBS V3	Red, NIR, TIR	1-layer Residual	Yes	0.050	5	Monthly	2003-2013 (11)
SSEBop	Red, NIR, TIR	1-layer Penman - Monteith	Yes	0.010	1	Monthly	2003-2013 (11)

Note: NIR = Near InfraRed; SWIR = Shortwave InfraRed; TIR = Thermal InfraRed; PMW = Passive Microwave; AMW = Active Microwave.

Evaporation Amsterdam Model (GLEAM V3.3b); the gap-filled MOD16A2 V6 from NASA's MODIS Adaptive Processing System (MODAPS) and others; the Surface Energy Balance System (SEBS V3) of the University of Twente and the Simplified Surface Energy Balance model (SSEBop) developed by the United States Geological Survey (USGS). ET data is available for the 2003–2013 period for all 172 sub-basins. However, ETmonitor only has data available from 2008 since it is a relatively new ET product. A summary of major characteristics for each dataset is provided in Table 1. Comparisons between annual ET values are based on Thai hydrological years, which begins on April 1st. All ET products are based on the principles of the surface energy balance, which partitions the net radiation (R_n) gained by the surface into three components as shown in Eq. (3):

$$R_n = G + H + \lambda E \quad (3)$$

where G is the soil heat flux, H is the sensible heat flux and λE is the latent heat flux. The entire discussion on ET modelling relates to the partitioning between H and λE, of which is the key differentiator amongst the seven ET models. Either H needs to be solved explicitly, followed by λE as a rest term, or soil moisture information should be used to assess λE directly without solving H. ALEXI, SEBS, SSEBop and ETMonitor use thermal infrared data to estimate λE and ET. All models use in addition visible and near-infrared spectral data. The CMRSET approach depends strongly on shortwave infrared data. Over against that, GLEAM and ETMonitor retrieve soil moisture from microwave data and have a data assimilation scheme to incorporate moisture measurements into a soil water balance model. MOD16A2 uses atmospheric moisture as a surrogate for soil moisture. This is a nice solution for periods with persistent cloud cover that affects the performance of thermal-based energy balance solutions. Hence, the spectral data used in the 7 ET models is similar for vegetation indices but differ mutually otherwise to solve the partitioning between H and λE.

All models are based on a reference ET_0 with absolute values (mm d^{-1}) to describe day-to-day variability. The Penman-Monteith formulation partitions H and λE by circumventing the need for surface temperature data (Penman, 1948; Monteith, 1965). Four models (MOD16A2, ETMonitor, SEBS and SSEBop) include a certain form of the Penman-Monteith equation. SSEBop uses the FAO standard version for reference ET_0 (Allen et al., 1998), while MOD16A2 and SEBS use a specific solution for the aerodynamic and bulk surface resistances to represent forests (MOD16A2) or cold and wet surfaces (SEBS). ETMonitor is based on a more-advanced two-layer Penman-Monteith equation (Shuttleworth and Wallace, 1985). The three remaining models (ALEXI, GLEAM and CMRSET) are based on the Priestley-Taylor equation for reference ET_0 which is a simplified solution of the Penman-Monteith equation valid for saturated land surfaces (Priestley and Taylor, 1972). There is one distinct difference: ALEXI uses Priestley and Taylor for potential crop transpiration, while GLEAM and CMRSET use it for reference ET_0. Another major difference in the 7 ET models is that GLEAM and ETMonitor use a vertical soil water balance approach to compute daily soil moisture.

Hence, every method has its own merits, and it is very interesting how these models perform in Thailand beyond their safety zone where they are developed and calibrated. The theoretical background of the seven ET products is presented in the supplementary materials.

2.2.5. Land use

The land use in Thailand can be classified into eight classes including forest, irrigated agriculture, rainfed agriculture, aquaculture, urban, water body, wetland and others. Forestlands, which comprises 35.32% of the total land area, have greater abilities at retaining incoming precipitation due to deeper root zone and thus higher soil moisture storage, resulting in greater occurrences of ET compared to other land uses (Sriwongsitanon and Taesombat, 2011). Around 3.44 million ha and 23.12 million ha of agricultural areas are irrigated and rainfed, of which respectively account 7.62% and 44.08% of the total land area, respectively.

2.2.6. ISODATA unsupervised classification

Iterative Self-Organising Data Analysis Technique (ISODATA) is an unsupervised classification method which iteratively clusters pixels into the nearest cluster. After initially arbitrarily assigning the means of each cluster, each pixel is assigned to the cluster with the minimum distance to its mean (Tou and Gonzalez, 1974). Based on the standard deviations of each cluster, those with

dissimilarities greater than the defined thresholds are split, or are otherwise merged. The mean of each cluster is recalculated to redefine cluster statistics. This process is iterated until the clusters are sufficiently stable – i.e. changes to cluster mean and spread are below the defined thresholds. Since the average separability values between clusters increases with the number of specified classes, the set of clusters which provide the highest minimum separability values was considered to be optimal for developing the output cluster images.

In this study, ISODATA was used to classify monthly average ET_{RS} values into different clusters based on similarities in ET signatures. The scheme can be hypothesized to cluster pixels based on commonness of physical conditions (land use, leaf area index, soil type, soil moisture) and water management measures. This extra information was inferred to provide an extra qualitative validation against the land use classes classified by the Land Development Department (LDD). The selected image and its clusters will be used to assist ET gradient map analysis (Richards, 1986).

3. Methodology

3.1. Assessing the accuracy of ET products on the annual scale

By using annual rainfall (P), streamflow (Q) and storage change ($TWSC_{GRACE}$) datasets, the bulk ET_{WB} rate for each sub-basin was determined using the water balance equation as shown in Eq. (1). Annual ET_{WB} values and ET_{RS} estimates were compared by aggregating the pixels covering each sub-basin. The ET product with the highest agreement to ET_{WB} was determined.

3.2. Assessing the accuracy of ET products on the monthly scale

As opposed to modelling ΔS on an annual time scale, there is lower confidence on monthly time scales, for the reasons outlined in Section 2.2.3. Thus, to assess the accuracies of these products on a monthly time scale, the resultant ET_{RS} values are input into an inverted version of the water balance to provide an estimate of storage change, as shown in Eq. (4).

$$\Delta S_{ET} = P - Q - ET_{RS} \tag{4}$$

The seven resulting ΔS_{ET} datasets were compared to storage changes as derived from GRACE (i.e. $TWSC_{GRACE}$) to deduce the most plausible ET product. Standard correlation indicators including slope, R^2 and Nash-Sutcliffe coefficient (NSE) were adopted.

3.3. Correlation between ET_{RS} estimates and land use classifications

To provide additional creditability to certain products, ET_{RS} data is parsed against land use data. Since 87.02% of Thailand is covered by forest, irrigated crops and rainfed agriculture, the ISODATA algorithm was used to classify monthly ET_{RS} values into these three predominant land use classes. Confusion matrices were used to compute the degree of correctness of classifying ET_{RS} pixels to the supposed land use class by dividing the sum of the diagonal elements by the total number of pixels, as shown in Eq. (5). The overall accuracies were further assessed by Cohen's Kappa coefficient (K) as expressed in Eq. (6) (Congalton, 1991).

$$\text{Degree of correctness} = \frac{100\sum_{i=1}^{r}(X_{ii})}{N} \tag{5}$$

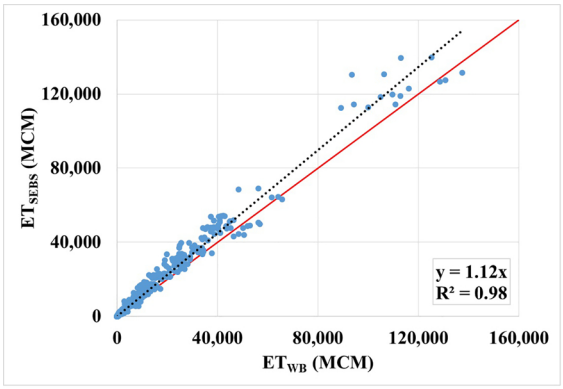
$$\text{Kappa coefficient} = \frac{N\sum_{i=1}^{r}(X_{ii}) - \sum_{i=1}^{r}(X_{i+}X_{+i})}{N^2 - \sum_{i=1}^{r}(X_{i+}X_{+i})} \tag{6}$$

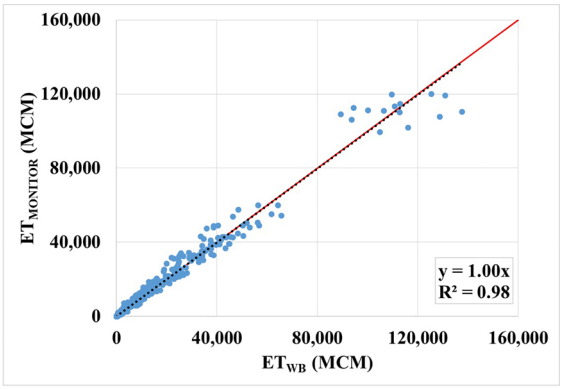
where x_{ii} is the total number of pixels in row *i*, column *i*, N is the total number of pixels, r is the number of rows (columns), xi + and x + i are the total rows and columns, respectively.

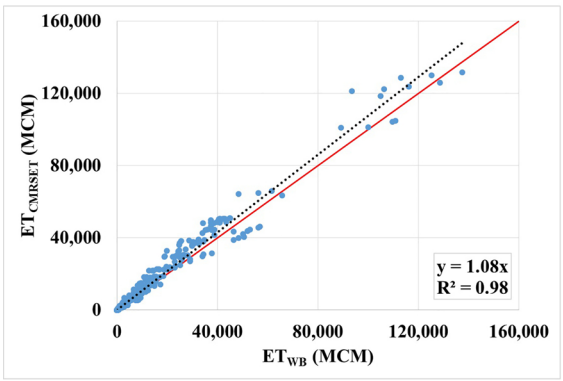
4. Results

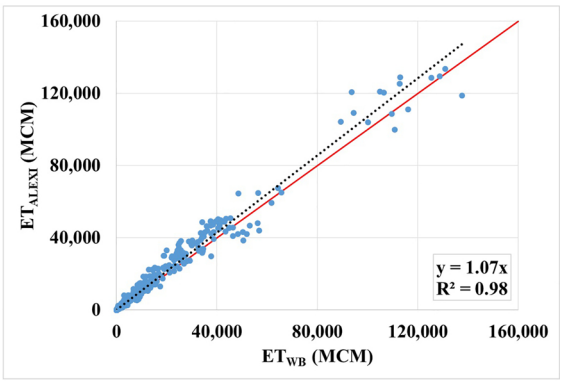
4.1. Accuracy of ET products on the annual scale

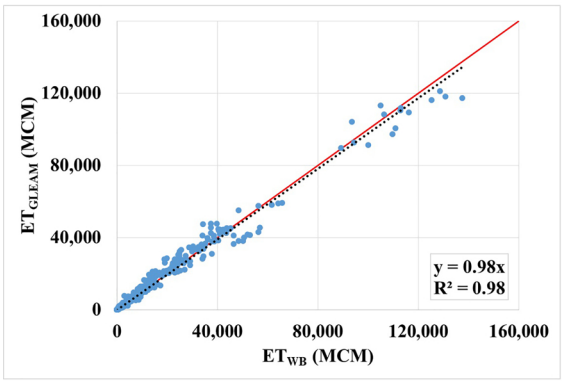
The correlation of annual ET estimates from each product (ET_{RS}) to ET_{WB} has been displayed in Fig. 2 as through a scatterplot. Each data point represents the annual ET averages for each sub-basin. To demonstrate the impact of sub-basin size on ET occurrence, each point was multiplied by its corresponding watershed area. ETMonitor, GLEAM V3.3b and MOD16A2 displayed near perfect correlation to ET_{WB}, requiring 0%, 2% and 2% of bias correction to conserve the water balance, respectively. Moreover, each product showed R^2 and NSE indicators of 0.97 or above. These products have not been calibrated by ground measurements in Thailand before, which makes these results rather encouraging. CMRSET and ALEXI required bias correction of 7–8%, despite the other two indicators yielding similar results to the best-performing products. Conversely, SEBS V3 and SSEBop produced large over-predictions, thus requiring corrections of 12% and 32%, respectively. While the physics of these latter models seem more rigorous, the performance under practical conditions shows that other aspects impact the reliability of ET predictions in an operational context.

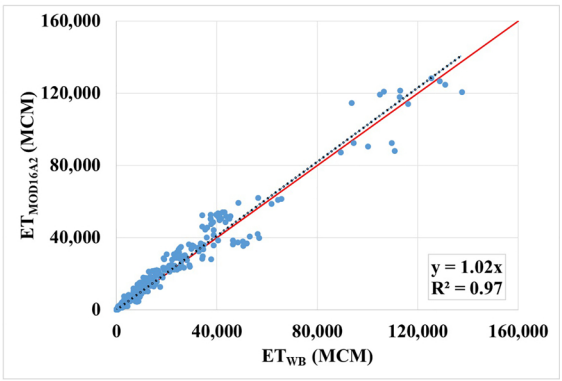


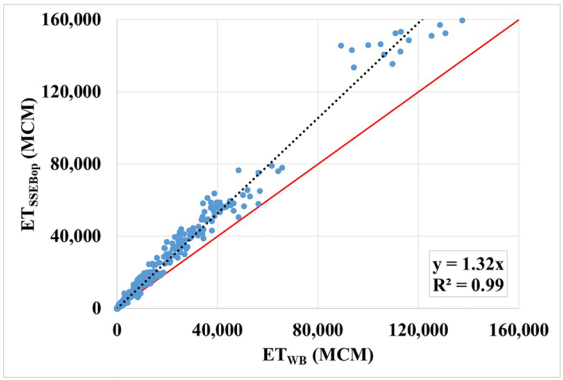












ET products	Spatial resolution (km)	NSE (%)	$ET_{Product} = a\ ET_{WB}$		
			a	|1-a|	R^2
CMRSET	5	97.38	1.08	0.08	0.98
ETMonitor	1	98.42	1.00	0.00	0.98
GLEAM V3.3b	25	98.68	0.98	0.02	0.99
MOD16A2	0.5	97.25	1.02	0.02	0.97
ALEXI	5	97.43	1.07	0.07	0.98
SEBS	5	96.48	1.12	0.12	0.98
SSEBop	1	88.15	1.32	0.32	0.99

Fig. 2. Correlation between annual ET_{WB} and annual ET_{RS} values from 7 ET products in 172 sub-basins from 2008 to 2012 without any a priori calibration or tuning.

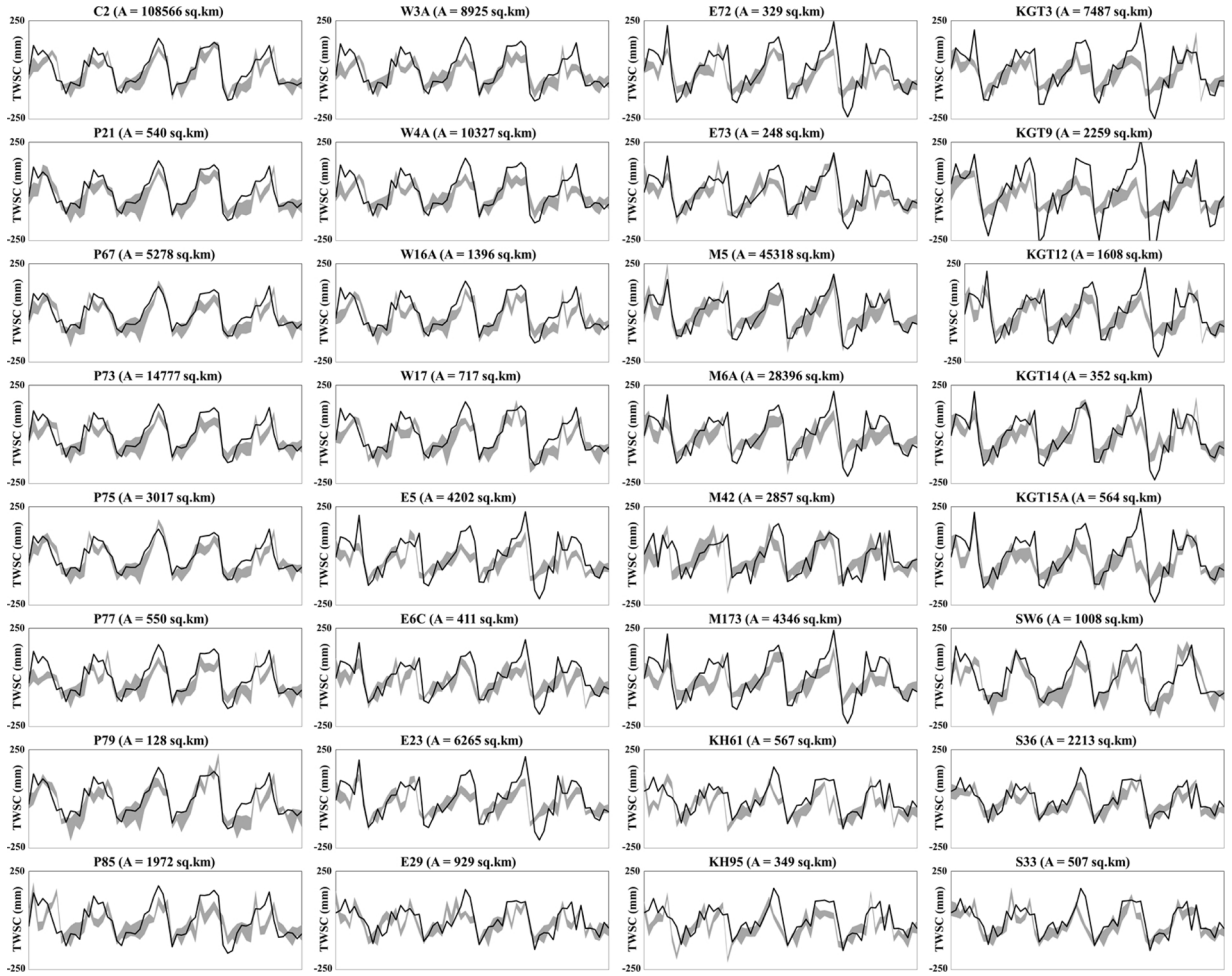


Fig. 3. Comparison between $TWSC_{GRACE}$ and monthly storage changes across 32 example sub-basins of various areas during 2008-2012. Black line represents $TWSC_{GRACE}$, while the grey area indicates variation of ΔS_{ET} from 7 ET products.

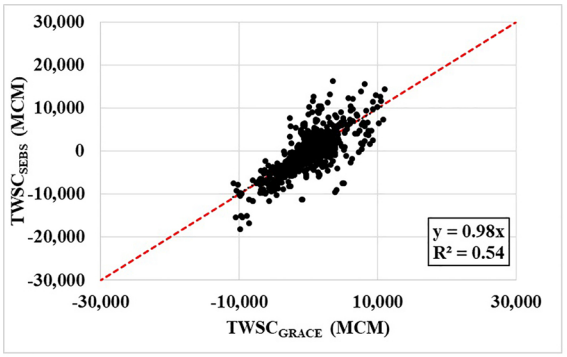
4.2. Accuracy of ET products on the monthly scale

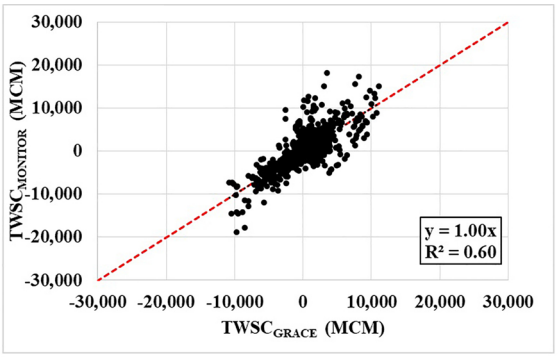
To deduce the ET product that most plausibly mimics monthly storage change dynamics, Fig. 3 compares monthly ΔS_{ET} and $TWSC_{GRACE}$ estimates for 32 example sub-basins which represent an entire range of sub-basins from ~100 km^2 to ~110,000 km^2 in size. Certain ΔS_{ET} products yielded temporal trends with more similarity in magnitude and graduality of changes than others, confirming the hypothesis that each ET product has its own cyclic behaviour. Further, Fig. 4 shows scatterplots between monthly ΔS_{ET} and $TWSC_{GRACE}$ from 2008 to 2012 across 172 sub-basins. To our surprise, all products demonstrated relatively strong adequacies at capturing temporal variations of storage changes (NSE > 0.51). CMRSET and MOD16A2 were the top-performing products (NSE > 0.70, R^2 > 0.67), despite having large systematic biases which can be easily adjusted. These results are indicative of the capabilities of GRACE to represent monthly storage changes in catchments of various sizes. Based on the criteria of showing highest NSE values, GLEAM V3.3b (0.69) and ETMonitor (0.67) were selected alongside CMRSET and MOD16A2 to undergo further analysis. Despite posting similar results, a significant factor that differentiates these best-performing products is their spatial resolutions.

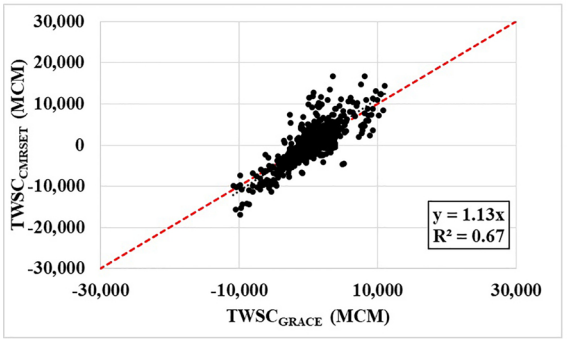
Possessing fine spatial resolutions of 0.5 km and 1 km, MOD16A2 and ETMonitor respectively provide much greater detail over CMRSET (5 km) and GLEAM V3.3b (25 km), thus creating more variability at the temporal scale and thus better expressing the bulk behaviour of composite terrain. This additionally limits their usage is key hydrological applications such as water accounting, irrigation management, as well as drought detection.

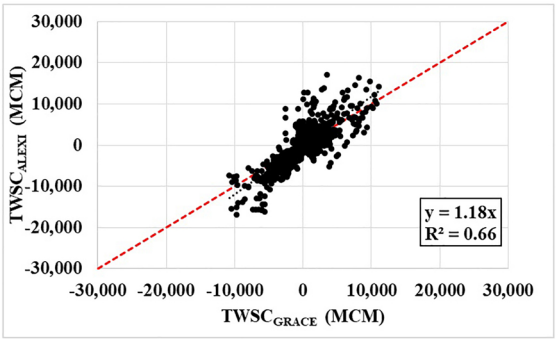
4.3. Correlations between ET_{RS} estimates and land use classifications

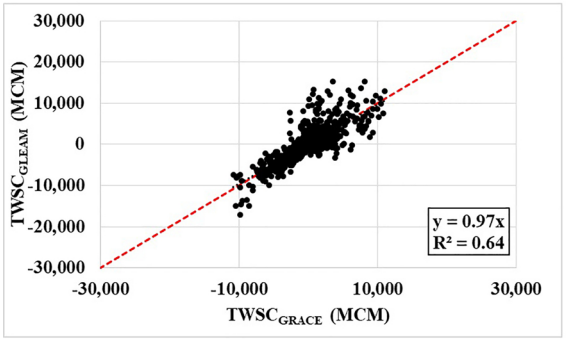
Given that CMRSET, MOD16A2, GLEAM V3.3b and ETMonitor most accurately captured monthly ET dynamics in Thailand, the ISODATA scheme was implemented on these four products to deduce the one which most correctly classifies the land use type of each pixel with respect to the LDD classifications.The classification of ET_{RS} pixels into four clusters provided the highest minimum

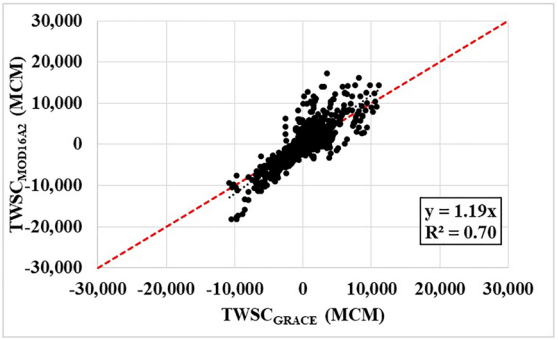


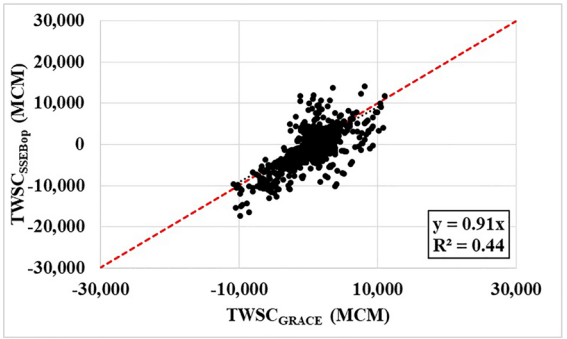












$TWSC_{WB}$	Spatial Resolution (km)	NSE (%)	$TWSC_{WB} = aTWSC_{GRACE}$		
			a	|1-a|	R^2
CMRSET	5	70.99	1.13	0.13	0.67
ETMonitor	1	67.02	1.00	0.00	0.60
GLEAM V3.3b	25	68.63	0.97	0.03	0.64
MOD16A2	0.5	70.35	1.19	0.19	0.70
ALEXI	5	65.25	1.18	0.18	0.66
SEBS V3	5	63.12	0.98	0.02	0.55
SSEBop	1	50.83	0.91	0.09	0.45

Fig. 4. Scatterplots between monthly ΔS_{ET} and $TWSC_{GRACE}$ from 2008 to 2012 across 172 sub-basins.

separability values, and was therefore considered to be optimal for producing the output cluster images. The four clusters appear to be in the order of decreasing average ET values. It could be easily deduced that the two lower clusters are well-watered and water-stressed croplands, respectively. Thus, in accordance to Fig. 5, clusters 3 and 4 represent irrigated and rainfed agricultural areas, respectively. In contrast, clusters 1 and 2 have the highest ET values and can thus be classed as forestlands and perineal vegetation. However, for Eastern and Southern Thailand, the high ET occurrences can be attributed to the greater annual rainfall over other regions, as shown in Fig. 6B. Therefore, the four-cluster classification was replaced with a three-cluster classification through merging clusters 1 and 2 into the forest/perineal land use (i.e. cluster 1) as shown in Fig. 6C–F. The confusion matrices have been presented in Table 2, while the Kappa coefficient (κ) and degree of correctness (DOC) are displayed in Table 3. With respect to the LDD land use

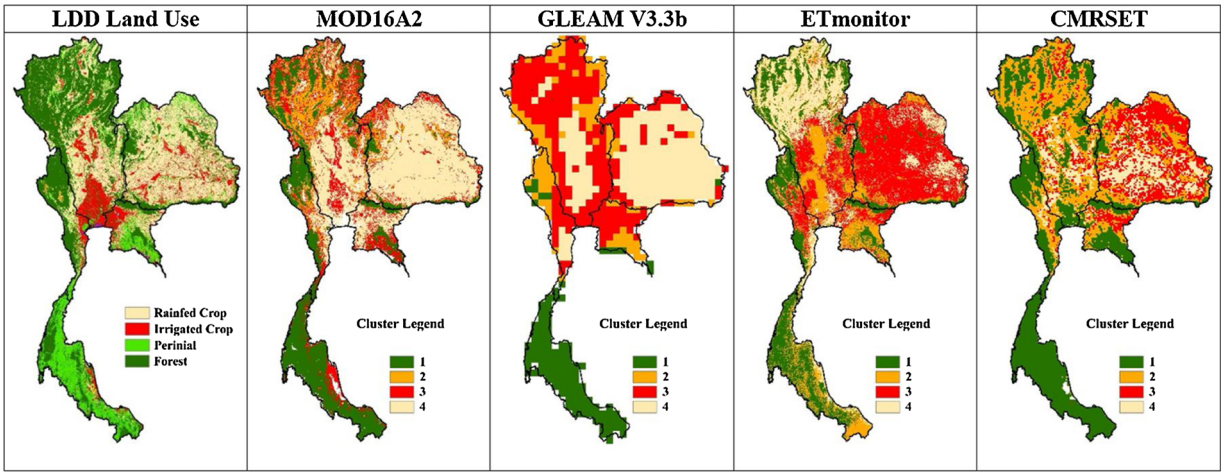


Fig. 5. The classification of ET_{RS} pixels into four clusters by four ET products, with reference to the LDD land use classification.

classification (Fig. 6A), the greatest agreement was demonstrated with the spatial variation of MOD16A2 ($\kappa = 0.52,\ DOC = 70.55\%$), followed by CMRSET (0.34 and 63.97%), GLEAM V3.3b (0.23 and 45.40%), and ETmonitor (0.15 and 45.97%). Assuming that land use is a dominant constraint for the occurrence of ET, the strong performance of MOD16A2 in this test concludes that the product is the most suitable product amongst the seven applicants.

5. Conclusion

Seven global ET products computed from spectral remote sensing measurements have been tested and evaluated by an independent and international science team. None of the products have been developed for Thailand, so this is a real independent check for humid tropics with forests, rainfed and irrigated crops. The ET products were compared to bulk ET which was calculated from the residual of basin rainfall, changes in terrestrial water storage and streamflow at the outlet of 172 sub-basins. The results showed all ET products to strongly resemble annual water balances (NSE > 0.88). It is quite impressive to witness these levels of accuracy with minimum bias correction for some of the products. In addition, monthly storage change were rather adequately captured by all ET products (NSE > 0.51). This also indicates the robustness of the GRACE satellite product for capturing seasonally-induced dynamics of storage changes. However, considering their relative performances and spatial resolutions, four of the seven products (CMRSET, MOD16A2, GLEAM V3.3b and ETMonitor) were deemed to be meritable to be further examined by assessing the ability to provide ET estimates consistent to the soil moisture conditions and thus land use classes. MOD16A2 was able to classify ET_{RS} pixels into either forest/perineal, irrigated- and rainfed-agricultural areas with most correspondence to the LDD land use classifications. In other comparison studies, MOD16A2 was usually not ranked as the best (Trambauer et al., 2013; Hu et al., 2015), but the scientific facts for the good agreement for the environmental conditions in Thailand cannot be disregarded. Apparently, the tropical savanna climate (Aw) amended by the Equatorial climate (Af) in the coast and monsoon climate (Am) in the Southern region are suitable for MOD16A2. Fortunately, the open data access provides great opportunities for this most-promising product to be utilised for various applications in Thailand, including in the fields of hydrology, agronomy and irrigation processes.

6. Discussions

This ground-breaking study significantly improves on previous water balance studies owing to our access to high-quality daily rainfall and runoff measurements. In addition, this study has demonstrated the competence of the GRACE satellite mission in capturing the seasonal dynamics of terrestrial water storage. Moreover, the comparison to established land use classifications provides the ultimate hurdle for testing the ability to provide realistic ET estimates. Therefore, this validation procedure that we have devised should provide a general yet vigorous test that could be adopted beyond our area of study. On another note, future research should focus on investigating the spatiotemporal variability of soil moisture which directly impact on ET fluxes. However, field measurements of soil moisture in large river basins is near impossible, necessitating the exploration of potential indirect remote sensing techniques.

CRediT authorship contribution statement

Nutchanart Sriwongsitanon: Conceptualization, Methodology, Writing - review & editing, Project administration. **Thienchart Suwawong:** Conceptualization, Investigation, Formal analysis, Validation. **Sansarith Thianpopirug:** Data curation, Investigation, Formal analysis, Validation. **James Williams:** Writing - review & editing. **Li Jia:** Supervision. **Wim Bastiaanssen:** Supervision, Visualization.

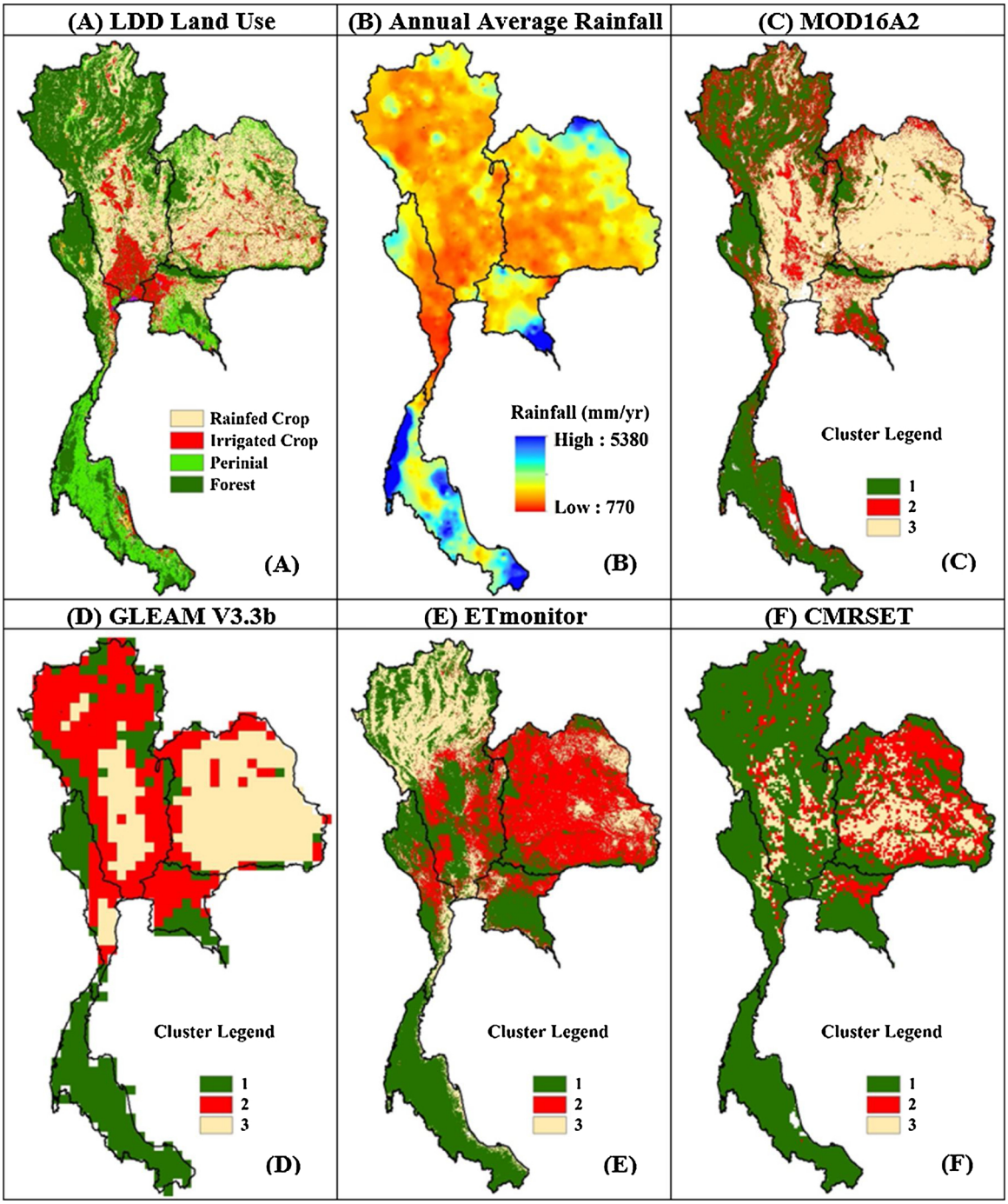


Fig. 6. The revised classification of ET_{RS} pixels into three clusters by four ET products (C–F), with reference to the LDD land use classification (A) and the spatial variation of annual average rainfall (B).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Table 2
Showing the correspondence between the ET clusters with the land use classifications as defined by the LDD for ETmonitor, GLEAM V3.3b, MOD16A2 and CMRSET.

Region	LDD Land use class	Cluster representing decreasing average ET value											
		ETmonitor			GLEAM V3.3b			MOD16A2			CMRSET		
		1	2	3	1	2	3	1	2	3	1	2	3
Eastern	Forest/Perineal	13673	750	3845	25792	228	4069	6995	38	140	14140	2138	3592
	Irrigated Crop	1689	1423	2724	58983	5210	25053	5497	910	1554	1117	739	3111
	Rainfed Crop	226	1131	283	4964	9803	13632	2792	1954	5084	59	421	91
Northern & Central	Forest/Perineal	43784	7618	12846	25792	228	4069	62359	99	4141	84927	12658	21422
	Irrigated Crop	6776	4161	16849	58983	5210	25053	21226	6058	10345	4142	876	8689
	Rainfed Crop	40933	3464	13312	4964	9803	13632	7360	8597	28234	2038	1703	12826
Northeastern	Forest/Perineal	13780	436	5240	5034	230	710	14585	5	829	25202	1708	13929
	Irrigated Crop	21271	3681	59716	12388	321	10032	8633	631	5553	9886	2548	37362
	Rainfed Crop	2858	910	13032	18935	4424	66315	14652	4286	71063	2674	771	26546
Southern	Forest/Perineal	45566	813	1197	44311	1271	1404	42566	290	668	46599	1540	1436
	Irrigated Crop	10	8	0	0	0	0	3251	1255	759	1	0	4
	Rainfed Crop	1822	822	303	0	0	0	13	5	3	1	23	0
Western	Forest/Perineal	20102	380	1065	15635	3	653	21361	97	388	25222	1635	2415
	Irrigated Crop	3814	460	2309	5635	1127	2550	3130	319	1092	322	126	430
	Rainfed Crop	2489	984	809	3082	556	896	1597	1225	2647	577	28	1321
Thailand	Forest/Perineal	136905	9997	24193	116564	1960	10905	147866	529	6166	196090	19679	42794
	Irrigated Crop	33560	9733	81598	135989	11868	62688	41737	9173	19303	15468	4289	49596
	Rainfed Crop	48328	7311	27739	31945	24586	94475	26414	16067	107031	5349	2946	40784

Table 3
Accuracy assessment of the classification of ET_{RS} pixels into the three LDD land use classes by four ET products.

Index	Region	ETmonitor	GLEAM V3.3b	MOD16A2	CMRSET
Kappa	E	0.23	0.08	0.29	0.17
	NC	0.03	0.08	0.43	0.28
	NE	0.10	0.18	0.40	0.23
	S	0.20	0.00	0.36	0.01
	W	0.21	0.21	0.42	0.32
	TH	***0.15***	***0.23***	***0.52***	***0.34***
Percentage Correct	E	59.74	30.21	52.03	58.92
	NC	40.91	30.21	65.12	66.07
	NE	25.22	60.54	71.76	45.01
	S	90.77	94.31	89.78	93.94
	W	65.94	58.59	76.37	83.14
	TH	***45.97***	***45.40***	***70.55***	***63.97***

Acknowledgements

The authors gratefully acknowledge Faculty of Engineering, Kasetsart University for financially supporting this research. We also appreciate Royal Irrigation Department and Thai Meteorology Department for providing the hydrological data. The team of Dr. Martha Anderson from the USDA (USA) provided the ALEXI data to one of the authors, which we acknowledge. We are indebted to Dr. Guerschman from CSIRO (Australia) who provided the CMRSET data. Finally, we gratefully acknowledge the Thailand Research Fund through the Royal Golden Jubilee Ph.D. program (Grant No. PhD/0077/2551).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrh.2020.100709.

References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J., 1986. An introduction to the European Hydrological System—Systeme Hydrologique Europeen,"SHE", 1: History and philosophy of a physically-based, distributed modelling system. J. Hydrol. 87 (1–2), 45–59.

Akter, A., Babel, M.S., 2012. Hydrological modeling of the Mun River basin in Thailand. J. Hydrol. 452, 232–246.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome 300 (9), D05109.

Allen, R.G., Tasumi, M., Trezza, R., 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—model. J. Irrig. Drain. Eng. 133 (4), 380–394.

Anderson, M.C., Norman, J.M., Diak, G.R., Kustas, W.P., Mecikalski, J.R., 1997. A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. Remote Sens. Environ. 60 (2), 195–216.

Bastiaanssen, W.G.M., Chandrapala, L., 2003. Water balance variability across Sri Lanka for assessing agricultural and environmental water use. Agric. Water Manag. 58 (2), 171–192.

Bastiaanssen, W.G.M., Van der Wal, T., Visser, T.N.M., 1996. Diagnosis of regional evaporation by remote sensing to support irrigation performance assessment. Irrig. Drain. Syst. 10 (1), 1–23.

Billah, M.M., et al., 2015. A methodology for evaluating evapotranspiration estimates at the watershed-scale using GRACE. J. Hydrol. 523, 574–586.

Bonan, G.B., Polard, D., Tompson, S.L., 1992. Effects of boreal forest vegetation on global climate. Nature 359 (6397), 716–718.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37 (1), 35–46.

Cooley;, S.S., Landerer, F.W., 2019. Gravity Recovery and Climate Experiment (GRACE) Follow-On (GRACE-FO) Level-3 Data Product User Handbook.

Costa, M.H., Foley, J.A., 1997. Water balance of the Amazon Basin: dependence on vegetation cover and canopy conductance. J. Geophys. Res. Atmos. 102 (D20), 23973–23989.

Courault, D., Seguin, B., Olioso, A., 2005. Review on estimation of evapotranspiration from remote sensing data: from empirical to numerical modeling approaches. Irrig. Drain. Syst. 19 (3–4), 223–249.

da Motta Paca, V.H., et al., 2019. The spatial variability of actual evapotranspiration across the Amazon River Basin based on remote sensing products validated with flux towers. Ecol. Process. 8 (1), 6.

De Silva, R.P., Dayawansa, N.D.K., Ratnasiri, M.D., 2007. A comparison of methods used in estimating missing rainfall data. J. Agric. Sci. 3 (2).

Dias, L.C.P., Macedo, M.N., Costa, M.H., Coe, M.T., Neill, C., 2015. Effects of land cover change on evapotranspiration and streamflow of small catchments in the Upper Xingu River Basin, Central Brazil. J. Hydrol. Reg. Stud. 4, 108–122.

Guerschman, J.P., et al., 2009. Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia. J. Hydrol. 369 (1–2), 107–119.

Ha, L.T., Bastiaanssen, W.G.M., van Griensven, A., van Dijk, A.I.J.M., Senay, G.B., 2017. SWAT-CUP for calibration of spatially distributed hydrological processes and ecosystem services in a Vietnamese river basin using remote sensing. Hydrol. Earth Syst. Sci. Discuss.

Hu, G., Jia, L., 2015. Monitoring of evapotranspiration in a semi-arid inland river basin by combining microwave and optical remote sensing observations. Remote Sens. 7 (3), 3056–3087.

Hu, G., Jia, L., Menenti, M., 2015. Comparison of MOD16 and LSA-SAF MSG evapotranspiration products over Europe for 2011. Remote Sens. Environ. 156, 510–526.

Jiang, D., et al., 2014. The review of GRACE data applications in terrestrial hydrology monitoring. Adv. Meteorol. 2014.

Kalma, J.D., McVicar, T.R., McCabe, M.F., 2008. Estimating land surface evaporation: a review of methods using remotely sensed surface temperature data. Surv. Geophys. 29 (4–5), 421–469.

Karimi, P., Bastiaanssen, W.G.M., 2015. Spatial evapotranspiration, rainfall and land use data in water accounting–part 1: review of the accuracy of the remote sensing data. Hydrol. Earth Syst. Sci. 19 (1), 507–532.

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

Long, D., Longuevergne, L., Scanlon, B.R., 2014. Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. Water Resour. Res. 50 (2), 1131–1151.

Martens, B., et al., 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model. Dev. 10 (5), 1903–1925.

McCabe, M., et al., 2015. The GEWEX LandFlux Project: Evaluation of Model Evaporation Using Tower-Based and Globally-Gridded Forcing Data.

Michel, D., et al., 2016. TheWACMOS-ET project–part 1: tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms. Hydrol. Earth Syst. Sci. 20 (2), 803–822.

Monteith, J.L., 1965. Evaporation and environment, 19. Cambridge University Press (CUP), Cambridge, pp. 205–234.

Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. Proc. R. Soc. Lond. A Math. Phys. Sci. 193 (1032), 120–145.

Poortinga, A., et al., 2017. A self-calibrating runoff and streamflow remote sensing model for ungauged basins using open-access earth observation data. Remote Sens. 9 (1), 86.

Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather. Rev. 100 (2), 81–92.

Ramoelo, A., et al., 2014. Validation of global evapotranspiration product (MOD16) using flux tower data in the African savanna, South Africa. Remote Sens. 6 (8), 7406–7423.

Rodell, M., Famiglietti, J.S., 1999. Detectability of variations in continental water storage from satellite observations of the time dependent gravity field. Water Resour. Res. 35 (9), 2705–2723.

Rodell, M., Famiglietti, J.S., 2002. The potential for satellite-based monitoring of groundwater storage changes using GRACE: the High Plains aquifer, Central US. J. Hydrol. 263 (1–4), 245–256.

Running, S., Mu, Q., Zhao, M., Moreno, A., 2019. MOD16A2GF MODIS/Terra Net Evapotranspiration Gap-Filled 8-Day L4 Global 500 M SIN Grid V006 [Data Set]. NASA EOSDIS Land Processes DAAChttps://doi.org/10.5067/MODIS/MOD16A2GF.006.

Sakumura, C., Bettadpur, S., Bruinsma, S., 2014. Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models. Geophys. Res. Lett. 41 (5), 1389–1397.

Savenije, H.H.G., 2010. HESS Opinions "Topography driven conceptual modelling (FLEX-Topo)". Hydrol. Earth Syst. Sci. Discuss. 14 (12), 2681–2692. https://doi.org/10.5194/hess-14-2681-2010.

Searcy, J.K., Hardison, C.H., 1960. Double-Mass Curves. US Government Printing Office.

Senay, G.B., et al., 2013. Operational evapotranspiration mapping using remote sensing and weather datasets: a new parameterization for the SSEB approach. JAWRA J. Am. Water Resour. Assoc. 49 (3), 577–591.

Shilpakar, R.L., Bastiaanssen, W.G.M., Molden, D.J., 2011. A remote sensing-based approach for water accounting in the East Rapti River Basin, Nepal. Himalayan J. Sci. 7 (9), 15–30.

Shuttleworth, W.J., Wallace, J.S., 1985. Evaporation from sparse crops-an energy combination theory. Q. J. R. Meteorol. Soc. 111 (469), 839–855.

Srinivasan, R., Ramanarayanan, T.S., Arnold, J.G., Bednarz, S.T., 1998. Large area hydrologic modeling and assessment part II: model application 1. JAWRA J. Am. Water Resour. Assoc. 34 (1), 91–101.

Sriwongsitanon, N., 2018. Monthly Water Accounting Across Thailand, Report Submitted to Department of Water Resources. Ministry of Natural Resources and Environment, Bangkok, Thailand (in Thai).

Sriwongsitanon, N., Taesombat, W., 2011. Effects of land cover on runoff coefficient. J. Hydrol 410 (3–4), 226–238.

Su, Z., 2002. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. Hydrol. Earth Syst. Sci. 6 (1), 85–100.

Sutanudjaja, E.H., et al., 2018. PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. Geosci. Model. Dev. 11 (6), 2429–2453.

Swenson, S., Wahr, J., 2006. Post-processing removal of correlated errors in GRACE data. Geophys. Res. Lett. 33 (8).

Tou, J.T., Gonzalez, R.C., 1974. Pattern Recognition Principles.
Trambauer, P., Maskey, S., Winsemius, H., Werner, M., Uhlenbrook, S., 2013. A review of continental scale hydrological models and their suitability for drought forecasting in (sub-Saharan) Africa. Phys. Chem. Earth Parts A/B/C 66, 16–26.
Twine, T.E., et al., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agric. For. Meteorol. 103 (3), 279–300.
Van Eekelen, M.W., et al., 2015. A novel approach to estimate direct and indirect water withdrawals from satellite measurements: a case study from the Incomati basin. Agric. Ecosyst. Environ. 200, 126–142.
Wahr, J., Swenson, S., Zlotnicki, V., Velicogna, I., 2004. Time-variable gravity from GRACE: first results. Geophys. Res. Lett. 31 (11).
Wang-Erlandsson, L., et al., 2016. Global Root Zone Storage Capacity From Satellite-Based Evaporation.
Wang, K., Dickinson, R.E., 2012. A review of global terrestrial evapotranspiration: observation, modeling, climatology, and climatic variability. Rev. Geophys. 50 (2).
Winsemius, H.C., Savenije, H.H.G., Bastiaanssen, W.G.M., 2008. Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins? Hydrol. Earth Syst. Sci. 12 (6), 1403–1413.
Yasin, H.Q., Clemente, R.S., 2014. Application of SWAT model for hydrologic and water quality modeling in Thachin River Basin, Thailand. Arab. J. Sci. Eng. 39 (3), 1671–1684.
Zwart, S.J., Bastiaanssen, W.G.M., de Fraiture, C., Molden, D.J., 2010. A global benchmark map of water productivity for rainfed and irrigated wheat. Agric. Water Manag. 97 (10), 1617–1627.