remote sensing

MDPI

Article

An Automated and Improved Methodology to Retrieve Long-time Series of Evapotranspiration Based on Remote Sensing and Reanalysis Data

Mojtaba Saboori [1,*], Yousef Mousivand [1], Jordi Cristóbal [2,3], Reza Shah-Hosseini [4] and Ali Mokhtari [5]

[1] Department of Geography, Kharazmi University, Tehran 1491115719, Iran
[2] Efficient Use of Water in Agriculture Program, Institute of Agrifood Research and Technology, Fruitcentre, Parc Científic i Tecnològic Agroalimentari de Lleida 23, 25003 Lleida, Spain
[3] Department of Geography, Autonomous University of Barcelona, 08193 Cerdanyola del Vallès, Spain
[4] School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran
[5] School of Life Sciences, Technical University of Munich, 85354 Freising, Germany
* Correspondence: mojtabasaboori1989@gmail.com

Abstract: The large-scale quantification of accurate evapotranspiration (ET) time series has substantially been developed in recent decades using automated approaches based on remote sensing data. However, there are still several model-related uncertainties that require precise assessment. In this study, the Surface Energy Balance Algorithm for Land (SEBAL) and meteorological data from the Global Land Data Assimilation System (GLDAS) were used to estimate long-term daily actual ET based on three endmember selection procedures: two land cover-based models, one with (WF) and the other without (WOF) morphological functions, and the Allen method (with the default percentiles) for 2270 Landsat images. Models were evaluated for 23 flux tower sites with four main vegetation cover types as well as different climate types. Results showed that endmember selection with morphological functions (WF_ET) generally performed better than the other endmember approaches. Climate-based classification assessment provided the clearest discrimination between the performance of the different endmember selection approaches for the humid category. For humid zones, the land cover-based methods, especially WF, appropriately outperformed Allen. However, the performance of the three approaches was similar for sub-humid, semi-arid and arid climates together; the Allen approach was therefore recommended to avoid the need for dependency on land cover maps. Tower-by-tower validation also showed that the WF approach performed best at 12 flux tower sites, the WOF approach best at 5 and the Allen approach best at 6, suggesting that the use of land cover maps alone does not explain the differences between the performance of the land cover-based models and the Allen approach. Additionally, the satisfactory error metrics results when comparing the EC estimations with EC measurements, with root mean square error (RMSE) $\approx$ 0.91 and 1.59 mm·day^{-1}, coefficient of determination (R^2) $\approx$ 0.71 and 0.41, and bias percentage (PBias) $\approx$ 2% and 60% for crop and non-crop flux tower sites, respectively, supports the use of GLDAS meteorological forcing datasets with the different automated ET estimation approaches. Overall, given that the thorough evaluation of different endmember selection approaches at large scale confirmed the validity of the WF approach for different climate and land cover types, this study can be considered an important contribution to the global retrieval of long time series of ET.

Keywords: evapotranspiration; SEBAL; endmember selection; Landsat; GLDAS; time series

Citation: Saboori, M.; Mousivand, Y.; Cristóbal, J.; Shah-Hosseini, R.; Mokhtari, A. An Automated and Improved Methodology to Retrieve Long-time Series of Evapotranspiration Based on Remote Sensing and Reanalysis Data. *Remote Sens.* **2022**, *14*, 6253. https://doi.org/10.3390/rs14246253

Academic Editor: Nicola Montaldo

Received: 23 October 2022
Accepted: 7 December 2022
Published: 9 December 2022

1. Introduction

In the coming decades, world food production will have to coexist with water's growing scarcity, extreme climatic events such as droughts or heatwaves [1], the limited availability of arable land and unpredictable costs, among many other considerations,

while still meeting the needs of a world population that is expected to have increased around 70% by 2050 [2]. Good and effective agricultural water management is therefore essential. In this regard, evapotranspiration (ET) is a fundamental variable. It is one of the key parameters in the water balance equation, making it an indispensable factor in the monitoring of crop water status and the optimization of irrigation scheduling. Currently, satellite remote sensing offers a unique perspective for operational ET monitoring that complements the in situ-based climate and hydrologic data and tools that are traditionally used, including for example the eddy covariance (EC), the Bowen ratio (BR) and lysimeters. According to the literature, satellite-based actual ET estimation can be employed to manage farm water and schedule irrigation [3], map biomass [4], detect water stress [5], calculate soil moisture [6], and study drought issues [7]. Thus, remote sensing helps specialists accelerate ET estimation at regional scales in a feasible and economic way.

The surface energy balance (SEB) approach, presented by [8], was the pioneer of remote-sensing-based ET modeling. Later, a surface energy balance algorithm for land (SEBAL) procedure was developed by [9] which expanded the one-source ET estimation methods. This method uses an iterative mechanism to calculate the sensible heat flux (H) equation based on the accurate selection of upper and lower boundaries, which are known as endmember pixels (i.e., hot (dry) and cold (wet) pixels), for ET rate estimation. Another SEB approach, the surface energy balance system (SEBS) model, was also developed by [10] for satellite-based ET estimation. This method is a one-source model that uses canopy cover (CC) for ET estimation. The endmember pixels extracted through land surface temperature (LST) and CC thresholds are used to calculate H for each pixel. However, the endmembers are still manually selected. Later, the METRIC (Mapping EvapoTranspiration at high Resolution with Internalized Calibration) approach was developed by [11], which still required endmember pixel selection. Other methods to estimate ET, including the LST-VI triangle [12], the simple surface energy balance index, S-SEBI [3,13], the modified SEBAL, M-SEBAL [14], the simplified surface energy balance for operational applications, SSEBop [15], and SEBAL advection, SEBAL-A [16] have also been developed.

Recently, the METRIC, SEBAL, and SEBS methods have been improved, yielding better results than the other one-source ET models [5,17,18]. However, they show different sensitivity to meteorological data. METRIC, for instance, requires high-quality hourly and daily meteorological data [19–22], which has a considerable impact on the reference ET (ETr) estimated by the Penman–Monteith equation [23]. For its part, SEBAL is less susceptible to meteorological inputs and more sensitive to LST and the near-surface temperature gradient, dT [24,25]. Although some study areas such as Europe and North America are equipped with highly developed meteorological networks, several geographic regions are meteorologically data-limited [26,27], hindering the ability to accurately estimate ET. Moreover, Blankenau et al. [28] reported that the spatial resolution of meteorological data is not a primary factor determining correspondence to station or satellite data. Recently, some studies have suggested using reanalysis meteorological datasets such as the ERA5-Land hourly dataset of the European Centre for Medium-Range Weather Forecasts (ECMWF) [29,30], the National Oceanic and Atmospheric Administration's (NOAA) climate forecast system version 2 (CFSv2) dataset [31], the Modern-Era Reanalysis for Research and Applications version 2 (MERRA-2) dataset [24], or the MERRA-1 dataset [32] as alternatives to ground-sourced meteorological data due to their high potential for ET estimation. However, the capability of the Global Land Data Assimilation System (GLDAS) as a meteorological input for SEB methods remains uncertain.

The accuracy of satellite-based one-source ET algorithms depends on cold and hot pixel selection conducted under expert supervision [33]. However, the process is severely time consuming, especially for a long time series of remote sensing data. As the manual selection of hot and cold pixels can lead to human errors limiting the performance of SEBAL, METRIC, and SEBS algorithms [34], several studies have tried to automate this process. Allen et al. [35] introduced a simple automated endmember pixel selection method yielding acceptable accuracy, although the final endmembers were visually extracted. Olmedo

et al. [36] developed the "Water" package based on the "CalcAnchor" function to automatically select cold and hot pixels. In this package, 5% of the upper and lower ranges of the LST images were regarded, respectively, as hot and cold pixels. In addition, the normalized difference vegetation index (NDVI), the leaf area index (LAI), the surface roughness length for momentum (Z_{om}), and the albedo were masked with exclusive thresholds for cold and hot pixel selection. Nonetheless, the model may lead to uncertainties as LAI and Z_{om} measurements are empirically estimated and dependent on ground data for model calibration. Bhattarai et al. [5] also developed an exhaustive automated model to select endmember pixels, which was mostly assessed in humid climate zones and exclusively using land cover maps. Ref. [37] utilized SSEBop calibrated for croplands using the National Land Cover Database (NLCD) and TopoWx gridded meteorological datasets that are limited to the conterminous United States. A supervised classification method was used by [38] to map vegetation and bare soil for endmember pixel selection. However, this method was also dependent on user supervision. Ref. [39] also applied the "Water" package, including the adjacency of endmember pixels, to meteorological data, Z_{om} and LAI for hot and cold pixel selection. Nonetheless, the entire area consisted of short crops with a height of less than one meter that was incompatible with trees, requiring local adjustment for regional applications. Ref. [40] implemented a simple method using vegetation and topographic index thresholding as well as endmember distance from flux towers, without the need for land cover maps. Although the ET estimates were satisfactory, the study was mostly conducted for crop flux tower sites. However, while several automated ET calibration methods have been provided to date, the validation and comparison of their long-term performance for different climates, land covers and temporal analyses remain challenging.

Although SEB methods have been shown to be accurate enough to be used for agricultural and water management practices, the need for background knowledge in the physics of radiation can hinder their applicability for end users [41], and an abundant number of layers of satellite images, land cover maps, terrain and climate data from various sources and platforms is required. Additionally, a large amount of layer pre-processing before running the models, as well as endmember selection for ET estimation, make the process time consuming, labor intensive and, thus, inefficient, especially for monitoring long time series of ET [42,43]. However, the cloud computing service Google Earth Engine (GEE) has addressed these limitations, and several automated ET mapping applications using one-source energy balance methods have now been developed [30,44].

Most of the above-mentioned studies were optimized for a particular land cover, climate, region, and endmember selection approach. Thus, specific percentiles need to be defined for different climatological and land cover conditions to obtain accurate ET estimations, which is really challenging and currently difficult to apply worldwide. It is therefore crucial to make the ET estimation process as independent as possible of endmember pixel selection calibration for regional applications. Meteorologically, data-limited conditions can hinder ET estimations in less developed countries in, for example, African and Asian regions currently facing water scarcity issues and recurring droughts and where the need for accurate ET estimates is critical for irrigation planning and water management. Finally, providing an automated method for GEE-based ET estimation for end users is essential to help mitigate the effects of extreme events and to improve water and agricultural management. With the above in mind, the present study aims to improve ET estimates using SEBAL in the GEE environment to:

- Provide and evaluate a novel endmember selection method;
- Compare the performance of different endmember selection methods under a wide range of land cover and climate conditions avoiding local calibration;
- Evaluate the validity of GLDAS reanalysis meteorological data in ET estimation.

2. Material and Methods

2.1. Study Area and Dataset

2.1.1. Climate Type Classification

This study used two systems of climate zone classification, the Köppen system [41,45] and the global aridity index [41,42,46], to precisely assess the performance of different automated ET estimation procedures for various climate conditions. The Köppen system classifies climate zones into five main categories (tropical, arid, temperate, continental and polar) with 30 subclasses based on seasonal precipitation and temperature. The global aridity index uses mean annual precipitation and temperature and the reference evapotranspiration (ET_0) obtained from the Hargreaves equation [43,47] to provide five climate classes (hyper-arid, arid, semi-arid, dry sub-humid and humid). The global aridity index was produced by [44,48] on the basis of WorldClim2 global climate data with around 1 km spatial resolution by dividing mean annual precipitation by mean annual ET_0 (accessible through https://cgiarcsi.community/v6, accessed on 23 March 2022). The Köppen classification data can also be downloaded from koeppen-geiger.vu-wien.ac.at (accessed on 23 March 2022). Note that the benefit of the global aridity index compared to the Köppen system is its use of ET in the classification.

2.1.2. Flux Tower Sites

A total of 23 flux tower stations in different regions with various land covers, climate and topographic conditions (Table 1 and Figure 1) and a total of 2270 Landsat images were used for model evaluation. The study areas covered nine main land cover types: cropland (CRO), short vegetation (open shrublands (OSH) and grasslands (GRA)), forest (evergreen needleleaf forests (ENF), deciduous broadleaf forests (DBF), mixed forests (MF), savanna (SAV), and woody savanna (WSA)), and wetland (WET), with a wide range of altitudes from −5 to 3030 m above sea level. The study also considered a wide range of climate conditions: humid subtropical (Cfa: mild with no dry season, hot summer), humid continental (Dfa: humid with severe winter, no dry season, hot summer), Mediterranean (Csa: mild with dry, hot summer), steppe (Bsk: warm winter, Bsh: very cold winter), marine west coast (Cfb: mild with no dry season, warm summer), warm summer continental (Dfb: significant precipitation in all seasons), and subarctic (Dfc: severe winter, cool summer), based on the Köppen classification and the global aridity index. Half-hourly or hourly EC data were obtained from the Ameriflux database (https://ameriflux.lbl.gov/, accessed on 23 March 2022) for ET evaluation. It should be noted that only high-quality EC data for each climate were selected. A total of 3 and 7 flux tower sites were selected for arid and semi-arid climates, respectively. Likewise, 9 and 4 flux tower sites were chosen for humid and sub-humid climates, respectively. Thus, 10 towers corresponded to dry and 13 to wet climatological conditions Finally, the Penman–Monteith equation [23] was used to compute ETr, which was used for instantaneous upscaling to daily ET using air temperature, relative humidity, solar radiation and wind speed input data from the flux towers.

2.1.3. Remote Sensing, Land Cover and Meteorological Inputs

The surface reflectance and the brightness temperature of a total of 2270 scenes of Landsat 5, 7, and 8 datasets, corresponding to Thematic Mapper™, Enhanced Thematic Mapper Plus (ETM+), Operational Land Imager (OLI), and Thermal Infrared Sensor (TIRS), were used to estimate ET using SEBAL. The albedo, NDVI, and the soil adjusted vegetation index (SAVI) were estimated from multispectral level-2 calibrated surface reflectance bands, while LST was retrieved from thermal infrared bands. A threshold of 10% cloud cover was considered to select appropriate imagery. Cloud-free pixels were extracted based on the CFMask method [49–51]. The scan line corrector (SLC) failure pixels on Landsat-7 ETM+ were not considered when falling on the flux tower station.

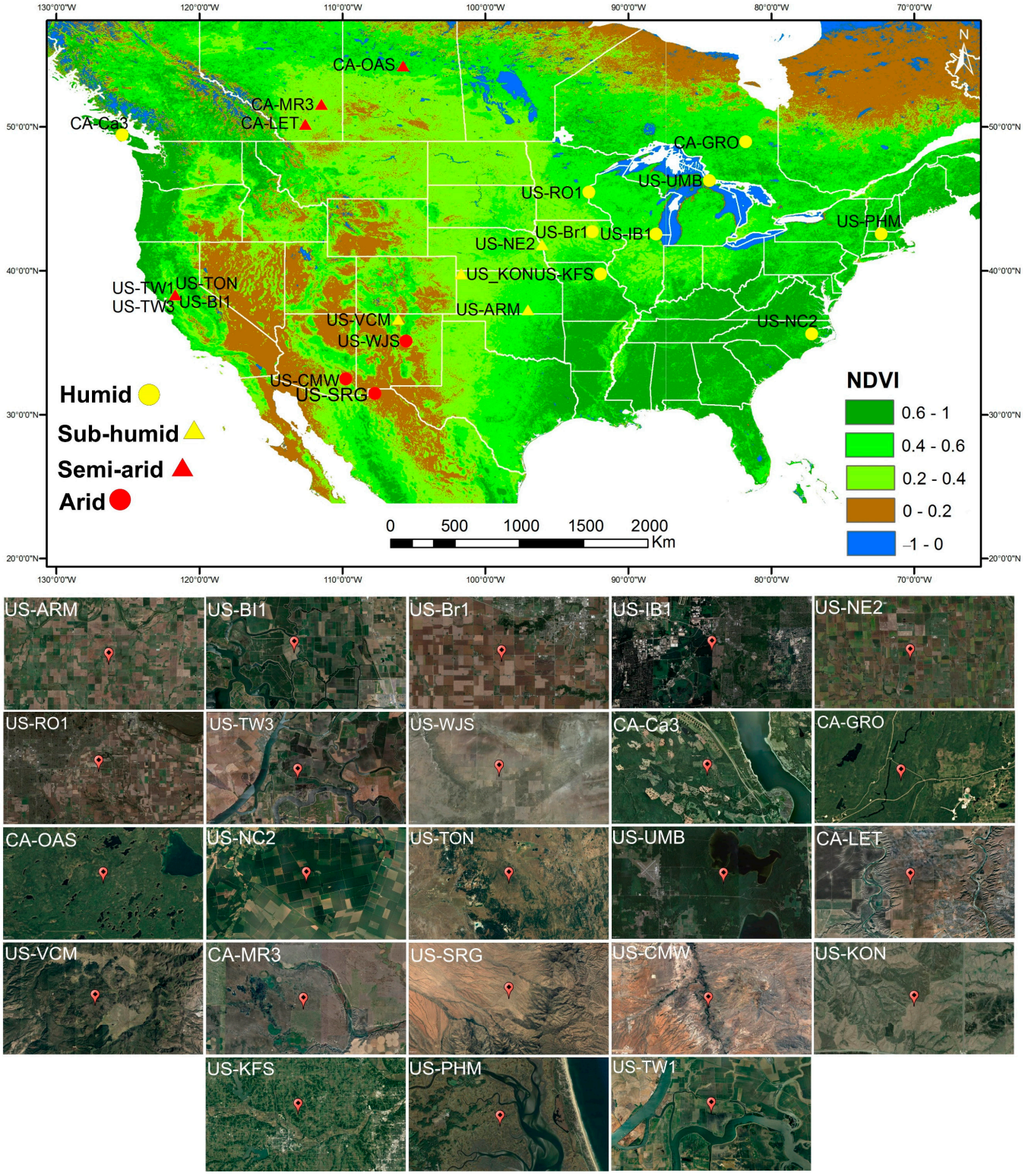

Figure 1. Location and climate types of flux tower sites used in this study.

Table 1. Data of study areas.

Station	Land Cover	Climate Köppen/Aridity Index	Average Temperature (°C)	Annual Precipitation (mm)	Latitude	Longitude	Elevation	Time Interval	Number of Data
US-ARM	CRO	Cfa/Sub-humid	14.76	843	36.6058	−97.4888	314	2006-19	141
US-BI1	CRO	Csa/Semi-arid	16	338	38.0992	−121.502	−2.7	2016-21	134
US-Br1	CRO	Dfa/Humid	8.95	842.33	41.9749	−93.6906	313	2005-11	90
US-IB1	CRO	Dfa/Humid	9.18	929.23	41.8593	−88.2227	226.5	2005-18	103
US-NE2	CRO	Dfa/Sub-humid	10.08	788.89	41.1649	−96.4701	362	2006-19	87
US-RO1	CRO	Dfa/Humid	6.4	879	44.7143	−93.0898	290	2004-17	115
US-TW3	CRO	Csa/Semi-arid	15.6	421	38.1152	−121.647	−4	2013-18	92
US-WJS	SAV	Bsk/Arid	15.2	361	34.4255	−105.862	1931	2007-19	175
CA-Ca3	ENF	Cfb/Humid	9.94	1676	49.5346	−124.9	164	2012-18	71
CA-GRO	MF	Dfb/Humid	1.3	831	48.2167	−82.1556	340	2004-12	49
CA-OAS	DBF	Dfc/Semi-arid	0.34	428.53	53.6289	−106.198	530	2002-10	74
US-NC2	ENF	Cfa/Humid	16.6	1320	35.803	−76.6685	5	2005-20	90
US-TON	WSA	Csa/Semi-arid	15.8	559	38.4309	−120.966	177	2008-18	182
US-UMB	DBF	Dfb/Humid	5.83	803	45.5598	−84.7138	234	2007-18	72
US-VCM	ENF	Dfb/Sub-humid	6.4	646	35.8884	−106.5321	3030	2014-20	109
US-CMW	DBF	Bsh/Arid	17	288	31.6637	−110.1777	1199	2014-18	97
CA-LET	GRA	Dfb/Semi-arid	5.36	398.4	49.7093	−112.94	960	2000-7	124
CA-MR3	GRA	Dfc/Semi-arid	-	325	50.8671	−111.905	712	2012-16	30
US-SRG	GRA	Bsk/Arid	17	420	31.7894	−110.828	1291	2008-18	139
US-KON	GRA	Cfa/Sub-humid	12.77	867	39.0824	−96.5603	417	2014-18	43
US-KFS	GRA	Cfa/Humid	12	1014	39.0561	−95.1907	310	2013-18	34
US-PHM	WET	Dfa/Humid	7.2	1262	42.7423	−70.8301	1262	2013-20	44
US-TW1	WET	Csa/Semi-arid	15.5	421	38.1074	−121.647	−5	2014-20	175

The GLDAS reanalysis data, a combination of in situ observations and satellite information, developed by data assimilation algorithms, was used as SEBAL meteorological input data. This dataset provides global meteorological information with a spatial and temporal resolution of 0.25° (≈27 km) and 3 h, respectively. Thus, air temperature (Ta), wind speed (WS), and shortwave solar radiation (SW) were selected from the GLDAS 3-hour data, which were calculated between 9 AM and 12 PM, concurrent with the Landast data acquisition time. Due to the lack of relative humidity data in the GLDAS dataset, this was calculated based on Shuttleworth [52].

Two global land cover maps, from the Copernicus Global Land Service (CGLS) and the European Space Agency (ESA), were used as inputs in the ET modeling. The CGLS maps with 100 m spatial resolution have 23 dynamic land covers and have been provided since 2015. Buchhorn et al. [53] evaluated the CGLS map, reporting an overall accuracy of approximately 80%. The highest accuracy (>85%) was for forests, bare/sparse vegetation, snow/ice, and permanent water, and the lowest (<65%) was for herbaceous wetlands, lichen/moss, and shrubs, while herbaceous vegetation, croplands and built-up covers revealed a moderate accuracy (70–85%). The ESA dataset provides a land cover map at 10 m resolution with 11 classes based on Sentinel-1 and Sentinel-2 data for 2020 [54]. Although no research has yet been carried out to evaluate ESA land cover map accuracy, its higher spatial resolution may be more compatible with the 30 m spatial resolution of the Landsat data. However, for more certainty, firstly, the ESA and CGLS maps were integrated (hereinafter referred to as the ESA-CGLS map), and then, the four main classes—cropland (CRO in the maps), forest (ENF, DBF, MF, SAV, WSA, and trees in the maps), short vegetation (GRA and OSH in the maps) and wetland (WET in the maps)—were reclassified and used as inputs in the endmember selection process.

2.2. Methods

2.2.1. SEBAL Algorithm: A Brief Description

ET estimation was conducted based on the SEBAL methodology developed by Bastiaanssen et al. [9] which calculates LE as a residual of the surface energy balance.

$$LE = Rn - G - H \quad (1)$$

where LE, Rn, G and H are latent heat flux (ET), net radiation, soil heat flux, and sensible heat flux in ($W \cdot m^{-2}$), respectively. Rn was calculated using the shortwave and longwave radiation balance [9,55], while G was estimated based on Bastiaanssen [56].

H was estimated as follows [56]:

$$H = \frac{\rho_a C_p dT}{r_{ah}} \tag{2}$$

where ρ_a, C_p, dT, and r_{ah} are air density, air specific heat (1004 $J \cdot kg^{-1} \cdot K^{-1}$), temperature (K) difference (T1–T2) between two heights (z1–z2), and aerodynamic resistance to heat transport ($s \cdot m^{-1}$), respectively. r_{ah} was computed as follows:

$$r_{ah} = \frac{1}{Ku_*}\left[ln\left(\frac{z_2}{z_1}\right) - \psi_{h(z_2)} + \psi_{h(z_1)}\right] \tag{3}$$

$$u_* = \frac{Ku_b}{ln\left(\frac{z_b}{z_{om}}\right) - \psi_m(z_b)} \tag{4}$$

where u_* is the friction velocity ($m \cdot s^{-1}$), u_b is the wind speed at the blending height ($z_b z_b$), k is the Von Karman constant (0.41), and ψ_h and ψ_m are the stability correction factors for momentum and heat, respectively. z_{om} is the surface roughness for momentum estimated according to [57,58]:

$$z_{om} = exp\left((5.3 \times NDVI) - 5.2\right) \tag{5}$$

SEBAL assumes the difference between the aerodynamic and air temperature (dT) at the reference height as being linearly aligned with LST. The linear coefficients are then extracted from hot and cold pixels with extreme hydrologic situations, chosen based on NDVI, LST and/or albedo from remote sensing imagery and user knowledge of a study area. As in Equation (2), H is computed based on two unknown parameters, dT and r_{ah}. The endmember pixels are utilized to facilitate the calculation. Thus, the accuracy of ET estimation in SEBAL is dependent on careful selection of the pixels for H as a contributing parameter in the ET equation (Equation (1)) [25,55]. More specifically, H was estimated using an iterative process modifying r_{ah} and u_*. In this process, the Monin–Obukhov atmospheric correction and stability functions were used to stabilize H [9,59] based on a manual endmember selection of "hot" and "cold" pixels. For the cold endmember, H was set to 0 with the maximum ET, thus satisfying $LE = Rn - G$, which occurs in well-irrigated and highly vegetated surfaces. With respect to the hot endmember, characterized by bare and dry areas, all energy was supposed to be H and, consequently, ET was considered to be 0, thus satisfying $H_{hot} = Rn - G$ and $LE_{hot} = 0$.

Finally, the evaporative fraction (EF) was calculated at satellite pass ET, which was converted to daily ET (ET_{24h}) according to Equations (6) and (7).

$$EF = \frac{LE}{Rn - G} \tag{6}$$

$$ET_{24h} = \frac{EF \times Rn_{24h}}{\lambda} \tag{7}$$

where λ and Rn_{24h} are the latent heat of vaporization and daily average net radiation, respectively (more information can be found in Bastiaanssen et al. [9]).

Moreover, while albedo was calculated based on surface reflectance [60], a single-channel algorithm was used to calculate LST as an input to SEBAL [61]:

$$LST = \frac{BT}{1 + \left[\frac{\lambda \cdot BT}{\rho}\right] \cdot ln\left(\varepsilon_0\right)} \tag{8}$$

where LST and brightness temperature (BT) are calculated based on Kelvin (K) and λ is the central band wavelength of emitted radiance (11.45 µm). ρ was computed as follows:

$$\rho = h \times \frac{c}{\sigma} \times \left(1.438 \times 10^{-2}\right) \tag{9}$$

where σ is the Boltzmann constant (1.38×10^{-23} J·K^{-1}), h is Planck's constant (6.626×10^{-34} J·s), and c is the speed of light (2.998×108 m·s^{-1}). ε_0 is the surface emissivity that was estimated by the following equation [62]:

$$\varepsilon_0 = 0.004 \times CC + 0.986 \tag{10}$$

where CC was calculated from $NDVI$ as follows:

$$CC = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \right)^2 \tag{11}$$

2.2.2. Automated Endmember Selection

To avoid a manual endmember selection as well as to determine the most efficient ET estimation method, three approaches were considered for the selection of cold and hot endmembers: (a) land cover-based method with morphological filters (WF), (b) land cover-based method without morphological filters (WOF) and, (c) the Allen method [35] excluding land cover maps.

For the WF method, the endmembers were selected based on the land cover map and morphological features as well as the cold and hot pixels' NDVI, albedo and LST characteristics through a three-step process (Figure 2):

(1) Image pre-processing and land cover extraction for ET estimation: To extract land cover categories for ET calculation, an area within 30 km of a meteorological station was selected as an area of interest (AOI) for endmember selection following [35,51]. Cloud and shadow areas were removed using the CFMask method. Considering that cloud boundaries usually remain in the CFMask output, pixels with LST less than 270 K were also removed in accordance with [43]. Water bodies were also masked by the normalized difference water index (NDWI) and the NDVI by defining $NDWI > -0.1$ [63] and $NDVI < 0$ thresholds in accordance with [32,40,63,64]. Unlike cropland, which includes vegetated and bare agricultural fields, the other three classes may lack bare soil for hot pixel selection. To address this issue, pixels belonging to bare soil were extracted in the forest, short vegetation and wetland classes, and afterward combined into each vegetation land cover (except cropland). To this end, non-vegetation regions, which can contain built-up, stony and bare areas, were selected using a vegetation index of $0 < NDVI < 0.2$. In addition, while urban areas were masked by the ESA-CGLS map, a vegetation index threshold of $0 < NDVI < 0.05$ was used to mask stony regions [65]. Consequently, the remaining non-vegetation areas were considered as bare soil.

(2) Implementation of morphological functions for each land cover: To screen out unfeasible endmembers for each land cover, erosion morphological filters were used to remove (a) single pixels resulting from the classification process which are usually confused with the intended class, (b) cloud boundaries and cloud shadow boundaries, which are cooler than sunlit land surfaces, remaining even after CFMask use, and (c) crop field edges and bare soil margins which might be bounded by trees, shrubs, built-up and stony areas. In addition, thermal bands were downscaled to 30 m to match multispectral data. However, this might lead to thermal pixel contamination, usually occurring at farmland edges [35] and other class boundaries. Land cover maps participating in the endmember selection allowed the use of homogeneity filters [5,35,51], preventing candidate pixels from lying at farmland margins. Nonetheless, in addition to decelerating image processing [51], it is uncertain that homogeneity filters can diminish misclassification errors as well as cloud and cloud shadow boundaries. Moreover, the homogeneity operators might result in less acceptable ET performance compared to Allen [51]. Thus, morphological filters were applied alone and assumed to help selecting endmembers away from cropland and cloud edges as well as misclassification errors.

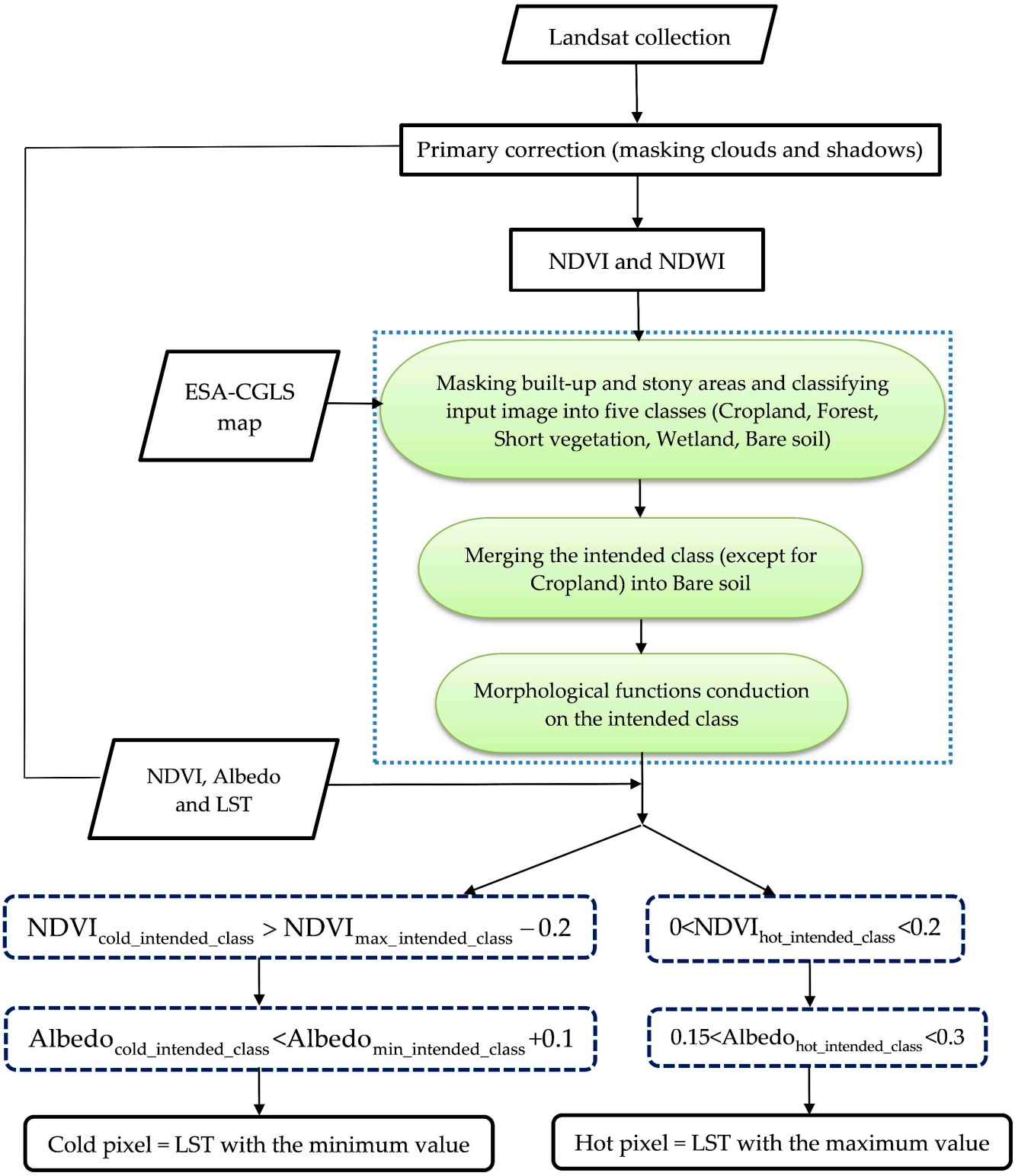

Figure 2. Endmember selection flowchart following the WF approach.

(3) Definition of specific thresholds for NDVI, albedo and LST to mask cold and hot endmembers: Cold pixels indicating optimum ET show higher NDVI values around the maximum. Additionally, albedo and LST are also close to the minimum. Accordingly, $NDVI_{cold\_intended\_class} > NDVI_{max\_intended\_class} - 0.2$ was regarded as the first cold pixel mask. Vegetated surfaces also have albedo with the lowest value, around 0.1 [55]. Therefore,

the albedo mask, $Albedo_{cold\_intended\_class} < Albedo_{min\_intended\_class} + 0.1$ was defined and added to the NDVI mask. Finally, LST with the minimum value was selected as a cold pixel based on the $Albedo_{cold\_intended\_class}$ output (Figure 2). Hot pixel selection had the same process with different thresholds. More specifically, while a hot pixel for a crop class must fall in fallow fields with the lowest moisture, this pixel for non-crop classes can also be situated in bare soil. Accordingly, the proper NDVI value for hot pixel selection was higher than zero and lower than 0.2 ($0 < NDVI_{hot\_intended\_class} < 0.2$). Given that built-up and stony areas were masked in step 1, the chance of an appropriate hot pixel selection in fallow or bare fields increased. However, the fallow and bare soil classes have relatively high albedo, and non-saline soils usually show an albedo lower than 0.3 followed by salty soil (usually $Albedo > 0.3$) [40,66]. Hence, a middle range of albedo can be an apt choice for hot pixel selection ($0.15 < Albedo_{hot\_intended\_class} < 0.3$). Subsequently, LST with the maximum value was selected as a hot pixel from the $Albedo_{hot\_intended\_class}$ output (Figure 2).

To further investigate the potential of the proposed algorithms (WOF and WF) for endmember selection, the results obtained with an automated approach developed by [35] based on statistical analysis were compared with those of the WF and WOF approaches. In Allen's approach, cold pixels fall in a set with the highest 5% of NDVI and the lowest 10% of LST. In contrast, hot candidates are selected from areas with the bottom 10% NDVI and the top 20% LST. Note that the Allen method [35] was conducted without considering land cover maps.

3. Results

3.1. Evaluation of Endmember Selection

The endmember selection was implemented based on three approaches: two land cover-based methods (WF, WOF) and the Allen method [35]. As previously described, morphological filters were applied exclusively for the WF method. According to visual interpretation, a 3×3 and 5×5 size can be deemed appropriate, while filter size selection is optional based on study area conditions (Figure 3).

Since the Allen method [35] considers all land covers and the top 5% of NDVI values, cold pixels of crop sites were usually selected over non-crop vegetated areas, especially throughout off-season periods and the early stages of the growing season (Figures 4 and A1). Furthermore, over non-crop land covers, densely vegetated farmland or water bodies might adversely affect cold pixel selection (LET, MR3, TW1 and PHM stations in Figure A1). Allen's method uses the bottom 10% NDVI values, which results in hot pixels mostly falling on built-up and stony covers. Occasionally, this happened in the WOF and WF approaches using land cover maps and different thresholds. Regarding WOF, the cold endmembers were predominantly located at farmland margins bounded by other vegetated land covers (ARM, BI1, and Br1) and even cloud shadow boundaries (NE2 in Figure 4 and UMB in Figure A1). Urban and stony areas were also viable for selection as hot pixels with WOF (Br1, IB1, and TW3 in Figure 4) as they were next to farmlands. Misclassification errors were the second source of weakness in WOF for both crop and non-crop flux towers (IB1, RO1, TW3, Ca3, GRO, and KON for cold pixel selection, and KFS, Ca3, NC2, PHM, TON and UMB for hot pixel selection). Unlike the uncertainties found in the Allen and WOF approaches, these were considerably reduced in WF by applying morphological kernels. Endmember pixels were then found over dense vegetation and fallow farmlands for crop sites (Figure 4), and bare or fallow and densely vegetated areas for non-crop flux stations (Figure A1). Moreover, endmembers selected by WF were far from crop field edges (with less contamination due to Landsat thermal pixels [35]), cloud edges and cloud shadow edges, resulting in more accurate ET estimation (Figures 4 and A1).

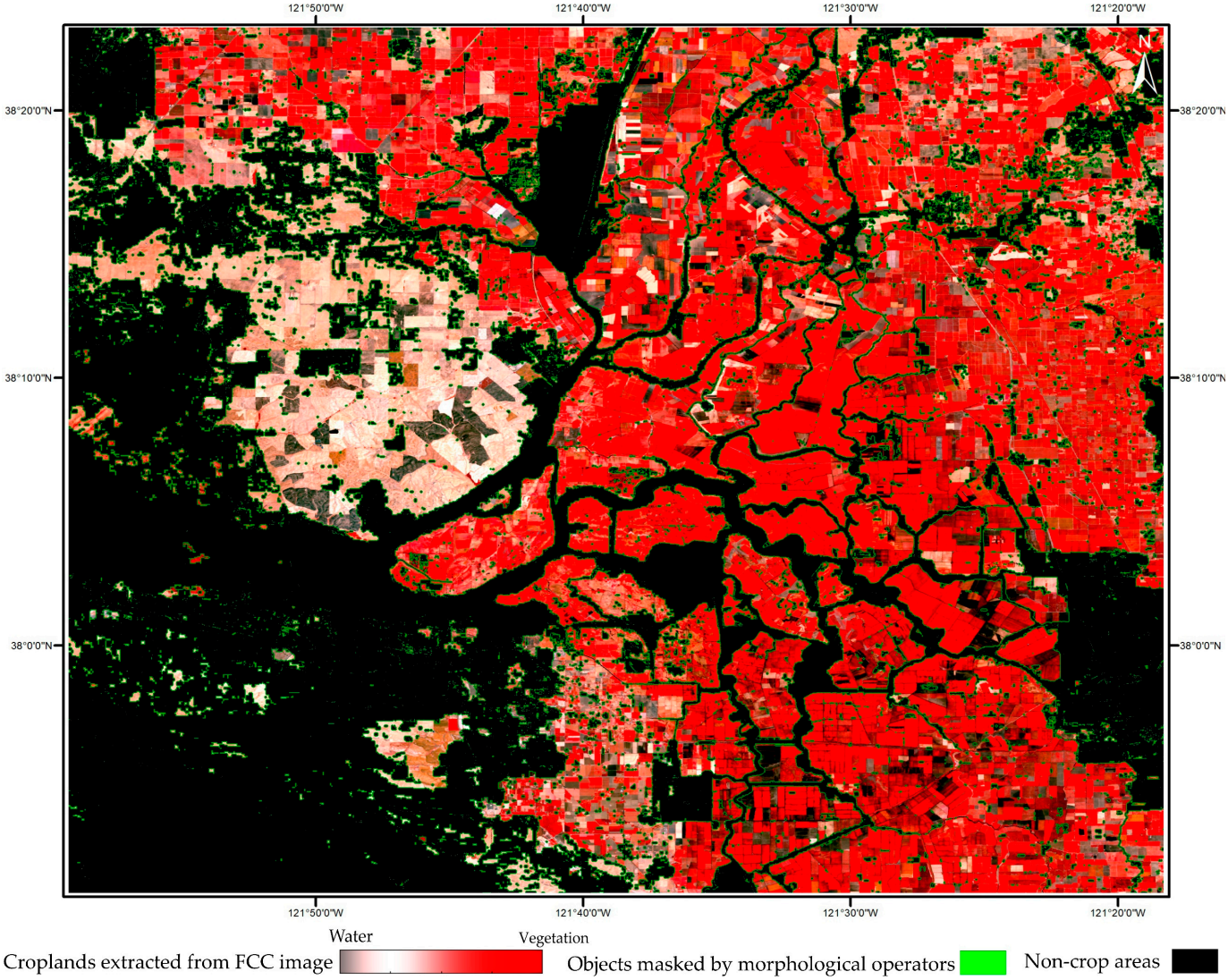

Figure 3. Output when using morphological operators (3 × 3) for the US-TW3 station. This map illustrates the crop class which was extracted from False Color Composite (FCC) imagery and subjected to changes with the morphological functions. The green areas include small polygons, single pixels or misclassification errors, linear objects and polygon margins, diminished by the functions of the WF method.

A comparison of the NDVI and LST values obtained using the WF, WOF and Allen approaches (Figures 4 and A1) showed that the correlation between NDVI and LST was positive or negative, whereas it was generally expected to be negative [67–69]. According to [70], the correlation in winter is mostly positive and in summers mostly negative and is thus impacted by seasonality. The correlation can also fluctuate during autumn and spring. However, this principle was contradicted in several cases (TW3, WJS, GRO, TON, and KFS) in which, despite being in summer, there was a positive correlation between NDVI and LST values for cold pixels of the Allen, WOF and WF approaches. Karnieli et al. [71] reported that the correlation is also sensitive to water and energy, meaning that it is positive when water is the primary parameter for vegetation growth (at lower altitudes and elevations and in midseason) and negative when energy is the key factor for plant growth (at higher altitudes and elevations and in the primary stages of the growing season). Thus, different conditions can lead to uncertainties with respect to the correlation between NDVI and LST, impeding the possibility of applying a constant relationship. On the other hand, mean NDVI values in both hot and cold endmembers differed considerably with the

Allen approach compared to the WF or WOF methods, despite having similar standard deviations (STD) (Table 2). Despite the differences in hot and cold NDVI values for the land cover-based and Allen approaches, similar ET estimation accuracies were obtained. Moreover, similar hot and cold NDVI values with the WF and WOF approaches resulted in different ET estimation performances. Additionally, the cold pixel LST difference between the WF and Allen approaches was less than 1 K, while in the case of hot pixels, this difference was about 2 K between WF and the other approaches and 4 K between the WOF and Allen approaches.

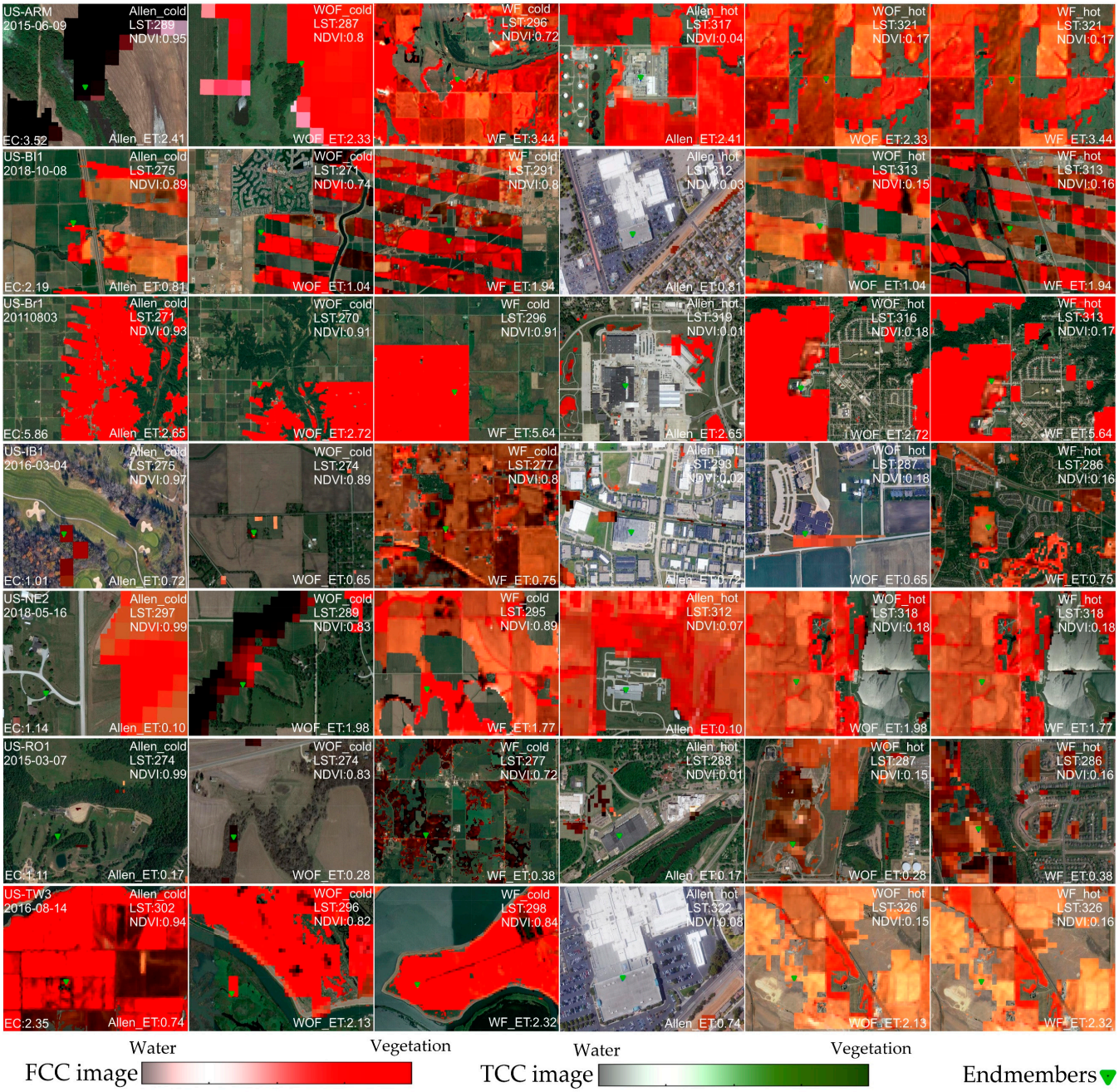

Figure 4. Cold and hot pixel locations and values alongside the statistics for one selected image of crop flux towers for different endmember selection methods using GLDAS meteorological data in SEBAL. The green triangles indicate endmember location on the intended land cover extracted from the FCC image, and over the background of Google Earth images. Each scenario's ET estimation performance is coded as Allen_ET, WOF_ET, and WF_ET.

Table 2. Overall mean and standard deviation (STD) of NDVI (dimensionless) and LST (K) with the different endmember selection methods for all flux tower sites (2270 images).

		NDVI_WF	NDVI_WOF	NDVI_ALLEN	LST_WF	LST_WOF	LST_ALLEN
Cold pixels	Mean	0.76	0.77	0.91	287.91	285.68	288.83
	STD	0.04	0.03	0.03	3.18	3.64	3.23
Hot pixels	Mean	0.17	0.17	0.06	308.77	310.72	306.94
	STD	0.01	0.01	0.01	6.65	5.82	5.35

3.2. Assessment of ET Estimates with the Different Endmember Selection Models

In this study, the root mean square error (RMSE), coefficient of determination (R^2) and bias percentage (PBias) error metrics [72] were used to analyze the automated ET estimation approaches based on temporal variations, land covers and climate types. Lower RMSE and PBias magnitudes and higher R^2 indicate stronger agreement with EC data, while the PBias sign is representative of underestimation (negative) and overestimation (positive). Model improvement assessment was performed in an absolute way by comparing error and agreement metrics as [5,30,51,73–75] suggested.

3.2.1. Overall Daily ET Evaluation

Evaluation of ET estimates by different endmember selection approaches against EC measurements showed that WF (WF_RMSE, WF_R^2 and WF_PBias) generally outperformed WOF (WOF_RMSE and WOF_R^2 and WOF_PBias) and Allen (Allen_RMSE and Allen_R^2 and Allen_PBias) (Table 3). The absolute mean difference between WF_RMSE and WOF_RMSE values across crop flux stations was 0.11 mm·day^{-1}, and that between WF_RMSE and Allen_RMSE for crop and non-crop flux stations was 0.19 and 0.08 mm·day^{-1}, respectively, thus yielding better results than WOF and Allen. In addition, the R^2 of the WF was slightly higher than the corresponding WOF and Allen values for all stations. Regarding PBias, all the automated ET estimation approaches for crop flux tower stations showed optimal magnitudes (less than ±10%) with slightly better performances with the WF and WOF approaches. Model overestimation was the dominant trend for all flux stations (Table 3).

Table 3. Statistical summary of ET time series calculated using the different endmember estimation procedures vs. EC flux stations (N: number of towers—images).

	Crop Flux Stations (N: 7–762)			Non-Crop Flux Stations (N: 16–1508)			All Flux Stations (N: 23–2270)		
	WF	**WOF**	**Allen**	**WF**	**WOF**	**Allen**	**WF**	**WOF**	**Allen**
RMSE (mm·day^{-1})	0.91	1.02	1.10	1.59	1.60	1.67	1.38	1.43	1.50
R^2	0.71	0.63	0.68	0.43	0.41	0.39	0.51	0.48	0.48
PBias (%)	2.46	−1.96	5.53	61	63	50	42	43	36

3.2.2. Evaluation of ET Estimation Results for Climate-Based (Global Aridity Index and Köppen Classification) and Land Cover-Based Trials

The performance of the different automated ET estimation approaches was evaluated by climate type—based on the global aridity index (Figure A3) and the Köppen climate classification (Figure A4)—and land cover type (Figure A5) for each flux tower station. A summary of the main results of these assessments is shown in Table 4. It should be noted that despite a stronger RMSE value (about 0.08 mm·day^{-1}) for the Allen compared to the WF and WOF approaches in arid/steppe climate zones, land cover-based endmember selection models, especially WF, outperformed Allen, with around 0.33 mm·day^{-1} difference for humid/humid continental climates. Moreover, warm summer continental was preferable to marine west coast for the highest PBias over all methods because of the more station proportion in the experiments. Overall, despite the meaningless underestimation with

WOF with respect to crop flux stations (Figure A5), overestimation was the dominant trend throughout all the analyses.

Table 4. Summary of ET estimation trials using the global aridity index, the Köppen climate classification and land cover type.

Experiment	Metric	Scenario with the Best Performance (Climate or Land Cover)	Scenario with the Worst Performance (Climate or Land Cover)
Aridity index climate-based trials	RMSE	Allen (arid)	Allen (humid)
	R^2	Allen (sub-humid)	Allen (arid)
	PBias	Allen (semi-arid)	WOF (arid)
Köppen climate-based trials	RMSE	Allen (Bsk, Bsh) + WF(Dfa)	Allen (Dfb) + WF (Cfb)
	R^2	Allen (Cfb) + WF (Dfa)	Allen (Dfc)
	PBias	Allen (Dfc) + WF (Dfa)	WOF and WF (Cfb)
Land cover-based trials	RMSE	WF (crop)	Allen (forest)
	R^2	WF (crop)	Allen and WOF (short veg.)
	PBias	WF (crop)	WOF (forest)

3.2.3. Evaluation of ET Estimation When Using Humid/Non-Humid Climate-Based Classifications

The trend of the metrics for each scenario, calculated using both land cover and aridity index climate types simultaneously, was plotted from the lowest (humid) to the highest (arid) aridity index value (Figure A6). Although, generally, WF outperformed WOF and Allen for flux towers in humid regions, the differences between the results of the scenarios for semi-arid, sub-humid and arid stations were completely inconsistent (Figure A6). Moreover, as there were minor metrics differences between the WF, WOF and Allen approaches in some analyses or flux tower ET assessments (Figures 5 and A2, Figures A3–A5, and Table 3), other contributing sub-factors must therefore have been driving the difference between automated model performance. This was explored through a division of climate zones into 'humid' and 'non-humid' based on the aridity index, which were then quantitatively assessed through the error metrics (Table 5).

In humid zones, for both crop and non-crop flux tower sites, there was a clear difference in performance between the Allen and the land cover-based model approaches, with a particularly superior performance of WF at crop flux tower sites. Differences between WF_RMSE and Allen_RMSE were around 0.34 and 0.30 mm·day^{-1} for crop and non-crop sites, respectively, compared to about 0.10 and 0.25 mm·day^{-1} between WOF_RMSE and Allen_RMSE. While the WF approach appropriately outperformed WOF for crop flux tower stations, the differences were less evident, although still in favor of WF, for non-crop stations. Overall, WF outperformed WOF and Allen with respect to all towers lying in humid zones.

Table 5. Evaluation of automated approaches for humid climate-based tower classification (N: number of towers—images).

		RMSE (mm·day^{-1})			R^2			PBias (%)		
Climate Zones	**Land Cover-Based Towers**	**WF**	**WOF**	**Allen**	**WF**	**WOF**	**Allen**	**WF**	**WOF**	**Allen**
Humid	Crop (N: 3–308)	0.94	1.16	1.27	0.71	0.54	0.64	8	1	26
	Non-crop (N: 6–360)	1.86	1.89	2.15	0.41	0.41	0.38	61	67	78
Non-humid	Crop (N: 4–454)	0.90	0.92	0.98	0.71	0.70	0.71	−2	−5	−10
	Non-crop (N: 10–1200)	1.43	1.43	1.39	0.44	0.41	0.40	60	60	32

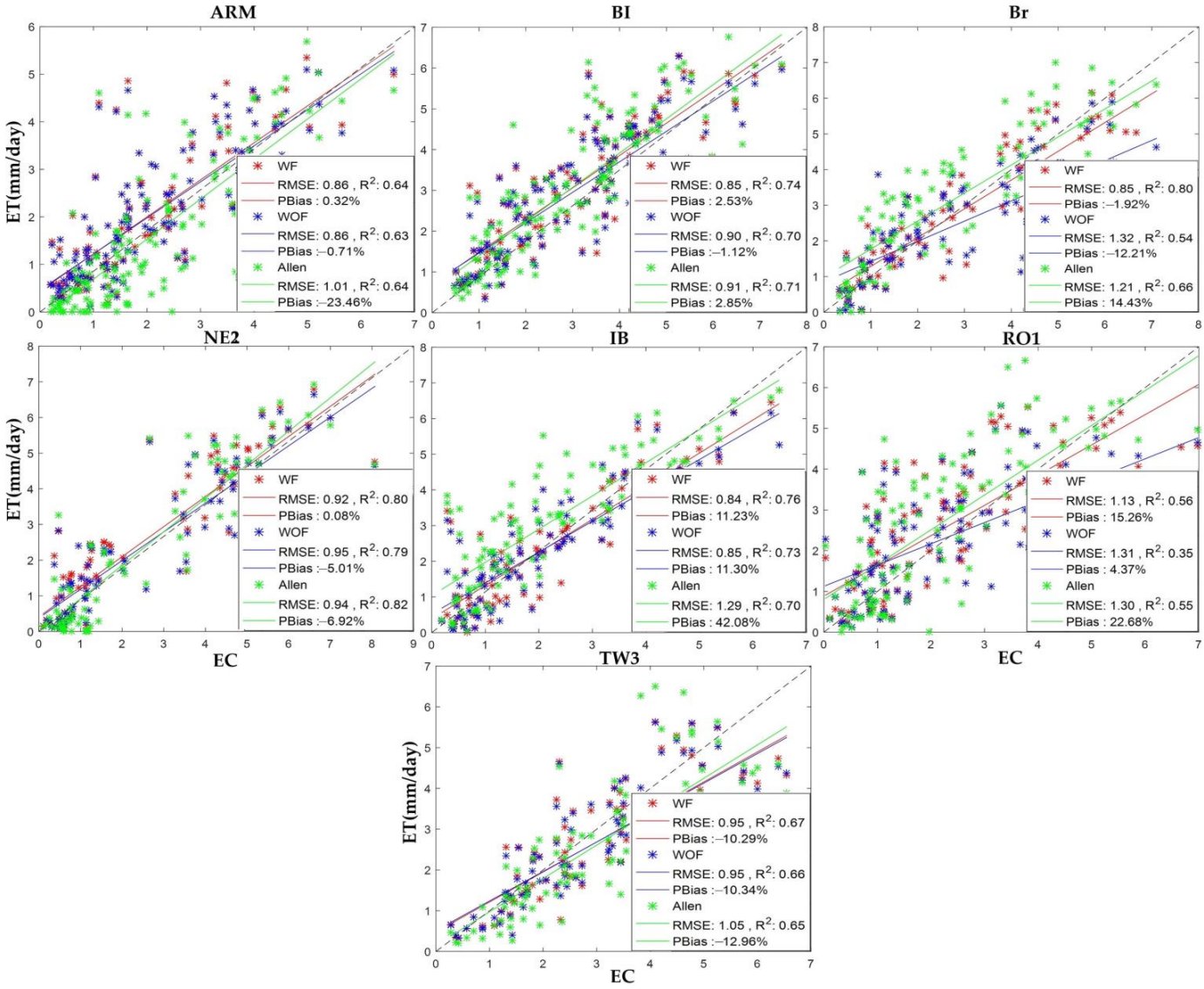

Figure 5. Daily actual ET estimated by the automated calibration methods vs. crop flux tower EC measurements.

In non-humid zones (sub-humid, semi-arid and arid), compared to humid zones, there were only small differences in performance between the automated models for both crop and non-crop land cover types. The two land cover-based models performed slightly better than the Allen method [35] at crop flux tower sites, while the Allen approach performed slightly better at non-crop stations. The only underestimation was observed for crop towers located in non-humid zones for all three automated procedures.

3.2.4. Overall Monthly ET Evaluation

In this evaluation, daily ET estimates were monthly aggregated and compared against EC monthly measurements (Figure 6).

All sites yielded the smallest RMSE values in January, November and December for all land covers and in February only for crop towers, following the lowest rate of greenness. Generally, the lowest R^2 values were found for all models and land covers in December, January and March, which might be attributable to the lower functionality of the Landsat data in cloudy or snowy conditions compared to the rest of the year.

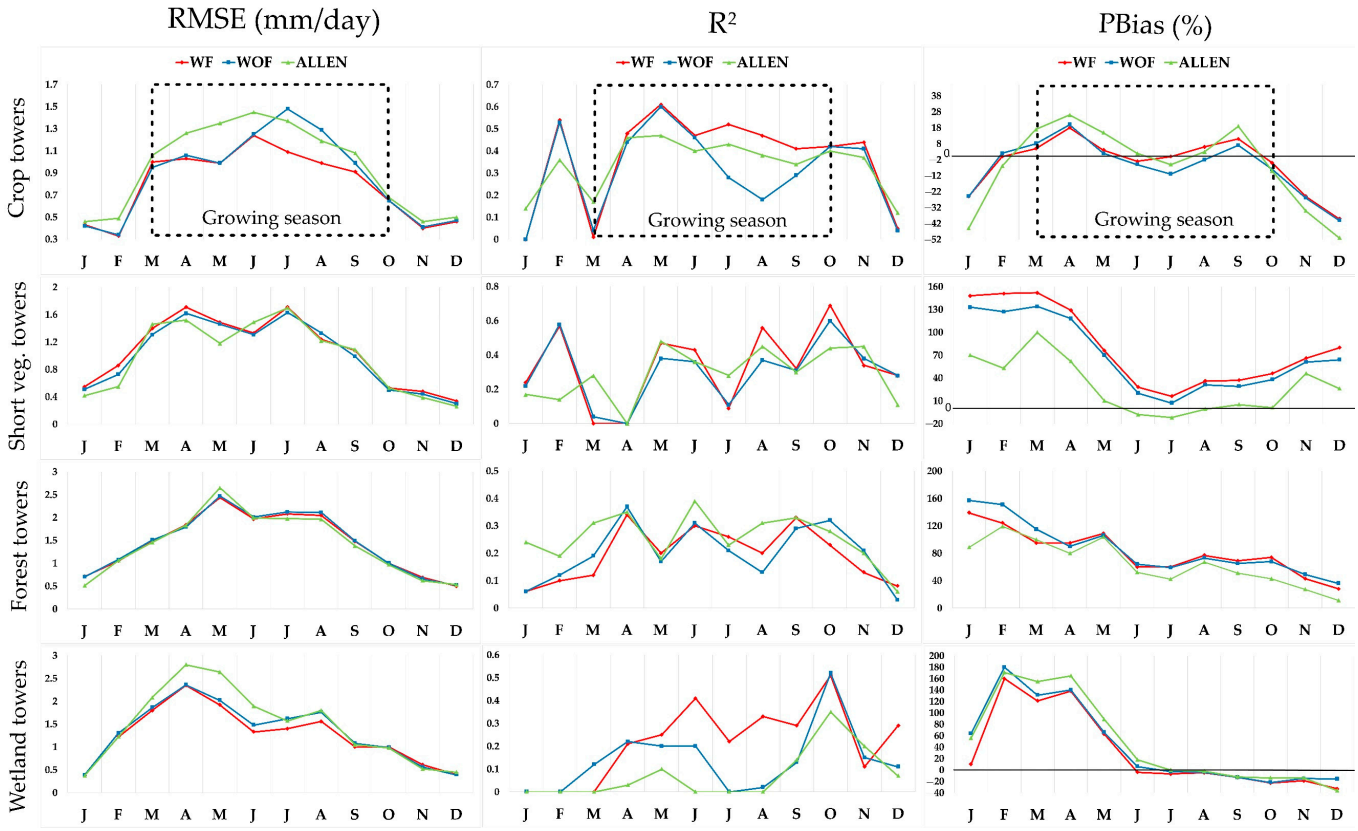

Figure 6. Monthly ET evaluation of WF, WOF and Allen models based on land cover type.

With respect to crop towers, the main difference between WF_RMSE and WOF_RMSE was found from June to October and between WF_RMSE and Allen_RMSE from March to October (during the growing season). In all cases, WF showed the highest agreement with EC data (Figure 6). Moreover, the RMSE results with all the methods behaved similarly across the off-season period. For short vegetation towers, there were two RMSE peaks in July and April with all three approaches. The main R^2 difference was found in March when the Allen approach gave slightly better results than WF and WOF. For forest flux towers, all three approaches performed similarly, with the highest RMSE values in May. All three approaches for wetland flux tower sites showed comparable trends, although the Allen approach yielded higher RMSE values from March to July. With respect to the R^2 values, the highest correlation was observed in October for the short vegetation and wetland flux tower sites. The months of the growing season with low correlations were July for the short vegetation flux tower sites and August for the other tower sites. An upward trend in correlation was observed from August to October for all scenarios and land covers, except short vegetation. The Allen approach in wetland flux tower sites yielded the lowest R^2 with around zero correlation from January to August. The peak correlation values were obtained, for all three approaches, in October for wetland and short vegetation flux tower sites, in March for crop flux tower sites, and in April for forest flux tower sites. Generally, the lowest absolute PBias magnitudes were recorded in June or July months with almost the maximum NDVI or LAI for all approaches. The growing season period between May to October yielded lower PBias (around ±10%) for crop flux tower sites, with zero bias in July for WF. Allen_PBias yielded similar results for crops, forests and wetland flux tower sites but outperformed the other approaches in wetland sites. WF and WOF for short vegetation and forest flux tower sites predominantly tended to overestimate ET (by between 10% and 160%). This tendency to overestimation was also observed for the Allen approach during the whole year at forest flux tower sites and during the first five and last three months of the year at short vegetation flux tower sites. Conversely, there was a tendency

at wetland flux tower sites toward overestimation during the first six months of the year and underestimation during the last three. For crop flux tower sites, there was a tendency toward underestimation throughout the off-season period for all three approaches as well as from June to August for the Allen approach and toward overestimation for the rest of the year.

To assess the performance of the scenarios across both different months and land cover type together, the month with the highest RMSE value for each land cover type (Figure 6) was targeted. In the case of forest flux tower sites, all three approaches showed the highest error in May and in the case of wetland flux tower sites in April. When using the Allen approach, May yielded the worst results for crop flux tower sites and June for short vegetation flux tower sites. WF in June for crop flux towers and in July for short vegetation flux towers, and WOF in July for both crop and short vegetation flux towers, showed the worst RMSE values. As previously described, candidate cold pixels for WOF over crops may be selected on the crop margins usually occupied by grass and shrubs, thus lowering the agreement. This adjacency effect negatively impacted the short vegetation performance of WOF for cold pixel selection. That is, crop and short vegetation sites have an adverse mutual effect on each other. The highest RMSE with the Allen approach coincided with that of WF in June for crop flux towers. This shows that the Allen approach was not impacted by the adjacency effect (Figures 5 and A2). More specifically, the adjacency effect can have a negative implication when using land cover maps which needs to be addressed by using, for example, morphological filters. However, if trees were abundant in each site, May might be assumed to be the month with the worst RMSE with the Allen approach for crop flux tower sites, as was observed across forest flux tower sites. In short, WF seems resistant to the adjacency effect that results from multiple land covers. As forest provides the highest NDVI, cold pixel selection is rarely vulnerable to multiple land cover conditions, and hence, the highest RMSE values with all three approaches were observed simultaneously for the forest flux tower sites. Moreover, because of the essence of wetlands, precise extraction is extremely challenging, and hence, its accuracy in classification maps is usually lower than for other land cover types. For this reason, the worst RMSE results with the WF and WOF approaches occurred in April. It should also be noted that there was no expansive forest over the wetland sites that were used, which meant that the Allen approach performed similarly.

4. Discussion

4.1. Daily ET Evaluation

The results (Figures 4 and A1) show that the morphological functions in the WF approach successfully diminished other vegetated areas falling at the boundary of crop classes. This helped the model to reduce the negative influence of cold pixel detection over crop flux tower sites, thus increasing model performance. Although this was largely not a concern with non-crop classes, built-up and stony areas which occasionally happened over the margins of bare soil adversely affected hot pixel selection at non-crop flux tower sites. In addition, misclassification errors leading to erroneous endmember pixel selection by WOF were addressed by WF.

Tower-by-tower assessment (Figures 5 and A2) showed that WF yielded the best results, followed by Allen and WOF, in 51%, 27% and 22% of the tower dataset, respectively. Moreover, the improvement in ET accuracy with the WF approach was particularly notable for crop land cover flux tower sites, where this approach yielded better error metrics than the others (Table 3). Taken together, these results suggest that the use of land cover maps alone does not sufficiently explain the differences between the performance of the land cover-based models and the Allen approach, especially across farmlands. It would seem that the key criterion that enabled the land cover-based endmember selection methods to perform better than the Allen approach was to alleviate the adjacency effect with multiple land covers, and, in the specific case of the WF approach, to take into account misclassification errors using morphological operators.

The reason WF showed weaker ET estimation performance than its counterparts in some non-crop flux tower sites might be related to the pruning of marginal pixels that in reality belong to forest, short vegetation or wetland areas by morphological functions. Unlike farmlands with certain textures and shapes that might lead to the growth of non-crop vegetation at the crop edges, non-crop covers do not show the same behavior. However, the possible selection of built-up and stony areas, which occasionally happen over the margins of bare soil, as hot pixels by WOF, in addition to misclassification errors detected as endmember pixels (Figures 4 and A1), can be diminished when using the morphological filters, compensating for erroneous pixel pruning issues. For this reason, WF_ET and WOF_ET estimations were fairly close at non-crop sites.

4.2. Evaluation of Endmember Statistics

The NDVI- and LST-based endmember selection method assessment (Figures 4 and A1, and Table 2) showed that the correlation is dependent on different conditions and that there is no clear pattern. Thus, not every pixel with a higher (or lower) NDVI and LST value is more (or less) appropriate for selection as a cold or hot endmember.

The comparison of overall average endmember NDVI values between the automated approaches (Table 2) showed that similar NDVI values can result in different ET estimation performance and vice versa. Hence, NDVI cannot be regarded as a discriminating factor between automated ET calibration approaches. The NDVI should therefore only be applied as a filter to minimize pixel candidates for the later steps of endmember selection. Additionally, averaged cold pixels for the 2270 Landsat images (Table 2) showed comparable results for the Allen and WF approaches, while there was a notable difference between their ET estimations. On the other hand, the same validation for hot pixels showed a value difference between the three approaches which could be the main reason for ET estimation disagreements. Consequently, the endmember selection methods, and hence ET estimation performance, were found to be more sensitive to the LST-based selection of hot pixels than cold pixels, which was also highlighted by Long et al. [25].

Finally, despite the high sensitivity of endmember selection approaches to hot pixel LST selection and the location of Allen hot pixels mostly over built-up and stony areas, Allen_ET accuracies were close to or even better than WOF. This suggests higher priority should be given to LST-based pixel consideration rather than the location or the nature of the object hosting the endmembers.

4.3. Land Cover-Specific ET Assessment

The land cover-specific analysis showed that farmlands and forests yielded the strongest and weakest ET estimation performances when compared to EC measurements for all scenarios (Figure A5 and Table 4). These disagreements might be due to errors in LAI and vegetation height estimates, which will require future research. Among trees, WJS delivered the best performance, followed by CMW, for the automated approaches. This can be attributed to the arid and steppe climate types having the lowest RMSE (Figures A3 and A4, and Table 4). Concerning short vegetation, this might be ascribed to the higher concentration of grass or shrubs than in croplands, acting the same as LAI and canopy. The contiguity of grasslands to croplands, causing the highest misclassification errors in cold pixel endmember selection, could also be the reason for the higher RMSE of short vegetation flux tower sites to crop sites. With respect to wetlands, areas that are usually flooded, water evaporation, as well as soil evaporation and plant transpiration, adds uncertainty to ET estimation. Wetlands were also often bounded by other vegetated covers making their proper classification difficult due to a similar phenology and structure to farmlands and grasslands, which negatively affected cold pixel endmember selection. For croplands, vegetation height above 2 m, as is the case for maize and corn, would adversely affect SEB model performance [32]. This may be due to the issues described above, especially high LAI and dense canopy, and it may be the reason for the less accurate results

for RO1 (Figure 5), the only WF crop trial with an RMSE above 1 mm·day^{-1}, compared to the other crop sites.

4.4. Land Cover-Climate ET Analysis

Generally, for all land covers, the RMSE of all methods decreased from humid to arid. Over all land covers except forest, there was an appropriate difference between Allen_RMSE and WF_RMSE in humid climates. This was also noticed between WOF_RMSE and WF_RMSE but only with respect to crop flux tower sites. In contrast, WF_RMSE had close and consistent values for crop and wetland flux tower sites with all climate types. This was also observed for forests except for arid climates. WF_RMSE and WOF_RMSE performed similarly with all climate types at short vegetation and forest flux tower sites.

It is complicated to make a strict interpretation of the R^2 and absolute PBias variation owing to the high inconsistency of the results (Figure A6). However, the PBias sign showed consistent overestimation with respect to the validations conducted based on climate type systems (Figures A4 and A5, and Table 5). Among land covers, the highest overestimation was found in forest flux tower sites, while crop flux tower sites showed optimal PBias for all scenarios. The only trial with a slight underestimation was with the WOF approach for crop flux tower sites.

4.5. Humid Climate- and Landscape-Based ET Assessment

With respect to almost all the humid stations (Br, IB1, Ca3, RO1, PHM, KFS, GRO, NC2, and UMB), a better RMSE performance was obtained with the land cover-based approaches, especially WF (Figures 5 and A2).

As for the non-humid stations (WJS, SRG and CMW were located in the arid zone), for the WJS station, WF_RMSE was similar to Allen_RMSE and better than WOF_RMSE, while for the SRG station, the Allen approach gave the lowest and the WF approach the highest RMSE value. In contrast, the Allen approach performed worse at the CMW station. The high accuracy of the Allen approach may be attributable to two factors: station location in an arid zone (as Allen endmember percentiles were defined for this climate [32]) and the steppe landscapes. More specifically, environments dominated by arid climates usually lack multiple land covers, inducing homogenous landscapes. Under these circumstances, cold pixel selection with the Allen approach, excluding land cover maps, is facilitated and, consequently, ET estimation performance is boosted. In this regard, WJS has an expansive bare area alongside a few scattered farmlands and sparse trees (low cover canopy of less than 30% (https://ameriflux.lbl.gov/, accessed on 23 March 2022)), and SRG and CMW wide and scattered grasslands. However, the worse Allen_RMSE performance compared to the land cover-based models at CMW was possibly due to its being on a river basin which can make cold pixel detection challenging with the Allen approach. In addition, with respect to the WJS station, the high misclassification errors might be the cause of the worse performance of WOF_RMSE increase compared to Allen_RMSE, since the adjacency effect is decreased over homogeneous environments.

The sub-humid stations are ARM, NE2, KON and VCM, and the semi-arid stations are BI1, TW3, TW1, LET, MR3, OAS, and TON. With respect to the crop stations (ARM, NE2, BI1 and TW3), the WF approach gave lower RMSE values than WOF or Allen. However, the RMSE results were inconsistent at the other stations. Our investigation suggests that stations bounded by farmlands, river basins or water bodies might be associated to better WF_RMSE or WOF_RMSE accuracies than Allen_RMSE. More specifically, the presence of rough surfaces leads to cold pixel selection with the Allen approach. For instance, in addition to crop flux tower stations, LET with farmlands, MR3 with river basin and croplands, and TW1 with wetlands and farmlands were filled, leading to a better ET estimation performance of the land cover-based models compared to the Allen approach. In contrast, TON lacked expansive croplands and water bodies, producing the best Allen_ET performance among the scenarios. However, the OAS and KON areas occupied with farmlands or water bodies neglected this principle and yielded a more

acceptable Allen_RMSE than its counterparts. Thus, there are still uncertainties that require further research in future studies.

In general, over all the trials, WOF_RMSE accuracy can be ascribed to the intensity of the adjacency effect or misclassification errors in a particular site, while Allen_RMSE is predominantly dependent on climate type or landscape homogeneity. Accordingly, the most challenging conditions for Allen can be the presence of both a humid climate and heterogeneous landscapes at sites such as RO1, Br, IB, UMB, and PHM. These sites yielded less agreement compared to the land cover-based models. In contrast, the coincidence of arid climate and homogenous areas can provide better RMSE results with the Allen approach, as in WJS and SRG. In comparison, WF_RMSE is resistant to the aforementioned conditions, while WOF_RMSE depends on the quality of the land cover map and the adjacency effect.

In short, in the comparison of the WF and WOF approaches, the biggest differences were found in humid zones for crop flux tower sites in contrast to comparable results that were achieved under the other conditions. The WF model outperformed the Allen approach for all land covers lying in humid zones, but Allen can be utilized instead of land-cover based approaches in non-humid zones with any land cover type regardless of local calibration of Allen endmember percentiles.

4.6. Monthly ET Evaluation

Since ET increases with the rise in LST cold pixels because of the decrease in H [25], the coincidence of maximum ET and LST cold pixels is expected. Our monthly analysis satisfied this expectation only for the WF model, with this being the case in July for all land covers except short vegetation in which the highest LST cold pixel and ET values were found in June and July, respectively (Figure 7). On the other hand, ET estimation can be a function of the contribution of plant NDVI, LAI or height in the calculation of Z_{om}, r_{ah}, and subsequently H, which are correlated with the canopy selected as the cold pixel. As the acquisition of accurate LAI information is extremely challenging, we consulted the results of other studies. According to [76], the highest LAI is observed around June for forest, short vegetation and wetland and, according to [77], in June for temperate cereals and in July for corn and soybean as agricultural products. Both these studies were conducted in North America. As can be seen in Figure 7, the highest ET and EC values for short vegetation flux tower sites coincided with maximum NDVI in July, whereas for the remaining land covers, they coincided in the same month with maximum LAI. This suggests that ET estimation for short vegetation is mostly sensitive to NDVI but to LAI for the other land cover classes. Moreover, monthly EC and measured ET had comparable trends for all land covers, especially crops (Figure 7). These results again confirm the higher accuracy of the WF compared to the WOF or Allen approaches for all land covers and months.

With respect to the lower performance in different months or crop flux tower sites, this might be due to a simultaneous increase in crop height, canopy and LAI, as well as in irrigation while the temperature is increasing compared to earlier months. Under these circumstances, soil evaporation may not be appropriately estimated due to active vegetation growth, usually resulting in both higher LAI and crop height. This might also be the case with forest flux tower sites. Regarding short vegetation flux tower sites, as July provided the maximum NDVI for this land cover, the high concentration of greenness might affect soil evaporation computing. The effect on cold pixel selection of flooded conditions and the volume of water bodies, which also add water evaporation and water stress to the common process of ET estimation, may principally take place in April (Figure 6) for wetland flux tower sites.

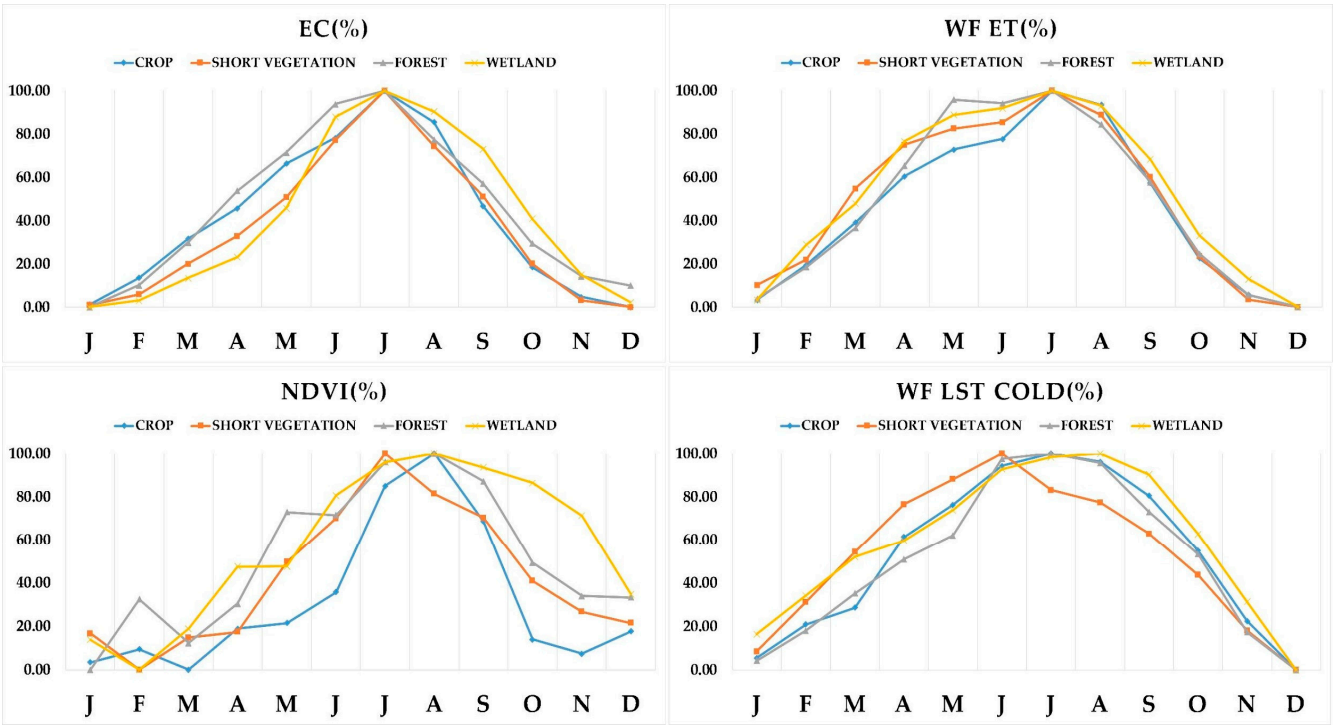

Figure 7. Monthly variation of ET and LST cold values obtained for WF, EC and NDVI at flux tower stations for each land cover. Note that the trend of NDVI at stations was the same as that for NDVI cold pixels.

4.7. Comparison with Previous Studies

Laipelt et al. [30] used SEBAL, the Allen endmember selection method and ERA5 reanalysis meteorological datasets to provide an interactive graphic user interface (GUI) named geeSEBAL. This program was evaluated using 224 Landsat images for 10 flux towers. However, as it is sensitive to different climate conditions and, thus, requires percentile calibration and parameterization for regional applications, the Allen approach may not be applicable worldwide. In addition, geeSEBAL was only optimized for humid tropical and subtropical conditions in Brazil, whereas the Allen method [35] was primarily calibrated for arid climate zones. Kayser et al. [51] assessed the use of land cover maps and homogeneity filters in geeSEBAL for around 300 Landsat data, five flux towers for two land cover types (crop and grassland) and humid subtropical zones in Brazil. In addition to the issues reported above for geeSEBAL, it was also reported in [51] that using indices thresholding alone in Allen had a greater effect on ET estimation performance than adding land cover maps and homogeneity operators for endmember selection, suggesting that homogeneity operators may not be as effective in ET estimation as expected. However, unlike [30,51], the study presented in this paper used a novel endmember selection model which is more reliable than the Allen method and is applicable worldwide. Our model has also been evaluated for more land cover types, climate zones, Landsat images and for a longer time series. Finally, regarding meteorological inputs, the ERA5 data provided in GEE and used by Laipelt et al. [30] have a 3-month latency, thus hampering near real-time ET mapping compared to the GLDAS reanalysis data used in our analyses.

Allen et al. [44] used a coded GUI interface called EEFlux and eeMETRIC, based on the METRIC ET estimation model, the Allen endmember selection method, and CFSv2/CFSR reanalysis meteorological datasets. The most important disadvantage was that use of METRIC as the algorithm requires high-quality or ground-sourced meteorological data [19,21] for accurate ET estimation. In addition to using the Allen method and hence the requirement of local calibration, EEFlux was optimized only for croplands and therefore needs to be evaluated for other vegetated covers [43]. In contrast, in this study, SEBAL together

with land cover maps, enhanced by morphological functions, and GLDAS meteorological data with a better spatial resolution were used to estimate ET for different climates and land covers, and they yielded better error agreement.

An ensemble model, namely OpenET, of six satellite-based procedures (eeMETRIC, geeSEBAL, PT-JPL, ALEXI/DisALEXI, SIMS, and SSEBop) was developed by Melton et al. [74]. OpenET was validated for around 5000 Landsat images of 48 crop flux tower sites. However, OpenET is restricted to and optimized for the western United States, and it only considered arid and semi-arid climates. Furthermore, the values of the error metrics (RMSE, R^2 and Bias) obtained for the OpenET analyses were close to our results. Given the complexity of ensemble models and the high computational time that they require compared to a single method such as SEBAL, the latter can be considered preferable.

Jaafar et al. [29] used a hybrid single-source energy balance (HSEB) model alongside reanalysis meteorological data and gridded land cover maps in GEE to estimate ET for different climates and land covers similar to this study. They reported an RMSE $\approx$ 1.40 and 1.24 mm·day^{-1} for crop and non-crop land covers, respectively, which are higher than those obtained for the same categories in this study of about 0.90 and 1.60 mm·day^{-1}, respectively. Jaafar et al. [29] results were also compared with those of a Two-Source (soil and vegetation) Energy Balance Model (TSEB) in [78] which was given the comparable performance with slightly better metrics in favor of TSEB (RMSE $\approx$ 1.30 mm·day^{-1}). The difference between the performances of [29,78], and our study can be attributed to different SEB models natures and their sensitivities to reanalysis meteorological data. Moreover, using the morphological operators can be effective in the performance.

An exhaustive land cover-based automated model for endmember selection was developed by Bhattarai et al. [5]. However, this method was mostly evaluated in humid climate zones and was not compared with methods being independent of land cover maps. SSEBop was employed by Senay et al. [37] to calculate ET over croplands. In addition to not considering other land covers for ET assessment, they used the National Land Cover Database (NLCD) and TopoWx gridded meteorological datasets that are confined to the conterminous United States. Silva et al. [38] automated SEBAL based on land cover maps and index thresholding. The main disadvantage of that study was using supervised classification algorithms to produce land cover maps being computational- and time-expensive, and user-dependent interference. Moreover, it was applied in Caatinga land cover (white forest) covered by semi-arid climate. Jaafar et al. [39] used the "Water" package being dependent on Z_{om} and LAI. These parameters are empirically estimated and controlled by ground-sourced data in contrast to our approach being independent of in situ data. Saboori et al. [40] implemented a simple method using vegetation and topographic index thresholding as well as endmember distance from flux towers without the need for land cover maps. Although the ET estimates were satisfactory, the study was mostly conducted for crop flux tower sites. Unlike most of the above-mentioned studies that were conducted for a short series of images or a specific method, land cover and climate, this study covered different land covers, endmember selection methods and climatological conditions using a long-term image dataset.

5. Conclusions

This study evaluated the performance of three automated procedures for long-term ET estimation when compared to EC measurements, using SEBAL, GLDAS climatological data with 2270 Landsat images and different climate, land cover and temporal conditions.

The comparison of the performance of the endmember selection methods showed that overall, the WF model yielded the best results, which was followed by the Allen and WOF approaches. It therefore seems that the use of land cover maps alone does not sufficiently explain the differences between the performance of the land cover-based models and the Allen approach. The performance of the Allen approach was found to be dependent on landscape homogeneity and climate type and that of the land cover-based procedures on land cover map quality and the adjacency effect. The use of land cover maps without

reforming processes such as morphological operations may therefore not be as effective for ET mapping as was expected. In addition, the most discriminating criterion to choose an appropriate endmember selection method was humid climate-based classification as opposed to land cover type or the Köppen classification system. More specifically, only humid zones (classified based on the global aridity index system) were extremely sensitive to the endmember selection methods. Using the Allen approach, without land cover maps, for non-humid climate zones was found to provide close ET estimation accuracies when compared to the land cover-based methods. Thus, given the problems of easy access to high quality land cover maps, the Allen method is suggested for non-humid climate zones.

The NDVI and LST endmember values changed with no observable consistent pattern when compared together. Therefore, comparing NDVI and LST value fluctuations is not recommended to determine the most suitable endmembers. In addition, the endmember selection methods and, consequently, ET estimation performance were insensitive to NDVI in contrast to LST, with higher hot than cold pixel selection dependency. Thus, an NDVI filter has little to no impact on ET estimation performance, and it is added to endmember selection only to reduce candidate pixels for the later steps of the process.

Generally, the results indicate that vegetated cover ET reaches its maximum value at the peak of greenness, mostly between June and August when the weakest SEBAL_RMSE results were obtained in the comparison of ET estimates against EC measurements. Forest and crop flux tower sites had, respectively, the worst and best SEBAL_RMSE results. In contrast, the best SEBAL_PBias (below $\pm$10%) results coincided with maximum greenness. Higher LAI, NDVI, canopy cover and plant height values therefore corresponded to weaker SEBAL_RMSE results and more robust SEBAL_PBias performance. Additionally, SEBAL_RMSE results decreased as humidity in the zone increased, with arid and humid zones giving, respectively, the best and worst SEBAL_RMSE results. WF_ET estimation for different land covers and climates generally outperformed the Allen and WOF approaches, as the WF method avoids the need for local calibration of endmember percentiles, which is the major weakness of the Allen method. This was particularly relevant for crop areas in which WF showed the best error performance in ET estimation, which is a factor that is crucial in irrigation planning and crop water management. Overall, the high accuracy in ET estimation yielded by the WF approach allows the successful mapping of ET at regional scale worldwide.

Finally, despite the coarse spatial resolution ($\approx$27 km) and 3 hourly data, the GLDAS meteorological forcing dataset used in the automated ET estimation approaches provided satisfactory estimations. Thus, such datasets can be an appropriate alternative to ground-sourced meteorological data, especially for areas with a lack or scarcity of meteorological data, for regional or continental ET mapping.

Author Contributions: Conceptualization, M.S.; methodology, M.S.; programming, M.S.; validation, all; format analysis, all; writing, M.S., J.C.; supervision, M.S., J.C. All authors have read and agreed to the published version of the manuscript.

Funding: Funds for J.C. participation in this research came from the project "Enhanced remote sensing ET estimates for agricultural drought monitoring through improvements in ET partitioning and heterogeneous crop biophysical parameters retrieval" [PID2021-127345OR-C31] and "Evaluation of the digital twin paradigm applied to precision irrigation (DigiSPAC)" [TED2021-131237B-C21] funded by the Spanish Ministry of Science and Innovation.

Data Availability Statement: Eddy covariance data was provided by AMERIFLUX portal (https://ameriflux.lbl.gov/). Landsat collections and reanalysis meteorological dataset were also secured by Google Earth Engine platform.

Acknowledgments: We would like to thank the Google Earth Engine and AMERIFLUX teams for freely providing satellite data collections and eddy covariance datasets.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

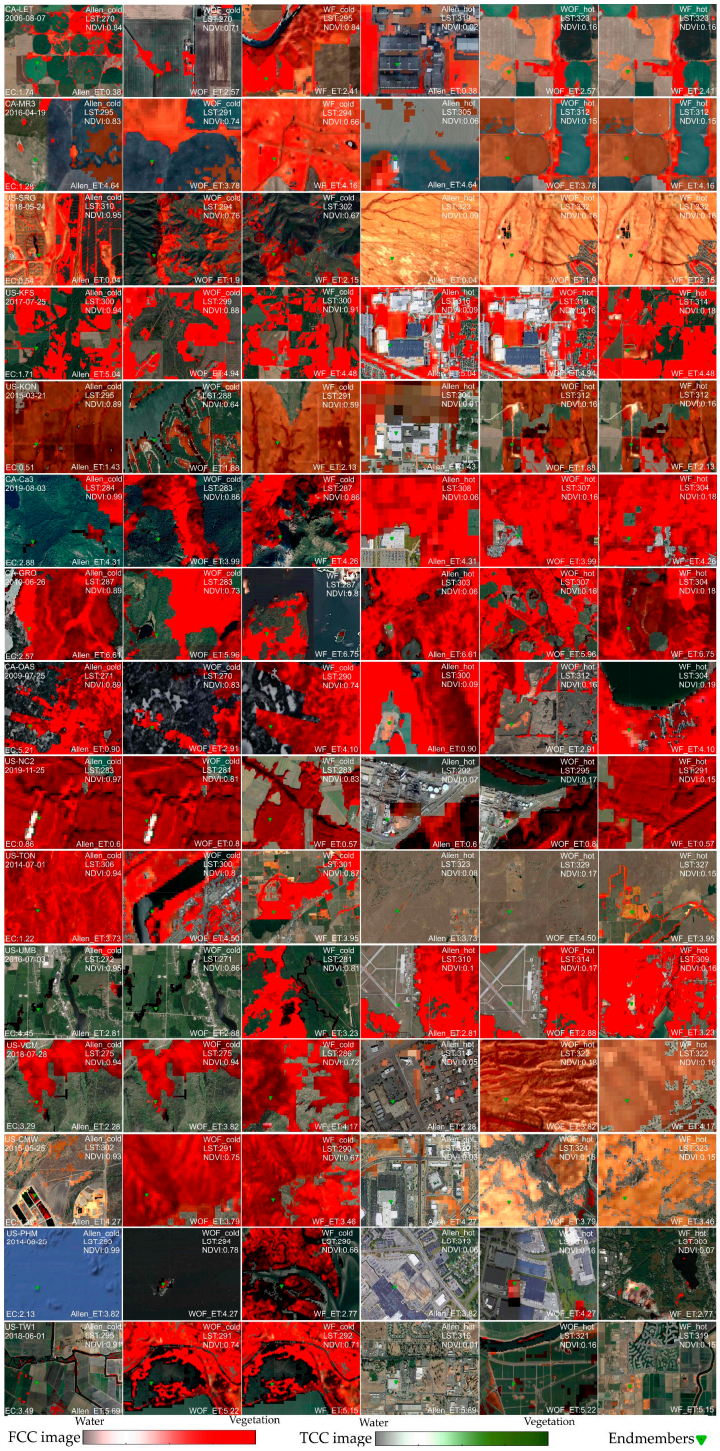

Figure A1. Location and values of cold and hot pixels alongside the statistics of one selected image of non-crop flux towers for the different endmember selection methods, generated using GLDAS meteorological data in SEBAL. Green triangles show endmember location on the intended land cover extracted from FCC, and with the background of Google Earth images. ET estimation performance in each scenario is coded as Allen_ET, WOF_ET, or WF_ET.

Appendix B

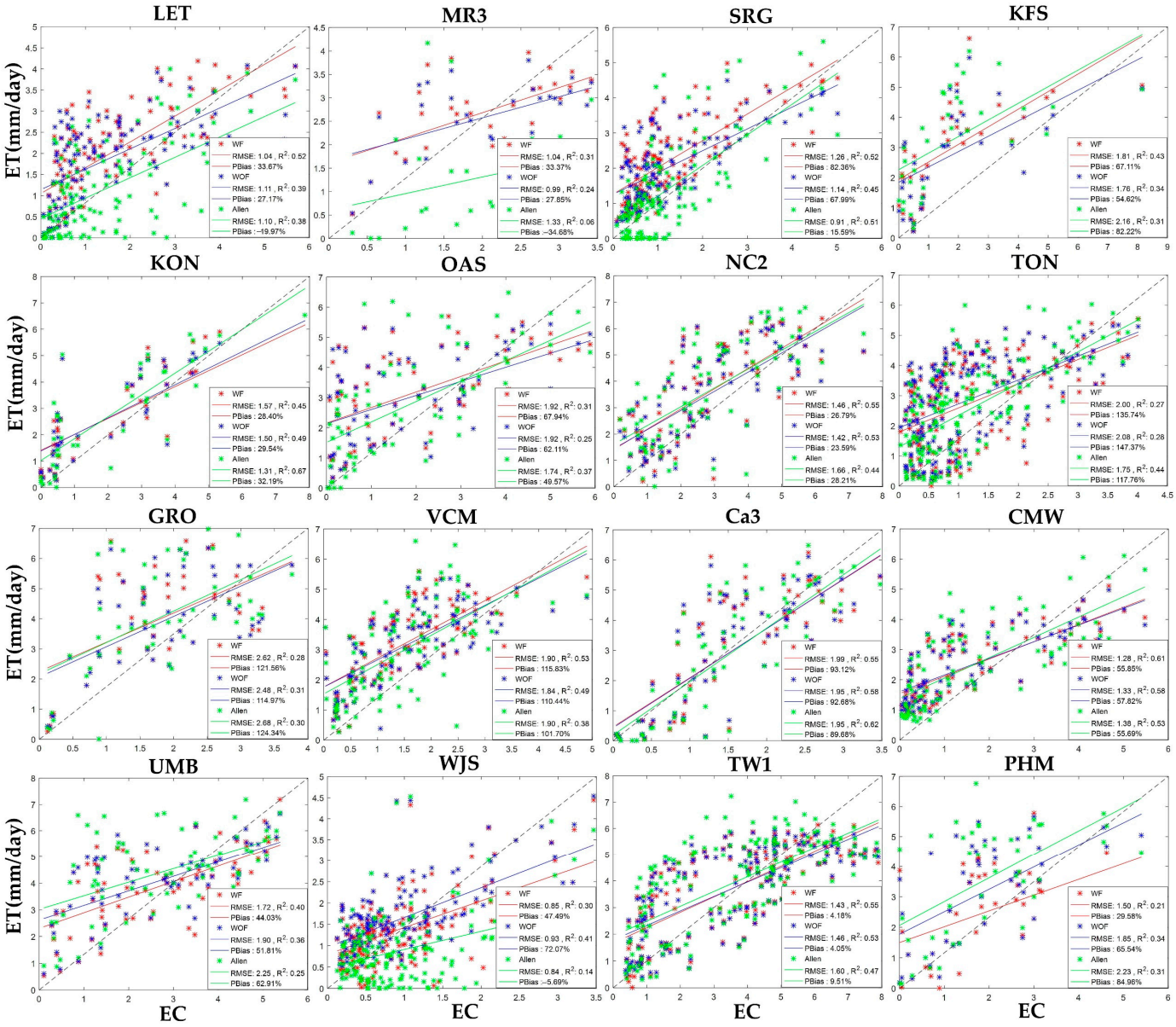

Figure A2. Daily actual ET estimated by the endmember selection models and using the GLDAS datasets vs. EC for the different crop flux tower sites.

Appendix C

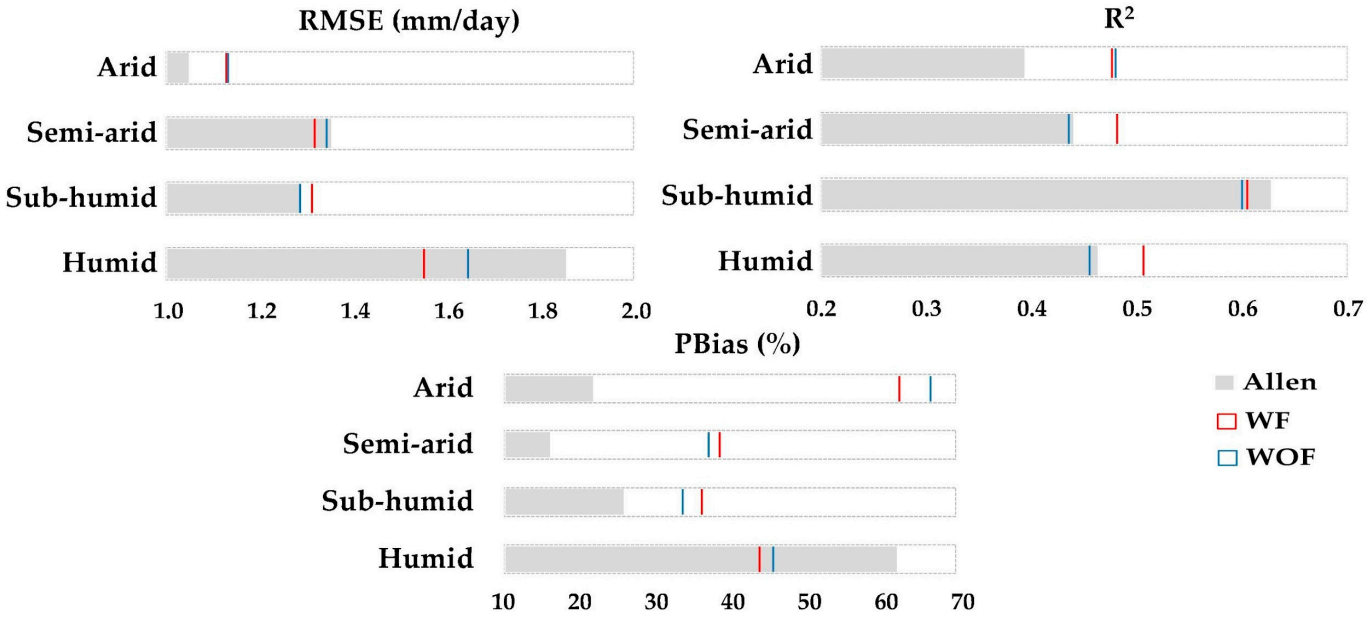

Figure A3. Evaluation of endmember selection procedures based on aridity index climate types.

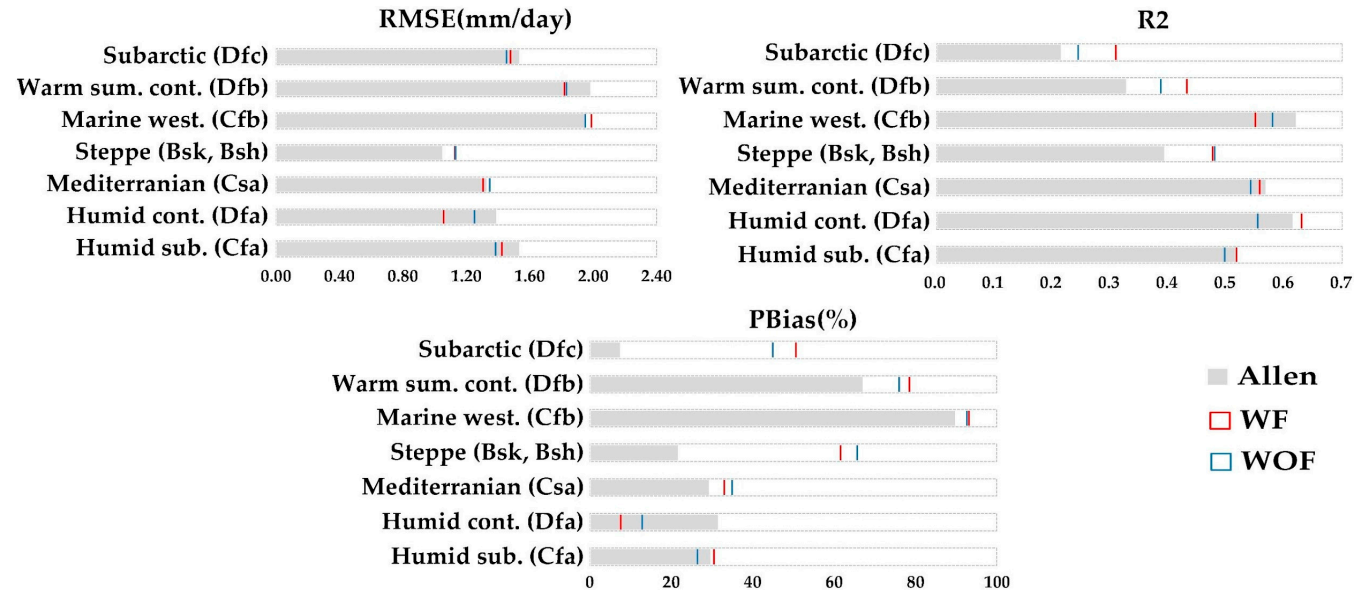

Figure A4. Evaluation of endmember selection procedures based on Köppen climate types.

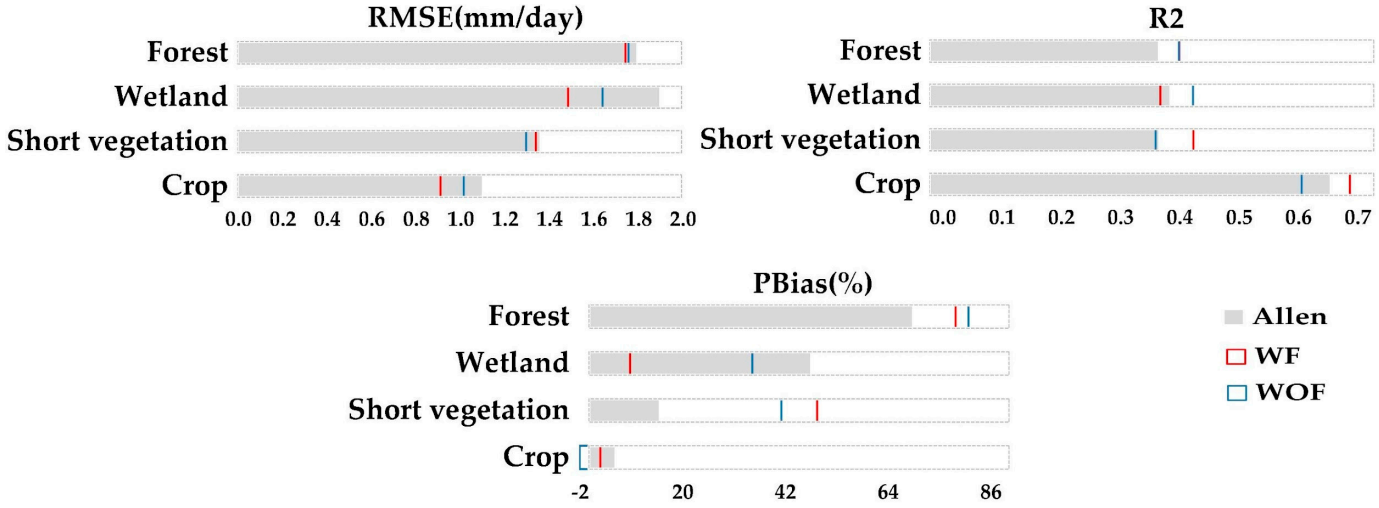

Figure A5. Evaluation of land cover-specific automated ET estimation performance.

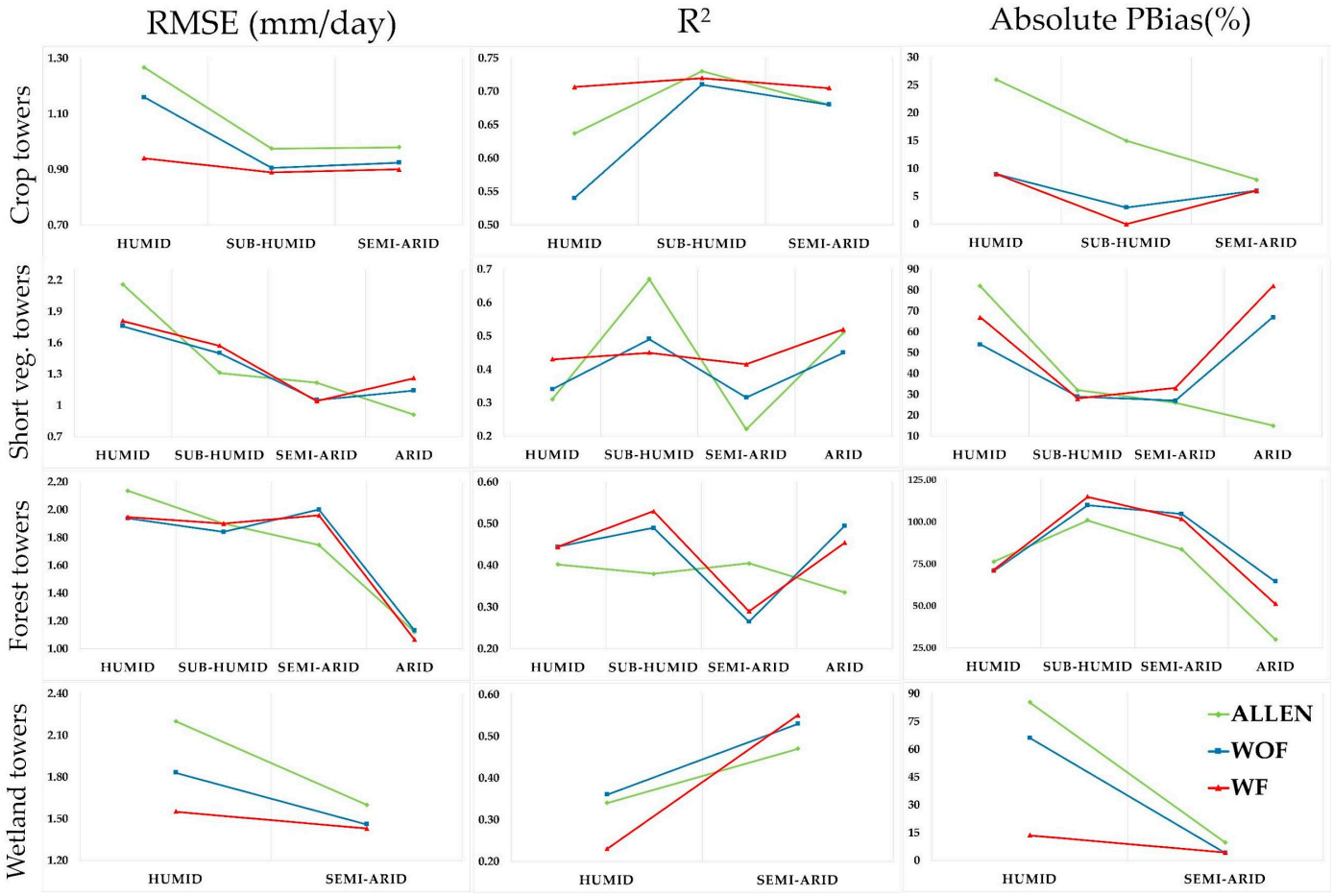

Figure A6. Trend of automated ET estimation performance based on both climate and land cover type.

References

1. Enenkel, M.; See, L.; Bonifacio, R.; Boken, V.; Chaney, N.; Vinck, P.; You, L.; Dutra, E.; Anderson, M. Drought and food security—Improving decision-support via new technologies and innovative collaboration. *Glob. Food Secur.* **2015**, *4*, 51–55. [CrossRef]
2. Alexandratos, N.; Bruinsma, J. *World Agriculture towards 2030/2050: The 2012 Revision*; Agricultural Development Economics Division, Food and Agriculture Organization of the United Nations: Rome, Italy, 2012; pp. 1–147.
3. Mokhtari, A.; Noory, H.; Vazifedoust, M.; Bahrami, M. Estimating net irrigation requirement of winter wheat using model- and satellite-based single and basal crop coefficients. *Agric. Water Manag.* **2018**, *208*, 95–106. [CrossRef]
4. Lobell, D.B.; Asner, G.P.; Ortiz-Monasterio, J.I.; Benning, T.L. Remote sensing of regional crop production in the Yaqui Valley, Mexico: Estimates and uncertainties. *Agric. Ecosyst. Environ.* **2003**, *94*, 205–220. [CrossRef]
5. Bhattarai, N.; Quackenbush, L.J.; Im, J.; Shaw, S.B. A new optimized algorithm for automating endmember pixel selection in the SEBAL and METRIC models. *Remote. Sens. Environ.* **2017**, *196*, 178–192. [CrossRef]
6. Dhungel, R.; Allen, R.G.; Trezza, R.; Robison, C.W.J.M.A. Evapotranspiration between satellite overpasses: Methodology and case study in agricultural dominant semi-arid areas. *Meteorol. Appl.* **2016**, *23*, 714–730. [CrossRef]
7. Eden, U. Drought Assessment by Evapotranspiration Mapping in Twente, the Netherlands. Master's Thesis, University of Twente, Enschede, The Netherlands, 2012.
8. Bartholic, J.; Namken, L.; Wiegand, C.J.A.J. Aerial thermal scanner to determine temperatures of soils and of crop canopies differing in water stress 1. *Agron. J.* **1972**, *64*, 603–608. [CrossRef]
9. Bastiaanssen, W.G.M.; Menenti, M.; Feddes, R.A.; Holtslag, A.A.M. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *J. Hydrol.* **1998**, *212–213*, 198–212. [CrossRef]
10. Su, Z. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrol. Earth Syst. Sci.* **2002**, *6*, 85–100. [CrossRef]
11. Allen, R.G.; Tasumi, M.; Trezza, R. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Model. *J. Irrig. Drain. Eng.* **2007**, *133*, 380–394. [CrossRef]
12. Jiang, L.; Islam, S. A methodology for estimation of surface evapotranspiration over large areas using remote sensing observations. *Geophys. Res. Lett.* **1999**, *26*, 2773–2776. [CrossRef]
13. Roerink, G.J.; Su, Z.; Menenti, M. S-SEBI: A simple remote sensing algorithm to estimate the surface energy balance. *Phys. Chem. Earth Part B Hydrol. Ocean. Atmos.* **2000**, *25*, 147–157. [CrossRef]
14. Almhab, A.A.; Busu, I. Estimation of Evapotranspiration with Modified SEBAL model using landsat-TM and NOAA-AVHRR images in arid mountains area. In Proceedings of the 2008 Second Asia International Conference on Modelling & Simulation (AMS), Kuala Lumpur, Malaysia, 13–15 May 2008; pp. 350–355.
15. Senay, G.B.; Bohms, S.; Singh, R.K.; Gowda, P.H.; Velpuri, N.M.; Alemu, H.; Verdin, J.P. Operational Evapotranspiration Mapping Using Remote Sensing and Weather Datasets: A New Parameterization for the SSEB Approach. *JAWRA J. Am. Water Resour. Assoc.* **2013**, *49*, 577–591. [CrossRef]
16. Mkhwanazi, M.; Chávez, J.L.; Andales, A.A. SEBAL-A: A Remote Sensing ET Algorithm that Accounts for Advection with Limited Data. Part I: Development and Validation. *Remote. Sens.* **2015**, *7*, 15046–15067. [CrossRef]
17. Elkatoury, A.; Alazba, A.A.; Abdelbary, A. Evaluating the performance of two SEB models for estimating ET based on satellite images in arid regions. *Arab. J. Geosci.* **2020**, *13*, 74. [CrossRef]
18. Wagle, P.; Bhattarai, N.; Gowda, P.H.; Kakani, V.G. Performance of five surface energy balance models for estimating daily evapotranspiration in high biomass sorghum. *ISPRS J. Photogramm. Remote. Sens.* **2017**, *128*, 192–203. [CrossRef]
19. Al Zayed, I.S.; Elagib, N.A.; Ribbe, L.; Heinrich, J. Satellite-based evapotranspiration over Gezira Irrigation Scheme, Sudan: A comparative study. *Agric. Water Manag.* **2016**, *177*, 66–76. [CrossRef]
20. Allen, R.; Irmak, A.; Trezza, R.; Hendrickx, J.M.H.; Bastiaanssen, W.; Kjaersgaard, J. Satellite-based ET estimation in agriculture using SEBAL and METRIC. *Hydrol. Process.* **2011**, *25*, 4011–4027. [CrossRef]
21. Bhattarai, N.; Shaw, S.B.; Quackenbush, L.J.; Im, J.; Niraula, R. Evaluating five remote sensing based single-source surface energy balance models for estimating daily evapotranspiration in a humid subtropical climate. *Int. J. Appl. Earth Obs.* **2016**, *49*, 75–86. [CrossRef]
22. Tasumi, M.; Trezza, R.; Allen, R.G.; Wright, J.L. Operational aspects of satellite-based energy balance models for irrigated crops in the semi-arid U.S. *Irrig. Drain. Syst.* **2005**, *19*, 355–376. [CrossRef]
23. Allen, R.G. *Crop Evapotranspiration—Guidelines for Computing Crop Water Requirements*; FAO Irrigation and drainage paper 56; FAO: Rome, Italy, 1998; Volume 56, pp. 60–64.
24. Laipelt, L.; Ruhoff, A.L.; Fleischmann, A.S.; Kayser, R.H.B.; Kich, E.D.M.; da Rocha, H.R.; Neale, C.M.U. Assessment of an Automated Calibration of the SEBAL Algorithm to Estimate Dry-Season Surface-Energy Partitioning in a Forest–Savanna Transition in Brazil. *Remote Sens.* **2020**, *12*, 1108. [CrossRef]
25. Long, D.; Singh, V.P.; Li, Z.-L. How sensitive is SEBAL to changes in input variables, domain size and satellite sensor? *J. Geophys. Res. Earth Surf.* **2011**, *116*, 2011JD016542. [CrossRef]
26. Fick, S.; Hijmans, R.J. WorldClim 2: Nouvelles surfaces climatiques de résolution spatiale de 1 km pour les zones terrestres mondiales. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [CrossRef]

27. Menne, M.J.; Durre, I.; Vose, R.S.; Gleason, B.E.; Houston, T.G. An Overview of the Global Historical Climatology Network-Daily Database. *J. Atmos. Ocean. Technol.* **2012**, *29*, 897–910. [CrossRef]
28. Blankenau, P.A.; Kilic, A.; Allen, R. An evaluation of gridded weather data sets for the purpose of estimating reference evapotranspiration in the United States. *Agric. Water Manag.* **2020**, *242*, 106376. [CrossRef]
29. Jaafar, H.; Mourad, R.; Schull, M. A global 30-m ET model (HSEB) using harmonized Landsat and Sentinel-2, MODIS and VIIRS: Comparison to ECOSTRESS ET and LST. *Remote. Sens. Environ.* **2022**, *274*, 112995. [CrossRef]
30. Laipelt, L.; Kayser, R.H.B.; Fleischmann, A.S.; Ruhoff, A.; Bastiaanssen, W.; Erickson, T.A.; Melton, F. Long-term monitoring of evapotranspiration using the SEBAL algorithm and Google Earth Engine cloud computing. *ISPRS J. Photogramm. Remote. Sens.* **2021**, *178*, 81–96. [CrossRef]
31. Jaafar, H.; Mourad, R. GYMEE: A Global Field-Scale Crop Yield and ET Mapper in Google Earth Engine Based on Landsat, Weather, and Soil Data. *Remote. Sens.* **2021**, *13*, 773. [CrossRef]
32. Biggs, T.W.; Marshall, M.; Messina, A. Mapping daily and seasonal evapotranspiration from irrigated crops using global climate grids and satellite imagery: Automation and methods comparison. *Water Resour. Res.* **2016**, *52*, 7311–7326. [CrossRef]
33. Feng, L. *Sensitivity Analysis of Hot/Cold Pixel Selection in SEBAL Model for ET Estimation*; Virginia Tech: Blacksburg, VA, USA, 2015.
34. Long, D.; Singh, V.P. A modified surface energy balance algorithm for land (M-SEBAL) based on a trapezoidal framework. *Water Resour. Res.* **2012**, *48*, 2011WR010607. [CrossRef]
35. Allen, R.G.; Burnett, B.; Kramber, W.; Huntington, J.; Kjaersgaard, J.; Kilic, A.; Kelly, C.; Trezza, R. Automated Calibration of the METRIC-Landsat Evapotranspiration Process. *JAWRA J. Am. Water Resour. Assoc.* **2013**, *49*, 563–576. [CrossRef]
36. Olmedo, G.; Ortega-Farias, S.; Fonseca-Luengo, D.; de la Fuente-Saiz, D.; Peñailillo, F.F. Water: Actual Evapotranspiration with Energy Balance Models. *R J.* **2017**, *8*, 352–369. [CrossRef]
37. Senay, G.B.; Schauer, M.; Friedrichs, M.; Velpuri, N.M.; Singh, R.K. Satellite-based water use dynamics using historical Landsat data (1984–2014) in the southwestern United States. *Remote. Sens. Environ.* **2017**, *202*, 98–112. [CrossRef]
38. Silva, A.M.; da Silva, R.M.; Santos, C.A.G. Automated surface energy balance algorithm for land (ASEBAL) based on automating endmember pixel selection for evapotranspiration calculation in MODIS orbital images. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *79*, 1–11. [CrossRef]
39. Jaafar, H.H.; Ahmad, F.A. Time series trends of Landsat-based ET using automated calibration in METRIC and SEBAL: The Bekaa Valley, Lebanon. *Remote Sens. Environ.* **2020**, *238*, 111034. [CrossRef]
40. Saboori, M.; Mokhtari, A.; Afrasiabian, Y.; Daccache, A.; Alaghmand, S.; Mousivand, Y. Automatically selecting hot and cold pixels for satellite actual evapotranspiration estimation under different topographic and climatic conditions. *Agric. Water Manag.* **2021**, *248*, 106763. [CrossRef]
41. Venancio, L.P.; Eugenio, F.C.; Filgueiras, R.; da França Cunha, F.; dos Argolo Santos, R.; Ribeiro, W.R.; Mantovani, E.C. Mapping within-field variability of soybean evapotranspiration and crop coefficient using the Earth Engine Evaporation Flux (EEFlux) application. *PloS ONE* **2020**, *15*, e0235620. [CrossRef] [PubMed]
42. de Oliveira Costa, J.; José, J.V.; Wolff, W.; de Oliveira, N.P.R.; Oliveira, R.C.; Ribeiro, N.L.; Coelho, R.D.; da Silva, T.J.A.; Bonfim-Silva, E.M.; Schlichting, A.F. Spatial variability quantification of maize water consumption based on Google EEflux tool. *Agric. Water Manag.* **2020**, *232*, 106037. [CrossRef]
43. Foolad, F.; Blankenau, P.; Kilic, A.; Allen, R.G.; Huntington, J.L.; Erickson, T.A.; Ozturk, D.; Morton, C.G.; Ortega, S.; Ratcliffe, I.; et al. Comparison of the automatically calibrated Google Evapotranspiration Application—EEFlux and the manually calibrated METRIC application. *arXiv* **2018**, arXiv:201807.0040.v1.
44. Allen, R.G.; Morton, C.; Kamble, B.; Kilic, A.; Huntington, J.; Thau, D.; Gorelick, N.; Erickson, T.; Moore, R.; Trezza, R. EEFlux: A Landsat-based evapotranspiration mapping tool on the Google Earth Engine. In Proceedings of the 2015 ASABE/IA Irrigation Symposium: Emerging Technologies for Sustainable Irrigation—A Tribute to the Career of Terry Howell, Sr. Conference Proceedings, Long Beach, CA, USA, 10–12 November 2015; pp. 1–11.
45. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [CrossRef]
46. Trabucco, A.; Zomer, R.J. Global aridity index and potential evapotranspiration (ET0) climate database v2. *CGIAR Consort Spat Inf.* **2018**, *10*, m9.
47. Hargreaves, G.H.; Samani, Z.A. Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.* **1985**, *1*, 96–99. [CrossRef]
48. Zomer, R.J.; Trabucco, A.; Bossio, D.A.; Verchot, L.V. Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agric. Ecosyst. Environ.* **2008**, *126*, 67–80. [CrossRef]
49. Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dilley, R.D., Jr.; Beckmann, T.; Schmidt, G.L.; Dwyer, J.L.; Hughes, M.J.; Laue, B. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* **2017**, *194*, 379–390. [CrossRef]
50. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
51. Kayser, R.H.; Ruhoff, A.; Laipelt, L.; Kich, E.D.M.; Roberti, D.R.; Souza, V.D.A.; Rubert, G.C.D.; Collischonn, W.; Neale, C.M.U. Assessing geeSEBAL automated calibration and meteorological reanalysis uncertainties to estimate evapotranspiration in subtropical humid climates. *Agric. For. Meteorol.* **2022**, *314*, 108775. [CrossRef]
52. Shuttleworth, W.J. *Terrestrial Hydrometeorology*; John Wiley & Sons: New York, NY, USA, 2012.

53. Buchhorn, M.; Lesiv, M.; Tsendbazar, N.-E.; Herold, M.; Bertels, L.; Smets, B. Copernicus Global Land Cover Layers—Collection 2. *Remote. Sens.* **2020**, *12*, 1044. [CrossRef]
54. Zanaga, D.; Van De Kerchove, R.; De Keersmaecker, W.; Souverijns, N.; Brockmann, C.; Quast, R.; Wevers, J.; Grosu, A.; Paccini, A.; Vergnaud, S.; et al. ESA WorldCover 10 m 2020 v100. In *ESA WorldCover Project 2020/Contains Modified Copernicus Sentinel Data (2020) Processed by ESA WorldCover Consortium*; Zenodo: Genève, Switzerland, 2021. [CrossRef]
55. Allen, R.; Tasumi, M.; Trezza, R.; Waters, R.; Bastiaanssen, W. Surface Energy Balance Algorithm for Land (SEBAL)–Advanced training and Users Manual. *Kimberly Ida. Implement.* **2002**, *1*, 98.
56. Bastiaanssen, W. SEBAL-based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey. *J. Hydrol.* **2000**, *229*, 87–100. [CrossRef]
57. Bastiaanssen, W.G.M. *Regionalization of Surface Flux Densities and Moisture Indicators in Composite Terrain: A Remote Sensing Approach under Clear Skies in Mediterranean Climates*; Wageningen University and Research: Wageningen, The Netherlands, 1995.
58. Wang, X.-G.; Wang, W.; Huang, D.; Yong, B.; Chen, X. Modifying SEBAL Model Based on the Trapezoidal Relationship between Land Surface Temperature and Vegetation Index for Actual Evapotranspiration Estimation. *Remote Sens.* **2014**, *6*, 5909–5937. [CrossRef]
59. Businger, J.A.; Wyngaard, J.C.; Izumi, Y.; Bradley, E.F. Flux-profile rrelationship in the atmospheric surface layer. *J. Atmos. Sci.* **1971**, *28*, 181–189. [CrossRef]
60. Tasumi, M.; Allen, R.G.; Trezza, R. At-Surface Reflectance and Albedo from Satellite for Operational Calculation of Land Surface Energy Balance. *J. Hydrol. Eng.* **2008**, *13*, 51–63. [CrossRef]
61. Artis, D.A.; Carnahan, W.H. Survey of emissivity variability in thermography of urban areas. *Remote Sens. Environ.* **1982**, *12*, 313–329. [CrossRef]
62. Sobrino, J.A.; Jiménez-Muñoz, J.C.; Paolini, L. Land surface temperature retrieval from LANDSAT TM 5. *Remote Sens. Environ.* **2004**, *90*, 434–440. [CrossRef]
63. Ji, L.; Zhang, L.; Wylie, B. Analysis of Dynamic Thresholds for the Normalized Difference Water Index. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1307–1317. [CrossRef]
64. Tasumi, M. Application of the SEBAL methodology for estimating consumptive use of water and stream flow depletion in the Bear River Basin of Idaho through remote sensing. *Append. C A Step By Step Guide Run. SEBAL* **2000**.
65. Deng, C.; Wu, C. BCI: A biophysical composition index for remote sensing of urban environments. *Remote Sens. Environ.* **2012**, *127*, 247–259. [CrossRef]
66. Fujimaki, H.; Shiozawa, S.; Inoue, M. Effect of salty crust on soil albedo. *Agric. For. Meteorol.* **2003**, *118*, 125–135. [CrossRef]
67. Guha, S.; Govil, H. Land surface temperature and normalized difference vegetation index relationship: A seasonal study on a tropical city. *SN Appl. Sci.* **2020**, *2*, 1661. [CrossRef]
68. Kumar, D.; Shekhar, S. Statistical analysis of land surface temperature–vegetation indexes relationship through thermal remote sensing. *Ecotoxicol. Environ. Saf.* **2015**, *121*, 39–44. [CrossRef]
69. Yue, W.; Xu, J.; Tan, W.; Xu, L. The relationship between land surface temperature and NDVI with remote sensing: Application to Shanghai Landsat 7 ETM+ data. *Int. J. Remote. Sens.* **2007**, *28*, 3205–3226. [CrossRef]
70. Sun, D.; Kafatos, M. Note on the NDVI-LST relationship and the use of temperature-related drought indices over North America. *Geophys. Res. Lett.* **2007**, *34*, 2007GL031485. [CrossRef]
71. Karnieli, A.; Agam, N.; Pinker, R.T.; Anderson, M.; Imhoff, M.L.; Gutman, G.G.; Panov, N.; Goldberg, A. Use of NDVI and Land Surface Temperature for Drought Assessment: Merits and Limitations. *J. Clim.* **2010**, *23*, 618–633. [CrossRef]
72. Yapo, P.O.; Gupta, H.V.; Sorooshian, S. Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data. *J. Hydrol.* **1996**, *181*, 23–48. [CrossRef]
73. Kalma, J.D.; McVicar, T.R.; McCabe, M.F. Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surv. Geophys.* **2008**, *29*, 421–469. [CrossRef]
74. Melton, F.S.; Huntington, J.; Grimm, R.; Herring, J.; Hall, M.; Rollison, D.; Erickson, T.; Allen, R.; Anderson, M.; Fisher, J.B.; et al. Openet: Filling a critical data gap in water management for the western united states. *JAWRA J. Am. Water Resour. Assoc.* **2021**. [CrossRef]
75. Elnashar, A.; Wang, L.; Wu, B.; Zhu, W.; Zeng, H. Synthesis of global actual evapotranspiration from 1982 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 447–480. [CrossRef]
76. Qu, Y.; Zhuang, Q. Modeling leaf area index in North America using a process-based terrestrial ecosystem model. *Ecosphere* **2018**, *9*, e02046. [CrossRef]
77. Levis, S.; Bonan, G.B.; Kluzek, E.; Thornton, P.; Jones, A.; Sacks, W.J.; Kucharik, C. Interactive Crop Management in the Community Earth System Model (CESM1): Seasonal Influences on Land–Atmosphere Fluxes. *J. Clim.* **2012**, *25*, 4839–4859. [CrossRef]
78. Jaafar, H.; Mourad, R.; Kustas, W.; Anderson, M. A Global Implementation of Single-and Dual-Source Surface Energy Balance Models for Estimating Actual Evapotranspiration at 30-m Resolution using Google Earth Engine. *Water Resour. Res.* **2022**, *58*, e2022WR032800. [CrossRef]