Improving field-scale crop actual evapotranspiration monitoring with Sentinel-3, Sentinel-2, and Landsat data fusion

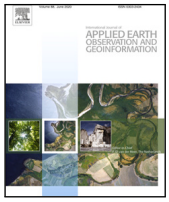
Radoslaw Guzinski, Héctor Nieto, Rubén Ramo Sánchez, Juan Manuel Sánchez, Ihab Jomaa, Rim Zitouna-Chebbi, Olivier Roupsard, Ramón López-Urrea

International Journal of Applied Earth Observation and Geoinformation 125 (2023) 103587

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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



Improving field-scale crop actual evapotranspiration monitoring with Sentinel-3, Sentinel-2, and Landsat data fusion

Radoslaw Guzinski [a,*], Héctor Nieto [b], Rubén Ramo Sánchez [c], Juan Manuel Sánchez [d], Ihab Jomaa [e], Rim Zitouna-Chebbi [f], Olivier Roupsard [g,h,i], Ramón López-Urrea [j]

[a] DHI, Denmark
[b] Institute of Agricultural Sciences (ICA-CSIC), Spain
[c] COMPLUTIG, Spain
[d] University of Castilla-La Mancha, Spain
[e] Lebanese Agricultural Research Institute, Lebanon
[f] Institut National de Recherche en Génie Rural, Eaux et Forêts, Tunisia
[g] CIRAD, UMR Eco&Sols, Senegal
[h] Eco&Sols, Univ Montpellier, CIRAD, INRAE, Institut Agro, IRD, France
[i] LMI IESOL, Centre IRD-ISRA de Bel Air, Senegal
[j] Desertification Research Centre (CIDE), CSIC-UV-GVA, Valencia, Spain

ARTICLE INFO

Keywords:
Irrigated agriculture
Remote sensing
Surface energy balance
Land surface temperature

ABSTRACT

One of the primary applications of satellite Land Surface Temperature (LST) observations lies in their utilization for modeling of actual evapotranspiration (ET) in agricultural crops, with the primary goals of monitoring and enhancing irrigation practices and improving crop water use productivity, as stipulated by Sustainable Development Goal (SDG) indicator 6.4.1. Evapotranspiration is a complex and dynamic process, both temporally and spatially, necessitating LST observations with high spatio-temporal resolution. Presently, none of the existing spaceborne thermal sensors can provide quasi-daily field-scale LST observations, prompting the development of methods for data fusion (thermal sharpening) of observations from various shortwave and thermal sensors to meet this spatio-temporal requirement. Previous research has demonstrated the effectiveness of combining shortwave-multispectral Sentinel-2 observations with thermal-infrared Sentinel-3 observations to derive daily, field-scale LST and ET estimates. However, these studies also highlighted limitations in capturing the distinct thermal contrast between cooler LST in irrigated agricultural areas and the hotter, adjacent dry regions. In this study, we aim to address this limitation by incorporating information on thermal spatial variability observed by Landsat satellites into the data fusion process, without being constrained by infrequent or cloudy Landsat thermal observations and while retaining the longwave radiance emission captured by the Sentinel-3 thermal sensor at its native resolution. Two approaches are evaluated, both individually and as a complementary combination, and validated against *in situ* LST measurements. The best performing approach, which leads to reduction in root mean square error of up to 1.5 K when compared to previous research, is subsequently used to estimate parcel-level actual evapotranspiration. The ET modeling process has also undergone various improvements regarding the gap-filling of input and output data, input datasets and code implementation. The resulting ET is validated using lysimeters and eddy covariance towers in Spain, Lebanon, Tunisia, and Senegal resulting in minimal overall bias (systematic underestimation of less than 0.07 mm/day) and a low root mean square error (down to 0.84 mm/day) when using fully global input datasets. The enhanced LST sharpening methodology is sensor agnostic and should remain relevant for the upcoming thermal missions while the accuracy of the modeled ET fluxes is encouraging for further utilization of observations from Sentinel satellites, and other Copernicus data, for monitoring SDG indicator 6.4.1.

* Corresponding author.
E-mail address: rmgu@dhigroup.com (R. Guzinski).

https://doi.org/10.1016/j.jag.2023.103587
Received 22 September 2023; Received in revised form 4 November 2023; Accepted 24 November 2023
Available online 29 November 2023

1. Introduction

Accurate modeling of actual evapotranspiration (ET) with satellite data requires observations in shortwave optical and thermal infrared spectral domains (García-Santos et al., 2022). For robust estimates, those observations should match the spatial resolution of the dominant landscape features (Kustas et al., 2004; Burchard-Levine et al., 2021) and be frequent enough to capture the main temporal dynamics. In agricultural landscapes, where the size of fields can be a couple of hectares and surface conditions can change from day to day, this implies a spatial resolution of tens of meters and preferably a daily temporal resolution. As none of the currently operational satellite sensors can provide such data, especially in the thermal infrared domain, various data fusion methods have been developed to meet this spatio-temporal requirement (Pu and Bonafoni, 2023). For the purpose of modeling ET, those data fusion methods can be split into two categories: semi-empirical shortwave-thermal land surface temperature (LST) sharpening methods exploiting physical relations between shortwave data with higher spatial resolution and LST with lower spatial but same or higher temporal resolutions (e.g. Gao et al., 2012; Bisquert et al., 2016a), and disaggregation approaches which are usually more complex and require the fusion to be performed at the level of low and high spatial resolution ET products which were derived with low and high spatial resolution LST respectively (e.g Norman et al., 2003; Guzinski et al., 2014). When considering the Copernicus Sentinel satellites the second category is generally not applicable since that constellation does not include a high resolution thermal sensor. However, combining data from Sentinel-2 satellites, providing shortwave observations with 10–20 m resolution and 5-day geometric revisit time at the equator, and Sentinel-3 satellites, acquiring daily LST observations with nominal resolution of 1 km, in a LST sharpening approach was previously shown to produce inputs highly suitable for field-scale ET estimations (Guzinski and Nieto, 2019; Guzinski et al., 2020, 2021).

One of the promising applications of satellite-derived ET is consistent estimation of actual evapotranspiration from field to continental scales which is critical for reliable monitoring and reporting of Sustainable Development Goals (SDG) indicator 6.4.1 (Change in Water Use Efficiency) (O'Connor et al., 2020). With this in mind the Food and Agriculture Organization of the United Nations (FAO) is running an online portal called WaPOR (wapor.apps.fao.org/, last accessed: 27.06.2023), which provides access to dekadal (10-day) ET estimates at spatial scales ranging from 30 m (referred to as Level 3 data by WaPOR) through 100 m (Level 2 data) to 250 m (Level 1 data) derived from satellite observations (FRAME Consortium, 2020). As of version 2 of the WaPOR portal (launched in June 2019 and the latest publicly available version as of June 2023) the satellite data was based on observations from Landsat, PROBA-V (replaced by Sentinel-2 since 2020) and Terra and Aqua satellites. A recent study (Guzinski et al., 2021) demonstrated the use of Copernicus data, especially ERA5 meteorological data (Hersbach et al., 2020) from Copernicus Climate Change Service and combined observations from Sentinel-2 (S2) and Sentinel-3 (S3) satellites, for consistent estimation of ET from 20 m to 300 m spatial resolutions with WaPOR-like specifications. That research showed that Copernicus data, together with advanced data fusion methods and physically-based models, greatly improved the accuracy of ET retrieval at scales in which Terra and Aqua (MODIS sensor) data was previously used. This guided FAO in the evolution of the WaPOR portal towards version 3 (https://data.apps.fao.org/wapor, last accessed: 25.10.2023), and especially in the adaption of the thermal-optical data fusion method and the use of ERA5 data (https://bitbucket.org/cioapps/wapor-et-look/wiki/Home, last accessed: 14.07.2023).

However, Guzinski et al. (2021) also exposed an inherent limitation of the thermal-shortwave data fusion approach, which is used to derive high-resolution LST by enhancing spatial resolution of S3 LST observations (acquired at around 1 km resolution) using S2 optical observations (acquired at 20 m spatial resolution). This is on one hand the inability of detecting short-term stress/wetting events, due to different physiological response time between shortwave and thermal bands. And on the other hand, the inability to reproduce very low LST values present in irrigated fields, thus leading to underestimation of ET from those fields. The reason for this is the difficulty in predicting values which are far beyond the LST range of the original Sea and Land Surface Thermal Radiometer (SLSTR) LST. The most extreme LST values will not be present in the SLSTR data since LST from different landscape features aggregate within the 1 km pixels. Other studies also reported similar findings when evaluating the same data fusion approach (Bellvert et al., 2020). The underestimation of ET in irrigated fields could represent a critical issue when it comes to SDG indicator 6.4.1 reporting since irrigated agriculture is a focus area of this indicator.

The main purpose of the current study is to improve the ability of S2–S3 data fusion approach to reproduce low LST values present in irrigated fields by incorporating Landsat LST into the data fusion methodology. It is however important to ensure that the inclusion of Landsat LST, with its lower revisit time, does not compromise the capability of the data fusion approach to produce high-resolution LST on all dates on which cloud-free S3 observations were acquired. In addition, energy must be conserved when re-aggregating sharpened LST back to S3 scale to ensure physical plausibility and consistency in ET retrieval across spatial scales. To our knowledge such fusion of Sentinel-2 shortwave observations with Sentinel-3 and Landsat thermal observations to derive high resolution LST estimates has not yet been reported in literature. Other smaller improvements to the sharpening approach were also developed and evaluated. Field measurements of LST from southeastern Spain from 2018 and 2019 were used for directly validating high-resolution LST produced through data fusion.

Similarly to Guzinski et al. (2021), the sharpened LST is subsequently used in two ET models: TSEB-PT (Kustas et al., 2016) and ET-Look (Bastiaanssen et al., 2012). In the previous research the validation of Copernicus-based ET was focused on sites in semi-arid Mediterranean climate. In this study the validation was extended to a site in Senegal with a distinct dry and wet seasons, with the interest of evaluating Copernicus data availability and performance of the proposed method also in climates with more prevalent cloud cover. This necessitates evolution of the input and output data gap-filling methods to allow ET modeling in cloudy climates and periods. Finally, in this study new ancillary data sources (in particular land-cover map) were introduced to mitigate some of the limitations observed in Guzinski et al. (2021).

Section 2 of this paper describes the method for modeling ET with Copernicus-based inputs with the focus on changes implemented in the current study. In Section 3 the field validation sites are introduced. The validation of sharpened LST against field measurements from southeastern Spain and the validation of ET estimates produced in this study and obtained from the WaPOR portal against *in situ* measurements are presented in Section 4. Section 5 interprets those results while Section 6 contains recommendations and conclusions reached as a result of this study.

2. Methods

The methods used to model ET with Copernicus inputs are as described in Guzinski et al. (2021) and summarized in the paragraph below, apart from the modifications presented in Sections 2.1–2.4.

Atmospherically corrected and cloud masked shortwave multispectral observations from Sentinel-2 satellites are used to determine surface biophysical properties (green and total leaf area index, albedo, etc.) at 20 m spatial resolution. The thermal infrared Sensor on board Sentinel-3 satellites provides LST observations that are sharpened from the native resolution of around 1 km to 20 m using Data Mining Sharpener (Gao et al., 2012; Guzinski and Nieto, 2019) with Sentinel-2 reflectance and digital elevation model (DEM) derived parameters used as explanatory variables. Meteorological forcing comes from the ERA-5 reanalysis (Hersbach et al., 2020) as provided by Copernicus

Climate Change Service and is subsequently corrected for topographical effects using a DEM at 300 m resolution before being resampled to 20 m resolution using bilinear interpolation. Finally, surface parameters that cannot be directly determined from Sentinel-2 data (e.g. vegetation height, canopy clumpiness or leaf angle orientation) are set using a land-cover map and a look-up table. All those data are used as inputs into two ET models: TSEB-PT (Kustas et al., 2016) and ETLook (Bastiaanssen et al., 2012). The output daily ET estimates are gap-filled and aggregated into 10-day (dekadal) composites in order to correspond to WaPOR ET product.

2.1. Improvements to thermal sharpening

One of the major challenges in modeling ET with Copernicus data is overcoming the limitation of low spatial resolution thermal data. In Guzinski and Nieto (2019), Guzinski et al. (2020, 2021) it was demonstrated that good results can be obtained when modeling high-resolution ET using Sentinel-3 LST sharpened with the Data Mining Sharpener (DMS) approach. However, certain limitations were also observed and they are addressed with the methodological modifications presented in the sections below.

2.1.1. Modifications to data mining sharpener

Two small modifications were incorporated in the Python implementation of DMS (https://github.com/radosuav/pyDMS, last accessed: 28.06.2023) used in this and previous studies. Firstly, in the DMS method a linear regression model is applied to all data points falling within one regression tree leaf node (see Gao et al., 2012; Guzinski and Nieto, 2019). This linear regression model was changed from Bayesian Ridge (https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.BayesianRidge.html, last accessed: 28.06.2023) to Ridge (https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Ridge.html#sklearn.linear_model.Ridge, last accessed: 28.06.2023). The original model was supposed to be more robust but it was found that it often resulted in linear interpolation with very low coefficient of determination and slope close to zero. This resulted in very limited interpolation and extrapolation of LST values assigned to a specific leaf node.

The second change, concerns the selection of pixels used for training the DMS model. The original approach was to select 80% of the most homogeneous Sentinel-3 resolution reflectance pixels (i.e. pixels from Sentinel-2 reflectance image resampled to Sentinel-3 grid) based on coefficient of variation (CV - standard deviation over mean) of the Sentinel-2 pixels falling within each Sentinel-3 pixel (see Gao et al., 2012; Guzinski et al., 2020). The inverse of CV was also used as a weighing factor during the model training, penalizing heterogeneous pixels and advancing homogeneous ones. However, it was observed that agricultural areas, and in particular irrigated agriculture in semi-arid climate, often exhibited the largest heterogeneity. Therefore those areas were often excluded from the model training set, even though they were of most interest in the later process of ET modeling. Therefore, the methodology was modified to use all available pixels during the training of the DMS model but to penalize the most heterogeneous 20% by further dividing their already low weight by a factor of two. This is expected to improve model fit in agricultural areas while not significantly degrading model performance in other areas through the introduction of extra noise in the training data.

2.1.2. Thermal contrast enhancement

One of the limitations of sharpening low-resolution thermal data is the dampened range of temperatures present in the sharpened high-resolution representation of LST (Guzinski and Nieto, 2019). This is because regression models generally output data with a similar value range to the input data and the input LST is dampened due to spatial aggregation at the 1 km spatial resolution. For example, a cold irrigated agricultural parcel next to a hot bare field falling within the same Sentinel-3 LST pixel would result in the sensed thermal radiance of that pixel being area weighted average of those two extremes. The DMS tries to address this limitation with a limited extrapolation allowed within the linear interpolation of values falling within one regression tree leaf node (see Gao et al., 2012 for details). However, it is still observed that the sharpened LST is overestimating the coldest temperatures and underestimating the hottest ones.

Another possibility of addressing this issue is to use ancillary dataset of higher resolution LST observation to enhance the contrast of the sharpened LST. In this study we used Landsat −7, −8 and −9 LST (acquired at spatial resolution between 60 m and 100 m but provided as a 30 m product) from Collection 2 L2 A product (https://www.usgs.gov/landsat-missions/landsat-collection-2-surface-temperature, last accessed: 14.07.2023) as high-resolution LST. Additional cloud masking was performed by removing any values 2 K lower than the air temperature. There are three [options] of using this ancillary data:

1. Using the high-resolution LST directly in the ET models instead of the sharpened LST
2. Using the high-resolution LST, either directly or as a spatial or temporal statistics (e.g. standard deviation) as an explanatory variable in the DMS model
3. Using the high-resolution LST during the post-processing of the DMS-produced sharpened LST

The first option was not suitable for the current study since it would lead to inconsistencies between the dates and areas with and without high-resolution LST observations and also would lead to lack of consistency between ET and LST products at 20 m and 300 m resolutions. The second and third options were implemented as described below and the results validated against field observations of LST in an agricultural area in southeastern Spain, with the results presented in Section 4.1.

LST exhibits high temporal variability due to its dependence on multiple factors including weather conditions, surface biophysical properties (e.g. vegetation cover), and soil water availability. At the same time, Landsat satellites do not provide daily observations of the land surface. Therefore, to include high-resolution LST as explanatory variable in the DMS model, the temporal statistics were calculated first under the assumptions that those are more robust to temporal changes than the absolute values of the LST observations. Those statistics consisted of standard deviation, minimum and maximum values observed in each high-resolution pixel over a period of ± 16 days centered on the date of Sentinel-3 acquisition which was undergoing sharpening. This compositing period was chosen to be as short as possible to limit the uncertainty due to the temporal variability of LST while still being sufficiently long to provide a reasonable probability of multiple cloud-free LST observations. Given the 16-day geometric revisit time of each Landsat satellite at the equator and the availability of two Landsat satellites (7 & 8 or 8 & 9) at a given time, this compositing period results in at least 5 possible observations (both cloudy and cloud-free) with more closer to the poles. The three statistics were selected to capture the expected range and variability of possible LST values while not to excessively increasing the number of explanatory variables used in DMS which would impact model complexity and execution time. They were used in addition to standard explanatory variables (Sentinel-2 reflectance, digital elevation model and solar illumination conditions). A limitation of this approach is that no sharpened LST values are produced in pixels in which no Landsat LST was observed during the compositing period.

Finally, the procedure of using high-resolution LST during the post-processing of sharpened LST to enhance the thermal contrast (spatial variability) within the sharpened scene is shown in Fig. 1. In the first step, a closest-pixel temporal aggregation of Landsat LST was performed using Landsat observations from ± 16 days centered on the date of Sentinel-3 acquisition. This produced a high-resolution LST

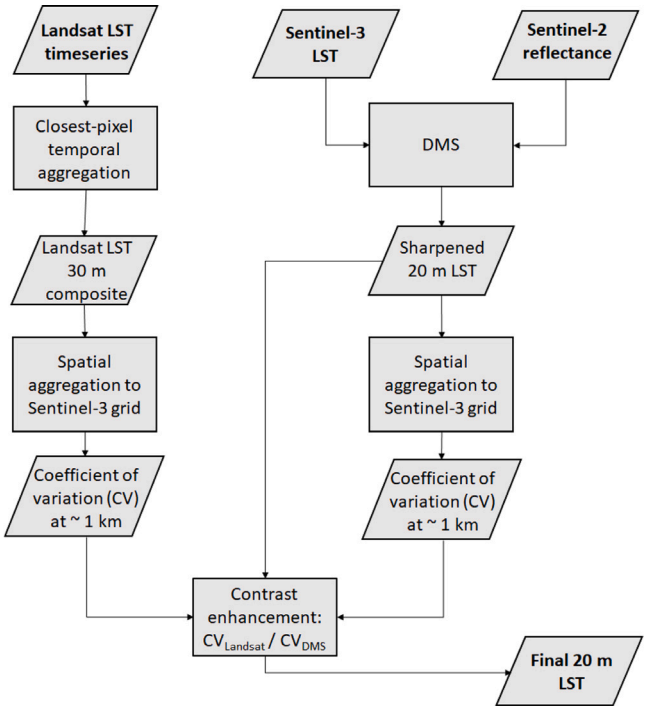


Fig. 1. Outline of the method for enhancing the thermal contrast (spatial variability) within a sharpened LST scene using the thermal contrast present within Landsat observations.

composite image with gaps due to clouds minimized. This composite was then resampled to the grid of the Sentinel-3 image being sharpened by computing a 1 km focal coefficient of variation from all Landsat 30 m LST pixels falling within each Sentinel-3 1 km pixel ($CV_{Landsat}$). Similarly, CV of the DMS sharpened LST (CV_{DMS}) was also calculated. Based on this, the contrast enhancement scaling factor S was calculated as

$$S = \max\left(\frac{CV_{Landsat}}{CV_{DMS}}, 1\right) \quad (1)$$

The scaling factor was resampled to Sentinel-2 resolution using bilinear interpolation, to avoid S3-pixel blocks, before being applied to the sharpened LST image:

$$LST_{out} = LST_{DMS,S3\_mean} + S * (LST_{DMS} - LST_{DMS,S3\_mean}) \quad (2)$$

where LST_{out} is the final sharpened and contrast enhanced high-resolution LST, LST_{DMS} is the LST after sharpening with DMS and $LST_{DMS,S3\_mean}$ is the mean of all LST_{DMS} pixels falling within original resolution Sentinel-3 LST pixel.

By applying the contrast enhancement in this way, similar to a standard rescaling correction (Yilmaz and Crow, 2013), the thermal energy of the original Sentinel-3 pixel is conserved while the within-pixel cooler areas become cooler, and warmer become warmer based on the spatial variability of Landsat LST. It can be assumed that spatial variability of Landsat LST within a 1 km area is more temporarily stable than the individual LST pixel values and therefore the Landsat acquisition does not have to fall on the same date as Sentinel-3 acquisition. Also if no valid Landsat LST observations are present during the compositing period for a given Sentinel-3 pixel then the contrast enhancement is not applied to that pixel (i.e. $S = 1$) and the output of DMS is used directly for this pixel.

2.2. Gap-filling of input and output data

In Guzinski et al. (2021) the focus was on semi-arid areas in Lebanon, Tunisia and Spain. Those Mediterranean areas are characterized by infrequent cloud cover for the majority of the year. Therefore, simple gap-filling approaches were chosen for both input and output data (temporal closest pixel composite and ratio of actual to reference ET respectively). In this research, additional focus area in the Senegal was added. In this area the cloud cover is more prevalent during the rainy season and therefore other gap-filling approaches need to be explored.

2.2.1. Gap-filling of input biophysical parameters

The gap-filling of input data focuses producing continuous land biophysical traits using Sentinel-2 observations. Those traits include leaf area index, spectral reflectance of leaves and soil, and fraction of LAI that is green. In Guzinski et al. (2021) a temporal closest pixel composite of biophysical parameters was created from all S2 observations falling within ± 10 days of the date on which ET was to be modeled. In this study this was modified to be based on linear interpolation of closest available observation before and after the central date within a ± 30 day window. With Sentinel-2 observations having a geometric revisit time of 5 days at the equator this results at least 6 observations before and after the central date in which a cloud-free pixel could be found. While the 60 day window seems large, changes in biophysical traits follow the phenological cycle and therefore usually undergo smooth transitions.

2.2.2. Gap-filling of evapotranspiration outputs

Despite gap-filling of the Sentinel-2 based inputs, gaps can still occur in the output ET maps due to clouds during the Sentinel-3 LST acquisition. In Guzinski et al. (2021), the gap-filling was performed using the ratio of actual to reference ET ($ratio_{ET}$) at the objective date (i.e. the date of the pixel that needs to be gap-filled) and closest valid actual ET output within 10 days of the objective date. In this study, $ratio_{ET}$ is still used. However, two modifications were applied to the method.

Firstly, instead of using temporarily closest available $ratio_{ET}$, a linear interpolation is performed between closest $ratio_{ET}$ before and after the objective date. This should better account for drying out of the soil. If there is only one available $ratio_{ET}$ within the gap-filling window (i.e. only either before or after the objective date) then it is used directly without linear interpolation.

Secondly, recent research (Delogu et al., 2021) suggests that rainfall should be considered when performing ET gap-filling for periods longer than a couple of days. We account thus for soil re-wetting by calculating a simple water balance (WB) for each date (i) within the gap-filling window as

$$WB_i = \max(0.75 * WB_{i-1} + PR_i - ET_{ref,i}, 0) \quad (3)$$

where PR_i is rainfall and $ET_{ref,i}$ is reference ET estimated using the standard FAO equation for short grass crop and terrain adjusted ERA5 meteorological forcing on day i and 0.75 approximates other water sinks (e.g. runoff or percolation to deeper layers) . If WB_i is larger than 2 (mm) on a given day then $ratio_{ET}$ for that day is set to maximum $ratio_{ET}$ observed within the gap-filling window. The linear interpolation of $ratio_{ET}$ is then performed on the values corrected using the water balance approach.

The gap-filling window was set as ± 15 days from the objective date. This is shorter than the gap-filling window selected for biophysical parameters for two reasons. Firstly, with both Sentinel-3 A & -3B satellites the thermal data is acquired with daily geometric revisit and therefore the duration of the gaps is expected to be shorter. Secondly, ET is more dynamic than the biophysical traits because it is more strongly influenced not only by phenology but also by meteorology (accounted for to a certain extent in $ratio_{ET}$) and water availability in the soil, which can change quickly due to rainfall, irrigation or drying out of the soil.

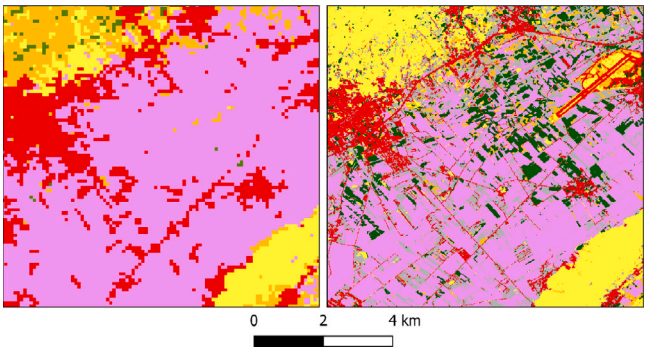


Fig. 2. Landcover maps from CGLC (left panel) and WorldCover 2020 (right panel) in Bekaa valley in Lebanon. Pink color indicates agriculture, red — urban areas, yellow — grassland/herbaceous vegetation, orange — shrubland and green — tree cover/forest.

2.3. New data sources

In Guzinski et al. (2021), the Copernicus Global Land Cover (CGLC) map was used to set surface parameters that could not be directly determined from Sentinel-2 or Sentinel-3 observations, such as vegetation structure/roughness. One of the two main limitations of this data was the 100 m spatial resolution, which would lead to box-shape artifacts in the output ET maps on the borders of very different land cover classes (such as agriculture and forest).

In this study we replaced CGLC by WorldCover land cover map, which has a 10 m spatial resolution (Zanaga et al., 2022). Both maps have global coverage and similar reported accuracies, with 77% overall accuracy for all classes and 80% producer and user accuracy for agriculture class for WorldCover (Zanaga et al., 2022) and 80% overall accuracy for all classes, 81% producer's accuracy and 73% user's accuracy for cropland class for CLGC (Tsendbazar et al., 2020). CGLC currently contains yearly updates from 2015 till 2019 while WorldCover maps are available for years 2020 and 2021. Despite not being a Copernicus product the use of WorldCover in this study is justified based on the fact that it is derived using Sentinel-1 and Sentinel-2 observation timeseries. However, WorldCover cannot overcome the second major limitation of CGLC, which is the presence of only one agricultural land cover class. This means that herbaceous fields are classified as agriculture and thus they will be assigned the same values for vegetation height and other structural parameters. Furthermore, in WorldCover map woody fields can be classified as agriculture, shrubland or tree cover (see Fig. 2 for an example from Lebanon), thus sometimes being assigned different structural parameters to herbaceous crops. However, it is not clear whether this is by accident or design and how consistent it is across different geographical regions.

The second change in ancillary data concerns replacement of SRTM DEM with Copernicus 30 m DEM (COP-DEM GLO-30 European Space Agency and Airbus, 2022). The DEM is used during atmospheric correction of Sentinel-2 reflectance, during sharpening of Sentinel-3 LST and in terrain correction of meteorological parameters. The change of DEM is expected to have minimal impact on the final ET maps but fits with the aim of using mainly Copernicus data in the study.

2.4. New WaPOR ETLook code

Similarly to Guzinski et al. (2021), in this study WaPOR ETLook model was run with the same Copernicus inputs as used in the TSEB-PT model. The TSEB-PT implementation (https://github.com/hectornieto/pyTSEB, last accessed: 29.06.2023) is practically unchanged (see Nieto et al., 2021 and Nieto et al., 2023 for comparison). However, the version and Python implementation of the ETLook model has changed between the studies. Previously, version 2 (v2) of the ETLook code which was available at https://bitbucket.org/cioapps/wapor-et-look.git repository as of June 10th 2020 was used. The code was further modified by Guzinski et al. (2021) to improve execution efficiency and allow the use of Copernicus data, with the actual ETLook code used available at https://github.com/DHI-GRAS/wapor-et-look/commits/copernicus_run (last accessed 29.06.2023).

In this study, ETLook code v3.2.4 which was available from https://bitbucket.org/cioapps/pywapor repository as of February 2nd 2023 was used. The code has undergone a lot of refactoring, bug fixing and efficiency improvements which is expected to result in higher accuracy outputs of ETLook model. The code allows running of ETLook v2 and v3 and we used v2 (same version as in Guzinski et al., 2021). The ETLook code has undergone further changes in this study to allow the use of Copernicus data, with the final code used available at https://github.com/DHI-GRAS/pywapor/tree/et4fao (last accessed 29.06.2023). When using the pyWaPOR code with Copernicus data in both phases of ET4FAO project, only the soil moisture and ET modeling functions (se_root and et_look respectively) were called since input data preparation was performed as described in this section.

3. Field validation sites

3.1. Southeast Spain

This validation site is located in and close to the Barrax experimental site and is the same as used in Guzinski et al. (2021). It contains instrumentation including large weighing lysimeters located in irrigated potato field (López-Urrea et al., 2016), irrigated festuca grassland López-Urrea et al. (2006), drip-irrigated vineyard (Sánchez et al., 2019) and eddy covariance (EC) towers located in drip-irrigated almond orchard (Sánchez et al., 2021) and rainfed wheat, codified as "ES-LTs" and "ES-FcO", respectively, in the European Fluxes Database Cluster (http://www.europe-fluxdata.eu/home/sites-list, last accessed: 01.08.2023).

3.2. Tunisia

The ET data from the field were collected at the El Koudia experimental station (36.546° N; 9.013°E, Fig. 3) of the National Institute of Field Crops (INGC). The site is located in the Governorate of Jendouba, northwest of Tunisia. This station is equipped with an EC station installed by the National Institute for Research in Rural Engineering, Water and Forests (INRGREF), which has expertise in this ET determination methodology and operates several other eddy covariance sites across Tunisia.

The field itself consists of Mio-Pliocene and Quaternary sedimentary rocks. The average annual temperature is 18°C, with peak temperatures reaching 35°C in July and August, and minimum temperatures of around 5–6°C from December to February. The average annual rainfall is 542 mm, with more than 70%–75% occurring between October and March. July has the lowest average precipitation with only 4 mm, while December has the highest with approximately 83 mm. The site features deep clay soils.

The eddy covariance station system were installed in a 2.4 ha plot (130 m x 185 m) located in the center of the farm, positioned at least 62 m away from the edge, in the middle of the plot. Energy balance was ensured by assigning residual energy to the latent heat flux (Twine et al., 2000).

The experimental site is situated in the Boussalem region, which is known for its predominant cultivation of herbaceous crops. During the winter season, the main crops cultivated in this experimental farm include durum wheat, faba bean, sugar beet, and rapeseed, utilizing conventional or direct sowing methods. In the summer, the fields are typically left fallow due to water scarcity for irrigation. Irrigation was not applied due to water issues during the summer, and during the winter rainfall predominantly supports the crops. Glyphosate herbicide is commonly used during sowing. The crop rotation for the different

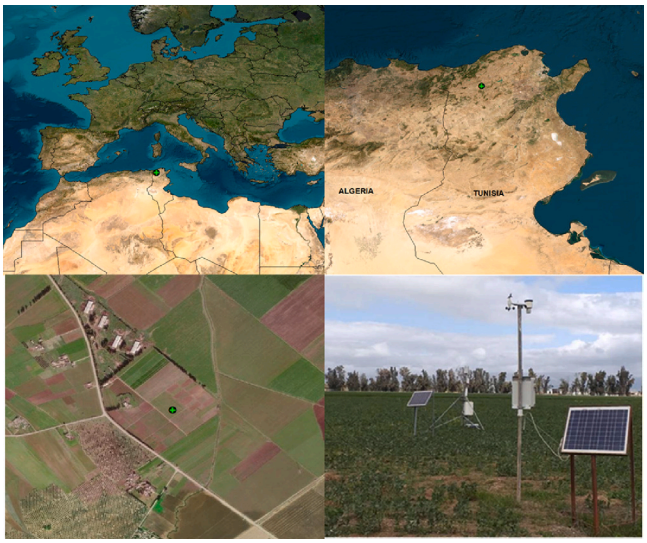


Fig. 3. Location of the validation site in Tunisia.

Fig. 4. Location of the validation site in Senegal.

Fig. 5. Location of the validation site in Lebanon.

seasons consisted of maize (July 19, 2020–November 3, 2020), faba bean (December 19, 2020–May 26, 2021), and sorghum (August 9, 2021–November 2, 2021).

This is a different Tunisian site to the one used in Guzinski et al. (2021). The main reason for this choice is that the new site features a flux tower over an arable land and thus with greater crop variability, as compared to the former rainfed olive site of Taous.

3.3. Senegal

The Niakhar site is situated in central Senegal within the Sereer region, in a rural area 150 km from Dakar. Established in 1962, the Niakhar observatory holds the distinction of being the oldest operational population observatory in Africa (Delaunay et al., 2018). Since 2018, a long-term collaborative observatory focused on agroforestry ecosystem services, such as food security and greenhouse gas fluxes, has been established by the UMR Eco&Sol research institute from France in association with Université Cheikh Anta Diop de Dakar (Rahimi et al., 2021).

The study area (Fig. 4) was established within active agro-silvo-pastoral bush fields managed by local farmers, where the dominant tree species is apple-ring acacia (*Faidherbia albida*) (Roupsard et al., 2020). Within a 15 ha area surrounding the flux tower, the density of *Faidherbia* trees was 6.8 trees per hectare, with a canopy cover of 5.14%. An understory in this area is present and consisted of a mosaic of crops including pearl millet, groundnut, watermelon, cowpea, and fallow areas. The Faidherbia-Flux site (https://lped.info/wikiObsSN/?Faidherbia-Flux, last accessed: 01.08.2023) is located at coordinates 14°29'44.916"N, 16°27'12.851"W, and is registered with FLUXNET as 'Sn-Nkr'.

The soil in the area is sandy and several meters deep, overlaying limestone bedrock. There is a brackish water table at approximately 6 m depth. The climate is characterized as soudano-sahelian, with a wet season spanning from June to October, followed by an eight-month dry season. Rainfall in the region decreased from 900 to 400 mm between 1950 and 1995 (the driest period), then partially recovered to around 500 mm by 2015. The seasonal distribution of rainfall shifted during this recovery period, with less rain occurring in the early part of the wet season and more towards the end.

The site is equipped with instruments and facilities for monitoring micro-meteorology, eddy-covariance fluxes of sensible heat, latent heat, and CO_2, soil water content, land surface and soil temperatures, NDVI (Normalized Difference Vegetation Index), soil respiration, sapflow, leaf area index (LAI), tree growth, fine root growth, crop productivity, and yield. Energy balance was ensured by assigning residual energy to the latent heat flux (Twine et al., 2000).

3.4. Lebanon

In Lebanon, the Tal Amara Research Station, operated by the Lebanese Agriculture Research Institute (LARI), serves as the site for ET field measurement and experiments. It is situated approximately 70 km northeast of Beirut (35.987927°E, 33.860117°N, Fig. 5).

The region, known as Bekaa Valley, experiences a typical Mediterranean climate, but its characteristics are influenced by two elevated mountain ranges flanking the Bekaa and Baalbeck-Hermel Governorates. The area is characterized by hot and dry summers, as well as relatively cold and rainy winters. Rainfall is concentrated between October and April, with minimal rainfall in May and September. On average, the annual rainfall is around 600 mm, with the majority (95%) occurring

between November and March. January stands out as the rainiest month, with an average of approximately 150 mm rainfall. Summers are completely dry, without any rain. The rainy season typically ends in mid-May and rarely starts again in mid-September. Temperature variations are significant throughout the year, with average summer temperatures hovering around 25°C and winter temperatures averaging around 6°C. The yearly average temperature stands at approximately 15°C. During July and August, temperatures can soar as high as 40°C, while certain winter days can experience freezing temperatures.

In this area, ET is measured thanks to the presence of a weighing lysimeter. However, due to travel restrictions due to COVID-19 the access to the lysimeter was heavily restricted throughout 2020 and 2021 leading to increased uncertainties in the measurements as time progressed. Repairs were resumed in August 2022, and thus *in situ* lysimeter data from this site between those dates should be treated with caution.

The cultivated land area encompassed approximately 9000 m^2. Ploughing and land leveling constituted routine procedures, accompanied by an intensive initial weed control practice. Prior to sowing in each season, the soil underwent moistening through a 10-hour sprinkler irrigation process, conducted approximately 20 days in advance. Faba bean crop was cultivated at the site from Dec. 3, 2020 till May 6, 2021 and this was followed by summer maize from June 17, 2021 till Oct. 8, 2021.

4. Results

4.1. Validation of thermal sharpening methods

During the summers of 2018 and 2019 a field campaign was conducted in "Las Tiesas" experimental farm at the Barrax site in Spain during which transect of LST were collected at the time of Sentinel-3 overpass, covering different crops and irrigation practices. A set of four hand-held infrared radiometers (IRTs) Apogee MI-210 were available for the LST transects, in addition to a multispectral radiometer CIMEL Electronique CE 312-2. Transects were conducted carrying back and forth the IRTs, at a rate of 5–10 registers/min, and covering as large area as possible within 10 min centered on the S3 overpass time. Ground LST were then obtained as the average value of the 50–100 punctual IRT measurements registered in every plot/site, covering an area representative of a grid of 3 × 3 Sentinel-2 20 m pixels. More details about the campaign and data can be found in Galve et al. (2022) and Sánchez et al. (2020). This field data was used to validate the modifications of the DMS sharpening method presented in Section 2.1.

The results of this exercise are presented in Fig. 6. The comparison of Sentinel-3 LST at its original resolution (around 1 km) against the field transects is shown if Fig. 6(a). If all the cases are considered it results in RMSE of 8.3 K. When only larger (≥ 1 ha) and not recently irrigated (closed circles in the plot) cases are evaluated this error reduces to close to 4.7 K. From now on those data points will be called "selected cases". Applying DMS as implemented in Guzinski et al. (2021) reduces the RMSE to 6.2 K for all cases and 2.5 K for selected cases (Fig. 6(b)). The correlation has also increased, especially when all cases are considered from 0.57 for S3 LST to 0.82 for DMS LST. However, it can be observed that the sharpened LST overestimates the coldest LST values and underestimates the warmest.

The improvements to the DMS method itself (described in Section 2.1.1) results in plots and statistics shown in Fig. 6(c). Compared to DMS model implementation in Guzinski et al. (2021), the RMSE for all cases is reduced by further 0.6 K. When recently irrigated cases are excluded the reduction in RMSE is 0.9 K. However, the reduction of RMSE in selected cases is only 0.2 K. This indicates that the largest improvements were in the smallest parcels, which by their nature would lie in the more heterogeneous pixels when resampled to Sentinel-3 resolution. In the rest of this section, this plot (Fig. 6(c)) is treated as a baseline against which to evaluate the impact of the different methods for inclusion of Landsat observations within in LST sharpening approach.

Using temporal Landsat LST parameters as explanatory variables in DMS (option (2) from Section 2.1.2) results in degradation of RMSE when considering both all cases and selected cases (Fig. 6(d)). However, the largest increase in RMSE (of 1.3 K) is obtained when the recently irrigated cases are excluded and it can be observed that for the recently irrigated cases (open circles) the error actually decreases. This implies that using Landsat data in this way is relevant for the main issue being addressed (overestimation of LST in irrigated agriculture) but this improvement is overshadowed by the general increase in noise in the sharpened LST. It might be that despite its infrequent acquisitions Landsat is able to capture at least one very low LST (present soon after irrigation) within the compositing period, in regions such as southern Spain where irrigation is frequent and cloud cover is minimal.

When Landsat LST is used for contrast-enhancement during the post-processing of sharpened LST (Fig. 6(e)) it results in reduction of RMSE of 0.3 K for all cases, 0.6 K for not recently irrigated cases but no significant change for selected cases compared to the improved DMS. This implies that contrast enhancement is especially useful for smaller parcels but does not have a significant effect on the LST of recently irrigated plots.

Finally, when Landsat LST is used both as explanatory variables in DMS and for contrast enhancement (Fig. 6(f)) the resulting RMSE for all cases decreases slightly compared to that of improved DMS. However, if the recently irrigated parcels are removed the error increases significantly by 0.8 K. This again shows that using Landsat LST as input to the DMS model can improve the accuracy of sharpened LST in the recently irrigated parcels. In addition, the usage of contrast enhancement also makes that the spatio-temporal variability of LST between the sharpened and the *in situ* has similar values (i.e. scale statistic with values closer to 1).

Considering all of the above results it was decided to use the contrast enhancement method (option (3) from Section 2.1.2 and Fig. 6(e)) during the production of sharpened LST to be used for ET modeling. This was based on three major factors. Firstly this method resulted in lowest RMSE when all the combinations of cases are considered. Secondly, as explained in Section 2.1.2, this method results in sharpened LST pixels even in the absence of Landsat LST. This is especially important in tropical regions where frequent cloud cover could result in very few Landsat LST observations. Thirdly, the LST dynamic range showed similar values to the dynamic range of the measurements (i.e. scale value closer to 1), and thus the over (under) estimation of cold (hot) pixels is reduced.

4.2. Validation of ET fluxes

The following results will evaluate both TSEB-PT and ETLook models, utilizing Copernicus data as inputs as well as WaPOR ET data obtained directly from the WaPOR portal. In all sites, 20 m Copernicus-based ET was used to validate ET. For cases where a flux tower exists, a buffer size of 5 × 5 pixels around the EC location coordinates was used to match the spatial extent of the EC footprint (i.e. ca. 100 m footprint). On the other hand a single pixel was used for the coordinates over the sites with lysimeters.

4.2.1. Southeastern Spain

In this region of Spain, there are no available data from WaPOR database, so the comparison only contains the results obtained with TSEB and ETLook using Copernicus data (in subsequent figures referred to as $TSEB_C$ and $ETLook_C$ respectively).

Table 1 presents the descriptive statistics of dekadal ET for both the observed data and all evaluated products. Meanwhile, Table 2 displays the error and agreement metrics between the observed values and the modeled values. Fig. 7 illustrates the overall results for all sites through

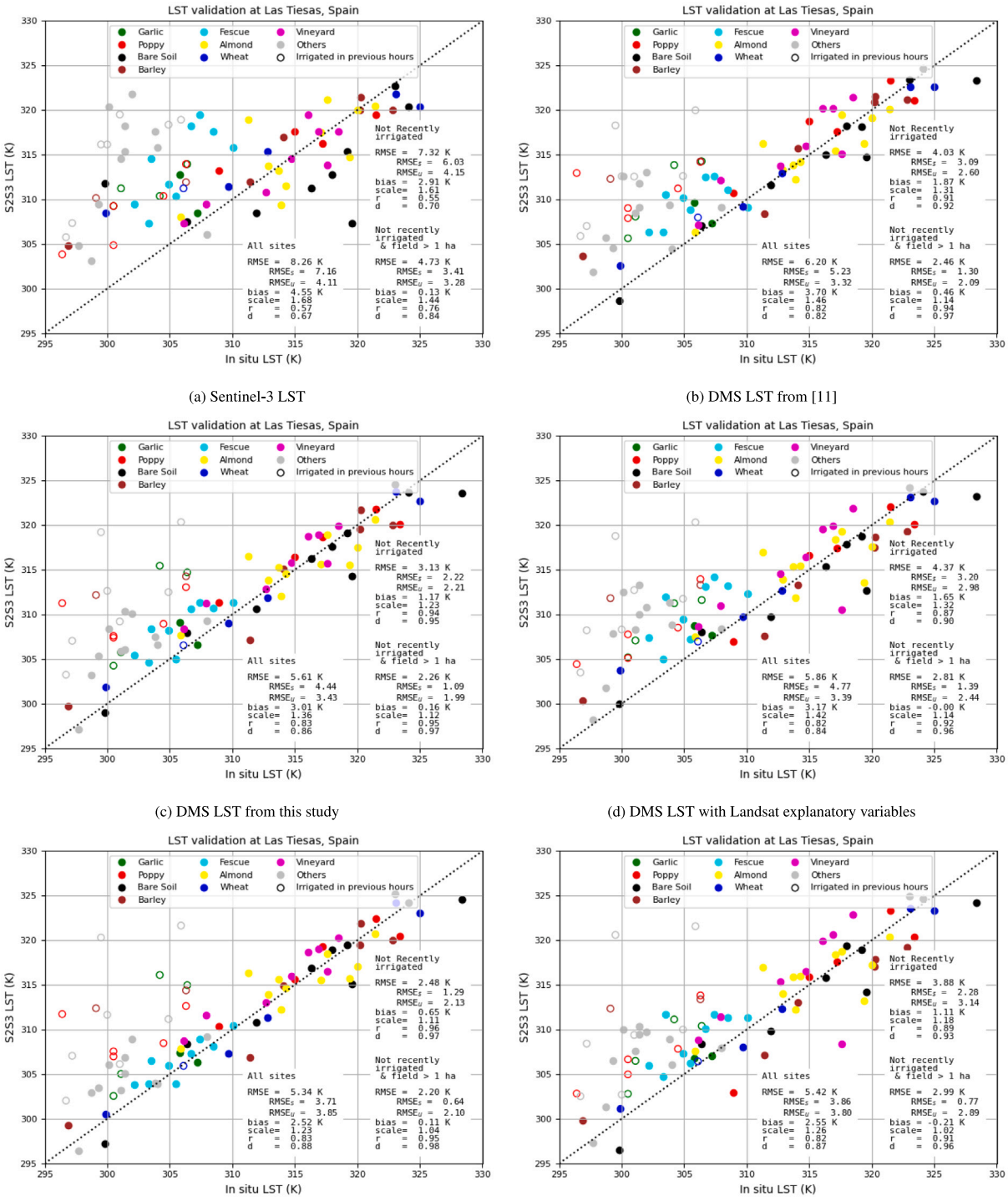


Fig. 6. Comparison of original resolution Sentinel-3 LST (a) and Sentinel-3 LST at 20 m resolution sharpened with different approaches (b–f) against *in situ* LST transects. Presented statistics are total, systematic and unsystematic root mean square error (RMSE, RMSE$_s$ and RMSE$_u$ respectively), mean bias (satellite LST - *in situ* LST), scale (standard deviation of *in situ* LST over standard deviation of satellite LST), Pearson Correlation coefficient (r) and Willmott's Index of Agreement (d). Statistics with "Not recently irrigated" in the header exclude the cases which were irrigated in the hours immediately before Sentinel-3 overpass (open circles in the plots).

a scatterplot representing the relationship between the observed values and the predictions of the models.

The ability to track ET temporal changes as well as the error deviations from the *in situ* measurement are shown in Fig. 8. In this figure, and for the sites with EC towers, an uncertainty band around the observed dekadal ET is displayed, due to energy closure imbalance.

Fig. 8(a) demonstrates the consistent underestimation of ET for almond crops throughout the entire time series. This difference has intensified compared to the results from (Guzinski et al., 2021), particularly during the summer months, and contrasts with the results obtained for other crops, which exhibit an opposite trend. This difference reaches up to 2 mm at certain points and is more pronounced in ETLook, reaching up to 2.5 mm. For this latter model, the change in

Table 1
Descriptive statistics for estimated and *in situ* dekadal ET (mm/day) for Barrax site. $\overline{Obs.}$ and $\overline{Pre.}$ show the mean and $\sigma_{Obs.}$ and $\sigma_{Pre.}$ the standard deviations of ET values for the *in situ* and model ET datasets respectively.

Site	Source	Model	N	$\overline{Obs.}$	$\overline{Pre.}$	$\sigma_{Obs.}$	$\sigma_{Pre.}$
All	Copernicus	TSEB-PT	119	3.39	3.32	1.92	1.72
All	Copernicus	ETLook			2.60		2.05
Festuca	Copernicus	TSEB-PT	36	5.08	4.91	1.66	1.17
Festuca	Copernicus	ETLook			4.58		1.48
Potato	Copernicus	TSEB-PT	14	4.69	4.92	1.96	1.87
Potato	Copernicus	ETLook			4.31		2.47
Grapevine	Copernicus	TSEB-PT	23	1.92	2.10	0.74	0.52
Grapevine	Copernicus	ETLook			1.29		0.48
Almond	Copernicus	TSEB-PT	26	2.53	1.88	0.75	0.59
Almond	Copernicus	ETLook			1.36		0.40
Wheat	Copernicus	TSEB-PT	20	2.22	2.61	1.31	0.81
Wheat	Copernicus	ETLook			0.95		0.77

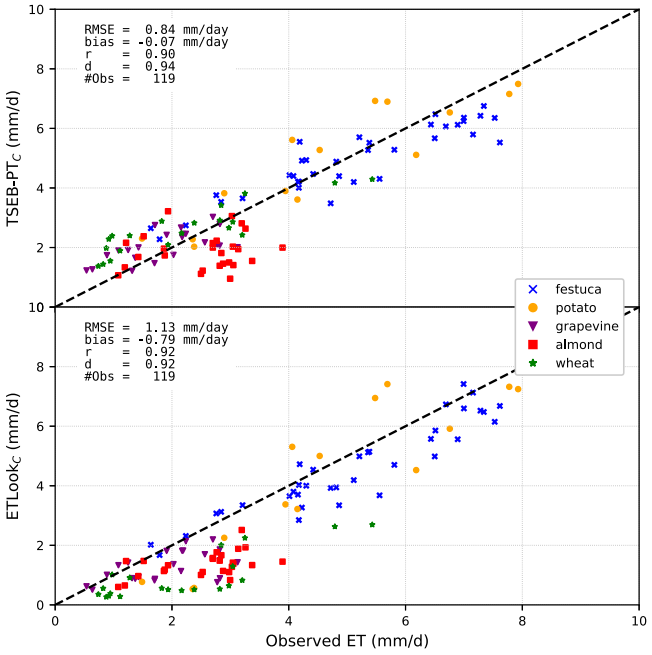


Fig. 7. Scatterplot between the *in situ* and estimated dekadal ET (mm/day) for crops in Spain, with Copernicus inputs and TSEB-PT model in top panel and ETLook model in bottom panel.

error has not been as drastic compared to Guzinski et al. (2021), which was up to 2.3 mm.

The estimation of TSEB and ETLook for the lysimeters in Spain significantly improves the results compared to Guzinski et al. (2021), especial for the grass site (Fig. 8(c)) where the majority of errors are below 1 mm/day. For the grapevine crop and potato, a similar effect is observed (Figs. 8(b) and 8(d)).

The validation results for the wheat crop (Fig. 8(e)) are quite similar in both studies for the TSEB model. The error has been reduced by 0.05 mm (9%), and the correlation has improved by 8% (r = 0.89 vs. 0.84). The same pattern is observed in ETLook, where the error has been reduced by 0.23 mm (13%), and the correlation has improved by 9% (0.81 vs. 0.72).

4.2.2. Tunisia

At this site also WaPOR Level 2 (100 m) data is available. The validation data was divided into three sets corresponding to the crops grown in the area within the study period: maize, beans, and sorghum. Validation statistics for each crop are shown in Table 3.

For the sorghum crop, WaPOR exhibits a very low and negative correlation (−0.07), while TSEB and ETLook show significantly higher correlations (0.88 and 0.80 respectively). The errors in ET estimation by TSEB are considerably lower (up to 70%) compared to the other two models.

These correlations are reduced for maize. The best results are obtained by TSEB and ETLook, with a Pearson coefficient of 0.60 and 0.49 respectively. Among these two models, TSEB achieves the lowest RMSE (0.51 mm/day), which is more than 59% lower than that shown by WaPOR. On the other hand, the results of ETLook exhibit a lower correlation (0.49) than TSEB with a slightly lower RMSE than that generated by WaPOR (1.18 mm/day).

Finally, the data obtained for the beans crop yield similar results to those for sorghum. In this case, the model with the lowest RMSE (0.59 mm/day) is TSEB, with a correlation of 0.92, followed by ETLook (r = 0.92 and RMSE = 0.81 mm/day) and WaPOR (r = 0.87 and RMSE = 1.01 mm/day).

The time series for this site including all the available data for the three crops is depicted in Fig. 9(a). In general terms, TSEB exhibits the best fit. Although the other models deviate slightly from the observed ET, TSEB maintains a consistent error rate around 1 mm/day, except for February 2021. The other models have a relatively higher error rate, especially for sorghum, ranging from 1.5–2 mm, and also particularly during the spring months.

4.2.3. Senegal

At this site the WaPOR data used corresponds to Level 2 (100 m) data. Table 3 shows the degree of fit between the validation data and the TSEB and ETLook models.

The model that performs best in this area is TSEB, which has an RMSE of 0.55 mm/day that is 50% lower than that of WaPOR and 47% lower than that of ETLook. The degree of fit in this case is slightly better for TSEB (1% better than WaPOR).

Fig. 9(b) depicts the complete time series data. It can be observed that the errors of TSEB do not exceed the 1 mm threshold throughout the analysis period, except for October 2021, where it drops to −1.8 mm/day. The other models show an overestimation of ET ranging from 1–2 mm/day, depending on the time of year, but still within the uncertainty energy closure gap.

4.2.4. Lebanon

For this area, ET data are available from the following products: TSEB, ETLook, Level 3 (30 m) WaPOR v2.

Table 3 shows the accuracy statistics of observed vs. predicted values for faba beans and maize crops . The correlation of faba beans ET is significantly higher (> 0.92) than that found for maize (< 0.40). In this case, the WaPOR product exhibits better correlation (0.98) compared to Copernicus TSEB and ETLook (0.92), but both models have lower error rates (1.14 mm/day and 1.02 mm/day respectively) compared to 1.55 mm/day of WaPOR.

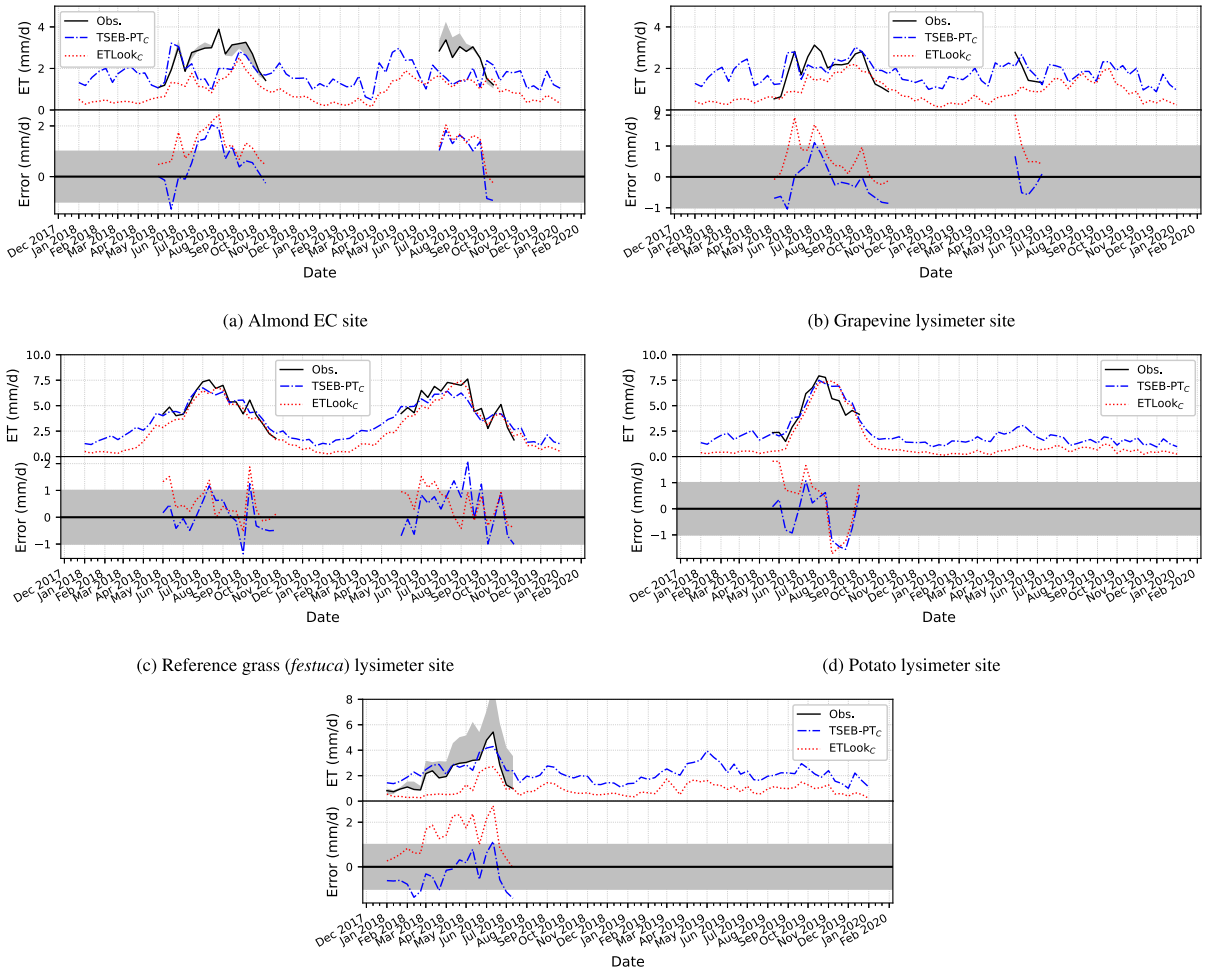
On the contrary, a different trend is observed for maize. In this case, only TSEB shows a positive correlation (0.40), while ETLook and WaPOR yield lower and negative results (−0.16, −0.32 respectively).

Using the entire available timeseries of validation data, Fig. 9(c) illustrates how the errors increase during the summer months, coinciding with the maize growing season, while for the rest of the year they are clustered around 1 mm/day. The models that show the poorest fit to the reference data is WaPOR, which consistently underestimates the ET throughout the entire analysis period.

Travel restrictions associated to COVID-19 lockdown in Lebanon made it impossible to access and maintain the lysimeter starting in early 2020 and thus the results for this site should be treated carefully.

Table 2
Error metrics between estimated and *in situ* dekadal ET (mm/day) for Barrax site. $\overline{bias}$ (mm/day) is the average bias, computed as the mean difference between the predicted and the observed, MAE is the mean absolute error (mm/day), RMSE (mm/day) is the Root Mean Square Error, which is decomposed between its unsystematic ($fMSE_u$) and systematic ($fMSE_s$) fractions ($fMSE_u + fMSE_s = 1$), a is the slope of the regression between the observed and the predicted, scale is the ratio between the standard deviation of the observed over the predicted, r is the Pearson Correlation coefficient between observations and predictions, and d is the Willmott's Index of Agreement.

Site	Source	Model	$\overline{bias}$	MAE	RMSE	$fMSE_s$	$fMSE_u$	a	scale	r	d
All	Copernicus	TSEB-PT	−0.07	0.69	0.84	0.20	0.80	0.81	1.11	0.90	0.94
All	Copernicus	ETLook	−0.79	0.92	1.13	0.48	0.52	0.98	0.93	0.92	0.92
Festuca	Copernicus	TSEB-PT	−0.17	0.64	0.79	0.62	0.38	0.64	1.42	0.91	0.92
Festuca	Copernicus	ETLook	−0.50	0.63	0.80	0.52	0.48	0.83	1.12	0.93	0.94
Potato	Copernicus	TSEB-PT	0.24	0.72	0.85	0.17	0.83	0.87	1.05	0.91	0.95
Potato	Copernicus	ETLook	−0.37	1.07	1.19	0.14	0.86	1.13	0.79	0.90	0.93
Grapevine	Copernicus	TSEB-PT	0.17	0.49	0.57	0.56	0.44	0.47	1.44	0.68	0.77
Grapevine	Copernicus	ETLook	−0.63	0.69	0.89	0.79	0.21	0.35	1.54	0.54	0.63
Almond	Copernicus	TSEB-PT	−0.65	0.93	1.11	0.72	0.28	0.09	1.28	0.11	0.46
Almond	Copernicus	ETLook	−1.17	1.19	1.34	0.93	0.07	0.27	1.86	0.51	0.51
Wheat	Copernicus	TSEB-PT	0.39	0.69	0.79	0.78	0.22	0.55	1.61	0.89	0.86
Wheat	Copernicus	ETLook	−1.27	1.27	1.50	0.91	0.09	0.48	1.69	0.81	0.65



(a) Almond EC site

(b) Grapevine lysimeter site

(c) Reference grass (*festuca*) lysimeter site

(d) Potato lysimeter site

(e) Winter wheat EC site

Fig. 8. Timeseries at the Spanish validation sites of: *in situ* and estimated dekadal ET (mm/day), with black line indicating *in situ* observations (with energy imbalance correction applied at the eddy-covariance sites) and grayed area indicating uncertainty of measured fluxes due to measured energy imbalance for the eddy-covariance sites (top panel); differences between *in situ* and estimated dekadal ET, with grayed area corresponding to errors within ±1 mm/day of the black line from the top panel (bottom panel).

4.2.5. Summary statistics in Africa and Middle East

Tables 3 and 4 provide a statistical description of the models, detailing the errors and deviations found in relation to the observed data. Overall, for all the analyzed regions in Africa and Middle East combined, TSEB has a deviation of the mean predicted value from the mean observed value of 5.5%, while ETLook driven with Copernicus data and WaPOR have deviations of 39% and 48%, respectively, indicating underestimation of the ET.

In general terms, the standard deviation of the observed data is greater than that of the predicted data. This may imply that the models are capturing a portion of the variability in the observed data, but they

Table 3
Error metrics between estimated and *in situ* dekadal ET (mm/day) for Africa and Middle East sites where WaPOR data is produced. $\overline{bias}$ (mm/day) is the average bias, computed as the mean difference between the predicted and the observed, MAE is the mean absolute error (mm/day), RMSE (mm/day) is the Root Mean Square Error, which is decomposed between its unsystematic ($fMSE_u$) and systematic ($fMSE_s$) fractions ($fMSE_u + fMSE_s = 1$), a is the slope of the regression between the observed and the predicted, scale is the ratio between the standard deviation of the observed over the predicted, r is the Pearson Correlation coefficient between observations and predictions, and d is the Willmott's Index of Agreement.

Site	Crop	Source	Model	$\overline{bias}$	MAE	RMSE	$fMSE_s$	$fMSE_u$	a	scale	r	d
Tunisia	faba	Copernicus	TSEB-PT	0.37	0.48	0.59	0.55	0.45	0.80	1.14	0.92	0.93
		Copernicus	ETLook	−0.68	0.70	0.81	0.70	0.30	0.94	0.99	0.92	0.88
		WaPOR	ETLook	−0.81	0.85	1.01	0.65	0.35	0.91	0.95	0.87	0.83
	maize	Copernicus	TSEB-PT	−0.38	0.39	0.51	0.81	0.19	0.39	1.53	0.60	0.63
		Copernicus	ETLook	−1.12	1.12	1.18	0.97	0.03	0.25	1.92	0.49	0.39
		WaPOR	ETLook	−1.19	1.19	1.25	0.98	0.02	0.18	2.10	0.37	0.37
	sorghum	Copernicus	TSEB-PT	−0.34	0.36	0.46	0.77	0.23	0.64	1.36	0.88	0.84
		Copernicus	ETLook	−1.04	1.04	1.12	0.97	0.03	0.40	2.00	0.80	0.54
		WaPOR	ETLook	−1.42	1.42	1.57	0.98	0.02	−0.02	3.06	−0.07	0.34
Senegal	faidherbia	Copernicus	TSEB-PT	0.10	0.43	0.55	0.18	0.82	0.83	1.09	0.91	0.95
		Copernicus	ETLook	−0.87	0.92	1.05	0.76	0.24	0.79	1.13	0.89	0.83
		WaPOR	ETLook	−0.97	1.03	1.12	0.78	0.22	0.85	1.05	0.90	0.82
Lebanon	faba	Copernicus	TSEB-PT	0.32	0.88	1.14	0.84	0.16	0.53	1.75	0.92	0.88
		Copernicus	ETLook	−0.54	0.84	1.02	0.63	0.37	0.71	1.30	0.92	0.92
		WaPOR	ETLook	−1.10	1.16	1.55	0.98	0.02	0.49	1.99	0.98	0.79
	maize	Copernicus	TSEB-PT	−2.41	2.41	2.51	0.99	0.01	0.17	2.42	0.40	0.36
		Copernicus	ETLook	−2.90	2.90	3.13	0.94	0.06	−0.16	1.01	−0.16	0.27
		WaPOR	ETLook	−3.20	3.20	3.36	0.98	0.02	−0.19	1.68	−0.32	0.25
All	All	Copernicus	TSEB-PT	−0.14	0.68	1.00	0.53	0.47	0.55	1.43	0.79	0.85
		Copernicus	ETLook	−1.01	1.08	1.36	0.75	0.25	0.63	1.32	0.83	0.79
		WaPOR	ETLook	−1.23	1.27	1.57	0.82	0.18	0.56	1.44	0.80	0.72

Table 4
Descriptive statistics for estimated and *in situ* dekadal ET (mm/day) for sites in Africa and Middle East where WaPOR data is produced. $\overline{Obs.}$ and $\overline{Pre.}$ show the mean and $\sigma_{Obs.}$ and $\sigma_{Pre.}$ the standard deviations of ET values for the *in situ* and model ET datasets respectively.

Site	Crop	Source	Model	N	$\overline{Obs.}$	$\overline{Pre.}$	$\sigma_{Obs.}$	$\sigma_{Pre.}$
Tunisia	faba	Copernicus	TSEB-PT	16	2.65	3.02	1.15	1.00
		Copernicus	ETLook			1.97		1.16
		WaPOR	ETLook			1.84		1.21
	maize	Copernicus	TSEB-PT	11	1.76	1.38	0.43	0.28
		Copernicus	ETLook			0.64		0.22
		WaPOR	ETLook			0.57		0.20
	sorghum	Copernicus	TSEB-PT	9	2.16	1.82	0.63	0.46
		Copernicus	ETLook			1.12		0.32
		WaPOR	ETLook			0.74		0.21
Senegal	faidherbia	Copernicus	TSEB-PT	38	2.13	2.22	1.29	1.18
		Copernicus	ETLook			1.25		1.13
		WaPOR	ETLook			1.15		1.21
Lebanon	faba	Copernicus	TSEB-PT	16	2.50	2.82	2.11	1.21
		Copernicus	ETLook			1.96		1.63
		WaPOR	ETLook			1.40		1.06
	maize	Copernicus	TSEB-PT	9	5.41	3.01	0.78	0.32
		Copernicus	ETLook			2.51		0.77
		WaPOR	ETLook			2.21		0.46
All	All	Copernicus	TSEB-PT	99	2.53	2.39	1.61	1.12
		Copernicus	ETLook			1.52		1.22
		WaPOR	ETLook			1.30		1.11

are not fully explaining all the temporal fluctuations or patterns present in the actual data. These trends are replicated for each individual crop as well.

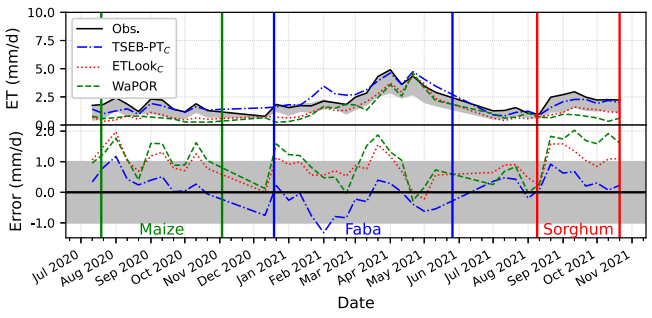
Overall, when considering the correlations, ETLook obtains the highest correlation (0.82), followed by WaPOR (0.83) and TSEB (0.79). Despite these results, the model that generates the lowest errors is TSEB, with the MAE being 47% lower than WaPOR, the RMSE 36% lower, and the unsystematic part of the error being 35% lower. ETLook falls between the other two models. If we look at the bias (predicted - observed) of the three models, in all cases it is negative, indicating an underestimation of ET. Comparing the three models, TSEB shows the lowest bias (−0.14 mm/day), which is 88% lower than WaPOR (−1.23 mm/day). These differences suggest that there is a systematic tendency to underestimate ET by WaPOR and ETLook, which is significantly less pronounced in the case of TSEB. This trend can be observed again in the time series graphs, where the error line of TSEB is always below the other two models.

5. Discussion

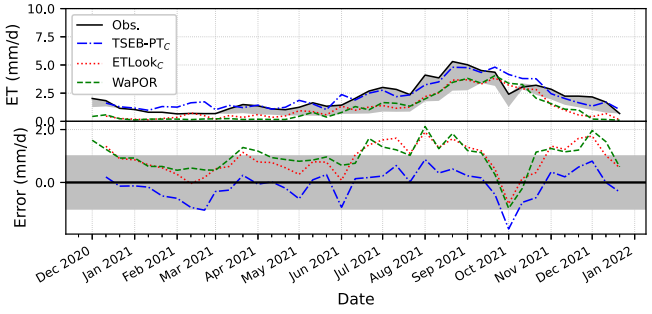
5.1. Influence of method changes on accuracy of modeled ET

The interpretation of results from the Barrax site in Spain allows for evaluation of the impact of the improvements presented in Section 2 on modeled ET. This is because exactly the same field measurements were used in Guzinski et al. (2021) and in this study. By comparing Table IV of Guzinski et al. (2021) with Table 2 of this study it can be seen that the overall improvement in accuracy of modeled ET is significant with bias reducing from −0.29 mm/day to −0.07 mm/day (from −1.39 mm/day to −0.79 mm/day) and RMSE reducing from 0.96 mm/day to 0.84 mm/day (from 1.69 mm/day to 0.79 mm/day) for TSEB-PT (ETLook) model. In certain crops this improvement is even more significant with e.g. the RMSE of festuca modeled with ETLook decreasing by 1.5 mm/day. There are three main sources of this improvement: the improvements in sharpening of LST, the switch from CGLC to WorldCover land cover map, and the use of new ETLook code. The impact of the gap-filling approach is expected to be minimal since the validation is performed mainly over the summer months with very infrequent cloud cover.

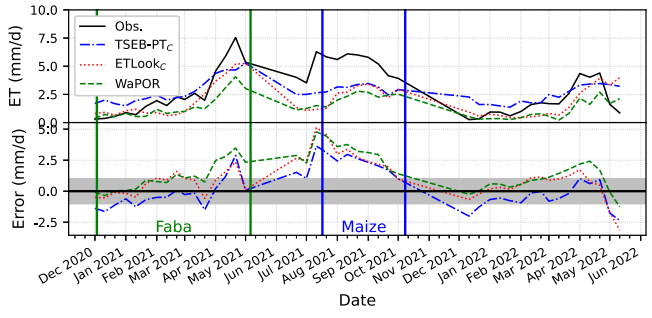
In CGLC map all the Spanish sites were located in agricultural class (class 40 in Table III of Guzinski et al., 2021). According to WorldCover, potato and wheat remain in agricultural class, festuca is assigned a grassland class (class 30 in Table III of Guzinski et al., 2021) and almond and vineyard are assigned a shrubland class (class 20 in Table III of Guzinski et al., 2021). The main influence of the land cover map on the TSEB-PT model is through setting the vegetation height. In agricultural class the vegetation starts with a height of 15 cm and reaches 150 cm once total LAI reaches 5. Similarly, grassland starts at 10 cm and reaches 100 cm once total LAI reaches 4. For shrubland, the vegetation height remains constant at 2 m throughout the growing season.



(a) EC site in Tunisia

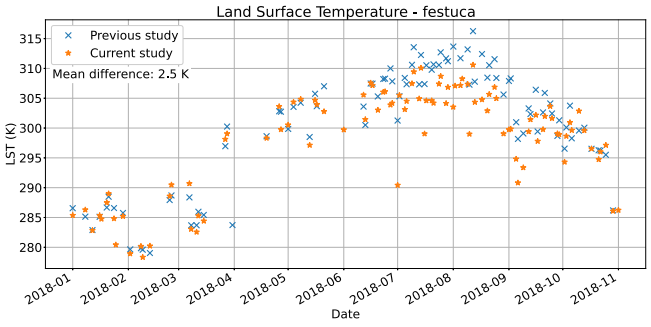


(b) EC site in Senegal

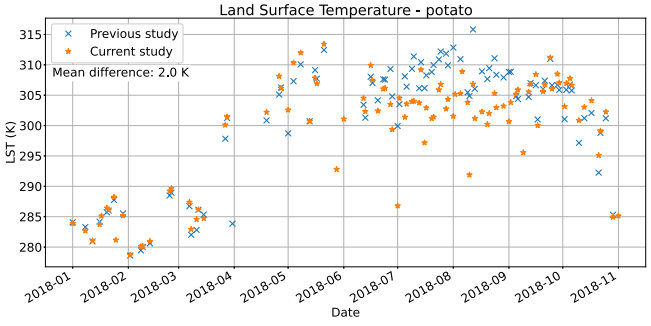


(c) Lysimeter site in Lebanon

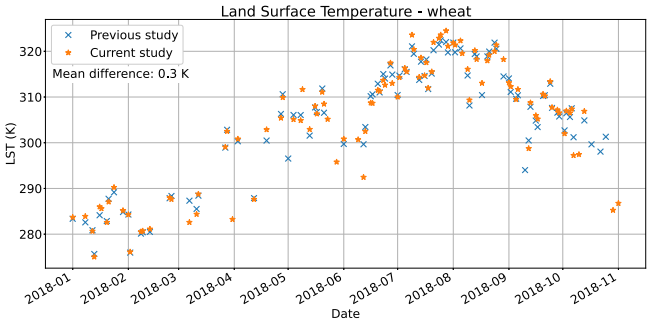
Fig. 9. Timeseries at the Tunisia, Senegal and Lebanon validation sites of: *in situ* and estimated dekadal ET (mm/day), with black line indicating *in situ* observations (with energy imbalance correction applied at the eddy-covariance sites) and grayed area indicating uncertainty of measured fluxes due to measured energy imbalance for the eddy-covariance sites (top panel); differences between *in situ* and estimated dekadal ET, with grayed area corresponding to errors within ±1 mm/day of the black line from the top panel (bottom panel). Vertical bars indicate the start and end of respective growing seasons.



(a) Festuca



(b) Potato



(c) Wheat

Fig. 10. Timeseries plots of Land Surface Temperature of selected crops parcels from Barrax area, derived using the method from (Guzinski et al., 2021) ("Previous study") and option (3) from Section 2.1.2 ("Current study").

Looking only at ET of potato and wheat modeled with TSEB-PT we can isolate the influence of the improvements in LST sharpening. In both cases the RMSE decreases slightly in the current study (from 0.99 to 0.85 for potato and from 0.84 to 0.79 for wheat) while change in bias is also small but mixed (from −0.54 to 0.24 for potato and 0.21 to 0.39 for wheat). The improvements in accuracy of festuca (RMSE decrease from 1.38 to 0.79 and bias decrease from −1.05 to −0.17) are also mainly due to changes in LST sharpening since the height parameterization is not significantly different. Since festuca and potato are frequently sprinkler irrigated, and therefore exhibit the coldest LST on the edge of the range which would be observed by Sentinel-3, it was expected that the changes to LST sharpening methodology would have the largest influence at those sites and this is indeed the case (Figs. 10(a) and 10(b)). On the other hand, wheat is rainfed and therefore with LST well within the range observed by Sentinel-3 and as expected the improved sharpening has the least influence on LST (Figs. 10(c)) and therefore on the modeled fluxes at this site. Grapevine and almond are both drip irrigated, so with quite high LST, and the WorldCover parameterization as shrubland fits better with reality than CGLC classification as agriculture. However here the results are mixed with improved ET accuracy for grapevine but degraded accuracy for almond. The worse accuracy of almond ET (as compared to other crops) could be partly caused by the discrepancy between vegetation height assigned by the landcover map (2 m) and the actual height of the almond trees (3–4 m).

Regarding ETLook estimated ET, it shows improvements between (Guzinski et al., 2021) and this study for all sites with especially significant improvements for festuca, potato and grapevine. ETLook is less influenced by vegetation height, and other land cover based parameters, than TSEB-PT and is also not as sensitive to LST as TSEB-PT. Therefore, most of those improvements can be associated with improved implementation of ETLook model as described in Section 2.4.

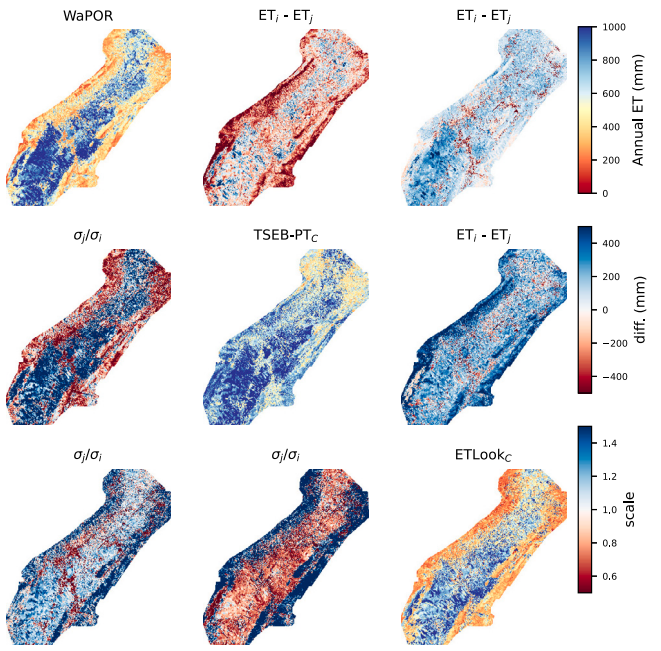


Fig. 11. Annual comparison of ET for products at WaPOR Level 3 resolution (20–30 m) in the Bekaa Valley of Lebanon. Maps in the diagonal show cumulative annual ET of each product. The off-diagonal shows products' intercomparison, both in ET differences (upper-right) and in standard deviations scale (lower-left). "i" and "j" represent products in row and column respectively.

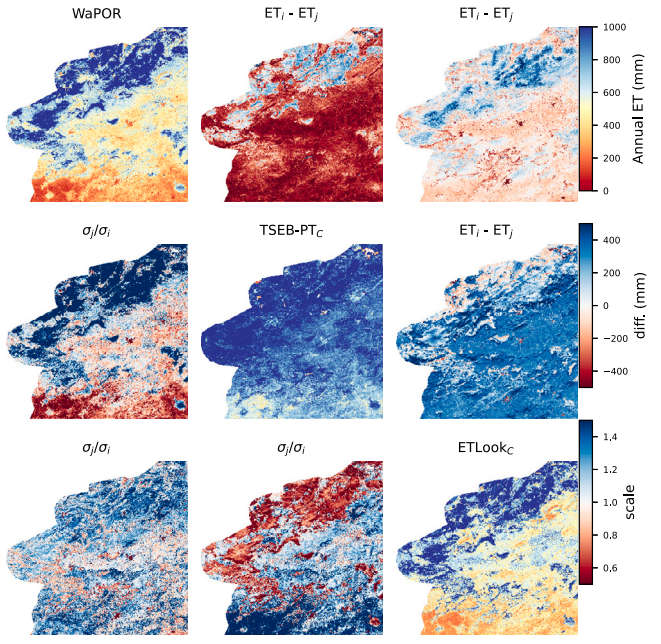


Fig. 12. Annual comparison of ET for products at WaPOR Level 2 resolution (100 m) for an area covering one Sentinel-2 tile (100 km by 100 km) in Tunisia. Maps in the diagonal show cumulative annual ET of each product. The off-diagonal shows products' intercomparison, both in ET differences (upper-right) and in standard deviations scale (lower-left). "i" and "j" represent products in row and column respectively. Due to the country boundary cropping performed by WaPOR, data for the western area of the image is not available.

5.2. Model intercomparison

As observed in Section 4, TSEB-PT generally produces the highest ET fluxes and WaPOR dataset contains the lowest values. When performing spatial comparison of the modeled ET maps in Lebanon (Fig. 11),

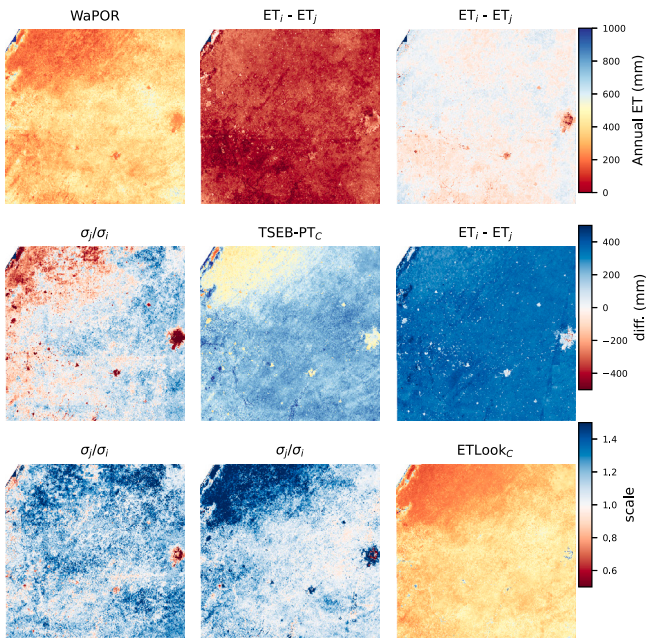


Fig. 13. Annual comparison of ET for products at WaPOR Level 2 resolution (100 m) for an area covering one Sentinel-2 tile (100 km by 100 km) in Senegal. Maps in the diagonal show cumulative annual ET of each product. The off-diagonal shows products' intercomparison, both in ET differences (upper-right) and in standard deviations scale (lower-left). "i" and "j" represent products in row and column respectively.

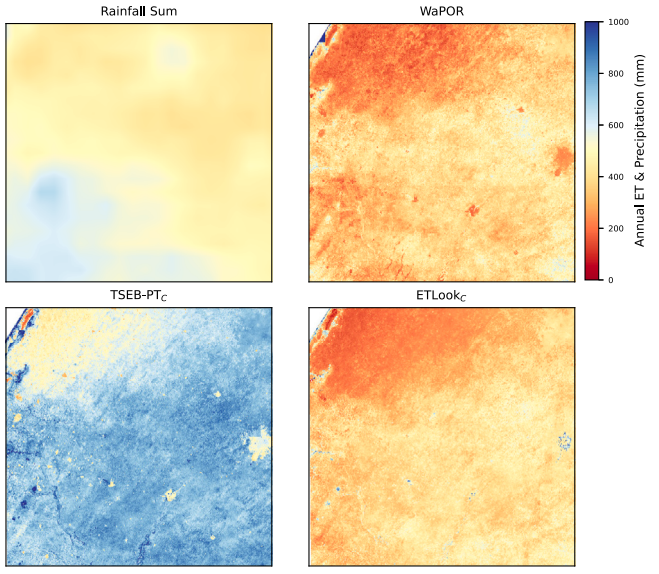


Fig. 14. Annual comparison of 2021 annual sums of CHIRPS v2.0 rainfall and the three ET products at WaPOR Level 2 resolution (100 m) for an area covering one Sentinel-2 tile (100 km by 100 km) in Senegal.

Tunisia (Fig. 12) and Senegal (Fig. 13) the same pattern can be observed. This is especially evident in rainfed natural and semi-natural vegetation, e.g the areas surrounding irrigation districts in Bekaa valley (Fig. 11), and can also be noticed during dry or low vegetation seasons (Figs. 8 and 9). It is unclear whether this is due to TSEB-PT consistently overestimating or ETLook consistently underestimating ET in such conditions. When looking at the timeseries plots of rainfed sites where field measurements were acquired (wheat in Spain — Fig. 8(e), maize and sorghum in Tunisia — Fig. 9(a), Faidherbia in Senegal — Fig. 9(b)) there is no obvious under- or over-estimation by any of the three ET datasets (the estimates usually lie within the measurement

uncertainty band shown as gray area in the top-panels of each sub-figure), apart from the wheat site in Spain where ETLook displays quite strong underestimation. At the other sites, there are periods when TSEB-PT is above the uncertainty band (e.g. winter and autumn 2021 in Senegal and winter/spring 2021 in Tunisia) and where ETLook models are below the uncertainty bands (e.g. winter/spring 2021 for both Copernicus ETLook and WaPOR in Senegal and autumn 2020 and 2021 in Tunisia for WaPOR).

Looking at annual rainfall sums versus ET sums in Senegal points to overestimation by TSEB-PT in semi-arid rainfed vegetation but does not provide clear indication of the magnitude of this overestimation or potential underestimation by ETLook (Fig. 14). The mapped area in Senegal is spread over four provinces with mean 2021 annual rainfall according to CHIRPS v2.0 dataset (Funk et al., 2015) of 497 mm (ranging from 396 mm in the north to 671 mm in the south), while both Copernicus ETLook and WaPOR annual ET sums are around 200–400 mm and TSEB-PT annual sum is around 500–700 mm. However, actual ET can sometimes exceed total rainfall even at annual basis as some plants (e.g. *Faidherbia albida* Roupsard et al., 1999) can extract water from quite deep into the ground. Therefore, the apparent overestimation of TSEB-PT should be investigated with longer timeseries and with validating the model against more EC data collected in semi-arid natural vegetated areas.

Another observation, which differs from the conclusions of Guzinski et al. (2021), is that the outputs of ETLook model run with Copernicus inputs and with WaPOR inputs (i.e. ET maps downloaded from WaPOR portal) are generally closer to each other than to TSEB-PT outputs. This is especially the case with WaPOR Level 2 data, as can be observed in the timeseries plots from Tunisia (Fig. 9(a)) and Senegal (Fig. 9(b)) and maps from Tunisia (Fig. 12) and Senegal (Fig. 13). Those maps compare both spatial variability (in upper-right panels as difference of the annual sums) and temporal variability (in bottom-left panels as ratio of annual standard deviations) of the three datasets and for both of those the agreement between the ETLook models is stronger than between ETLook and TSEB-PT data for most of the mapped areas. At WaPOR Level 3 the differences between $ETLook_C$ and WaPOR are larger, for example as observed in timeseries plot in Lebanon (Fig. 9(c)) where during the peaks of growing seasons the ET from $ETLook_C$ is closer to $TSEB\text{-}PT_C$ than to WaPOR. However, the spatial and temporal variability of WaPOR and Copernicus ETLook agree to a large extent when looking at the whole Bekaa valley (Fig. 11). This improvement in data agreement is mainly due to the use of new ETLook code (see Section 2.4) which tries to replicate the code used in WaPOR data production as closely as possible. This is an encouraging result for further integration of Copernicus data within the WaPOR production chain while maintaining a consistent timeseries of ET estimates.

Finally, while the improvements in gap-filling methods are performing well, they are still not robust enough for the most cloud-prone areas. Firstly, the timeseries plots of modeled ET at the Spanish sites from (Guzinski et al., 2021) and this study show similar temporal patterns, especially in autumn, winter and spring where the influence of changes in gap-filling method would be stronger than the influence of change in the LST sharpening method. Secondly, in Tunisia and Senegal where measurements were acquired throughout the year the agreement between modeled and measured ET remains similar during dry and wet months. Especially at the Niakhar site in Senegal, where the peak of vegetation growth (and therefore ET) coincides with the rainy season (June to October), the agreement between modeled and measured fluxes remains high during the cloudy periods. Also, the agreement between $ETLook_C$ and WaPOR data is very high throughout the year at the Tunisian and Senegalese validation sites, despite WaPOR using different gap-filling methods. All this demonstrates that the improvements in gap filling methods of both biophysical parameters and ET (i.e. using the simplified water balance model) are robust and improve the dekadal ET estimates.

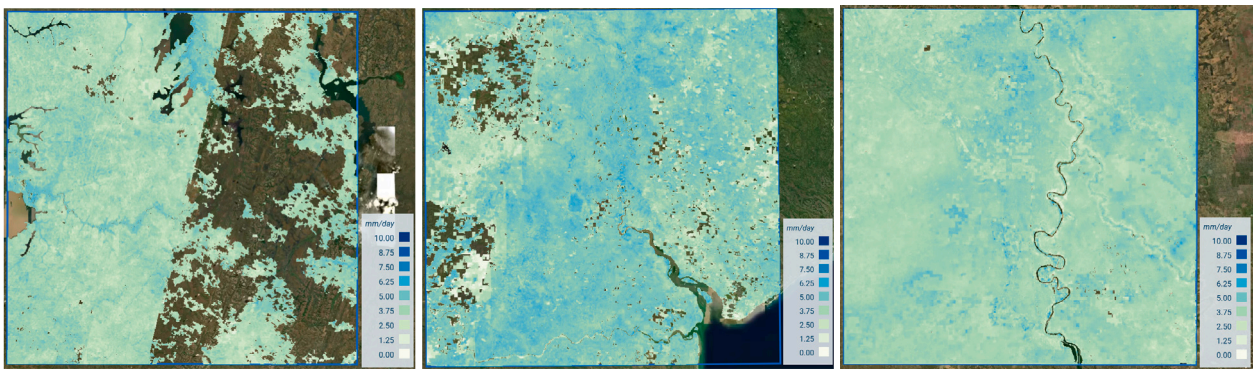
However, the improved gap-filling method still leaves some gaps and also can result in blocky artifacts with the size of Sentinel-3 pixel during the most cloudy periods. This can be observed on the maps produced as part of the same study (shown in Fig. 15) and displayed on an online portal (et4fao.dhigroup.com, last accessed 29.06.2023), especially in tropical areas (e.g. Rwanda in the second dekade of November 2021 or Mozambique in the last dekade of December 2021) but also in other regions during the rainy season (e.g. in Sudan in the second dekade of July 2021). The gaps present in Rwanda or Mozambique are mainly due to the lack of cloud free Sentinel-2 observations needed for biophysical parameterization of the surface. Especially in Rwanda the distribution of the gaps displays clear pattern related to Sentinel-2 orbits (less gaps in the western half, more in the eastern half) which could indicate issues with Sentinel-2 data in addition to cloud cover. In any case, using more robust interpolation techniques, e.g. based on phenological cycles and outliers filtering (such as Whittaker smoother (Eilers, 2003) used in the WaPOR method) could help to address those issues. Another possibility is the data fusion of Sentinel-2 and Sentinel-3 shortwave optical observations or biophysical products. On the other hand, the artifacts visible in Sudan are mainly due to the gap filling of the daily ET maps. This issue could be addressed by a more advanced implementation of water balance model. However, a balance needs to be established between producing seamless gap-filled maps which are visually pleasing and the physical accuracy of those maps. Those two aspects are not always in agreement and it might be preferable to leave certain gaps and artifacts within the maps in order to provide ET estimates with acceptable accuracy.

5.3. Comparison with previous studies

Most existing studies evaluating LST sharpening methods are based on low-resolution thermal observations acquired by MODIS sensor on Aqua and Terra satellites or by VIIRS sensor on Suomi NPP satellite. Only a few studies use the low spatial resolution thermal observations from the SLSTR sensor onboard the Sentinel-3 satellites. In Wang et al. (2022) the Sentinel-3 LST was sharpened to 300 m using a DEM and vegetation indices calculated from Sentinel-3 shortwave observations and compared to Landsat −8 LST, while in Huryna et al. (2019) Sentinel-3 LST was sharpened using Sentinel-2 NDVI and also compared to Landsat −8 LST. Both studies indicated suitability of Sentinel-3 thermal data for use in thermal sharpening methods.

Similarly, not many studies compare high resolution LST, either acquired at such resolution or sharpened, against ground-based transects in heterogeneous agricultural landscapes. In Galve et al. (2022), Landsat −8 LST derived with different retrieval algorithms was compared to transects of LST ground measurements from the same "Las Tiesas" experimental farm as used in the current study. The results showed RMSE of around 2 K and absolute values of bias ranging from 0.1 K to 1.2 K depending on the retrieval algorithm used. Ground measurements of LST from the same area were also used to validate a method for thermal sharpening of MODIS LST using Sentinel-2 vegetation indices (Sánchez et al., 2020). When recently irrigated cases were excluded that study reported RMSE of sharpened LST of 2.2 K to 2.7 K and bias of 0.2 K to 0.4 K, depending on the sharpening method used. The results from those studies are comparable to the RMSE of 2.5 K and bias of 0.7 K which were obtained when validating DMS sharpened LST with contrast enhancement (Fig. 6(e)) with recently irrigated cases excluded.

Sharpened LST has been used to estimate high resolution ET in agricultural areas in number of studies. One of the earlier studies used synthetic low resolution LST data (i.e. Landsat observations resampled to 960 m) sharpened to higher resolution to assess the impact of the sharpening on ET fluxes produced with TSEB model (Agam et al., 2008). The results indicated that while in rainfed agriculture the use of sharpened LST reduced errors in ET by 30%–60% and significantly improved correlation, in irrigated agriculture the impact was marginal. A later study, in preparation for Sentinel-2 and Sentinel-3, used MODIS



(a) Rwanda, second dekade of November 2021 (b) Mozambique, last dekade of December 2021 (c) Sudan, second dekade of July 2021

Fig. 15. Example maps taken from online portal (et4fao.dhigroup.com) showing artifacts in average daily ET from a specified dekade due to persistent cloudiness during satellite data acquisition.

LST sharpened with Spot 5 shortwave imagery in a simplified TSEB model to obtain daily ET with RMSE of 1.2 mm/day and bias of −0.4 mm/day when validated against flux measurements from the "Las Tiesas" experimental farm (Bisquert et al., 2016b). A more recent study, combined thermal observations from Landsat, ECOSTRESS and VIIRS sensors, all sharpened to 30 m resolution using Sentinel-2 and Landsat shortwave reflectance, with a complex ET modeling and fusion scheme to obtain daily high resolution ET estimates (Xue et al., 2022). When validated against EC measurements in irrigated agriculture the resulting RMSE and bias at a weekly timestep (1.01 mm/day and −0.24 mm/day respectively) were of comparable magnitude to the results at the validation site in southeastern Spain from the current study for both TSEB-PT$_C$ (0.84 mm/day and −0.07 mm/day respectively) and ETLook$_C$ (1.13 mm/day and −0.79 mm/day respectively) at dekadal timestep.

WaPOR evapotranspiration has been used for assessing irrigation performance (Chukalla et al., 2022), evaluating potential improvements in crop water use productivity (Seijger et al., 2023; Hajirad et al., 2023) and other agricultural applications. In Blatchford et al. (2020) the WaPOR actual ET was validated against EC tower measurements spread across Africa and southern Spain. This validation was performed using Level 1 (250 m) product and achieved an overall RMSE of 1.2 mm/day and bias of 0.5 mm/day. In irrigated agricultural areas the accuracies were degraded (average RMSE of 1.5 mm/day and bias of 1.4 mm/day) which could be due to the heterogeneous nature of those landscapes. This compares to RMSE of 1.6 mm/day and bias of −1.23 mm/day obtained when comparing WaPOR Level 2 and Level 3 data against measurements in agricultural areas in Tunisia, Lebanon and Senegal (Table 3). ETLook$_C$ obtained RMSE of 1.4 mm/day and bias of −1.0 mm/day while TSEB-PT$_C$ achieved RMSE of 1.0 mm/day and bias of 0.1 mm/day at those same sites.

5.4. Relevance for upcoming thermal missions

In the period up to 2030 it is expected that at least four large satellite missions focusing on thermal LST observations with high spatio-temporal resolution will be launched. Those include LSTM from Copernicus Expansion (Koetz et al., 2018), Indian-French TRISHNA (Lagouarde et al., 2018), and SBG (Stavros et al., 2023) and evolution of Landsat missions (Wu et al., 2019) from NASA. While all those missions focus on improving both spatial and temporal resolution of thermal observations, none of them will by itself achieve daily geometric revisit time. Therefore, data fusion between different thermal and shortwave sensors will still be necessary to obtain daily representation of LST at field scale.

The method presented in Section 2.1.2 is sensor agnostic and therefore should be applicable with the above listed (and other) missions. The improved temporal resolution of future missions should also improve the feasibility of using temporal statistics of high-resolution LST as explanatory variables in the DMS model (option (2) in Section 2.1.2), which showed the best ability to reproduce the low LST of recently irrigated parcels (see Section 4.1). However, certain new issues could arise, such as differences between local overpass times of the different satellites (LSTM and other missions are planned to have overpass in the early afternoon local time). The first step to evaluating the new possibilities and issues is to use the LST product from the ECOSTRESS sensor (https://lpdaac.usgs.gov/products/eco2lstev001/, last accessed: 14.07.2023) in the different data fusion approaches. Due to its placement on the International Space Station, ECOSTRESS provides LST observations at varying temporal resolutions and overpass times. This will be the focus of upcoming research.

6. Conclusions

Reliable estimation of actual crop evapotranspiration at parcel scale requires LST observations with spatial resolution on the order of tens of meters and short revisit period. None of the currently operational satellite thermal sensors can fulfill this requirement. To overcome this limitation a number of thermal-optical data fusion methods have been developed. In Guzinski et al. (2021) high-resolution LST derived with fusion of Sentinel-3 and Sentinel-2 observations with Data Mining Sharpener was used in an ET modeling framework to demonstrate the feasibility of using Copernicus data for the purpose of SDG indicator 6.4.1 monitoring. However that study also exposed certain limitations both within the thermal sharpening method and within the ET modeling approach.

With the methodological improvements developed and implemented in this study it was possible to achieve negligible bias (−0.07 mm/day) and reasonably low RMSE (0.84 mm/day) when evaluating the modeled fluxes against field measurements taken in Spanish irrigated and rainfed agricultural sites. Extending the validation to sites in Africa and Middle East also resulted in acceptable bias (−0.14 mm/day) and RMSE (1.00 mm/day) despite high uncertainty of measurements at the Tal Amara site in Lebanon. Those results improve on the findings from (Guzinski et al., 2021), in large part due to improvements in the LST sharpening which led to decrease in RMSE of up to 1.5 K when comparing sharpened LST against *in situ* measurements.

Further investigation is required to determine the reasons why using Landsat LST as a explanatory variable within the DMS model improves the accuracy of LST in just-irrigated cases but degrades it otherwise. This degradation could be caused by the long temporal aggregation window (± 16 days) needed due to the infrequent Landsat acquisitions. If this is the case, improvements would be expected with new generation of thermal sensors with high spatio-temporal resolution such as proposed Copernicus LSTM mission.

Further validation should be performed in irrigated and not-irrigated agriculture in tropical regions to better understand the benefits and limitations of different gap filling methods for input and output data.

This study confirmed the main finding from Guzinski et al. (2021), namely the high suitability of Copernicus data for the production of WaPOR-like ET products with 20 m spatial resolution and 10-day temporal resolution. Those findings already led to incorporation of some of the methods (e.g. thermal sharpening using DMS) and data sources (e.g. ERA5 meteorological data or Sentinel-2 observations) into the recently released version 3 of the WaPOR portal. This highlights the utility of Copernicus satellite observations and services for monitoring SDG goals and indicators at continental and global scales.

CRediT authorship contribution statement

Radoslaw Guzinski: Conceptualization, Methodology, Software, Validation, Writing – original drat. **Héctor Nieto:** Conceptualization, Methodology, Software, Validation, Writing – original drat. **Rubén Ramo Sánchez:** Methodology, Validation, Writing – original drat. **Juan Manuel Sánchez:** Resources, Data Curation, Writing – review & editing. **Ihab Jomaa:** Resources, Data Curation. **Rim Zitouna-Chebbi:** Resources, Data Curation. **Olivier Roupsard:** Resources, Data Curation, Writing – review & editing. **Ramón López-Urrea:** Resources, Data Curation, Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank European Space Agency (ESA) for funding this study through ET4FAO project (contract no. 4000130120/20/I-DT). Additional funding was provided by Sat-ET4Drought project PID2021-127345OR-C32, funded by MCIN/AEI DOI:10.13039/501100011033 and FEDER. Field instrumentation and measurements from Spain were supported by the Spanish Ministry of Science and Innovation, MCIN/AEI, together with Next Generation EU/PRTR funds (projects PID2020-113498RB-C21, PID2021-123305 OB-C31 and TED2021-130405B-I00). Faidherbia-Flux site in Senegal received supports from EU-LEAP-Agri (RAMSES II), EU-DESIRA (CASSECS), EU-H2020 (SUSTAINSAHEL), AGROPOLIS and TOTAL Foundations (DSCATT), CGIAR (GLDC).

References

Agam, N., Kustas, W.P., Anderson, M.C., Li, F., Colaizzi, P.D., 2008. Utility of thermal image sharpening for monitoring field-scale evapotranspiration over rainfed and irrigated agricultural regions. Geophys. Res. Lett. 35 (2), http://dx.doi.org/10.1029/2007GL032195, URL https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2007GL032195, 2007GL03219.

Bastiaanssen, W.G.M., Cheema, M.J.M., Immerzeel, W.W., Miltenburg, I.J., Pelgrum, H., 2012. Surface energy balance and actual evapotranspiration of the transboundary Indus Basin estimated from satellite measurements and the ETLook model. Water Resour. Res. 48 (11), http://dx.doi.org/10.1029/2011WR010482, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011WR010482.

Bellvert, J., Jofre-Ċekalović, C., Pelechá, A., Mata, M., Nieto, H., 2020. Feasibility of using the two-source energy balance model (TSEB) with sentinel-2 and sentinel-3 images to analyze the spatio-temporal variability of vine water status in a vineyard. Remote Sens. 12 (14), 2299. http://dx.doi.org/10.3390/rs12142299, URL https://www.scilit.net/article/4a0123189bed8f0d26c4f1513e33375a, Publisher: MDPI.

Bisquert, M., Sánchez, J.M., Caselles, V., 2016a. Evaluation of disaggregation methods for downscaling MODIS land surface temperature to landsat spatial resolution in barrax test site. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (4), 1430–1438. http://dx.doi.org/10.1109/JSTARS.2016.2519099.

Bisquert, M., Sánchez, J.M., López-Urrea, R., Caselles, V., 2016b. Estimating high resolution evapotranspiration from disaggregated thermal images. Remote Sens. Environ. 187, 423–433. http://dx.doi.org/10.1016/j.rse.2016.10.049, URL http://www.sciencedirect.com/science/article/pii/S0034425716304291.

Blatchford, M.L., Mannaerts, C.M., Njuki, S.M., Nouri, H., Zeng, Y., Pelgrum, H., Wonink, S., Karimi, P., 2020. Evaluation of WaPOR V2 evapotranspiration products across Africa. Hydrol. Process. 34 (15), 3200–3221. http://dx.doi.org/10.1002/hyp.13791, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.13791.

Burchard-Levine, V., Nieto, H., Riaño, D., Migliavacca, M., El-Madany, T.S., Guzinski, R., Carrara, A., Martín, M.P., 2021. The effect of pixel heterogeneity for remote sensing based retrievals of evapotranspiration in a semi-arid tree-grass ecosystem. Remote Sens. Environ. 260, 112440. http://dx.doi.org/10.1016/j.rse.2021.112440, URL https://www.sciencedirect.com/science/article/pii/S0034425721001589.

Chukalla, A.D., Mul, M.L., Van Der Zaag, P., Van Halsema, G., Mubaya, E., Muchanga, E., Den Besten, N., Karimi, P., 2022. A framework for irrigation performance assessment using WaPOR data: The case of a sugarcane estate in Mozambique. Hydrol. Earth Syst. Sci. 26 (10), 2759–2778. http://dx.doi.org/10.5194/hess-26-2759-2022, URL https://hess.copernicus.org/articles/26/2759/2022/.

Delaunay, V., Desclaux, A., Sokhna, C., 2018. Niakhar, mémoires et perspectives. Recherches pluridisciplinaires sur le changement en afrique. Bull. Soc. Pathol. Exotique 111, 259–261. http://dx.doi.org/10.3166/bspe-2018-0049.

Delogu, E., Olioso, A., Alliès, A., Demarty, J., Boulet, G., 2021. Evaluation of multiple methods for the production of continuous evapotranspiration estimates from TIR remote sensing. Remote Sens. 13 (6), 1086. http://dx.doi.org/10.3390/rs13061086, URL https://www.mdpi.com/2072-4292/13/6/1086, Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

Eilers, P.H.C., 2003. A perfect smoother. Anal. Chem. 75 (14), 3631–3636. http://dx.doi.org/10.1021/ac034173t, URL https://pubs.acs.org/doi/10.1021/ac034173t.

European Space Agency, Airbus, 2022. Copernicus DEM. http://dx.doi.org/10.5270/ESA-c5d3d65, URL https://spacedata.copernicus.eu/collections/copernicus-digital-elevation-model, Institution: European Space Agency.

FRAME Consortium, 2020. WaPOR Data Manual - Evapotranspiration v2. Food and Agriculture Organization of the United Nations, URL https://bitbucket.org/cioapps/wapor-et-look/downloads/FRAME_ET_v2_data_manual_finaldraft_v2.2.pdf.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. Sci. Data 2 (1), 150066. http://dx.doi.org/10.1038/sdata.2015.66, URL https://www.nature.com/articles/sdata201566.

Galve, J.M., Sánchez, J.M., García-Santos, V., González-Piqueras, J., Calera, A., Villodre, J., 2022. Assessment of land surface temperature estimates from landsat 8-TIRS in a high-contrast semiarid agroecosystem. Algorithms intercomparison. Remote Sens. 14 (8), 1843. http://dx.doi.org/10.3390/rs14081843, URL https://www.mdpi.com/2072-4292/14/8/1843.

Gao, F., Kustas, W.P., Anderson, M.C., 2012. A data mining approach for sharpening thermal satellite imagery over land. Remote Sens. 4 (11), 3287–3319. http://dx.doi.org/10.3390/rs4113287, URL http://www.mdpi.com/2072-4292/4/11/3287.

García-Santos, V., Sánchez, J.M., Cuxart, J., 2022. Evapotranspiration acquired with remote sensing thermal-based algorithms: A state-of-the-art review. Remote Sens. 14 (14), 3440. http://dx.doi.org/10.3390/rs14143440, URL https://www.mdpi.com/2072-4292/14/14/3440, Number: 14 Publisher: Multidisciplinary Digital Publishing Institute.

Guzinski, R., Nieto, H., 2019. Evaluating the feasibility of using sentinel-2 and sentinel-3 satellites for high-resolution evapotranspiration estimations. Remote Sens. Environ. 221, 157–172. http://dx.doi.org/10.1016/j.rse.2018.11.019, URL http://www.sciencedirect.com/science/article/pii/S0034425718305285.

Guzinski, R., Nieto, H., Jensen, R., Mendiguren, G., 2014. Remotely sensed land-surface energy fluxes at sub-field scale in heterogeneous agricultural landscape and coniferous plantation. Biogeosciences 11 (18), 5021–5046. http://dx.doi.org/10.5194/bg-11-5021-2014, URL https://www.biogeosciences.net/11/5021/2014/.

Guzinski, R., Nieto, H., Sánchez, J.M., López-Urrea, R., Boujnah, D.M., Boulet, G., 2021. Utility of copernicus-based inputs for actual evapotranspiration modeling in support of sustainable water use in agriculture. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 11466–11484. http://dx.doi.org/10.1109/JSTARS.2021.3122573, Conference Name: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

Guzinski, R., Nieto, H., Sandholt, I., Karamitilios, G., 2020. Modelling high-resolution actual evapotranspiration through sentinel-2 and sentinel-3 data fusion. Remote Sens. 12 (9), 1433. http://dx.doi.org/10.3390/rs12091433, URL https://www.mdpi.com/2072-4292/12/9/1433, Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

Hajirad, I., Mohammadi, S., Dehghanisanij, H., 2023. Determining the critical points of a basin from the point of view of water productivity and water consumption using the wapor database. In: The 7th International Electronic Conference on Water Sciences. MDPI, p. 86. http://dx.doi.org/10.3390/ECWS-7-14322, URL https://www.mdpi.com/2673-4931/25/1/86.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horanyi, A., Munoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., al., e., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049. http://dx.doi.org/10.1002/qj.3803, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803.

Huryna, H., Cohen, Y., Karnieli, A., Panov, N., Kustas, W.P., Agam, N., 2019. Evaluation of TsHARP utility for thermal sharpening of sentinel-3 satellite images using sentinel-2 visual imagery. Remote Sens. 11 (19), 2304. http://dx.doi.org/10.3390/rs11192304, URL https://www.mdpi.com/2072-4292/11/19/2304.

Koetz, B., Bastiaanssen, W., Berger, M., Defourney, P., Bello, U.D., Drusch, M., Drinkwater, M., Duca, R., Fernandez, V., Ghent, D., Guzinski, R., Hoogeveen, J., Hook, S., Lagouarde, J., Lemoine, G., Manolis, I., Martimort, P., Masek, J., Massart, M., Notarnicola, C., Sobrino, J., Udelhoven, T., 2018. High spatio- temporal resolution land surface temperature mission - A copernicus candidate mission in support of agricultural monitoring. In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. pp. 8160–8162. http://dx.doi.org/10.1109/IGARSS.2018.8517433.

Kustas, W.P., Li, F., Jackson, T.J., Prueger, J.H., MacPherson, J.I., Wolde, M., 2004. Effects of remote sensing pixel resolution on modeled energy flux variability of croplands in iowa. Remote Sens. Environ. 92 (4), 535–547. http://dx.doi.org/10.1016/j.rse.2004.02.020, URL http://www.sciencedirect.com/science/article/pii/S003442570400183X.

Kustas, W.P., Nieto, H., Morillas, L., Anderson, M.C., Alfieri, J.G., Hipps, L.E., Villagarcía, L., Domingo, F., Garcia, M., 2016. Revisiting the paper "using radiometric surface temperature for surface energy flux estimation in Mediterranean drylands from a two-source perspective". Remote Sens. Environ. 184, 645–653. http://dx.doi.org/10.1016/j.rse.2016.07.024, URL http://www.sciencedirect.com/science/article/pii/S0034425716302814.

Lagouarde, J.-P., Bhattacharya, B., Crebassol, P., Gamet, P., Babu, S.S., Boulet, G., Briottet, X., Buddhiraju, K., Cherchali, S., Dadou, I., Dedieu, G., Gouhier, M., Hagolle, O., Irvine, M., Jacob, F., Kumar, A., Kumar, K.K., Laignel, B., Mallick, K., Murthy, C., Olioso, A., Ottle, C., Pandya, M.R., Raju, P.V., Roujean, J.-L., Sekhar, M., Shukla, M.V., Singh, S.K., Sobrino, J., Ramakrishnan, R., 2018. The Indian-French Trishna mission: Earth observation in the thermal infrared with high spatio-temporal resolution. In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Valencia, pp. 4078–4081. http://dx.doi.org/10.1109/IGARSS.2018.8518720, URL https://ieeexplore.ieee.org/document/8518720/.

López-Urrea, R., Martín de Santa Olalla, F., Fabeiro, C., Moratalla, A., 2006. Testing evapotranspiration equations using lysimeter observations in a semiarid climate. Agricult. Water Manag. 85 (1), 15–26. http://dx.doi.org/10.1016/j.agwat.2006.03.014.

López-Urrea, R., Martínez-Molina, L., de la Cruz, F., Montoro, A., González-Piqueras, J., Odi-Lara, M., Sánchez, J.M., 2016. Evapotranspiration and crop coefficients of irrigated biomass sorghum for energy production. Irrigat. Sci. 34 (4), 287–296. http://dx.doi.org/10.1007/s00271-016-0503-y.

Nieto, H., Guzinski, R., Graae, P., Jonas, ClaireBrenner, Mike, Gabrielmini, 2023. hectornieto/pyTSEB: v2.2. Zenodo, http://dx.doi.org/10.5281/ZENODO.8134956, URL https://zenodo.org/record/8134956.

Nieto, H., Guzinski, R., Graae, P., Sølvsteen, J., 2021. hectornieto/pyTSEB: ET4FAO. Zenodo, http://dx.doi.org/10.5281/ZENODO.4761365, URL https://zenodo.org/record/4761365, Language: en.

Norman, J.M., Anderson, M.C., Kustas, W.P., French, A.N., Mecikalski, J., Torn, R., Diak, G.R., Schmugge, T.J., Tanner, B.C.W., 2003. Remote sensing of surface energy fluxes at 10^1-m pixel resolutions. Water Resour. Res. 39 (8), 1221. http://dx.doi.org/10.1029/2002WR001775, URL http://onlinelibrary.wiley.com/doi/10.1029/2002WR001775/abstract.

O'Connor, B., Moul, K., Pollini, B., de Lamo, X., Allison, H., Albrecht, F., Guzinski, R., Larsen, H., McGlade, J., Paganini, M., 2020. Compendium of Earth Observation contributions to the SDG Targets and Indicators. European Space Agency, Paris, France, URL https://eo4society.esa.int/wp-content/uploads/2021/01/EO_Compendium-for-SDGs.pdf.

Pu, R., Bonafoni, S., 2023. Thermal infrared remote sensing data downscaling investigations: An overview on current status and perspectives. Remote Sens. Appl.: Soc. Environ. 29, 100921. http://dx.doi.org/10.1016/j.rsase.2023.100921, URL https://linkinghub.elsevier.com/retrieve/pii/S2352938523000034.

Rahimi, J., Ago, E.E., Ayantunde, A., Berger, S., Bogaert, J., Butterbach-Bahl, K., Cappelaere, B., Cohard, J.-M., Demarty, J., Diouf, A.A., Falk, U., Haas, E., Hiernaux, P., Kraus, D., Roupsard, O., Scheer, C., Srivastava, A.K., Tagesson, T., Grote, R., 2021. Modeling gas exchange and biomass production in West African Sahelian and Sudanian ecological zones. Geosci. Model Dev. 14 (6), 3789–3812. http://dx.doi.org/10.5194/gmd-14-3789-2021, URL https://gmd.copernicus.org/articles/14/3789/2021/.

Roupsard, O., Audebert, A., Ndour, A.P., Clermont-Dauphin, C., Agbohessou, Y., Sanou, J., Koala, J., Faye, E., Sambakhe, D., Jourdan, C., le Maire, G., Tall, L., Sanogo, D., Seghieri, J., Cournac, L., Leroux, L., 2020. How far does the tree affect the crop in agroforestry? New spatial analysis methods in a Faidherbia parkland. Agric. Ecosyst. Environ. 296, 106928. http://dx.doi.org/10.1016/j.agee.2020.106928, URL https://www.sciencedirect.com/science/article/pii/S0167880920301134.

Roupsard, O., Ferhi, A., Granier, A., Pallo, F., Depommier, D., Mallet, B., Joly, H.I., Dreyer, E., 1999. Reverse phenology and dry-season water uptake by *faidherbia albida* (Del.) A. Chev. in an agroforestry parkland of Sudanese west Africa. Funct. Ecol. 13 (4), 460–472. http://dx.doi.org/10.1046/j.1365-2435.1999.00345.x, URL https://besjournals.onlinelibrary.wiley.com/doi/10.1046/j.1365-2435.1999.00345.x.

Sánchez, J.M., Galve, J.M., González-Piqueras, J., López-Urrea, R., Niclòs, R., Calera, A., 2020. Monitoring 10-m LST from the combination MODIS/Sentinel-2, validation in a high contrast semi-arid agroecosystem. Remote Sens. 12 (9), 1453. http://dx.doi.org/10.3390/rs12091453, URL https://www.mdpi.com/2072-4292/12/9/1453.

Sánchez, J.M., López-Urrea, R., Valentín, F., Caselles, V., Galve, J.M., 2019. Lysimeter assessment of the simplified two-source energy balance model and eddy covariance system to estimate vineyard evapotranspiration. Agricult. Forest Meteorol. 274, 172–183. http://dx.doi.org/10.1016/j.agrformet.2019.05.006.

Sánchez, J.M., Simón, L., González-Piqueras, J., Montoya, F., López-Urrea, R., 2021. Monitoring crop evapotranspiration and transpiration/evaporation partitioning in a drip-irrigated Young almond orchard applying a two-source surface energy balance model. Water 13 (15), 2073. http://dx.doi.org/10.3390/w13152073, URL https://www.mdpi.com/2073-4441/13/15/2073, Number: 15 Publisher: Multidisciplinary Digital Publishing Institute.

Seijger, C., Chukalla, A., Bremer, K., Borghuis, G., Christoforidou, M., Mul, M., Hellegers, P., Van Halsema, G., 2023. Agronomic analysis of wapor applications: Confirming conservative biomass water productivity in inherent and climatological variance of wapor data outputs. Agric. Syst. 211, 103712. http://dx.doi.org/10.1016/j.agsy.2023.103712, URL https://linkinghub.elsevier.com/retrieve/pii/S0308521X23001178.

Stavros, E.N., Chrone, J., Cawse-Nicholson, K., Freeman, A., Glenn, N.F., Guild, L., Kokaly, R., Lee, C., Luvall, J., Pavlick, R., Poulter, B., Schollaert Uz, S., Serbin, S., Thompson, D.R., Townsend, P.A., Turpie, K., Yuen, K., Thome, K., Wang, W., Zareh, S.-K., Nastal, J., Bearden, D., Miller, C.E., Schimel, D., 2023. Designing an observing system to study the surface biology and geology (SBG) of the earth in the 2020s. J. Geophys. Res.: Biogeosci. 128 (1), http://dx.doi.org/10.1029/2021JG006471, URL https://onlinelibrary.wiley.com/doi/10.1029/2021JG006471.

Tsendbazar, N.-E., Tarko, A., Li, L., Herold, M., Lesiv, M., Fritz, S., Maus, V., 2020. Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015–2019: Validation Report. Tech. Rep., Zenodo, http://dx.doi.org/10.5281/ZENODO.3938974, URL https://zenodo.org/record/3938974, Version Number: Dataset v3.0, doc issue 1.0.

Twine, T., Kustas, W., Norman, J., Cook, D., Houser, P., Meyers, T., Prueger, J., Starks, P., Wesely, M., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agricult. Forest Meteorol. 103 (3), 279–300. http://dx.doi.org/10.1016/S0168-1923(00)00123-4.

Wang, Z., Sui, L., Zhang, S., 2022. Generating daily land surface temperature downscaling data based on sentinel-3 images. Remote Sens. 14 (22), 5752. http://dx.doi.org/10.3390/rs14225752, URL https://www.mdpi.com/2072-4292/14/22/5752.

Wu, Z., Snyder, G., Vadnais, C., Arora, R., Babcock, M., Stensaas, G., Doucette, P., Newman, T., 2019. User needs for future Landsat missions. Remote Sens. Environ. 231, 111214. http://dx.doi.org/10.1016/j.rse.2019.111214, URL https://linkinghub.elsevier.com/retrieve/pii/S0034425719302275.

Xue, J., Anderson, M.C., Gao, F., Hain, C., Knipper, K.R., Yang, Y., Kustas, W.P., Yang, Y., Bambach, N., McElrone, A.J., Castro, S.J., Alfieri, J.G., Prueger, J.H., McKee, L.G., Hipps, L.E., del Mar Alsina, M., 2022. Improving the spatiotemporal resolution of remotely sensed ET information for water management through landsat, sentinel-2, ECOSTRESS and VIIRS data fusion. Irrigat. Sci. http://dx.doi.org/10.1007/s00271-022-00799-7.

Yilmaz, M.T., Crow, W.T., 2013. The optimality of potential rescaling approaches in land data assimilation. J. Hydrometeorol. 14 (2), 650–660. http://dx.doi.org/10.1175/JHM-D-12-052.1, URL http://journals.ametsoc.org/doi/10.1175/JHM-D-12-052.1.

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.-E., Xu, P., Ramoino, F., Arino, O., 2022. ESA WorldCover 10 m 2021 v200. Zenodo, http://dx.doi.org/10.5281/ZENODO.7254221, URL https://zenodo.org/record/7254221.