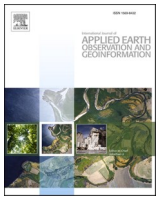
International Journal of Applied Earth Observation and Geoinformation 126 (2024) 103607

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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



A framework for disaggregating remote-sensing cropland into rainfed and irrigated classes at continental scale

Afua Owusu [a,*], Stefanie Kagone [b], Mansoor Leh [c], Naga Manohar Velpuri [c], Murali Krishna Gumma [d], Benjamin Ghansah [a], Paranamana Thilina-Prabhath [c], Komlavi Akpoti [a], Kirubel Mekonnen [e], Primrose Tinonetsana [c], Ismail Mohammed [d]

[a] International Water Management Institute, Accra, Ghana
[b] ASRC Federal Data Solutions, Contractor to U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, USA
[c] International Water Management Institute, Colombo, Sri Lanka
[d] International Crops Research Institute for the Semi-Arid Tropics, Hyderabad, India
[e] International Water Management Institute, Addis Ababa, Ethiopia

ARTICLE INFO

Keywords:
Cropland partitioning
Budyko model
Irrigated cropland
Rainfed cropland
Accuracy assessment

ABSTRACT

Agriculture consumes the largest share of freshwater globally; therefore, distinguishing between rainfed and irrigated croplands is essential for agricultural water management and food security. In this study, a framework incorporating the Budyko model was used to differentiate between rainfed and irrigated cropland areas in Africa for eight remote sensing landcover products and a high-confidence cropland map (HCCM). The HCCM was generated for calibration and validation of the crop partitioning framework as an alternative to individual cropland masks which exhibit high disagreement. The accuracy of the framework in partitioning the HCCM was evaluated using an independent validation dataset, yielding an overall accuracy rate of 73 %. The findings of this study indicate that out of the total area covered by the HCCM (2.36 million km^2), about 461,000 km^2 (19 %) is irrigated cropland. The partitioning framework was applied on eight landcover products, and the extent of irrigated areas varied between 19 % and 30 % of the total cropland area. The framework demonstrated high precision and specificity scores, indicating its effectiveness in correctly identifying irrigated areas while minimizing the misclassification of rainfed areas as irrigated. This study provides an enhanced understanding of rainfed and irrigation patterns across Africa, supporting efforts towards achieving sustainable and resilient agricultural systems. Consequently, the approach outlined expands on the suite of remote sensing landcover products that can be used for agricultural water studies in Africa by enabling the extraction of irrigated and rainfed cropland data from landcover products that do not have disaggregated cropland classes.

1. Introduction

Sustainable agricultural water management (AWM) can be informed by accurate mapping of the location and extent of croplands (Thenkabail et al., 2010). However, comparative analyses of many remote-sensing land use and landcover (LULC) products show high disagreement in croplands globally (Fritz et al., 2010; Pérez-Hoyos et al., 2017; Tchuenté et al., 2011). Fritz et al. (2010) found that the total cropland area worldwide differs by approximately 20 % between remote-sensing LULC products. In a comparison of cropland classes from nine LULC products, Pérez-Hoyos et al. (2017) also found that in over 30 % of the analyzed countries in Africa and the Americas, there was high disagreement whereby only three or fewer LULC products agreed on the classification of areas as cropland. Compared to other landcover classes, there is higher disagreement in cropland classification, particularly in regions with low cropping density and high landscape fragmentation and where major crops have similar seasonal development as the surrounding natural landscapes, as is the case in the savanna regions of Africa (Pittman et al., 2010). The result of this disparity between cropland classes in LULC products is uncertainty in their application for various assessments ranging from water accounting to food production anomaly studies (Karimi et al., 2013; Patle et al., 2023; Pérez-Hoyos et al., 2017; Thenkabail et al., 2010).

In addition to the poor correlation between remote-sensing LULC

* Corresponding author.
E-mail address: a.owusu@cgiar.org (A. Owusu).

https://doi.org/10.1016/j.jag.2023.103607
Received 19 July 2023; Received in revised form 28 November 2023; Accepted 4 December 2023
Available online 15 December 2023

cropland classes, many of these products also fail to distinguish between rainfed and irrigated agriculture but rather lump them together as one cropland class (Buchhorn et al., 2021; Friedl & Sulla-Menashe, 2019; Karra et al., 2021; Latham et al., 2014). Yet, agriculture consumes the largest share of freshwater globally, and therefore, meeting the Sustainable Development Goal (SDG) of 'Zero Hunger' and other associated SDGs, is helped by understanding of the pathways, fluxes, and footprints of water in croplands (Falkenmark & Rockström, 2006; Hoekstra & Mekonnen, 2012; Thenkabail et al., 2010). For water pathways, a distinction is made between blue and green water to understand the different sources of water for their management (Falkenmark & Rockström, 2006). Blue water refers to the water resources available from surface and groundwater that is typically made available to plants through irrigation, while green water refers to the water from precipitation that infiltrates the soil and is naturally available to plants for biomass production (Falkenmark & Rockström, 2006; Johansson et al., 2016; Mekonnen & Hoekstra, 2011). The principal water flux, after precipitation, for this biomass production by plants is evapotranspiration (ET). Over the continents, annual ET volumes are estimated to be over 1.5 times annual runoff (Rodell et al., 2015), yet, owing to its gaseous state, ET is relatively difficult to quantify directly, even at discrete locations (Senay et al., 2016; Sposito, 2017). As such, indirect methods combining satellite-, ground- and model-based methods have been applied to investigate blue and green ET fluxes over different land uses and landcover (LULC) (Chukalla et al., 2015; Kiptala, 2016; Msigwa et al., 2021; Senay et al., 2016; Simons et al., 2016; van Eekelen et al., 2015; Velpuri & Senay, 2017). For example, Chukalla et al. (2015) used the AquaCrop model and the global consumptive water footprint accounting standard to partition blue and green ET fluxes and investigate the water footprint of various irrigation techniques. Velpuri & Senay (2017) used two ET models, one based on the energy balance and the other, the water balance, to partition annual ET for different LULCs in the United States (US). They found that green ET makes up 70 % of total ET while blue ET makes up 30 %, with the annual green and blue ET proportion for croplands being 87 % and 13 %, respectively (Velpuri & Senay, 2017). Forests were found to consume the highest blue ET among non-water landcover classes, approximately 350 mm/year, while for irrigated croplands, annual blue ET was about 250 mm/year (Velpuri & Senay, 2017). These findings are comparable to that of Msigwa et al. (2021), who concluded that high blue water fluxes were attributed to dense forests and irrigated plantations in the Kikuletwa catchment in Tanzania.

One of the models used by Msigwa et al. (2021) in their study of blue and green water fluxes in the Kikuletwa basin was the Budyko model (Budyko, 1974). This model has been applied widely at both large and fine scales for water-energy balance studies owing to its parsimony and accuracy (Reaver et al., 2022; Sposito, 2017). Examples of such applications range from simple estimations of total ET (Gerrits et al., 2009; Zhang et al., 2008), to calibration of global hydrological models (Greve et al., 2020), bias-correction of precipitation data (Beck et al., 2020) and evaluation of the accuracy of remote sensing (RS) derived ET at the basin scale (Weerasinghe et al., 2020). Like Msigwa et al. (2021), as well as Patle et al. (2023), who applied the model to partition blue and green ET, Singh et al. (2022) also successfully utilized a parameterized version in assessing the changes in blue and green ET fractions across India at sub-basin scale by introducing the dimensionless shape parameter, 'ω'. This parameter was initially defined by Fu (1981) to account for catchment characteristics. Other adaptations to the Budyko model include parameters to represent water storage capacity in the root zone (Milly, 1993), vegetation type (Wang & Tang, 2014; Zhang et al., 2001), and then the depth and frequency of rainfall events considering soil properties (Porporato et al., 2004; Sankarasubramanian & Vogel, 2002). Some parameterized versions also aim to account for non-steady state conditions when the change in storage becomes substantial due to additional water from irrigation, groundwater, or water transfer (Du et al., 2016; Greve et al., 2016; Han et al., 2011). The basis of many of these parameterizations is that a basin's water balance is not only affected by climate, as originally derived in the "top-down" Budyko model but also depends on catchment conditions such as vegetation, terrain and soil properties. However, Wang et al. (2016) highlight the difficulty in determining the parameters for each catchment and the small contribution they make to improving the efficiency of the Budyko model. Sposito (2017), also concludes that physical interpretations of model parameters for parametric Budyko models may be "spurious." This is because applications of these parameterized models are catchment specific and cannot be extended beyond these settings or even across different periods within the same catchments (Sposito, 2017). They are therefore in agreement with Gentine et al. (2012), Greve et al. (2015), Padrón et al. (2017) as well as Reaver et al. (2022) who conclude from theoretical analysis as well as empirical tests in over 700 catchments that the non-parametric Budyko model provides an "overarching framework for understanding catchment hydrology in terms of energy and water balances".

In this study, the Budyko model is applied to partition total ET into green and blue fluxes and thereby identify rainfed and irrigated croplands in Africa. This cropland partitioning is done using remote sensing precipitation (P), actual and reference ET, and different LULC products for continental Africa. The specific objectives are to (1) assess the level of agreement between different LULC products for Africa; (2) generate an alternative composite cropland map which overcomes the high disagreement and errors inherent in individual crop masks of LULC products; (3) To demonstrate a scalable approach to partition croplands into irrigated and rainfed classes using the Budyko model; and (4) To apply this approach in partitioning croplands to the individual LULC products and thereby demonstrate its applicability. The novelty of this study lies in the scale at which croplands are being partitioned into rainfed and irrigated areas using a climate-based approach and remote-sensing data inputs. Africa is the second largest continent, extending from latitudes 37°N to 34°S and longitudes 51°E to 17°W, with diverse agricultural practices and climate ranging from Mediterranean to tropical. As a result, it is challenging to find approaches suitable for all areas at the continental scale. Yet, knowing whether a cropland is irrigated or rainfed is crucial for water use evaluation and estimating food production. Accurately distinguishing between the two types of croplands therefore has implications on sustainable management of highly variable, scarce, and transboundary water resources as well as food security on the continent (Chandrasekharan et al., 2021; Li & Troy, 2018). Many of the previous ET partitioning studies have been done at relatively small catchment (Msigwa et al., 2021; van Eekelen et al., 2015) or country (Gumma et al., 2011; Velpuri & Senay, 2017) scales. The methods used in these small-scale studies were considered but were ultimately abandoned in preference for the Budyko model mainly due to their data requirements. In contrast, the Budyko model requires only three climate variables that are readily available at continental scales through remote-sensing products.

2. Methodology

An overview of the methodology followed in this study is presented in Fig. 1, with details on the Budyko model, the datasets, and methods presented in the following sections.

2.1. Theory - Budyko model

The Budyko model (Budyko, 1974) empirically represents a river basin's long-term water and energy balance under steady-state conditions. Given that long-term changes in a basin's storage are negligible $(\Delta\overline{S} \approx 0)$, then mean precipitation $(\overline{P})$ in that basin is partitioned into either discharge $(\overline{Q})$ or (green) evapotranspiration $(\overline{ET})$:

$$\overline{P} = \overline{Q} + \overline{ET}_{green} \quad (1)$$

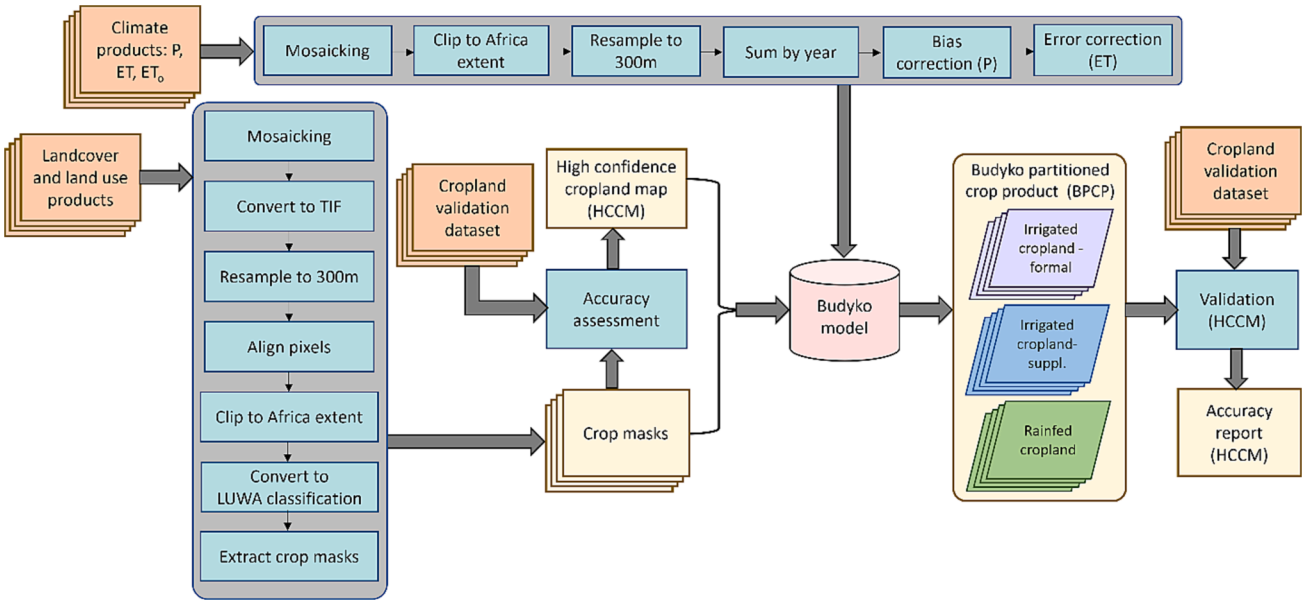


Fig. 1. Methodology followed in deriving rainfed and irrigated cropland areas from remote-sensing landcover products. NB: LUWA- Land Use for Water Accounting; TIF- Tag Image File format.

From observations in several catchments, Budyko (1974) proposed that the mean (green) evapotranspiration ($\overline{ET}$) of a catchment is a function of the available energy (i.e., potential evapotranspiration ($\overline{ET_o}$)), and available water (i.e., precipitation ($\overline{P}$)), or more specifically, its aridity index (ϕ) where:

$$\boldsymbol{\phi} = \overline{\boldsymbol{ET}_o}/\overline{\boldsymbol{P}} \tag{2}$$

Budyko thus derived the explicit functional relationship between (green) $\overline{ET}$ and ϕ as:

$$\overline{\boldsymbol{ET}}_{\boldsymbol{green}} = \overline{\boldsymbol{P}}\sqrt{(1-\boldsymbol{e}^{-\boldsymbol{\phi}}).\boldsymbol{\phi}\mathbf{tanh}1/\boldsymbol{\phi}} \tag{3}$$

Thus, considering available energy and water in a location, it is possible to determine any additional water input (blue $\overline{ET}$) if the measured actual $\overline{ET}$ exceeds green $\overline{ET}$ as:

$$\overline{\boldsymbol{ET}}_{\boldsymbol{blue}} = \overline{\boldsymbol{ET}}_{\boldsymbol{actual}} - \overline{\boldsymbol{ET}}_{\boldsymbol{green}} \tag{4}$$

In cropland areas, these are areas where formal or supplemental irrigation exists over the study period.

2.2. Datasets

2.2.1. Climate variables: P, ET and ET_o

Partitioning of ET using the Budyko model requires only three climatic inputs: precipitation, actual ET and potential ET (ET_o) (Eqs. (1)–(4)). In this study, the Multi-Source Weighted-Ensemble Rainfall (MSWEP) (Beck et al., 2019) precipitation, operational Simplified Surface Energy Balance (SSEBop) (Senay et al., 2020) actual ET, and Water Productivity through Open access of Remotely sensed derived data (WaPOR) (FAO, 2018) potential ET datasets served as inputs to the Budyko model (Table 1). The choice of MSWEP and SSEBop was informed by their good performance across Africa in continental-scale reviews by Mekonnen et al. (2023) and Weerasinghe et al. (2020) respectively. With respect to ET_o, WaPOR ET_o, which is grounded in the combined temperature- and radiation-based Penman-Monteith (Monteith, 1965) equation, was used. The simplified Food and Agriculture Organization's (FAO) version of the Penman-Monteith equation for a reference grass surface (Allen et al., 1998) is presented below:

$$\mathbf{ETo} = \frac{0.408\Delta(\mathbf{R_n} - \mathbf{G}) + \frac{900}{\mathbf{T}}\gamma\mathbf{u}_2\delta\mathbf{e}}{\Delta + \gamma(1 + 0.34\mathbf{u}_2)} \tag{5}$$

where ET_o = Potential evapotranspiration, as water volume

Table 1
Remote-sensing climate, and landcover and land use (LULC) products used in this study.

	Product	Spatial resolution* (m)	Temporal resolution/ Years	Source	Reference
Climate products	MSWEP precipitation	~ 11,000	monthly	GloH2O	Beck et al., 2019
	SSEBop actual evapotranspiration (ET)	~1,000	monthly	USGS	Senay et al., 2020
	WaPOR potential evapotranspiration (ET_o)	~19,000	monthly	FAO	FAO, 2018
Landcover and land use (LULC) products	MODIS LC	500	2003–2020	NASA	Friedl and Sulla-Menashe, 2019
	CCI LC V2	300	2000–2021	ESA	ESA CCI, 2015
	Sentinel-2	10	2017–2021	ESRI	Karra et al., 2021
	Copernicus	100	2015–2019	ESA	Buchhorn et al., 2021
	GCEP30	30	2016–2020	NASA	Thenkabail et al., 2021
	GLC-SHARE	1,000	2013	FAO	Latham et al., 2014
	GlobCover	300	2005,2009	ESA	Bicheron et al., 2008
	WaPOR	250	2009–2020	FAO	Latham et al., 2014

* Approximate spatial resolution in metre at the equator.

evapotranspired (mm day^{-1}), Δ = rate of change of saturation specific humidity with air temperature (Pa K^{-1}), R_n = net irradiance (MJ m^{-2} day^{-1}), G = ground heat flux (MJ m^{-2} day^{-1}), T = air temperature at 2 m (K), γ = psychrometric constant ($\gamma \approx 66$ Pa K^{-1}), u_2 = wind speed at 2 m height (m s^{-1}), and δe = vapor pressure deficit (kPa). ET_o is considered a less ambiguous variable than actual ET and therefore, remote-sensing ET_o products perform as well as the equations they are built on (Fisher et al., 2011). There is also little variation among different products.

2.2.2. LULC products

Eight common LULC products were used in disaggregating croplands into rainfed and irrigated croplands for Africa (Table 1). Detailed descriptions of each LULC product can be found in Appendix A.

2.2.3. Validation dataset

An independent validation dataset of 955 (554 irrigated and 401 rainfed) cropland locations (Fig. 2) was used in validation. This ground-truth dataset was curated by International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) from field visits and high-resolution satellite imagery during 2015—2019.

2.3. Methods

2.3.1. Preprocessing – Climate products

All three climate products were preprocessed by mosaicking, clipping to an Africa extent, resampling to 300 m using nearest neighbor method, and then aggregating to the annual time step (Fig. 1). The spatial resolution of 300 m was chosen in keeping with the median and modal resampled spatial resolution of the eight LULC products. The MSWEP data were bias-corrected using the annual bias estimate of 5 % for Africa compared to in situ rain gauge observations (Mekonnen et al., 2023). Furthermore, an error correction was done on actual ET given that natural landscapes such as grasslands should only show green water use. Therefore ET is at most, equal to precipitation (ET $\leq$ P). A mean annual error in grassland pixels (i.e., where actual ET > precipitation) at sub-catchment scale was therefore used as a correction factor for the SSEBop actual ET data.

2.3.2. Preprocessing – Landcover products

The specific preprocessing steps, legend translation and maps of the eight LULC products used in this study can be found in Appendix B. Generally, this preprocessing involved reprojection to the WGS84 geographic coordinate system, resampling to 300 m resolution using nearest neighbor method, harmonization to the Land Use for Water Accounting (LUWA) classification system (Karimi et al., 2013) and then the creation of cropland masks (Fig. 1). LUWA is a simple classification system with eight main classes that is widely used for water accounting (Karimi et al., 2013; Patle et al., 2023). It is based on the FAO's land-cover classification system, and it was chosen to simplify the landcover classes and thereby make map interpretation and comparison easier. The eight main LUWA classes are Tree-1, Shrub-2, Grass-3, Crop-4, Miscellaneous (Misc.)-5, Wetland-6, Water-7, and Barren-8 (Fig. 3). The 300 m spatial resolution was chosen to minimize the effects of resampling on

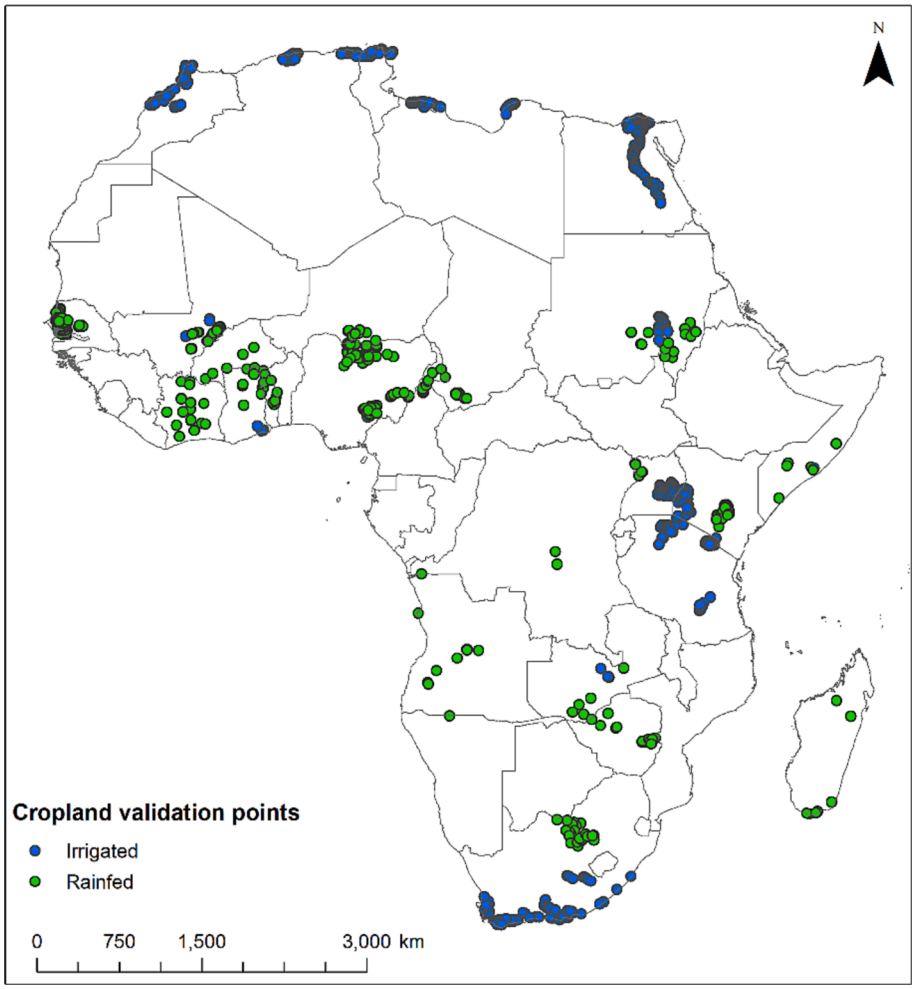


Fig. 2. Independent validation dataset of 955 rainfed and irrigated cropland locations across Africa.

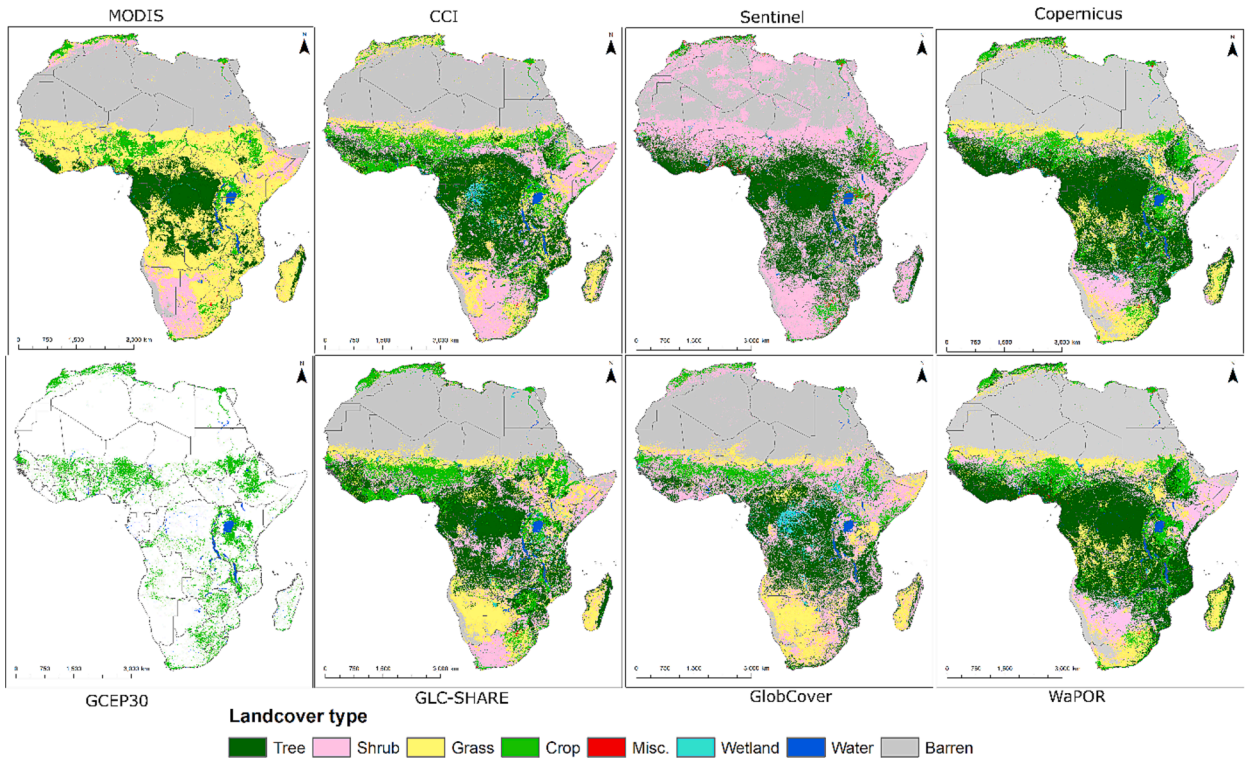


Fig. 3. Landcover products after legend translation to Land Use for Water Accounting (LUWA) classification.

the LULC products as this was the modal and median (rounded up to the nearest 100 m) resolution of the eight products. The crop masks for the eight landcover products are shown in Fig. 4. Appendix C of the supplementary material presents a map (Figure C1) showing how each product compares to country level data of cultivated area from the FAO Aquastat. The Sentinel cropland area, is generally the smallest relative to

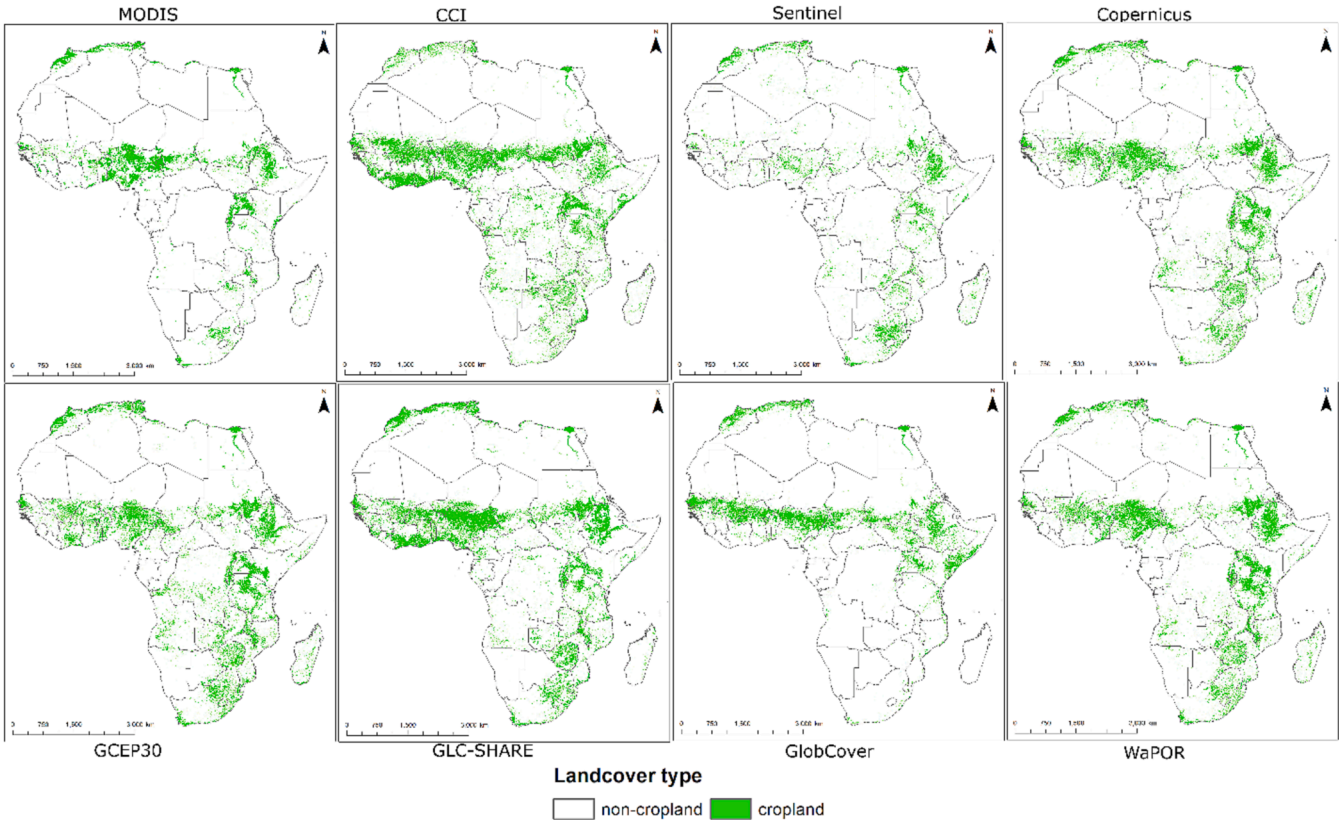


Fig. 4. Crop masks of eight landcover products.

the other landcover datasets while CCI cropland is the largest.

2.3.3. Accuracy assessment of cropland masks and derivation of high-confidence cropland map (HCCM)

Apart from apparent variations in spatial and temporal resolutions of LULC products (Table 1), these products also differ in the base imagery and ancillary input data, the methodology followed in their derivation, as well as the size and quality of training and validation data (Thenkabail et al., 2021). The result is high uncertainty due to the disagreement between these products (Pérez-Hoyos et al., 2017; Thenkabail et al., 2021). To overcome this uncertainty, following an analysis of the level of agreement of the eight LULC products for the year 2019, a high-confidence cropland mask (HCCM) was created for cropland partitioning (Fig. 1). 2019 was chosen as the reference year because it was the most recent year for which six out of eight products had data. The most recent year was used for the two products without data for that year, GLC-SHARE and GlobCover.

In deriving the HCCM, two assessments were carried out on the eight crop masks. The first was a point-to-pixel accuracy assessment using the independent cropland validation dataset. The second compares the area of each crop mask to country-level data on cultivated area from FAO's Aquastat data portal (FAO, 2021). For the first assessment, the validation database did not include non-cropland classes, so it was only possible to determine the overall accuracy of each crop mask. This is the percentage of correctly predicted cropland areas (i.e., true positive (TP)) by the LULC products out of the 955 validation points. Following confirmation that the Global Cropland-Extent Product at 30-m Resolution (GCEP30) was produced using the validation dataset, GCEP30 was excluded from the derivation of the HCCM (Thenkabail et al., 2021). The top three highest-scoring cropland masks from both assessments were used to derive the HCCM based on 66 % agreement between them. Note that the HCCM is not considered a new product as it is derived simply based on the agreement between three existing products. It is used in this study for calibration and validation of the cropland partitioning framework to overcome the errors inherent in individual cropland products.

2.3.4. Budyko-partitioned cropland product (BPCP)

The Budyko model was applied to the HCCM to derive green and blue ET for the croplands (Fig. 1). An annual blue ET threshold above which cropland was considered irrigated was set as the maximum of either the mean blue ET of grassland pixels at the sub-catchment scale or an infinitesimal blue ET of 0.01 mm. The mean blue ET of grassland pixels threshold caters to 'natural' blue ET, for example, from flood waters (Balana et al., 2019; Motsumi et al., 2012), whereas the infinitesimal blue ET threshold accounts for the evaporative fraction asymptotically approaching unity ($\overline{ET}/\overline{P} \rightarrow 1$) for large values of climate aridity (ϕ) in the Budyko model (Carmona et al., 2016). A further distinction was then made between irrigated cropland areas with high blue ET ($\geq$100 mm), i.e., areas under formal irrigation, and then areas with relatively low blue ET (<100 mm), i.e., supplemental irrigation. This threshold of at most 100 mm for supplemental irrigation was the average value based on research on supplemental irrigation for sorghum (Fox & Rockström, 2003) and maize (Gadédjisso-Tossou et al., 2018), two crops grown in nearly all ecological zones of Africa. The research found that yields in these rainfed crops were stabilized by applying minimal amounts of water to bridge dry spells, specifically 60–90 mm for sorghum and 150 mm for maize. Note that because annual values are used to identify irrigated areas, both ET from double cropping and ET outside the growing season are accounted for.

2.3.5. Validation of BPCP

Following cropland partitioning of the HCCM using the Budyko model, validation was carried out using the same validation database of cropland locations, this time differentiated into rainfed and irrigated cropland locations (Fig. 1). Out of the 955 validation points, 861 points that fell within the HCCM cropland area were used in generating a confusion matrix. For this accuracy assessment, a confusion matrix was generated to show both accuracies and errors, thereby providing a complete summary of the predictions made by the Budyko model as compared to the actual ground truth. Overall accuracy was calculated as the percentage of correctly predicted rainfed and irrigated areas (i.e., true negative (TN) and true positive (TP), respectively) by the Budyko model, out of the 861 validation points. Denoting false negatives and false positives as *FN* and *FP* respectively, the values for precision are the proportion of true positive predictions out of all positive predictions made by the model ($TP/(TP + FP)$); sensitivity is the proportion of true positive (irrigated) predictions out of all actual positive instances ($TP/(TP + FN)$); and specificity is the proportion of true negative (rainfed) predictions out of all actual negative instances ($TN/(TN + FP)$). A Cohen's Kappa score, which is a measure of the classification agreement between the Budyko-partitioned cropland map and the validation dataset, compared to the agreement that could have been expected by mere chance, was also calculated (McHugh, 2012).

2.3.6. Application of Budyko model to individual cropland masks

Applying the Budyko model to partition a composite cropland mask based on highly accurate crop masks, aims to produce an alternative disaggregated product that overcomes the errors inherent in individual crop masks of LULC products. However, each LULC product has been developed and validated using different approaches and databases (Karra et al., 2021; Latham et al., 2014; Tsendbazar et al., 2020; Zanaga et al., 2021), and as such, their crop masks may be preferred for specific applications or study areas. In such situations, the framework followed in partitioning the HCCM can be applied to the crop masks created from any individual LULC product. This has been done for the original eight cropland masks of the LULC products used in this study (Fig. 1).

3. Results

3.1. Agreement between landcover products

Following reclassification to LUWA classes, the mean agreement between the eight LULC products across all LUWA classes for the reference year 2019 ranges from 50.9 % to 68.6 % (Table 2). GCEP30 has the lowest agreement with the other products, an average of approximately 50.9 % because only three LULC types are classified in GCEP30: water, cropland, and a broad class for all non-croplands. WaPOR and Copernicus products have the highest average correlation with the other products (68.6 % and 67.2 %, respectively). WaPOR's and Copernicus' products also have a high correlation with each other (87.7 %), possibly because WaPOR landcover maps were based on a previous version (version 2) of the Copernicus product used in this study (version 3). Excluding their pairwise correlation, the mean correlation of WaPOR and Copernicus with the other products is slightly reduced to 65.4 % and 63.8 %, respectively. The lowest pairwise similarity was 33.9 % between GCEP30 and Sentinel LULC products.

Besides the small number of classes used in this study, the fair agreement between the LULC products is mainly driven by the relatively high agreement ($\geq$62 %) on barren (Sahara and Namib deserts), trees (central Africa) and to a smaller areal extent, water (African Great Lakes and other permanent lakes) classes (Fig. 5). In contrast, there is a low agreement ($\leq$50 %) in the classification of areas as shrubs, grass, cropland and wetland. For croplands, the exception to this generally low agreement occurs in small clusters across the Nile basin. For reference, a map of administrative regions in Africa (Figure C2), obtained from IGAD Climate Prediction and Application Center (ICPAC: https://geoportal.icpac.net/layers/geonode%3Aafr_g2014_2013_0) and another of major river basins (Figure C3) from HydroSHEDS (https://www.hydrosheds.org/products/hydrobasins) are provided in Appendix C in the Supplementary material.

Table 2
Correlation matrix for each possible dataset pair to compare the overall agreement (in %) for Africa. 2019 was chosen as the reference year.

Agreement %	MODIS LC	CCI V2	Sentinel	Copernicus	GCEP30	GLC-SHARE	WaPOR	GlobCover	Mean
MODIS LC									**54.6**
CCI LC V2	57.4								**61.7**
Sentinel-2	43.3	56.3							**51.1**
Copernicus	63.0	67.8	59.3						**67.2**
GCEP30	37.1	51.2	33.9	64.6					**50.9**
GLC-SHARE	60.5	66.1	51.7	65.1	61.4				**62.6**
WaPOR	63.6	68.2	59.1	87.7	72.8	65.9			**68.6**
GlobCover	54.4	64.9	54.4	63.0	35.5	67.4	62.9		**57.5**

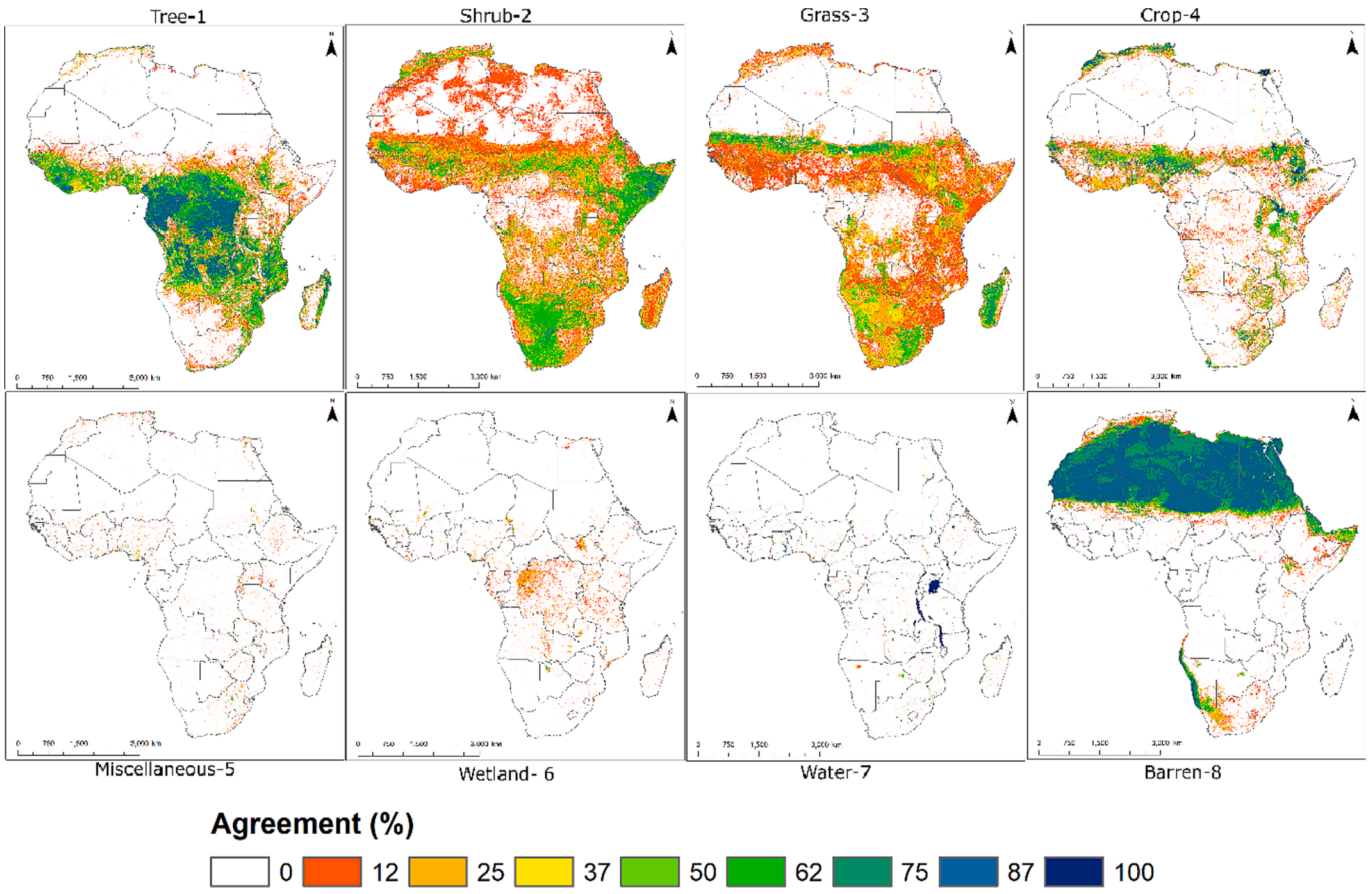


Fig. 5. Degree of agreement (%) of the eight remote sensing landcover products for the eight landcover for Water Accounting (LUWA) classes.

3.2. High confidence cropland mask (HCCM)

The crop masks of three LULC products, WaPOR landcover (WaPOR), Copernicus Global Land Service-Land Cover 100 Version 3 (Copernicus), and Global Land Cover-SHARE (GLC-SHARE), were the highest performing products in both assessments, in each case scoring more than 80 % (Table 3 and Fig. 6). These three products were therefore used to create a high-confidence cropland map (HCCM) with a threshold agreement of 66 % used to identify 'true' cropland areas. This threshold was chosen so that the total cropland area of the HCCM was comparable to the total cultivated area in Africa of 2.80 million km^2 according to FAO's Aquastat data for 2019 (FAO, 2021).

Table 3
Results of independent accuracy assessment of crop masks from landcover products (NB: GCEP30 was excluded from the accuracy assessment as the validation dataset used in this study was also used in producing GCEP30).

Landcover product	Correct points	Accuracy assessment (%)
WaPOR	868	91
Copernicus	807	85
GLC-SHARE	779	82
MODIS LC	604	63
Sentinel-2	542	57
CCI LC V2	504	53
GlobCover	475	50

As expected, there is high agreement between the three cropland masks found to have the highest accuracies following the twofold assessment (Fig. 7a). Relatively large areas of 100 % agreement between these three crop masks occur in the Nile basin in Egypt, Sudan, and Ethiopia; in the North West Coast river basin in Morrocco; in the Niger basin in Nigeria; and in the Rift Valley basin in Ethiopia. In contrast, along West Africa's coast, only one crop mask, and GLC-SHARE, identifies cropland pixels.

The HCCM is shown in Fig. 7b. The total cropland area in this mask is 2.36 million km^2, approximately 8 % of the total area of Africa. This area is also 84 % of the total cultivated area in Africa of 2.8 million km^2, according to FAO's Aquastat data for 2019 (FAO,2021).

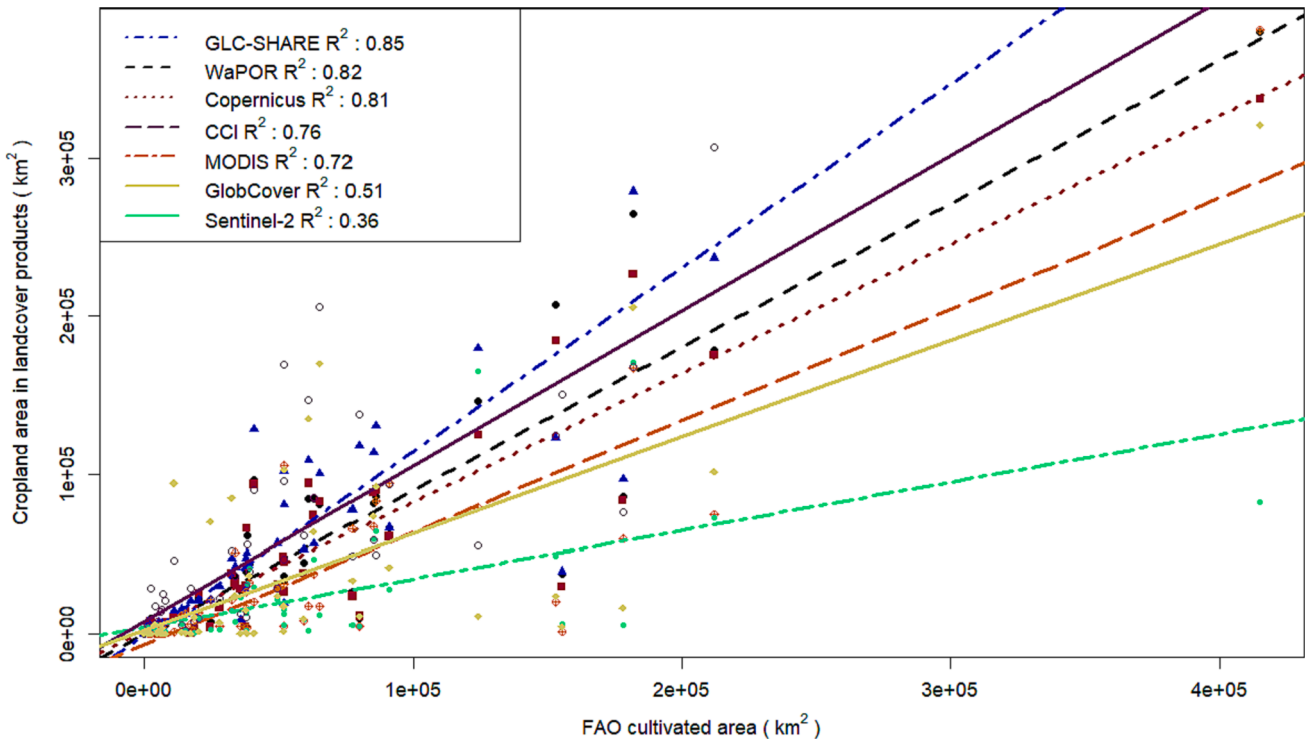


Fig. 6. Country-level FAO Aquastat data on cultivated area versus seven remote-sensing landcover products (where R2 is the the coefficient of determination).

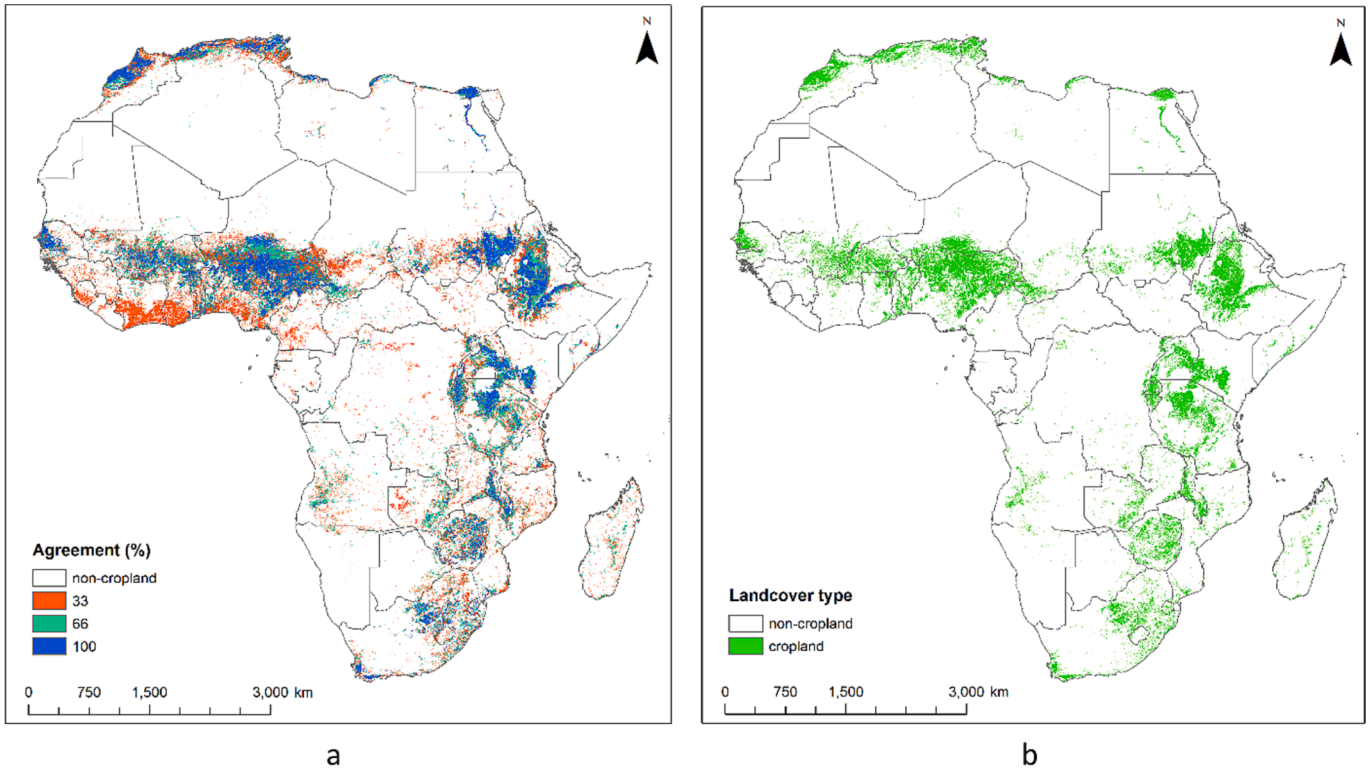


Fig. 7. **a.** Degree of agreement (in %) of three cropland masks with highest scores in two-fold assessment. **b.** High confidence cropland mask (HCCM) based on ≥ 66 % agreement between the three cropland masks with highest scores in two-fold assessment.

3.3. Budyko-partitioned cropland product (BPCP)

The disaggregated cropland map created by partitioning the HCCM into rainfed and irrigated areas using the crop partitioning framework is shown in Fig. 8, while the confusion matrix generated after comparison with the independent validation dataset is presented in Table 4. An area of 461,000 km^2, which is 19 % of the HCCM cropland area is identified as irrigated. Distinguishing between cropland areas of formal irrigation

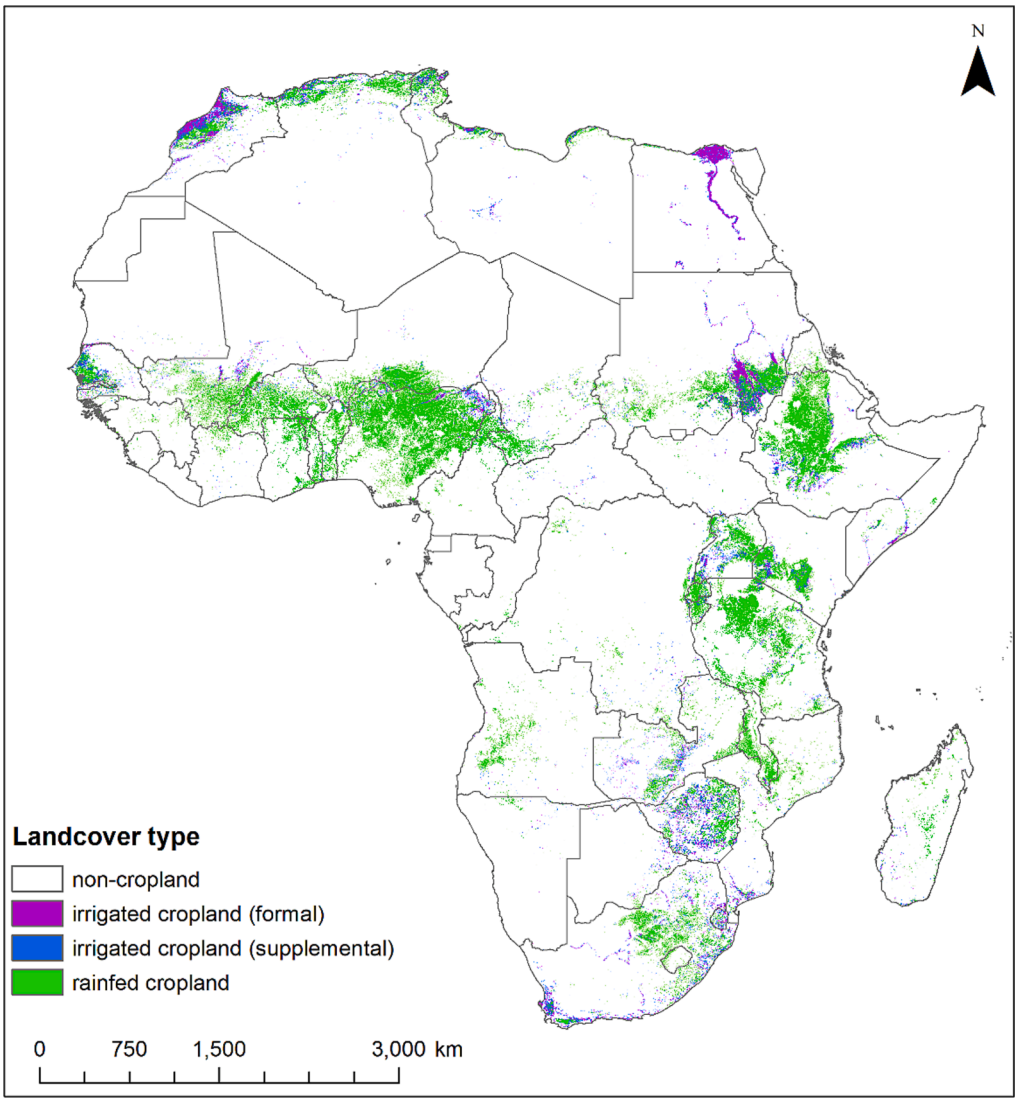


Fig. 8. Disaggregated cropland product for Africa showing irrigated and rainfed areas as identified using the Budyko model. A distinction is made between cropland areas with blue ET $\geq$ 100 mm (formal irrigation) and cropland areas with blue ET $\leq$ 100 mm (supplemental irrigation). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Confusion matrix of rainfed and irrigated cropland product derived after partitioning HCCM using the Budyko model.

Class	Rainfed	Irrigated	Total	User's accuracy	Commission errors
Rainfed	321	27	348	92 %	8 %
Irrigated	208	305	513	59 %	41 %
Total	529	332	861		
Producer's accuracy (%)	61 %	92 %			
Omission errors (%)	39 %	8 %			
Overall accuracy					73 %
Cohen's kappa					0.48

(i.e., blue ET $\geq$ 100 mm) and areas with supplemental irrigation in this Budyko-partitioned cropland product (BPCP), it is found that 239,000 km^2 (51.9 %) is under formal irrigation. In contrast, the remaining 222,000 km^2 can be classified as areas under supplemental irrigation. Large continuous areas where formal irrigation occurs can be found in the North West Coast basin in Morocco, the Nile Valley and Delta in Egypt, and the Nile basin in Sudan. Other notable areas where irrigation occurs include scattered croplands in the Zambezi basin in Malawi and Zambia, the inland delta of the Niger basin in Mali, the Lake Chad basin in north-eastern Nigeria, and South Africa.

Following validation, an overall accuracy of 73 % was obtained. The precision, sensitivity, and specificity values were respectively 92 %, 59 %, and 92 %. Cohen's Kappa score of 0.48 was also obtained.

The total and intensive irrigated cropland area per country (Fig. 9) and total for Africa (Table 5) for the BPCP is presented along with the FAO Aquastat statistics on the irrigated area (only) and areas under AWM. The latter FAO data captures areas under flood recession agriculture and cultivated inland valley bottoms in addition to formal irrigated areas. Additionally, the irrigated areas for the remote-sensing products, WaPOR and Landsat-Derived Global Rainfed and Irrigated-Cropland Product (LGRIP), are also compared. The continental total irrigated area of BPCP (461,000 km^2), is larger than the FAO irrigated area (101,000 km^2), the area under AWM (196,000 km^2), and WaPOR continental irrigated cropland area (275,000 km^2). However, the continental total irrigated area of BPCP is comparable to that of LGRIP, which is 494,000 km^2. The area under formal irrigation for BPCP in Africa is comparable to the WaPOR irrigated cropland area. Considering the irrigated area per country, Egypt is the only country where all the

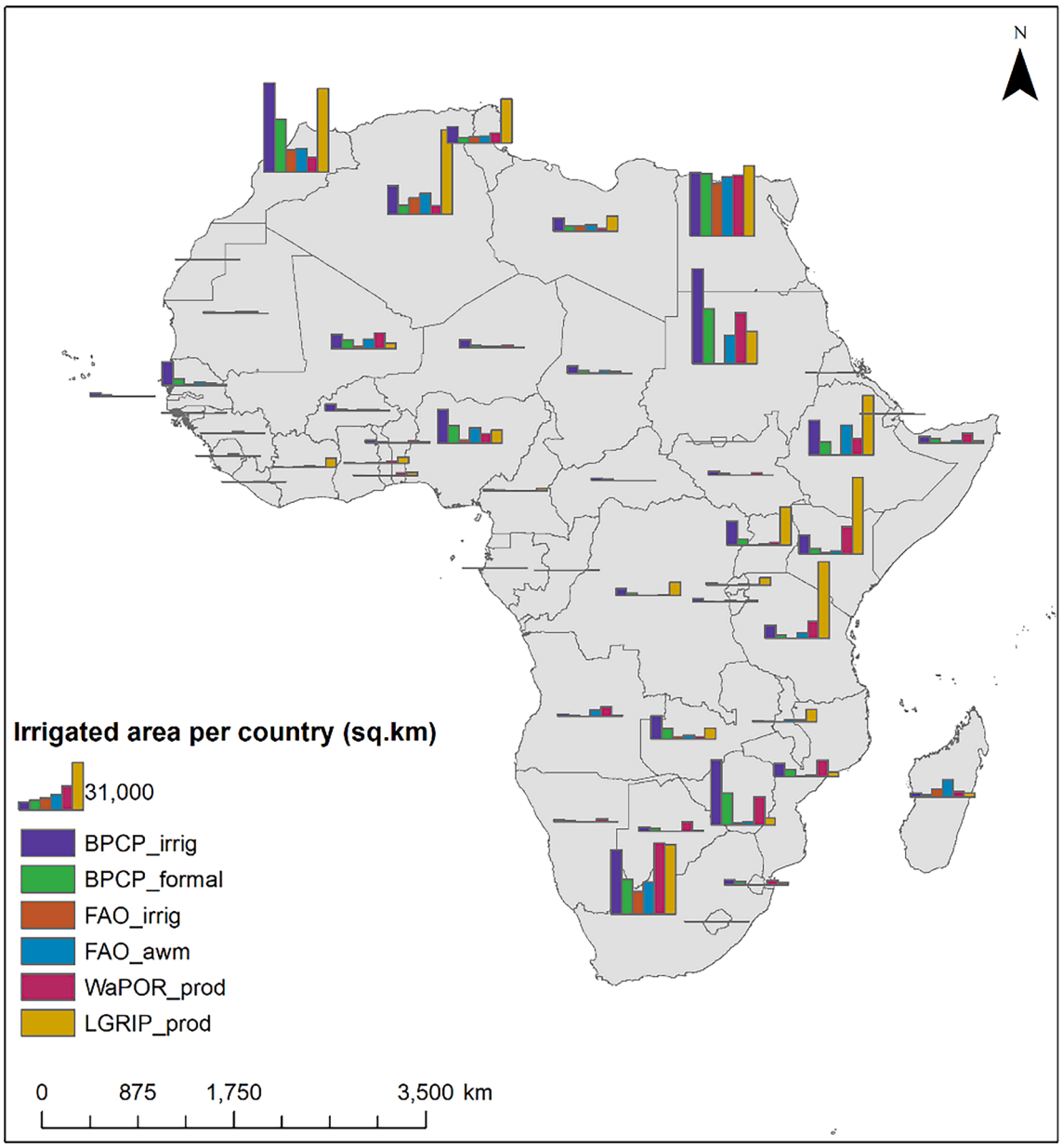


Fig. 9. Total and formal irrigated cropland area per country for Budyko-partitioned cropland product (BPCP) compared to FAO's Aquastat data on actual irrigated area and area under agricultural water management, as well as an irrigated area for WaPOR and LGRIP landcover products. (Notation in legend. irrig- total irrigated area; awm- agricultural water management; prod- product).

Table 5
Total irrigated cropland area for Africa for Budyko-partitioned cropland product (BPCP) compared to FAO's Aquastat data on actual irrigated area and area under agricultural water management, as well as an irrigated area for WaPOR and LGRIP landcover products.

Product	Total (km^2)
BPCP- total irrigated area	460,842
BPCP- formal irrigation	239,308
FAO- irrigated area	101,173
FAO- AWM	196,132
WaPOR product	275,108
LGRIP product	493,683

products agree. For South Africa, the total BPCP irrigated area is similar to both WaPOR and LGRIP.

3.4. *Partitioning individual crop masks using Budyko*

The Budyko model was used to partition the crop masks from the original eight LULC products in this study (Fig. 10). A range of 19 % to 30 % of the total cropland area in each product was found to be irrigated, and approximately half of the irrigated area for each product was under formal irrigation (Table 6). After partitioning an undifferentiated WaPOR crop mask (i.e., with irrigated and rainfed areas combined into one (crop) class) using the Budyko model, the irrigated area found was 499,000 km^2 out of a cropland area of 2.49 million km^2. This is 1.8 times larger than the WaPOR remote-sensing product's irrigated area of 275,000 km^2.

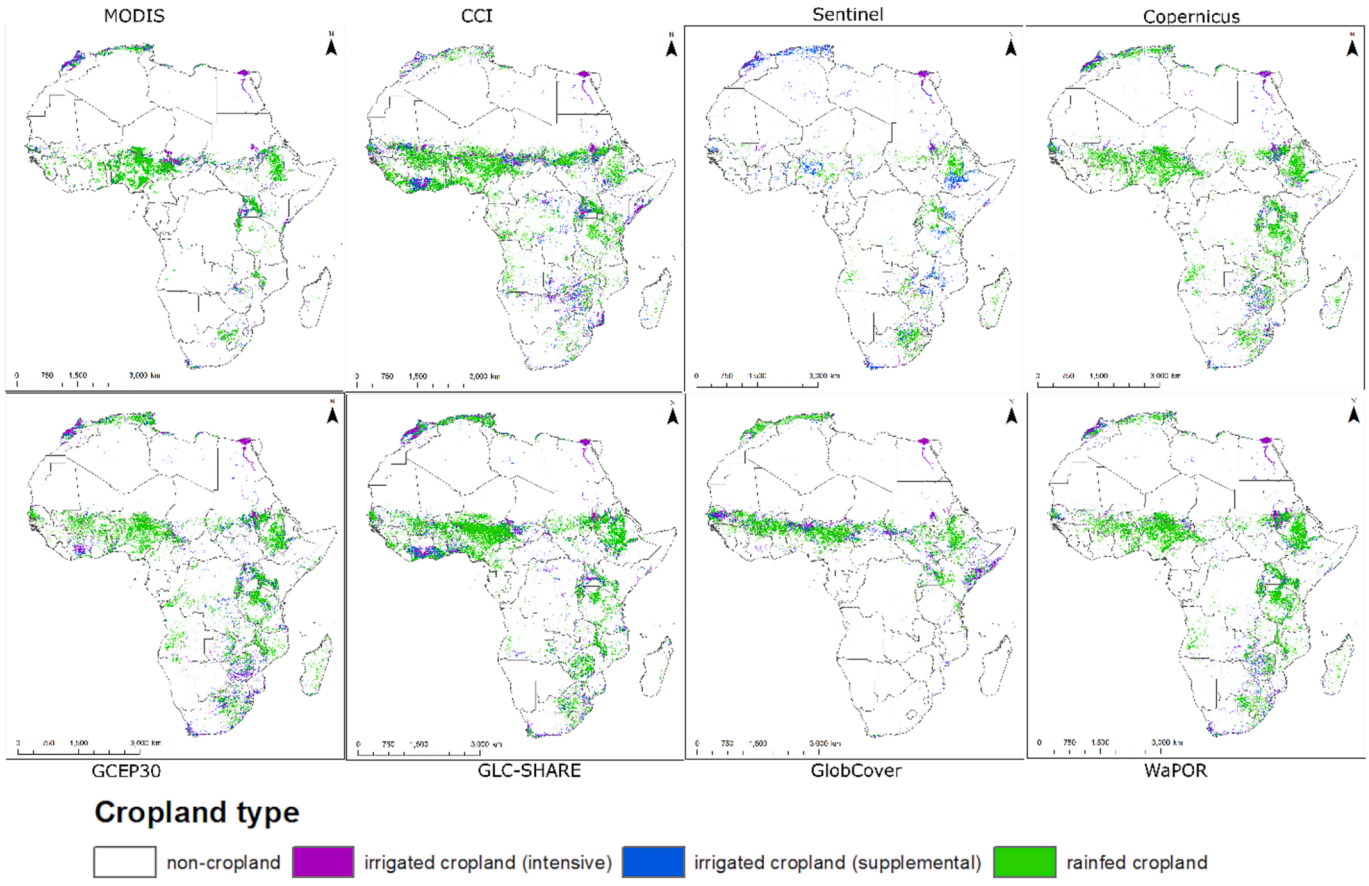


Fig. 10. Crop masks of eight landcover products for the reference year 2019 partitioned into rainfed and irrigated cropland using the Budyko model.

Table 6
Total and formal irrigated area of eight crop masks for Africa after partitioning using Budyko model.

	Product	Total cropland (km^2)	Budyko-Irrigated cropland (km^2)	Percentage Irrigated area (%)	Budyko- Formal irrigation (km^2)	Percentage Formal irrigation (%)
1	MODIS LC	1,613,328	416,071	25.8	233,915	14.5
2	CCI LC V2	3,165,210	893,880	28.2	472,640	14.9
3	Sentinel-2	1,052,578	314,770	29.9	188,897	17.9
4	Copernicus	2,337,732	453,483	19.4	232,033	9.9
5	GCEP30	2,689,838	725,249	30.0	378,860	14.0
6	GLC-SHARE	3,181,397	801,805	25.2	411,236	12.9
7	GlobCover	1,825,740	518,830	28.4	301,523	16.5
8	WaPOR	2,493,455	499,109	20.0	261,484	10.5

4. Discussion

The objectives of this study were to determine the level of concurrence between different LULC datasets, create a composite cropland map that addresses the challenges posed by the disparities found within individual crop masks from LULC products, and demonstrate a scalable approach for classifying croplands into two categories: irrigated and rainfed. On the first objective, a fair agreement is found between the eight preprocessed LULC products used in this study, given that the average similarity of each product compared to the other seven was at least 50 %. This fair agreement is primarily observed in barren areas, trees, and water classes. In contrast, there was a relatively low overlap in shrubs, grasses, and cropland classes. Specifically on cropland classification, Pérez-Hoyos et al. (2017) found that a full agreement between nine LULC products occurred in only 2 % of Africa, with the highest agreement in northern Morrocco, Algeria and Tunisia as well as the Nile Valley in Egypt. These areas of high cropland agreement coincide with 100 % of agreement between the top three crop masks (WaPOR, Copernicus and GLC-SHARE). Both studies, that of Pérez-Hoyos et al. (2017) and this present study, therefore align with the observation that areas of near-perfect agreement between cropland classes of LULC products occur in well-developed irrigated areas near large reservoirs (Thenkabail et al., 2010; Wei et al., 2020). Conversely, Pérez-Hoyos et al. (2017) found that areas of high cropland disagreement of up to 61.8 % occurred in Namibia and Angola as well as the Sahelian regions of Mauritania, northern Burkina Faso and Niger. These are areas where remote-sensing approaches fail to accurately discriminate between croplands and natural vegetation such as shrubs and grasses. This is evidenced in the high disagreement and overlap between these three LULC classes in this study (Fritz et al., 2010; Pittman et al., 2010).

The utility of different LULC products needs to be assessed for specific applications. However, considering the differences between individual LULC and cropland products, using multiple maps or composite maps based on the degree of agreement may overcome the uncertainty

associated with using one map for continental-scale studies. For example, in this study, the HCCM, which was derived based on the agreement between three existing products found to have high accuracy, was used for calibration and validation of the cropland partitioning framework. Even with this approach, legend translation of the original landcover products, especially for mosaic classes and the different spatial resolutions of the original datasets, introduces errors into the composite product (Fritz et al., 2010; Tchuenté et al., 2011). To illustrate, in this study, the LULC products were resampled from resolutions ranging from 10 m (Sentinel) to 1000 m (GLC-SHARE) to a uniform resolution of 300 m and then reclassified into a condensed eight-class LUWA classification system. Such preprocessing involved subjective decision-making based on literature as well as the knowledge and judgment of the interpreters (Fritz et al., 2010; Pérez-Hoyos et al., 2017; Tchuenté et al., 2011). It was thus subject to errors and information loss stemming from a misinterpretation of original classes, propagation of the inherent errors and inconsistencies of the original LULC products, and oversimplification of landcover characteristics. In particular, the reclassification from an average of 17 classes in the eight LULC products to a simple 8-class LUWA classification, as opposed to the typically extended 80-class LUWA classification, while making it easier to compare the various LULC products, resulted in the relatively high average agreement between LULC products.

The continental irrigated cropland area of BPCP (461,000 km^2) is relatively high compared to the FAO statistics on irrigated areas (101,000 km^2) and areas under AWM (196,000 km^2). Furthermore, when the Budyko model was applied to partition an undifferentiated WaPOR cropland product, the Budyko model identified substantially more irrigated cropland areas (499,000 km^2) as compared to the actual WaPOR remote-sensing product (275,000 km^2). These findings align with that of Msigwa et al. (2021), who concluded that the Budyko model identified high blue ET even for rainfed croplands. However, it is likely that the Budyko model is identifying blue ET from informal, farmer-led, supplemental irrigation that is widespread across Africa (Adams, 1992; Balana et al., 2019; Beekman et al., 2014; Drechsel & Keraita, 2014; Woodhouse et al., 2016). As highlighted by Woodhouse et al. (2016), a "blind spot" in the narrative and statistics on irrigation in Africa is the failure to recognize that relatively minor adjustments made by farmers to supply water, for instance, to cope with dry spells in the wet season, qualify as irrigation. As such, between purely rainfed agriculture on the one hand and the large-scale, fully irrigated technological schemes typically held up as models on the other hand, there is a continuum of in-between agricultural water management that is well established across Africa (Adams, 1992; Beekman et al., 2014; Woodhouse et al., 2016). Examples of such partial irrigation systems may include water collection directly from streams to nearby farms using watering cans or motorized pumps in Ghana (Akpoti et al., 2022; Drechsel & Keraita, 2014); furrow irrigation through the diversion of streams in mountainous regions in southern and eastern Africa (Beekman et al., 2014); flood recession agriculture in the floodplains of seasonally flooded rivers in the upper Volta basin (Balana et al., 2019); and the use of shallow groundwater in valley bottoms all across Africa (Woodhouse et al., 2016). The FAO's statistics on AWM captures two of these informal irrigation types: cultivated inland valley bottoms and flood recession agriculture. Still, this official data reported by countries to the FAO overlooks other informal supplemental irrigation.

The original, non-parametric Budyko model is a simple, lumped model for water-energy balance studies (Gentine et al., 2012; Gerrits et al., 2009; Greve et al., 2015; Mianabadi et al., 2020; Padrón et al., 2017; Reaver et al., 2022; Sposito, 2017; Wang et al., 2016). This is so even under anthropogenically induced climate change because the co-evolution of landscape conditions and hydrological processes occurs with these changes (Mianabadi et al., 2020; Wang et al., 2016). As such, although the scatter of data points from the overall trend in the Budyko curve indicates the influence of environmental factors beyond climate, climate is seen as the "primary control" of long-term water partitioning as soils, plants, and land features are themselves connected to climate and 'co-evolve' with it (Chen & Sivapalan, 2020). This makes the Budyko model especially well-suited for continental-scale studies because the peculiar landscape characteristics of each location are accounted for in the model. In its application to partition croplands across Africa in this study, the high score of 92 % for both precision and specificity implies that the Budyko model has a low rate of false positives and is sufficiently accurate when identifying a cropland pixel as irrigated. However, there is often a trade-off between precision and sensitivity, evidenced by the relatively low sensitivity score of 59 %. As such, there is a higher rate of false negatives whereby some positive instances (irrigated croplands) are missed and incorrectly labeled as rainfed. This trade-off is considered acceptable for this study to ensure that the identification of irrigated pixels by the Budyko model is trustworthy and precise. Based on the scale established by Landis & Koch (1977) the Cohen's Kappa score of 0.48 in this study is considered moderate. It is worth highlighting that some LULC products, such as WaPOR landcover and the recently released LGRIP (Teluguntla et al., 2023) use remote-sensing techniques to identify irrigated and rainfed cropland. For example, with LGRIP, using the GCEP30 as a base map, machine learning models were used to classify global agroecological zones and then identify pixels in these classes using ground data, spectral information, and image interpretation (Teluguntla et al., 2023). The methodology presented in this study, however expands the suite of remote sensing products that can be used for agricultural water studies as it makes it possible to derive irrigated and cropland products for the many remote-sensing LULC that do not have disaggregated cropland classes but may be preferred for certain locations or studies.

Despite its good performance in this study, applying the Budyko model has limitations. For example, owing to the unfeasibility of $\phi = 0$ when $ET/ET_o \rightarrow 1$, Budyko cannot accurately represent the physical limit in the relation between ϕ and ET/ET_o (Carmona et al., 2016). The uncertainty is also associated with the estimation of model parameters such as the aridity index. This study used remote-sensing-based climatic inputs that also have their limitations across Africa, partly stemming from the unavailability of in situ measurements (Mekonnen et al., 2023; Weerasinghe et al., 2020). Furthermore, it is possible that despite a tenfold increase in the spatial resolution used in this continental application compared to Msigwa et al. (2021), from 30 m to 300 m, steady-state conditions are still not met. Studies indicate that a much larger spatial scale of $> 10{,}000$ km^2 would improve the effectiveness of the Budyko model (Mianabadi et al., 2020; Wang et al., 2016, Wang et al., 2022). However, the main limitation of this study is the spatial scale of input landcover datasets, which at 300 m exceeds the average farm size in Africa estimated to be less than 1 ha (Giller et al., 2021). Thus, using even larger spatial resolutions would result in further loss of detail and accuracy with respect to the applicability of the framework to croplands in Africa.

5. Conclusion

Understanding the distribution of rainfed and irrigated areas provides valuable insights for agricultural water management and food security. However, identifying and disaggregating croplands in Africa is complex due to its extensive geographical expanse, diverse climates, and agricultural practices. This study first evaluated the agreement among eight land use and landcover datasets for Africa and then developed a high-confidence cropland map (HCCM) to overcome the disparities within individual crop masks of the LULC products. Subsequently, the Budyko model was used to distinguish between rainfed and irrigated cropland areas for the HCCM and the eight remote-sensing landcover products, demonstrating a scalable framework for partitioning cropland.

The assessment of agreement between all landcover classes of the eight products showed a fair pairwise agreement of approximately 50 %, mainly driven by high agreement on the classification of areas as barren areas, trees, or water. Alternatively, there was high uncertainty and

overlap in areas classed as shrubs, grasses, and crops. The results of cropland partitioning of the HCCM using the Budyko model show that an area of 461,000 km^2, approximately 19 % of the total HCCM area of 2.36 million km^2, is irrigated. For the eight remote-sensing landcover products, the irrigated areas ranged from 19 % to 30 % of their respective cropland areas. For the HCCM and individual landcover products, approximately half of the irrigated areas were identified as cropland under formal irrigation, defined in this study as irrigated areas where blue evapotranspiration exceeds 100 mm. Compared to FAO's irrigated area and areas under agricultural water management, the area of the HCCM identified as irrigated by the Budyko model is relatively large. This is likely due to the Budyko model capturing blue ET from all 'excess' water applied to crops as part of informal, farmer-led, supplemental irrigation widespread across Africa but not well-captured in the formal data reported by countries to the FAO. Using an independent validation dataset, the overall accuracy of the Budyko model in partitioning the HCCM was 73 %. High precision and specificity scores confirm that the Budyko model effectively identifies irrigated areas while making relatively few errors in misidentifying rainfed areas as irrigated.

The Budyko model performs well in partioning between irrigated and rainfed areas in Africa, making it a well-suited model for continental-scale water-energy balance investigations. The cropland partitioning framework demonstrated in this study thus expands the range of landcover products that can be used for studying agricultural water management on a continental scale. Ultimately, this contributes to the development of sustainable and resilient agricultural systems in Africa by enhancing the understanding of irrigation and rainfed patterns.

CRediT authorship contribution statement

Afua Owusu: Methodology, Software, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Validation, Visualization. **Stefanie Kagone:** Software, Formal analysis, Investigation, Data curation, Writing – review & editing, Visualization. **Mansoor Leh:** Conceptualization, Methodology, Writing – review & editing. **Naga Manohar Velpuri:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Project administration. **Murali Krishna Gumma:** Validation. **Benjamin Ghansah:** Data curation. **Paranamana Thilina-Prabhath:** Software. **Komlavi Akpoti:** . **Kirubel Mekonnen:** Writing – review & editing. **Primrose Tinonetsana:** Writing – review & editing. **Ismail Mohammed:** Validation.

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.

Data availability

Data citation: Owusu, A., Kagone, S., Leh, M., and Velpuri, N.M., 2023, Rainfed and Irrigated Cropland Areas for Africa: U.S. Geological Survey data release, https://doi.org/10.5066/P9N4R7SF.

Acknowledgements

The authors would like to acknowledge the land use/landcover dataset providers and developers for producing and making their datasets free of charge. This study is part of the International Water Management Institute's Digital Innovations for Water Secure Africa (DIWASA) project. We gratefully acknowledge The Leona M. and Harry B. Helmsley charitable trust for their financial support. The author's views are their own and do not necessarily reflect the views of the funding agency but do represent the views of the U.S. Geological Survey. We thank two anonymous reviewers and the editor for providing valuable feedback to improve the quality of this publication. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the authors or the U.S. Government.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2023.103607.

References

Adams, W.M., 1992. Wasting the rain: rivers, people and planning in Africa. Routledge, New York.

Akpoti, K., Higginbottom, T.P., Foster, T., Adhikari, R., Zwart, S.J., 2022. Mapping land suitability for informal, small-scale irrigation development using spatial modelling and machine learning in the Upper East Region. Ghana. Sci. Total Environ. 803, 149959 https://doi.org/10.1016/J.SCITOTENV.2021.149959.

Allen, R., Pereira, L., Raes, D., Smith, M., 1998. Crop evapotranspiration - Guidelines for computing crop water requirements (No. Paper 56), Evapotranspiración del cultivo Guías para la determinación de los requerimientos de agua de los cultivos. ESTUDIO FAO RIEGO Y DRENAJE 56., FAO Irrigation and Drainage. Rome.

Balana, B.B., Sanfo, S., Barbier, B., Williams, T., Kolavalli, S., 2019. Assessment of flood recession agriculture for food security in Northern Ghana: An optimization modelling approach. Agr. Syst. 173, 536–543. https://doi.org/10.1016/J.AGSY.2019.03.021.

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., Van Dijk, A.I.J.M., McVicar, T.R., Adler, R.F., 2019. MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. Bull. Am. Meteorol. Soc. 100, 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1.

Beck, H.E., Wood, E.F., McVicar, T.R., Zambrano-Bigiarini, M., Alvarez-Garreton, C., Baez-Villanueva, O.M., Sheffield, J., Karger, D.N., 2020. Bias correction of global high-resolution precipitation climatologies using streamflow observations from 9372 catchments. J. Clim. 33, 1299–1315. https://doi.org/10.1175/JCLI-D-19-0332.1.

Beekman, W., Veldwisch, G.J., Bolding, A., 2014. Identifying the potential for irrigation development in Mozambique: Capitalizing on the drivers behind farmer-led irrigation expansion. Phys. Chem. Earth 76–78, 54–63. https://doi.org/10.1016/J.PCE.2014.10.002.

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., Bontemps, S., Leroy Marc, Achard, F., Herold, M., Ranera, F., Arino, O., 2008. GlobCover - Products Description and Validation Report, Available at: https://publications.jrc.ec.europa.eu/repository/handle/JRC49240, Last accessed on 28-09-2023.

Buchhorn, M., Smets, B., Bertels, L., Roo, B. De, Lesiv, M., Tsendbazar, N.-E., Li, L., Tarko, A., 2021. Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual. 10.5281/ZENODO.4723921.

Budyko, M., 1974. Climate and Life. Academic Press, New York, NY, USA.

Carmona, A.M., Poveda, G., Sivapalan, M., Vallejo-Bernal, S.M., Bustamante, E., 2016. A scaling approach to Budyko's framework and the complementary relationship of evapotranspiration in humid environments: Case study of the Amazon River basin. Hydrol. Earth Syst. Sci. 20, 589–603. https://doi.org/10.5194/HESS-20-589-2016.

Chandrasekharan, K.M., Subasinghe, C., Haileslassie, A., 2021. Mapping irrigated and rainfed agriculture in Ethiopia (2015-2016) using remote sensing methods. International Water Management Institute (IWMI), Colombo, Sri Lanka. 10.5337/2021.206.

Chen, X., Sivapalan, M., 2020. Hydrological Basis of the Budyko Curve: Data-Guided Exploration of the Mediating Role of Soil Moisture. Water Resour. Res. 56 https://doi.org/10.1029/2020WR028221.

Chukalla, A.D., Krol, M.S., Hoekstra, A.Y., 2015. Green and blue water footprint reduction in irrigated agriculture: effect of irrigation techniques, irrigation strategies and mulching. Hydrol. Earth Syst. Sci. 19, 4877–4891. https://doi.org/10.5194/hess-19-4877-2015.

Drechsel, P., Keraita, B., 2014. Irrigated urban vegetable production in Ghana: characteristics, benefits and risk mitigation. International Water Management Institute (IWMI)., Colombo. 10.5337/2014.219.

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

ESA CCI, 2015. Land Cover CCI, Product User Guide V2, Available at: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf, Last accessed: 28-09-2023 .

Falkenmark, M., Rockström, J., 2006. The new blue and green water paradigm: Breaking new ground for water resources planning and management. J. Water Resour. Plan. Manag. 132, 129–132. https://doi.org/10.1061/(asce)0733-9496(2006)132:3(129).

FAO, 2018. Database Methodology: Level 1 Data Using Remote Sensing in Support of Solutions to Reduce Agricultural Water Productivity Gaps. Rome.

FAO, 2021. AQUASTAT - FAO's Global Information System on Water and Agriculture. accessed 5.4.23. https://www.fao.org/aquastat/en/.

Fisher, J.B., Whittaker, R.J., Malhi, Y., 2011. ET come home: Potential evapotranspiration in geographical ecology. Glob. Ecol. Biogeogr. 20, 1–18. https://doi.org/10.1111/J.1466-8238.2010.00578.X.

Fox, P., Rockström, J., 2003. Supplemental irrigation for dry-spell mitigation of rainfed agriculture in the Sahel. Agric Water Manag 61, 29–50. https://doi.org/10.1016/S0378-3774(03)00008-8.

Friedl, M., Sulla-Menashe, D., 2019. MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006 . NASA EOSDIS Land Processes DAAC. URL 10.5067/MODIS/MCD12Q1.006 (accessed 1.25.23).
Fritz, S., See, L., Rembold, F., 2010. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. Int. J. Remote Sens. 31, 2237–2256. https://doi.org/10.1080/01431160902946598.
Fu, B., 1981. On the calculation of the evaporation from land surface. Chin. J. Atmos. Sci. 5, 23–31. https://doi.org/10.3878/J.ISSN.1006-9895.1981.01.03.
Gadédjisso-Tossou, A., Avellán, T., Schütze, N., 2018. Potential of Deficit and Supplemental Irrigation under Climate Variability in Northern Togo. West Africa. Water (basel) 10, 1803. https://doi.org/10.3390/W10121803.
Gentine, P., D'Odorico, P., Lintner, B.R., Sivandran, G., Salvucci, G., 2012. Interdependence of climate, soil, and vegetation as constrained by the Budyko curve. Geophys. Res. Lett. 39 https://doi.org/10.1029/2012GL053492.
Gerrits, A.M.J., Savenije, H.H.G., Veling, E.J.M., Pfister, L., 2009. Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model. Water Resour. Res. 45 https://doi.org/10.1029/2008WR007308.
Giller, K.E., Delaune, T., Silva, J.V., van Wijk, M., Hammond, J., Descheemaeker, K., van de Ven, G., Schut, A.G.T., Taulya, G., Chikowo, R., Andersson, J.A., 2021. Small farms and development in sub-Saharan Africa: Farming for food, for income or for lack of better options? Food Secur 13, 1431–1454. https://doi.org/10.1007/S12571-021-01209-0/FIGURES/7.
Greve, P., Burek, P., Wada, Y., 2020. Using the Budyko Framework for Calibrating a Global Hydrological Model. Water Resour Res 56, e2019WR026280. 10.1029/2019WR026280.
Greve, P., Gudmundsson, L., Orlowsky, B., Seneviratne, S.I., 2015. Introducing a probabilistic Budyko framework. Geophys. Res. Lett. 42, 2261–2269. https://doi.org/10.1002/2015GL063449.
Greve, P., Gudmundsson, L., Orlowsky, B., Seneviratne, S.I., 2016. A two-parameter Budyko function to represent conditions under which evapotranspiration exceeds precipitation. Hydrol. Earth Syst. Sci. 20, 2195–2205. https://doi.org/10.5194/HESS-20-2195-2016.
Gumma, M.K., Thenkabail, P.S., Hideto, F., Nelson, A., Dheeravath, V., Busia, D., Rala, A., 2011. Mapping Irrigated Areas of Ghana Using Fusion of 30 m and 250 m Resolution Remote-Sensing Data. Remote Sens. (Basel) 3, 816–835. https://doi.org/10.3390/RS3040816.
Han, S., Hu, H., Yang, D., Liu, Q., 2011. Irrigation impact on annual water balance of the oases in Tarim Basin, Northwest China. Hydrol. Process. 25, 167–174. https://doi.org/10.1002/HYP.7830.
Hoekstra, A.Y., Mekonnen, M.M., 2012. The Water Footprint of Humanity. Proceedings of the National Academy of Sciences 109, 3232–3237. https://doi.org/10.1073/pnas.1109936109.
Johansson, E.L., Fader, M., Seaquist, J.W., Nicholas, K.A., 2016. Green and blue water demand from large-scale land acquisitions in Africa. Proc. Natl. Acad. Sci. 113, 11471–11476. https://doi.org/10.1073/pnas.1524741113.
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, 2459–2472. https://doi.org/10.5194/hess-17-2459-2013.
Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J.C., Mathis, M., Brumby, S.P., 2021. Global Land Use/Land Cover with Sentinel 2 and Deep Learning. International Geoscience and Remote Sensing Symposium (IGARSS) 2021-July, 4704–4707. 10.1109/IGARSS47720.2021.9553499.
Kiptala, J.K., 2016. Managing Basin Interdependencies in a Heterogeneous, Highly Utilized and Data Scarce River Basin in Semi-Arid Africa: The case of the Pangani River Basin, Eastern Africa (PhD Thesis). Delft Unciversity of Technology, Delft. 10.4233/UUID:156DA0D0-086A-4A5C-B944-C7FE921E11C7.
Landis, J.R., Koch, G.G., 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics 33, 159–174. https://doi.org/10.2307/2529310.
Latham, J.S., Cumani, R., Rosati, I., Bloise, M., 2014. 2014 Global Land Cover SHARE Global Land Cover SHARE. (GLC-SHARE) database Beta-Release Version 1.0-2014, Rome.
Li, X., Troy, T.J., 2018. Changes in rainfed and irrigated crop yield response to climate in the western US. Environ. Res. Lett. 13, 064031 https://doi.org/10.1088/1748-9326/aac4b1.
McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 22, 276. 10.11613/bm.2012.031.
Mekonnen, M.M., Hoekstra, A.Y., 2011. The green, blue and grey water footprint of crops and derived crop products. Hydrol. Earth Syst. Sci. 15, 1577–1600. https://doi.org/10.5194/HESS-15-1577-2011.
Mekonnen, K., Velpuri, N.M., Leh, M., Akpoti, K., Owusu, A., Tinonetsana, P., Hamouda, T., Ghansah, B., Thilina-Prabhath, P., Munzimi, Y., 2023. Accuracy of satellite and reanalysis rainfall estimates over Africa: A multi-scale assessment of eight products for continental applications. J. Hydrol.: Reg. Stud. 49, 101514 https://doi.org/10.1016/J.EJRH.2023.101514.
Mianabadi, A., Davary, K., Pourreza-Bilondi, M., Coenders-Gerrits, A.M.J., 2020. Budyko framework; towards non-steady state conditions. J Hydrol (amst) 588, 125089. https://doi.org/10.1016/J.JHYDROL.2020.125089.
Milly, P.C.D., 1993. An analytic solution of the stochastic storage problem applicable to soil water. Water Resour. Res. 29, 3755–3758. https://doi.org/10.1029/93WR01934.
Monteith, J.L., 1965. Evaporation and environment. Symp. Soc. Exp. Biol. 19, 205–234.
Motsumi, S., Magole, L., Kgathi, D., 2012. Indigenous knowledge and land use policy: Implications for livelihoods of flood recession farming communities in the Okavango Delta, Botswana. Phys. Chem. Earth 50–52, 185–195. https://doi.org/10.1016/J.PCE.2012.09.013.
Msigwa, A., Komakech, H.C., Salvadore, E., Seyoum, S., Mul, M.L., van Griensven, A., 2021. Comparison of blue and green water fluxes for different land use classes in a semi-arid cultivated catchment using remote sensing. J. Hydrol.: Reg. Stud. 36 https://doi.org/10.1016/J.EJRH.2021.100860.
Padrón, R.S., Gudmundsson, L., Greve, P., Seneviratne, S.I., 2017. Large-Scale Controls of the Surface Water Balance Over Land: Insights From a Systematic Review and Meta-Analysis. Water Resour. Res. 53, 9659–9678. https://doi.org/10.1002/2017WR021215.
Patle, P., Singh, P.K., Ahmad, I., Matsuno, Y., Leh, M., Ghosh, S., 2023. Spatio-temporal estimation of green and blue water consumptions and water and land productivity using satellite remote sensing datasets and WA+ framework: A case study of the Mahi Basin. India. Agric Water Manag 277, 108097. https://doi.org/10.1016/J.AGWAT.2022.108097.
Pérez-Hoyos, A., Rembold, F., Kerdiles, H., Gallego, J., 2017. Comparison of Global Land Cover Datasets for Cropland Monitoring. Remote Sens. (Basel) 9, 1118. https://doi.org/10.3390/RS9111118.
Pittman, K., Hansen, M.C., Becker-Reshef, I., Potapov, P.V., Justice, C.O., 2010. Estimating Global Cropland Extent with Multi-year MODIS Data. Remote Sens. (Basel) 2, 1844–1863. https://doi.org/10.3390/RS2071844.
Porporato, A., Daly, E., Rodriguez-Iturbe, I., 2004. Soil water balance and ecosystem response to climate change. Am. Nat. 164, 625–632. https://doi.org/10.1086/424970.
Reaver, N.G.F., Kaplan, D.A., Klammler, H., Jawitz, J.W., 2022. Theoretical and empirical evidence against the Budyko catchment trajectory conjecture. Hydrol. Earth Syst. Sci. 26, 1507–1525. https://doi.org/10.5194/hess-26-1507-2022.
Rodell, M., Beaudoing, H.K., L'Ecuyer, T.S., Olson, W.S., Famiglietti, J.S., Houser, P.R., Adler, R., Bosilovich, M.G., Clayson, C.A., Chambers, D., Clark, E., Fetzer, E.J., Gao, X., Gu, G., Hilburn, K., Huffman, G.J., Lettenmaier, D.P., Liu, W.T., Robertson, F.R., Schlosser, C.A., Sheffield, J., Wood, E.F., 2015. The Observed State of the Water Cycle in the Early Twenty-First Century. J. Clim. 28, 8289–8318. https://doi.org/10.1175/JCLI-D-14-00555.1.
Sankarasubramanian, A., Vogel, R.M., 2002. Annual hydroclimatology of the United States. Water Resour. Res. 38, 19–111. https://doi.org/10.1029/2001WR000619.
Senay, G.B., Friedrichs, M., Singh, R.K., Velpuri, N.M., 2016. Evaluating Landsat 8 evapotranspiration for water use mapping in the Colorado River Basin. Remote Sens. Environ. 185, 171–185. https://doi.org/10.1016/j.rse.2015.12.043.
Senay, G.B., Kagone, S., Velpuri, N.M., 2020. Operational Global Actual Evapotranspiration: Development, Evaluation, and Dissemination. Sensors 20, 1915. https://doi.org/10.3390/s20071915.
Simons, G., Bastiaanssen, W., Ngô, L.A., Hain, C.R., Anderson, M., Senay, G., 2016. Integrating Global Satellite-Derived Data Products as a Pre-Analysis for Hydrological Modelling Studies: A Case Study for the Red River Basin. Remote Sens. (Basel) 8, 279. https://doi.org/10.3390/RS8040279.
Singh, V., Singh, P.K., Jain, S.K., Jain, S.K., Cudennec, C., Hessels, T., 2022. Examining evaporative demand and water availability in recent past for sustainable agricultural water management in India at sub-basin scale. J. Clean. Prod. 346, 130993 https://doi.org/10.1016/j.jclepro.2022.130993.
Sposito, G., 2017. Understanding the Budyko Equation. Water 2017, Vol. 9, Page 236 9, 236. 10.3390/W9040236.
Tchuenté, A.T.K., Roujean, J.L., de Jong, S.M., 2011. Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. Int. J. Appl. Earth Obs. Geoinf. 13, 207–219. https://doi.org/10.1016/J.JAG.2010.11.005.
Teluguntla, P.R., Thenkabail, P., Oliphant, A., Gumma, M., Aneece, I., Foley, D., McCormick, R., 2023. Landsat-Derived Global Rainfed and Irrigated-Cropland Product 30 m V001 . NASA EOSDIS Land Processes DAAC [WWW Document].
Thenkabail, P.S., Hanjra, M.A., Dheeravath, V., Gumma, M., 2010. A Holistic View of Global Croplands and Their Water Use for Ensuring Global Food Security in the 21st Century through Advanced Remote Sensing and Non-remote Sensing Approaches. Remote Sensing 2010, Vol. 2, Pages 211-261 2, 211–261. 10.3390/RS2010211.
Thenkabail, P.S., Teluguntla, P.G., Xiong, J., Oliphant, A., Congalton, R.G., Ozdogan, M., Gumma, M.K., Tilton, J.C., Giri, C., Milesi, C., Phalke, A., Massey, R., Yadav, K., Sankey, T., Zhong, Y., Aneece, I., Foley, D., 2021. Global cropland-extent product at 30-m resolution (GCEP30) derived from Landsat satellite time-series data for the year 2015 using multiple machine-learning algorithms on Google Earth Engine cloud: U.S. Geological Survey Professional Paper 1868. US Geological Survey, Reston, Virginia. 10.3133/PP1868.
Tsendbazar, N.-E., Tarko, A., Li, L., Herold, M., Lesiv, M., Fritz, S., Maus, V., 2020. Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Validation Report. 10.5281/ZENODO.3938974.
van Eekelen, M.W., Bastiaanssen, W.G.M., Jarmain, C., Jackson, B., Ferreira, F., van der Zaag, P., Saraiva Okello, A., Bosch, J., Dye, P., Bastidas-Obando, E., Dost, R.J.J., Luxemburg, W.M.J., 2015. A novel approach to estimate direct and indirect water withdrawals from satellite measurements: A case study from the Incomati basin. Agr Ecosyst Environ 200, 126–142. https://doi.org/10.1016/J.AGEE.2014.10.023.
Velpuri, N.M., Senay, G.B., 2017. Partitioning Evapotranspiration into Green and Blue Water Sources in the Conterminous United States. Sci. Rep. 7, 1–12. https://doi.org/10.1038/s41598-017-06359-w.
Wang, D., Tang, Y., 2014. A one-parameter Budyko model for water balance captures emergent behavior in darwinian hydrologic models. Geophys. Res. Lett. 41, 4569–4577. https://doi.org/10.1002/2014GL060509.
Wang, C., Wang, S., Fu, B., Zhang, L., 2016. Advances in hydrological modelling with the Budyko framework: A review. Prog. Phys. Geogr. 40, 409–430. https://doi.org/10.1177/0309133315620997/ASSET/IMAGES/LARGE/10.1177_0309133315620997-FIG1.JPEG.

Wang, F., Xia, J., Zou, L., Zhan, C., Liang, W., 2022. Estimation of time-varying parameter in Budyko framework using long short-term memory network over the Loess Plateau. China. J Hydrol (amst) 607, 127571. https://doi.org/10.1016/J.JHYDROL.2022.127571.

Weerasinghe, I., Van Griensven, A., Bastiaanssen, W., Mul, M., Jia, L., 2020. Can we trust remote sensing ET products over Africa? Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-2019-233.

Wei, Y., Lu, M., Wu, W., Ru, Y., 2020. Multiple factors influence the consistency of cropland datasets in Africa. Int J Appl Earth Obs Geoinformation 89. https://doi.org/10.1016/J.JAG.2020.102087.

Woodhouse, P., Veldwisch, G.J., Venot, J.P., Brockington, D., Komakech, H., Manjichi, Â., 2016. African farmer-led irrigation development: re-framing agricultural policy and investment? J. Peasant Stud. 44, 213–233. https://doi.org/10.1080/03066150.2016.1219719.

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N.-E., Ramoino, F., Arino, O., 2021. ESA WorldCover 10 m 2020 v100. 10.5281/ZENODO.5571936.

Zhang, L., Dawes, W.R., Walker, G.R., 2001. Response of mean annual evapotranspiration to vegetation changes at catchment scale. Water Resour. Res. 37, 701–708. https://doi.org/10.1029/2000WR900325.

Zhang, L., Potter, N., Hickel, K., Zhang, Y., Shao, Q., 2008. Water balance modeling over variable time scales based on the Budyko framework – Model development and testing. J Hydrol (amst) 360, 117–131. https://doi.org/10.1016/J.JHYDROL.2008.07.021.