ET4FAO D4.2
Final report

V1.0
January 13, 2022

Contract No. 4000130120/20/I-DT
"Increasing Crop Water Use Efficiency at Multiple Scales Using Sentinel Evapotranspiration [ET4FAO]"

Submitted by

DHI GRAS

In cooperation with:

COMPLUTIG
EMPRESA DE BASE TECNOLÓGICA DE LA UNIVERSIDAD DE ALCALÁ

DOCUMENT RELEASE SHEET

Authors:	Radoslaw Guzinski (DHI GRAS), Héctor Nieto (COMPLUTIG)
Approval (internal)	Torsten Bondo (DHI GRAS)
Approval (ESA)	Benjamin Koetz
Distribution	Public distribution

CHANGE RECORD

Date	**Page(s)**	**Change record**
13.01.2021	–	Change distribution to "Public distribution"
14.07.2021	–	Version 1 submitted to ESA
13.07.2021	–	Draft version sent to partners

This document has been prepared under the DHI Business Management System certified by Bureau Veritas to comply with ISO 9001 (Quality Management)

Contents

Preface

The purpose of ET4FAO project was to demonstrate the feasibility of consistent monitoring of evapotranspiration (ET), from field to national scales, using Copernicus data sources. ET modeled using Copernicus data is compared against methods and data from FAO-run WaPOR portal.

This is the final deliverable of the project. It summarizes the methods used during the project, presents validation and inter-comparison of produced ET datasets and contains recommendations on the use of Copernicus data for ET monitoring in the context of Sustainable Development Goals.

Chapter 1

Introduction

1.1 Introduction

Food and Agriculture Organization (FAO) is the custodian agency of Sustainable Development Goal (SDG) indicator 6.4.1 - change in water use efficiency over time. This indicator is composed of three parts covering water use in three sectors of the economy: agriculture; industry, mining and power production; and services. Out of the three, agriculture is by far the largest consumer of water. Fortunately, it is also the component which is most suitable for monitoring through satellite remote sensing thus enabling consistent, independent and reliable estimates across national boundaries. Estimating water use efficiency in agriculture consists of estimating agricultural water input (irrigation) and agricultural output (e.g. yield). In the ET4FAO project the focus is on estimating actual evapotranspiration (ET) through Earth observation (EO), which is an essential information for deriving agricultural water use efficiency at significant spatial and temporal scales.

To encourage the use of EO data in SDG indicator 6.4.1 reporting, FAO is running a portal called WaPOR (`https://wapor.apps.fao.org`) which contains all the required EO-based products. Currently WaPOR portal uses satellite observations from Terra and Aqua, Proba-V (replaced by Sentinel-2 since 2020) and Landsat satellites to estimate ET at field to continental scales. However, the Sen-ET project (`https://www.esa-sen4et.org/`) has demonstrated that by utilizing the optical observations acquired at high spatial resolution (10 m - 60 m) by Sentinel-2 (S2) satellites together with thermal observations acquired by Sentinel-3 (S3) satellites at medium spatial resolution (1 km) it is possible to derive high quality ET estimates at field scales. Since Sentinel satellites form part of the Copernicus program, they have the advantage to being run as operational services, meaning a guaranteed long term continuity, redundancy and high data quality. In addition all Copernicus data is available freely and openly.

In the ET4FAO project we evaluated the feasibility of replacing current data sources used in WaPOR by Copernicus. At the same time we compared the outputs of the current ET model used by WaPOR (ETLook) by a model which was previously shown to perform well with Copernicus data (TSEB-PT). The schematic of the different configurations of data sources and ET models which were run in this project is presented in Fig. 1.1.

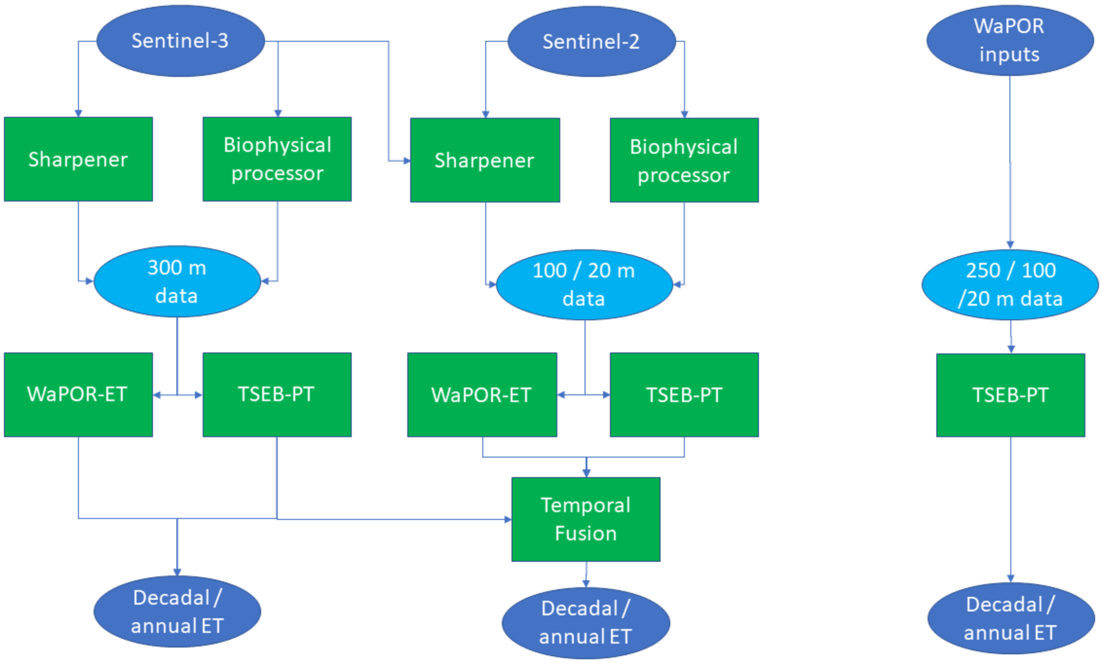


Figure 1.1: Schematic of configurations of data sources and ET models which are run in ET4FAO

1.2 Output products

WaPOR portal provides four ET-related products: actual evaporation (E), actual transpiration (T), interception (I) and actual evapotranspiration and interception (ETI). E represents water transfer from the top-soil into the atmosphere. T represents water transfer from the root-zone, through vegetation, into the atmosphere. I represents ponded rainwater intercepted by the leaves. Finally, ETI is the sum of the previous three products.

All four products and provided in dekadal (three dekades per month) time-steps with the value representing the average daily water transfer (in mm/day) during the dekade. The products are produced at three spatial levels (Level 1 - Continental, Level 2 - National, Level 3 - Local), depending on which satellite data is used as input, as shown in Table 1.1. In ET4FAO, all four product types are derived from Copernicus data and all four are also be produced at three spatial resolutions. However those resolutions differ slightly due to spatial characteristics of the satellite sensors used. Since Sentinel-3 thermal observations are used at all three spatial scales, the estimates of surface energy fluxes, and thus of evapotranspiration, should be consistent from local to continental scales.

In ET4FAO the ET estimates covering the whole of Lebanon and Tunisia in year 2019 were produced at the three spatial scales. They can be viewed online at `https://et4fao.dhigroup.com`. This was compared to Local scale estimates from WaPOR portal from Bekaa valley in Lebanon and Continental and National products covering the whole of Lebanon and Tunisia. In addition, ET estimates were produced in an agricultural area in southern Spain to allow validation against a number of field measurements stations (flux towers and lysimeters) in irrigated and non-irrigated agricultural fields in a climate similar to that of Lebanon and Tunisia.

Table 1.1: Spatial scales and main satellite data sources used at those scales in WaPOR portal and ET4FAO project

Spatial Scale	Spatial resolution and sensors used currently in WaPOR	Spatial resolution and sensors when using Copernicus data (ET4FAO)
Level 1 - Continental	250 m – MODIS on Terra and Aqua	300 m – SLSTR and OLCI on Sentinel-3A/B
Level 2 - National	100 m – MODIS on Terra and Aqua and Vegetation on PROBA-V (MSI on Sentinel-2 since 2020)	100 m – SLSTR on Sentinel-3 A/B and MSI on Sentinel-2 A/B
Level 3 - Local	30 m – OLI on Landsat 8 and ETM+ on Landsat 7 and TM on Landsat 5	20 m – SLSTR on Sentinel-3 A/B and MSI on Sentinel-2 A/B

1.3 Report structure

Chapter 2 of this report describes the data sources and pre-processing methods of data coming from the Copernicus program as well as data currently used in WaPOR portal. In addition, the theory behind the two ET models is laid out. This is followed by Chapters 3 and 4 where the different model runs are first validated against in-situ measurements and then inter-compared against each other and global ET products. Chapter 5 compares the results obtained in this study with other studies relying on WaPOR ET, while Chapter 6 contains recommendations and conclusions reached as a result of this project.

Chapter 2

Data preparation and evapotranspiration modeling

2.1 Copernicus Data

The main purpose of the ET4FAO project was to demonstrate the feasibility of using Copernicus data for country-wide operational, accurate and consistent estimation of evapotranspiration at different spatial resolutions, raging from 20 m to 300 m. For this purpose three sets of Copernicus data were used: meteorological, Sentinel (earth observation) and ancillary. The sources and pre-processing steps for those different data are described in the subsequent sections.

2.1.1 Meteorological Data

Meteorological inputs, which are critical for accurate estimation of ET, are based on ERA5 reanalysis dataset [1] produced by European Center for Medium Range Weather Forecasts and distributed through the Copernicus Climate Data Store (`https://cds.climate.copernicus.eu`). The ERA5 data contains surface meteorological parameters covering the whole Earth on 30 km grid and hourly temporal resolution and going back to 1950. New data is distributed with a 3-month delay, due to stringent quality checks. However a dataset called ERA5T (T for preliminary Near Real Time) is distributed with a five day delay and is also of very high quality. Therefore, ERA5 is used for historical analysis while ERA5T is used in near-real time processing.

Both instantaneous and daily parameters were derived from ERA5 data. Instantaneous parameters are used to drive the ET model and included air temperature, vapor pressure, wind speed, surface pressure, and instantaneous solar irradiance. All instantaneous data was temporarily interpolated to the time of Sentinel-3 SLSTR acquisition over the area of interest. Daily parameters are used to extrapolate and interpolate the instantaneous estimates of ET and include solar irradiance, precipitation and reference ET. They were integrated over a 24 hours period starting at midnight local time.

Instantaneous air temperature was based on 2 m air temperature ERA5 field. Due to the low spatial resolution of the meteorological data it was assumed that the air temperature that better represents meteorological conditions at that resolution is the temperature at atmospheric blending height (set to be 100 m above ground), where the impact of local surface conditions on those parameters is not so direct. The resolution of air temperature was additionally enhanced by using a 300 m resolution DEM (see Section 2.1.3) and ICAO standard lapse rate to correct for temperature changes due to changes in elevation between 2m above the geopotential height at which ERA5 temperature was produced and the 100m above DEM surface elevation.

Vapor pressure was derived from 2 m dew point temperature ERA5 field. Similarly to air temperature it was assumed to represent conditions at blending height and was corrected for changes in elevation using DEM and standard dew point lapse rate.

Wind speed was based on 100 m east-west and north-south ERA5 fields. Apart from trigonometric calculation to obtain total magnitude of wind speed no other preprocessing was performed.

Surface pressure is based on ERA5 field of the same name. Similarly to air temperature, it was corrected for changes due to varying elevation using a DEM.

The final instantaneous parameter was surface solar irradiance. Clear sky conditions were assumed since this parameter is only used at a time and place where thermal observations of the surface by Sentinel-3 satellite were possible. In addition, solar irradiance was also corrected by elevation, incidence angle and terrain shading. Firstly, irradiance on a horizontal surface was estimated using aerosol optical thickness, total column water vapor and air temperature ERA5 fields [2]. Subsequently this was corrected for elevation and terrain orientation to estimate irradiance on a tilted surface [3].

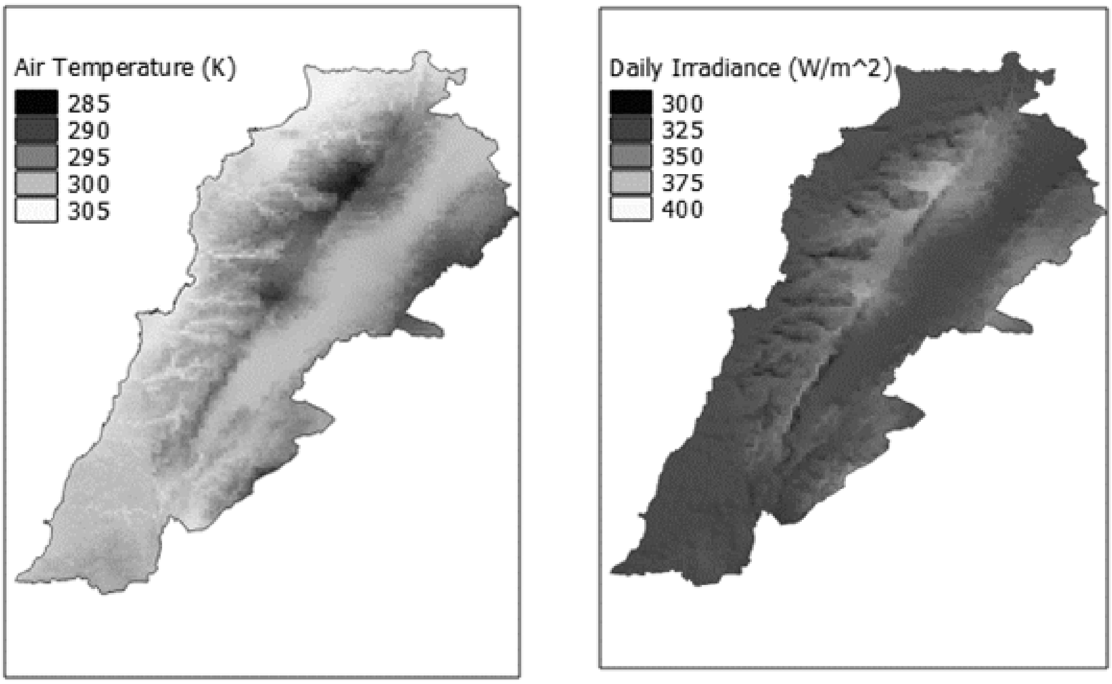


Figure 2.1: Example results of correcting air temperature and daily solar irradiance for topographical features in Lebanon.

Daily solar irradiance is calculated using 24 hour integration of ERA5 downward surface irradiance product and clear sky irradiance estimated as described above. At each hourly timestep a cloudiness factor is first estimated using the ratio of surface to horizontal clear sky irradiance. Based on this, beam and diffuse radiation components are derived and corrected for terrain orientation. Finally all the mean daily value is calculated.

Daily precipitation is based on a sum of two ERA5 products: large scale precipitation and convective precipitation. Those products are first added at hour time-steps, followed by summation over 24 hours. No additional corrections are performed.

Finally, the reference ET is estimated using the FAO56 model. It is calculated using daily mean values of air temperature, dew point temperature, wind speed, surface pressure and surface solar radiation ERA5 products. All the relevant terrain corrections are applied to those products before they are used in the FAO56 model.

2.1.2 Sentinel Data

Data acquired by three sensors on two types of Sentinel satellites is essential for ET model inputs. The sensors are MSI on Sentinel-2 satellites and SLSTR and OLCI on Sentinel-3 satellites.

MSI on Sentinel-2 provides high-resolution (20 m in our case) multi-spectral observations of Earth's surface which can be used to characterize surface biophysical properties. Those properties are used to model ET at 20 m and 100 m spatial resolutions. In addition, 20 m reflectance is used to sharpen the SLSTR LST, which is also needed to model ET at 20 m and 100 m resolutions. L1C S2 product was downloaded from Amazon Web Services (`https://registry.opendata.aws/sentinel-2/`) or from CREODIAS (`https://finder.creodias.eu/`) data stores. This product was used as input to Sen2Cor atmospheric correction model [4] as well as to Fmask cloud masking model [5].

The output top-of-canopy (TOC) reflectance from Sen2Cor was used within biophysical processor available in SNAP software (`https://step.esa.int/main/`) and subsequent Python scripts to derive leaf area index (LAI), fraction of vegetation which is green (f_g), fractional vegetation cover (FVC), and leaf broadband reflectance and transmittance and soil broadband reflectance as described in Section 2.3.1 of Guzinski et al. (2020)[6] based on the methodology developed in Sen-ET project.

A simple temporal compositing scheme was used to reduce data gaps due to clouds in S2 observations. For each date on which ET was to be modeled (i.e. thermal SLSTR data was available) all S2 images falling within 10 days were selected. The cloud free pixels were iteratively picked from the selected images starting with the ones closest to the target date. This was performed for TOC reflectance (Fig 2.2) and all derived biophysical parameters.

S3 Synergy SY_2_SYN product, which combines surface reflectance from shortwave optical bands on OLCI and SLSTR instruments, was used to characterize surface biophysical properties at 300 m spatial resolution and to sharpen SLSTR LST to 300 m. This product was retrieved from CREODIAS data store. The original idea was to derive biophysical parameters directly from SY_2_SYN through a radiative transfer model. However, this led to persistent bias between the S2 and S3 biophysical parameters which would introduce inconsistencies between the ET modeled at different spatial resolutions using different data. Therefore an alternative method was developed in which a machine-learning (ML) model (random forest) was trained using a S2 biophysical parameter resampled to 300 m and S3 Synergy reflectance. The model was then applied to S3 Synergy reflectance to derive the biophysical parameters at 300 m resolution. In order to preserve observation geometry between S2 and S3 acquisitions (i.e. close to nadir view) a 10-day minimum view zenith angle (VZA) composite was created from SY_2_SYN products before training and

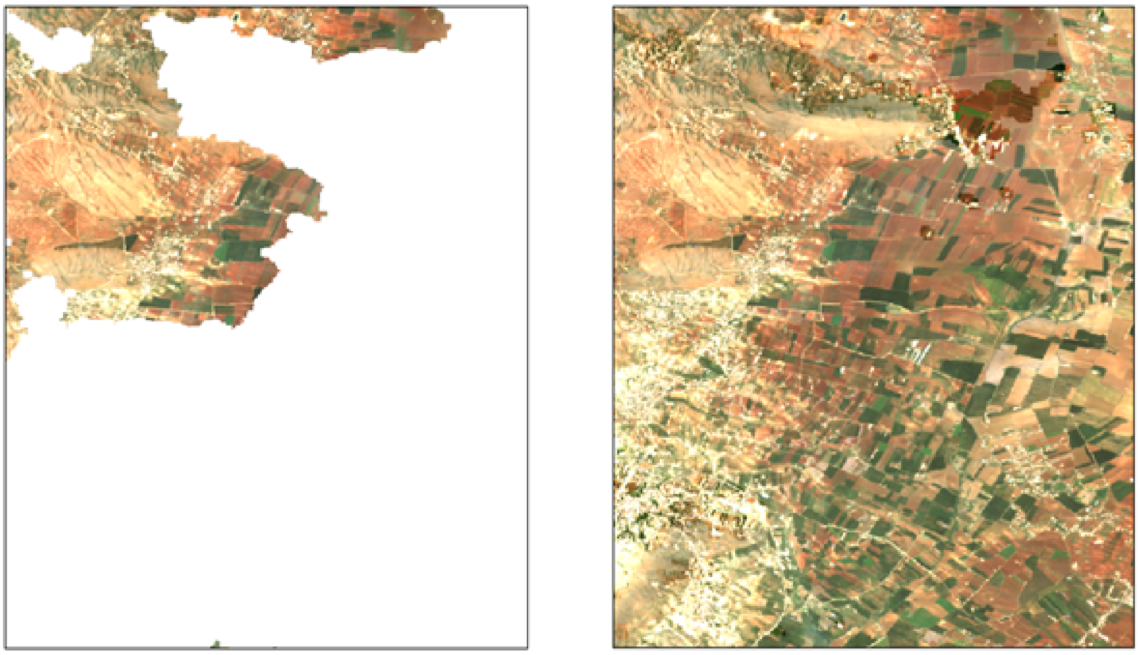


Figure 2.2: Example of input and output images of the S2 temporal compositing method

applying the biophysical ML model. Results of this approach are shown in Fig 2.3 and Fig. 2.4. It is worth noting that this approach trains the model from any available S2 tile within the SYN footprint, not requiring processing all S2 tiles included in the SYN scene. The trained random forest can then be applied to the whole SYN scene without any significant loss of accuracy.

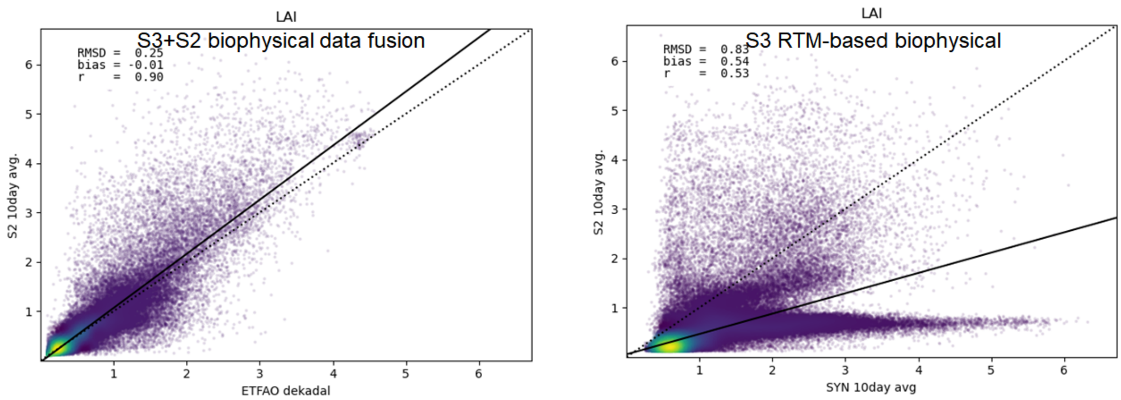


Figure 2.3: Scatter plot of S2 LAI versus S3 Synergy LAI derived using S2-S3 biophysical data fusion and directly using radiative transfer model and SY_2_SYN reflectance values

S3 SLSTR L2 (land surface temperature - LST) product is required to characterize energy fluxes at the land surface, of which ET is one. The product was retrieved from CREODIAS data store. Invalid pixels were identified using mask_in layer of the L2 file and following rules:

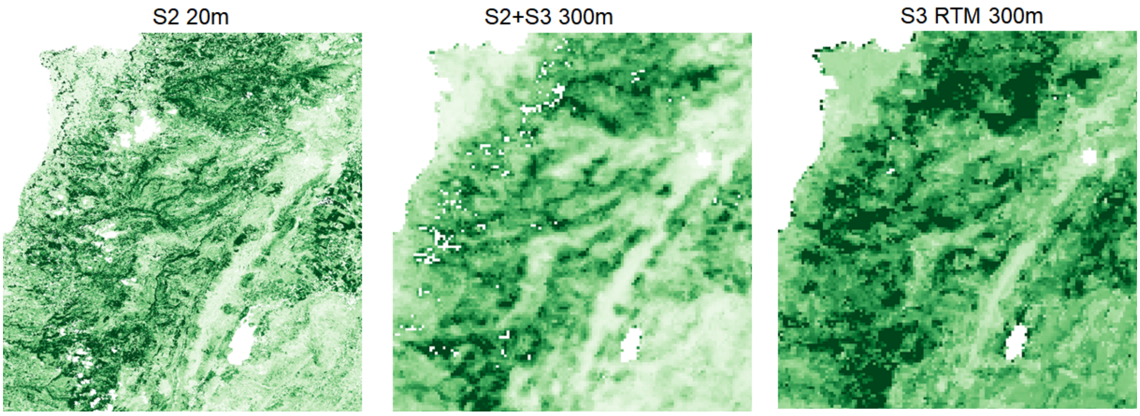


Figure 2.4: Maps of S2 LAI versus S3 Synergy LAI derived using S2-S3 biophysical data fusion and directly using radiative transfer model and SY_2_SYN reflectance values on 19 June 2019 in Lebanon

- LST $<$ 273.15 K
- LST - air temperature $<$ -2 K
- view zenith angle $>$ 45 degrees

Minimum view zenith angle (VZA) compositing was used in case of multiple SLSTR observations of the same area on the same day (i.e. by both Sentinel-3A and Sentinel-3B). This was done under the assumptions that smaller VZA would result in more accurate LST retrieval due to both shorter atmospheric path and reduced thermal directional effects.

Data Mining Sharpener (DMS) was used to sharpen the minimum VZA LST composite with spatial resolution of around 1 km, using shortwave reflectance and DEM at higher spatial resolution as described in 2.4 of [6]. Temporal composite of S2 TOC reflectance with 20 m resolution and centered on the date of S3 overpass was used in case of sharpening to 20 m. When sharpening to 300 m, SY_2_SYN product product acquired at the same time as LST was used. SLSTR VZA and meteorological inputs which were temporarily interpolated to the time of S3 overpass were resampled to 20 m or 300 m resolutions using bilinear resampling.

2.1.3 Ancillary Data

Two ancillary sources of data were used: Copernicus Global Land Cover (CGLC) for year 2019 and Shuttle Radar Topography Mission (SRTM) DEM.

The CGLC map was selected as it is also used as input to current WaPOR products. In fact, the maps used in Lebanon and Tunisia were downloaded directly from WaPOR portal to guarantee consistency. Land cover map is used to assign ET model parameters which are difficult to estimate directly from other satellite data. Those parameters, and values assigned to different land cover classes, are listed in Table 2.1. TSEB-PT model (see Chapter 2.3) requires all of the parameters, apart from stomatal resistance. ETLook model requires only maximum vegetation height and stomatal conductance and the values for those two parameters were extracted directly from the WaPOR documentation. Values of other parameters were set as in [6].

CGLC-LC	$h_{C,min}$ (m)	$h_{C,max}$ (m)	PAI_{max} (-)	f_C (-)	w_C/h_C (-)	l_w (m)	χ	r_{st} (s/m)
20	2	2	0	1	1	0.05	1	175
30	0.1	1	4	1	1	0.02	0.5	175
40	0.15	1.5	5	1	1	0.02	0.5	125
41	0.15	1.5	5	1	1	0.02	0.5	125
42	0.15	1.5	5	1	1	0.02	0.5	125
43	0.15	1.5	5	1	1	0.02	0.5	125
50	10	10	0	0	0	0	0	400
60	0.1	0.1	0	0	0	0	0	100
70	0.1	0.1	0	0	0	0	0	100
80	0.1	0.1	0	0	0	0	0	100
81	0.1	0.1	0	0	0	0	0	100
90	2	2	5	1	1	0.1	1	150
100	0.3	0.3	0	1	1	0.005	1	180
111	10	10	0	0.8	2	0.05	1	300
112	5	5	0	0.8	1	0.15	0.7	180
113	10	10	0	0.8	2	0.05	1	300
114	8	8	0	0.8	1	0.15	0.7	190
115	8	8	0	0.8	1.5	0.1	0.8	200
116	5	5	0	0.8	1.5	0.1	0.8	180
121	5	5	0	0.3	1.5	0.05	0.8	250
122	4	4	0	0.3	1	0.15	0.7	180
123	5	5	0	0.3	1.5	0.05	0.8	250
124	4	4	0	0.3	1	0.15	0.7	200
125	5	5	0	0.3	1.5	0.1	0.8	180
126	3	3	0	0.3	1.5	0.1	0.8	250
200	0.1	0.1	0	0	0	0	0	100

Table 2.1: Land cover based Look-Up-Table for ancillary parameters used in ET models. CGLC-LC is the land cover code for the Copernicus Global Land Cover legend (`https://land.copernicus.eu/global/products/lc`); $h_{C,min}$ is the minimum canopy height; $h_{C,max}$ is the maximum canopy height occurring when Plant Area Index (PAI) reaches PAI_{max}; f_C is fraction of the ground occupied by a clumped canopy (fC = 1 for a homogeneous canopy); w_C/h_C is canopy shape parameter, representing the canopy width to canopy height ratio; l_w is the average leaf size; χ Campbell [7] leaf angle distribution parameter; r_{st} is minimum stomatal resistance.

SRTM DEM was selected because it is used as default in Sen2Cor algorithm and also it is the DEM used in the WaPOR methodology. The DEM is used for three main purposes: during Sen2Cor atmospheric correction, to correct meteorological parameters for terrain effects (elevation and illumination conditions), and in the DMS thermal sharpening model to add elevation and illumination conditions as predictor variables.

2.2 WaPOR Data

WaPOR portal provides ET maps at three different resolutions: continental scale at 250 m resolution (called Level 1), national scale at 100 m resolution (called Level 2), local scale at 30 m resolution (called Level 3). While meteorological data used at the three levels is the same (GEOS-5 for weather, CHIRPS for precipitation and MSG for solar irradiance), the satellite data differs. For Level 1, both shortwave-optical and thermal data are obtained from MODIS sensor on board of Terra and Aqua satellites. At Level-2, thermal data is still acquired by MODIS but shortwave-optical comes from PROBA-V observations (Sentinel-2 since 2020). For Level-3 modeling both types of data come from Landsat satellites. The different data sources are summarized in Figures 2.5 to 2.7, which reproduce tables from Section 3 of "WaPOR Data Manual, Evapotranspiration v2.2" (`https://bitbucket.org/cioapps/wapor-et-look/downloads/FRAME_ET_v2_data_manual_finaldraft_v2.2.pdf`) [8].

Input data components	Type of input	Sensor	Data product	Comment
Precipitation	Model		CHIRPS v2, CHIRP	
Surface albedo	Sensor	MODIS	MOD09GA, MOD09GQ	
Weather data (temp, specific humidity, wind speed, air pressure, aerosol optical depth)	Model		MERRA/GEOS-5	MERRA used prior to start of GEOS-5 (21-2-2014)
NDVI	Sensor	MODIS	MOD09GQ	
Land Surface Temperature	Sensor	MODIS	MOD11A1, MYD11A1	Used to derive Soil Moisture Stress
Elevation, slope and aspect	Static		SRTM	Elevation, slope and aspect are derived from the DEM
Transmissivity	Model	MSG		Transmissivity is derived from MSG shortwave radiation products
Land Cover	Static		WaPOR L1 Land Cover Classification	If L1 LC was not yet available, preliminary LC obtained from ESA GlobCover.

Figure 2.5: WaPOR input data sources for the production of evapotranspiration data components (E,T, and I) at Level 1. Extracted from [8].

Pre-processing of WaPOR-like input data used in this study follows closely the methodology described in [8]. The only difference is pre-processing of shortwave optical Landsat data. While section 4.1 of WaPOR Data Manual describes using SMAC model [9] for atmospheric correction and manual cloud masking, in ET4FAO we relied on newly released Col-

Input data components	Type of input	Sensor	Data product	Comment
Precipitation	Model		CHIRPS v2, CHIRP	
Surface albedo	Sensor	Proba-V		PROBA-V data are available from March 2014, for earlier dates the Level 1 Surface albedo based on MODIS MOD09GQ, MOD09GA is resampled to 100m
Weather data (temp, specific humidity, wind speed, air pressure, aerosol optical depth)	Model		MERRA/GEOS-5	MERRA used prior to start of GEOS-5 (21-2-2014)
NDVI	Sensor	Proba-V		PROBA-V data are available from March 2014, data for earlier dates uses MODIS MOD09GQ, resampled to 100m
Land Surface Temperature	Sensor	MODIS	MOD11A1, MYD11A1	Used to derive Soil Moisture Stress
Elevation, slope and aspect	Static		SRTM	Elevation, slope and aspect are derived from the DEM
Transmissivity	Model	MSG		Transmissivity is derived from MSG shortwave radiation products
Land Cover	Static		WaPOR L2 Land Cover Classification	If L2 LC was not yet available, preliminary LC obtained from ESA GlobCover.

Figure 2.6: WaPOR input data sources for the production of evapotranspiration data components (E,T, and I) at Level 2. Extracted from [8].

Input data components	Type of input	Sensor	Data product	Comment
Precipitation	Model		CHIRPS v2, CHIRP	
Surface albedo	Sensor	Landsat 5 TM Landsat 7 ETM+ Landsat 8 OLI	L1TP	
Weather data (temp, specific humidity, wind speed, air pressure, aerosol optical depth)	Model		MERRA/GEOS-5	MERRA used prior to start of GEOS-5 (21-2-2014)
NDVI	Sensor	Landsat 5 TM Landsat 7 ETM+ Landsat 8 OLI	L1TP	
Land Surface Temperature	Sensor	Landsat 5 TM Landsat 7 ETM+ Landsat 8 OLI	L1TP	Used to derived Soil Moisture Stress
Elevation, slope and aspect	Static		SRTM	Elevation, slope and aspect are derived from the DEM
Transmissivity	Model	MSG		Transmissivity is derived from MSG shortwave radiation products
Land Cover	Static		WaPOR L3 Land Cover Classification	If L3 LC was not yet available, preliminary LC obtained from ESA CCI 20m LC (for 2016).

Figure 2.7: WaPOR input data sources for the production of evapotranspiration data components (E,T, and I) at Level 3. Extracted from [8].

lection 2 of Landsat L2 products (`https://www.usgs.gov/core-science-systems/nli/landsat/landsat-collection-2`). In addition, sections 4.1 and 4.2 of [8] describe slightly different approaches for gap-filling and smoothing the final NDVI and albedo products. In ET4FAO we used the method described for albedo also for NDVI, i.e. Savitzky-Golay filtering [10].

2.3 Evapotranspiration modeling

2.3.1 TSEB-PT

TSEB-PT stands for Two Source Energy Balance - Priestley Taylor [11] and it is the ET model which was shown in Sen-ET project to produce most accurate land-surface energy fluxes compared to two other approaches. As the name implies, the model considers vegetation and soil as two sources of land-surface energy fluxes. The energy transfer from the sources into the atmosphere are estimated separately, although they are linked (in an analogy to electrical circuits) by resistances to heat transfer which are arranged in series network. The model estimates net radiation of both canopy and soil (Rn_C and Rn_S respectively), sensible heat flux of both canopy and soil (H_C and H_S respectively) and ground heat flux of only soil (G) as shown in Fig. 2.8. Since energy balance must hold, latent heat flux, which is the energy used in evapotranspiration, of both canopy and soil (LE_C and LE_S respectively) is calculated as the residual of the other fluxes:

$$LE_C = R_{n,C} - H_C \tag{2.1a}$$
$$LE_S = R_{n,S} - H_S - G \tag{2.1b}$$

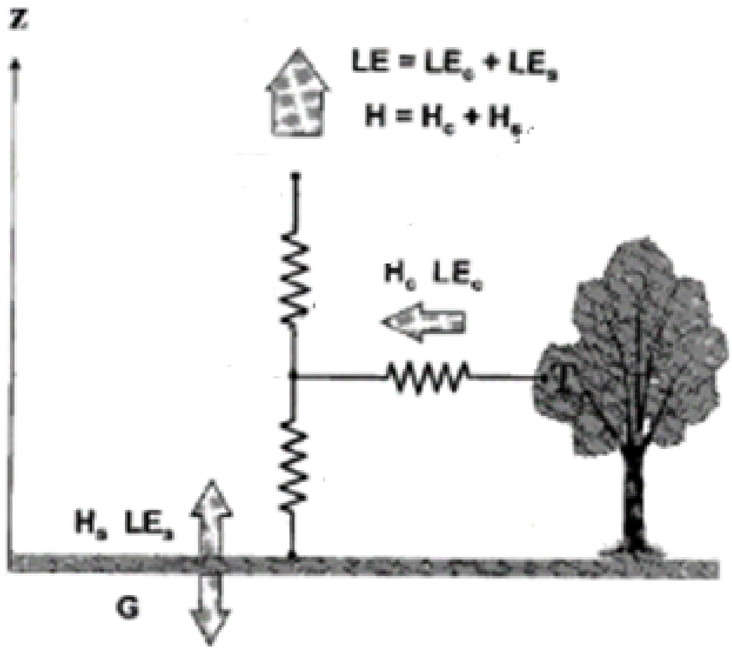


Figure 2.8: Simplified schematic of the TSEB-PT model, based on [12]

Since the model has to estimate energy fluxes of both soil and canopy based on a single bulk LST measurement, it initially takes the assumption that vegetation is transpiring at potential rate based on the Priestley-Taylor equation. This first guess transpiration is iteratively reduced within the model in case unrealistic fluxes are obtained (e.g. negative latent heat fluxes for either soil or canopy during daytime). TSEB-PT estimates instantaneous fluxes at the time of thermal image acquisition in units of Wm^{-2}. They are extrapolated to daily ET (mm/day) based on the assumption that ratio of latent heat flux to solar irradiance remains invariant during day-time hours [13]. More details on TSEB-PT model and the way it was applied in Sen-ET project can be found in Sections 2.1 and 2.1.2 of [6].

2.3.2 ETLook

ETLook model [14] is used in the WaPOR portal and the model equations are described in detail in Section 5 of "WaPOR Data Manual, Evapotranspiration v2.2" [8]. Similarly to TSEB-PT, the ETLook model considers evaporation from the soil and transpiration form vegetation as two separate fluxes. It also takes energy balance at the land surface into consideration. However, unlike TSEB-PT it does not estimate latent heat flux as a residual of the other energy fluxes. Instead it assumes that both E and T are transferring water into the atmosphere at rates based on Penman–Monteith equation and modulated using a number of stress factors. For E the only stress factor is based on top-soil moisture, while T is impacted by air temperature stress, vapour pressure stress, radiation stress and root-zone soil moisture stress.

As described above, ETLook requires soil moisture to estimate ET. However, soil moisture is not included in either WaPOR or Copernicus inputs. Instead it is estimated based on LST and vegetation fractional cover (FVC). Following the method of Yang et al. [15] a trapezoid is constructed in the LST-FVC space with corner values estimated based on theoretical considerations. The soil moisture of a given pixel is then estimated based on the relative location of LST and FVC within the theoretical trapezoid.

2.3.3 Gap-filling

Both TSEB-PT and ETLook produce daily ET estimates on days with thermal data acquisitions and for pixels which are not obscured by clouds during the satellite overpass. The WaPOR portal delivers decadal (10-day) composites of ET which contain the average daily value of evapotranspiration during the compositing period. If the average was derived only from estimates obtained during sunny conditions then this would lead to an overestimation. Therefore, it is important to gap-fill the timeseries and to take the conditions present during cloudy periods into account.

For outputs of both models the gaps in the daily ET maps due to cloudy conditions were filled using maps from adjacent dates (with up to 10-day temporal displacement) and an assumption that ratio of reference to actual ET remains steady over short periods. Since reference ET depends only on meteorological parameters it is possible to estimate it also during cloud periods. This approach also takes the changing meterological conditions during the cloudy periods (e.g. reduced solar irradiance and reduced air temperature) into account. At the same time, fine spatial details are retained since actual ET from adjacent dates, estimated at higher spatial resolution than reference ET, is also used during gap filling. This can be seen in Fig. 2.9. The figure also shows the sensitivity of the method

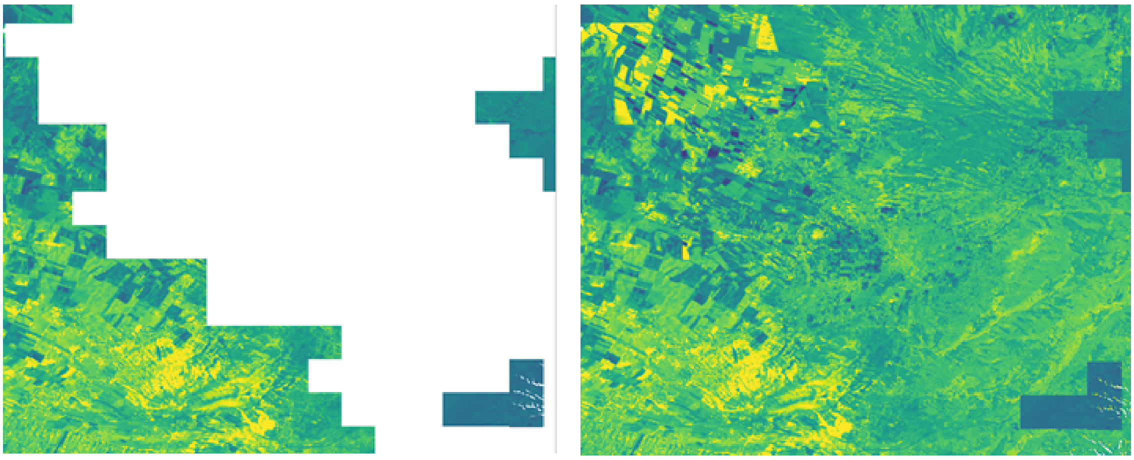


Figure 2.9: An example input and output of the gap-filling method.

to undetected clouds which affect the actual ET estimation.

2.4 Input and resource requirements

2.4.1 Input requirements

Despite relying on different approaches for estimating ET, both TSEB-PT and ETLook models require very similar set of inputs. This is true regardless of whether WaPOR or Copernicus inputs are used. This set of inputs is quite standard and would remain the same for most commonly used ET model. It consists of shortwave optical observations to characterize the state of land surface and vegetation, thermal infrared observations to derive LST, near-surface (e.g. air temperature, wind speed) and total column (e.g. transmissivity) meteorological data, land cover map to set parameters which cannot be determined directly from shortwave optical satellite observations (e.g. obstacle / vegetation height) and a digital elevation model (DEM) to perform terrain corrections of meteorological data.

The differences between TSEB-PT and ETLook models, and between WaPOR and Copernicus inputs, are in the type and number of higher-level parameters which are estimated based on the set of inputs and in the way in which those estimations are done. Both TSEB-PT and ETLook require basic vegetation parameterization in the form on Leaf Area Index, fractional vegetation cover and albedo (all three based on shortwave optical observations), as well as vegetation / obstacle height (based on land-cover map). In addition TSEB-PT needs parameters such as fraction of vegetation which is green (based on optical observations) or leaf angle distribution (based on land cover map). Both models require an estimation of LST and both require the same meteorological parameters, with the difference that TSEB-PT uses instantaneous values at the time of satellite overpass, while ETLook relies mainly on daily averages.

Regarding differences between WaPOR and Copernicus input preparation, the former use mostly simpler methods to derive the higher level products. For example, both fractional vegetation cover and leaf area index are based on relationship with NDVI, while when processing Copernicus data they are derived from neural-network inversion of a radiative transfer model. However, if it was determined that computational speed is more important than potential improvements in accuracy then the same simpler methods could be applied to Copernicus data.

2.4.2 Resource requirements

The production of input data and ET maps in ET4FAO project was conducted in a cloud environment, with each major step running in its own virtualized environment based on Docker images. The three most resource intensive steps are derivation of biophysical parameters from Sentinel-2 observations, sharpening of LST data and the actual ET estimation. The first step uses SNAP software and Python scripts and is allocated 3 virtual CPUs and 30 GB of memory and requires about 30 minutes to process a full Sentinel-2 scene. The LST sharpening step is based on Python scripts and is also allocated 3 virtual CPUs and 30 GB of memory and requires around 20 minutes to sharpen LST with coverage of one Sentinel-2 scene. The ET modeling is performed using both TSEB-PT and ETLook models. Both models are assigned 1 virtual CPU and 30 GB of RAM. Due to the more complex nature of the TSEB-PT model, and iterative computational steps, it requires around 20 minutes to derive ET for one Sentinel-2 tile compared to less than 10 minutes for ETLook.

Chapter 3

Validation against in situ data

3.1 Introduction

Focus of this assessment is put on numerical and statistical analysis, trying to avoid applying any expert knowledge criteria in order to avoid any possible bias by the authors. *In situ* observations, despite that they do not lack uncertainties, are therefore considered our "gold" standard.

Even tough Lebanon is one of the main study regions of ET4FAO, technical and economic issues prevented access to *in situ* validation data for the study period. For that reason, in addition to a site in Tunisia other sites in southeast Spain, with a similar climate to that of Tunisia and Lebanon, and containing various crops and irrigation systems, have been included as alternative.

Fig 3.1 shows the daily ET (mm/day) for all the sites included in this task. In summary one site corresponds to Tunisia (rainfed olive) whereas other 5 sites, either in irrigated or rainfed systems, are located in Albacete, Spain. Fig 3.2 shows size and surroundings for all the sites included in validation.

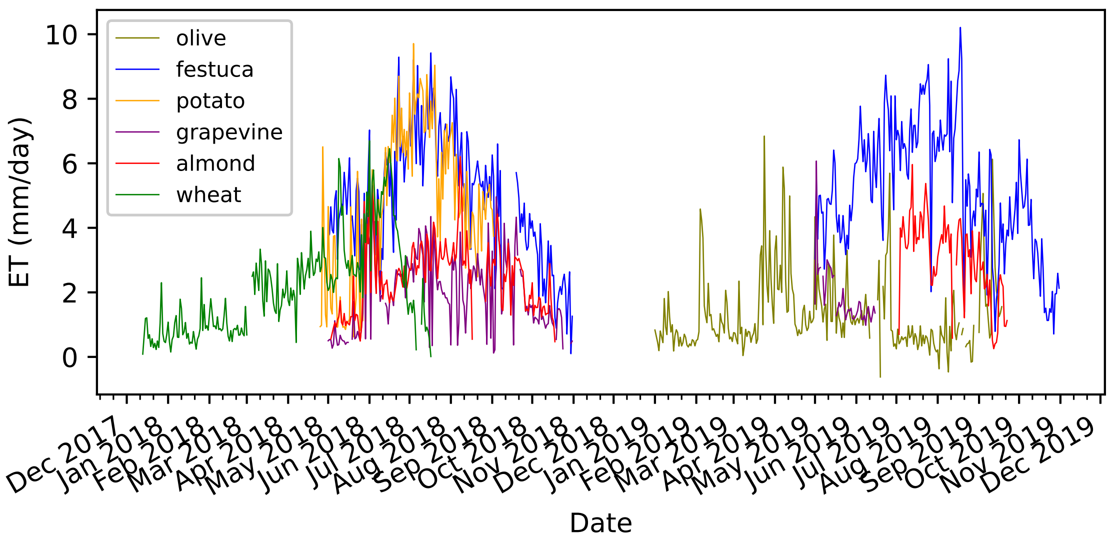
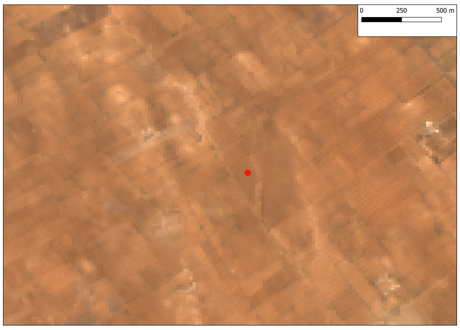
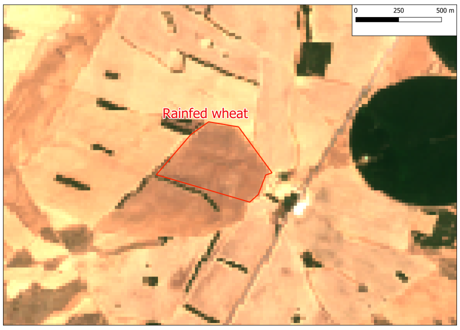


Figure 3.1: Daily observations of ET (mm/day) in all *in situ* sites used in this study



(a)

(b)



(c)

Figure 3.2: Overview maps of Tunisian olive orchard (3.2a) and Spanish potato, grapevine, festuca and almond (3.2b) and wheat (3.2c) field validation sites, with Sentinel-2 images from July used as background.

3.1.1 Tunisia

One site consisting of an eddy-covariance (EC) tower located in an rainfed olive grove is located in Tunisia (10.60153° E, 34.93111° N). The tower setup is as described in Chebbi et al. [16] although the location is slightly changed. Trees are planted 25 m apart, with a very low tree cover fraction of (5% to 10%). A 10 m tower was installed over a tree and a 3m tower over the bare soil, the net radiation and soil heat flux components being thus computed as the area average values. Data has been provided by CESBIO (France) and the Olive Institute (Tunisia), and consists in 30 minutes flux data (net radiation and ground, sensible and latent heat fluxes). Additional postprocessing of the EC data has consisted in detecting and removing outliers for the half-hourly EC flux time series [17], followed by a gap filling based on evaporative fraction.

3.1.2 Southeast Spain

These sites are located in the province of Albacete (Spain) and are part of the University of Castilla-La Mancha (UCLM) and the Agro-Technological Institute of Albacete Province (ITAP). Most of the sites are located in Las Tiesas Experimental Farm near Barrax, which has been used as long-term site for Calibration/Validation operations in several ESA-related activities.

Irrigated potato

This site consist of weighting lysimeter of 6.21 m^2 area in a field with rotating crops (2.10130° W, 39.06081° N), in which potato was planted 2018. ET at 15 minute timesteps, from May to October, was calculated from weight differences before and after the period and then aggregated at hourly timesteps. These records are visually and manually checked for consistence, in particular during rainfall and irrigation events, flagged out outliers and then gap-filled based on either reference ET or net radiation flux ratio. More information on this lysimeter can be found in [18].

Irrigated grass

This site is adjacent to the potato field and aims to represent a reference grass layer, mainly composed of perennial *Festuca* species, in which an anologous weighting lysimeter of 6.21 m^2 area is placed at coordinates 2.10009° W, 39.06046° N. In order to keep the surface the closest to the reference conditions described by FAO56 document, this site is frequently sprinkler irrigated and clipped to a height of ca. 0.12 m. Similarly, ET at 15 minute timesteps was calculated from weight differences before and after the period and then aggregated at hourly periods. These records are visually and manually checked for consistence, outliers flagged out and gap-filled based on either reference ET or net radiation flux ratio. Data is processed and available between May and October 2018 and 2019. More information on these lysimeters and the management of this site can be found in [18].

Irrigated vineyard

A third lysimeter located at coordinates 2.10104° W, 39.05972° N is installed under a 4 ha vineyard *cv. Tempranillo* with drip irrigation. This lysimeter occupies a surface of 9 m^2 (3x3 m) containing two grapevines. Data from May to October 2018 was available, and, despite for some technical issues encountered in summer 2019, data from May to 15 June 2019 was also produced. More information on this lysimeter and the vineyard characteristics can be found in [19].

Irrigated almond

Also located in Las Tiesas Experimental farm, an Eddy Covariance tower was deployed in a young drip-irrigated almond orchard (2.08965° W, 39.04228° N). Flux data has been fully processed and provided as daily ET estimates, from May to October 2018 and July-September 2019, both with uncorrected energy closure and corrected closure assigning all residual to latent heat flux.

Rainfed wheat, Spain

This site is not part of Las Tiesas farm, but is located several kilometers southeast of Albacete, in Orán (1.85970° W, 38.82337° N). It consists of an eddy-covariance (EC) tower over a rainfed cereal field. Flux data has been fully processed and provided as daily ET estimates, from January to July 2018, when winter wheat was grown and harvested. Daily ET is provided either with uncorrected energy closure or corrected closure assigning all residual to latent heat flux or assuming the preservation of the Bowen Ratio. However,

Figure 3.3: The EC tower in the almond drip irrigated site in Las Tiesas Experimental Farm, Albacete

the data provider and curator recommended using the data without any residual correction due to some possible unaccounted artifacts during the residual correction.

3.2 Results

These results show the validation of both TSEB-PT and ETLook models using as inputs Copernicus data as well as WaPOR-like data. Both models use the same remote sensing inputs, namely LST, LAI or NDVI, and surface albedo. Meteorological forcing for Copernicus uses ECWMF ERA-5 reanalysis data while for WaPOR-like products the standard products used in its workflow are implemented. On the other hand, for all products ancillary canopy parameters are derived from Copernicus Global Land Cover map.

For Barrax, level 3 is used due to the small size of the experimental plots and a buffer of 3x3 pixels around the lysimeter was extracted as representative of the lysimeter readings, considering possible geolocation uncertainties. In this case, since no WaPOR product is available, Landsat Collection 2 was downloaded and processed according to the WaPOR methodology. On the other hand, a larger buffer of 5x5 pixels, accounting for the typical EC flux footprint is extracted for the almond and wheat sites. Finally, since in the Tunisian olive EC site WaPOR data is available at level 2 (100m) all models and data sources are validated at 100m, using a buffer of 3x3 pixels (i.e. 300m size) around the flux tower.

Table 3.1 shows the descriptive statistics of dekadal ET for the observed all products evaluated, while error and agreement metrics between the observed and the modelled are shown in Table 3.2. It is worth noting that these statistics and results are considering the energy closure correction of EC data by assigning the residual part to latent heat flux [20], with exception of the wheat field which uses residual-uncorrected ET as suggested by the

Figure 3.4: The EC tower in the rainfed cereal site near Orán, Albacete

data provider. However, figures 3.6, 3.10 and 3.11 show the timeseries with the agreement for all possible EC residual corrections.

Overall results for all sites in Table 3.2 are also depicted in Fig 3.5 as a scatterplot between the observed and the models' predictions. All products are able to track the spatio-temporal flux variability in a similar way, with correlation coefficients between the observed and predicted between 0.85 and 0.91. Nevertheless, Copernicus-based products, despite of using sharpened LST from Sentinel-3, yield slightly better agreement between the observed and the predicted, with higher correlation (0.90 and 0.91) compared to WaPOR-like products (0.85 and 0.87). The lower temporal resolution of Landsat data in WaPOR-like products could probably affect the temporal interpolation and gap-filling processes that are required before producing the dekdadal ET composites.

On the other hand, ETLook model tends to show significant larger error metrics (using either Copernicus or WaPOR-like input data), with a systematic underestimation showed by a mean bias on dekadal ET higher than 1 mm/day, whereas TSEB-PT showed smaller bias and closer to 0 (0.3 mm/day and 0.6 mm/day when using respectively Copernicus and WaPOR-like data). This also results in lower RMSE in TSEB-PT as compared to ETLook. Willmott's Index of Agreement (d) [21] tries to summarize in a single value the agreement and error metrics, and as such, due to the larger RMSE and bias of ETLook, TSEB-PT shows consistently better values of d, both for all individual sites and when all sites are pooled together.

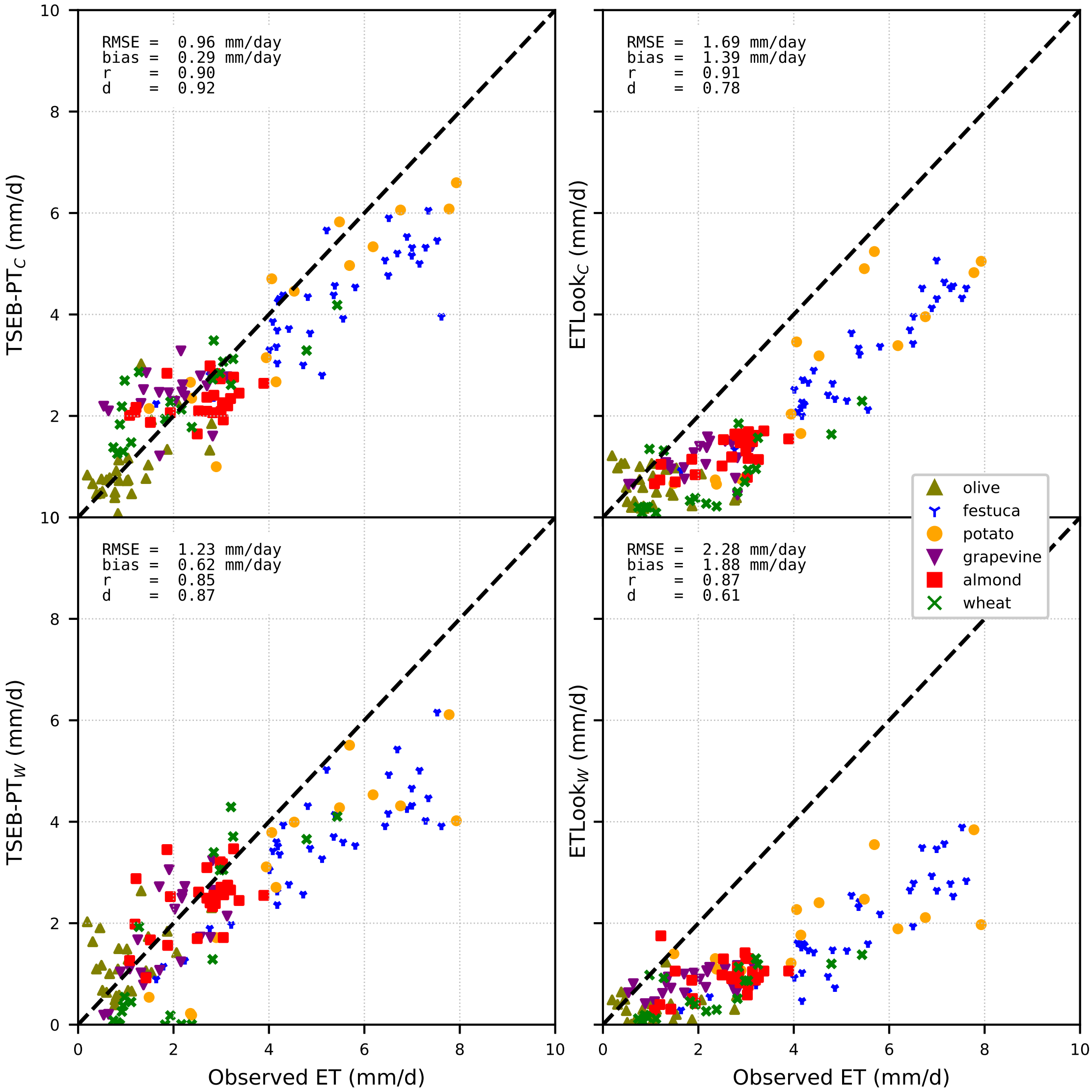


Figure 3.5: Scatterplot between the *in situ* dekadal ET and estimated dekadal ET for all sites. For ETLook$_W$ only the olive site uses actual WaPOR product, downloaded at level 2, the other sites uses ETLook run with WaPOR-like inputs

Table 3.1: Descriptive statistics for estimated dekadal ET (mm/day) and in situ dekadal ET (mm/day). $\overline{Obs.}$ and $\overline{Pre.}$ represent the average ET of in situ and model datasets, $\sigma_{Obs.}$ and $\sigma_{Pre.}$ are the standard deviations of ET values for respectively the insitu and model ET datasets. Only the olive site uses actual WaPOR product, downloaded at level 2, the other sites uses either ETLook or TSEB-PT model run with WaPOR-like inputs

Site	Source	Model	N	$\overline{Obs.}$	$\overline{Pre.}$	$\sigma_{Obs.}$	$\sigma_{Pre.}$
All	Copernicus	TSEB-PT	137	3.04	2.76	2.00	1.48
		ETLook	137	3.04	1.66	2.00	1.32
	WaPOR-Like	TSEB-PT	145	2.97	2.35	1.97	1.44
	WaPOR-Like*	ETLook	145	2.97	1.09	1.97	0.87
Olive	Copernicus	TSEB-PT	26	1.08	0.99	0.67	0.64
		ETLook	26	1.08	0.61	0.67	0.34
	WaPOR-Like	TSEB-PT	26	1.08	1.29	0.67	0.63
	WaPOR	ETLook	26	1.08	0.29	0.67	0.28
Festuca	Copernicus	TSEB-PT	33	5.33	4.27	1.50	1.02
		ETLook	33	5.33	3.11	1.50	1.08
	WaPOR-Like	TSEB-PT	36	5.08	3.53	1.66	1.19
		ETLook	36	5.08	1.85	1.66	0.98
Potato	Copernicus	TSEB-PT	14	4.69	4.14	1.96	1.72
		ETLook	14	4.69	2.89	1.96	1.71
	WaPOR-Like	TSEB-PT	14	4.69	3.22	1.96	1.83
		ETLook	14	4.69	2.02	1.96	0.82
Grapevine	Copernicus	TSEB-PT	20	2.05	2.47	0.71	0.45
		ETLook	20	2.05	1.13	0.71	0.34
	WaPOR-like	TSEB-PT	23	1.92	1.80	0.74	0.88
		ETLook	23	1.92	0.84	0.74	0.22
Almond	Copernicus	TSEB-PT	24	2.60	2.30	0.73	0.34
		ETLook	24	2.60	1.27	0.73	0.34
	WaPOR-like	TSEB-PT	26	2.53	2.43	0.75	0.63
		ETLook	26	2.53	0.92	0.75	0.35
Wheat	Copernicus	TSEB-PT	20	2.22	2.42	1.31	0.78
		ETLook	20	2.22	0.76	1.31	0.66
	WaPOR-like	TSEB-PT	20	2.22	1.52	1.31	1.62
		ETLook	20	2.22	0.63	1.31	0.45

When looking at the sites individually, ETLook consistently shows larger systematic underestimation and standard errors than TSEB-PT. However, ETLook run with Copernicus also demonstrates a better capability to track temporal changes of ET (with larger correlation coefficient than TSEB-PT) for the almond and festuca sites. It is known from SEN-ET project and previous studies that TSEB-PT is quite sensitive to correct determination of canopy structure/roughness. In particular, Copernicus Global Land Cover product flags both almond and grapevine sites as croplands, which, based on the actual look-up table derived from SEN-ET and WAPOR, assigns a maximum canopy height of 2m when LAI is at 5. Therefore, canopy height is strongly underestimated in these two sites yielding larger uncertainties in TSEB-PT. ETLook on the other hand, even though

Table 3.2: Error metrics between estimated dekadal ET (mm/day) and in situ dekadal ET (mm/day). $\overline{bias}$ (mm/day) is the average bias, computed as the mean difference between the observed and the predicted, 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 Person Correlation coefficient between observations and predictions, and d is the Wilmott's Index of Agreement. Only the olive site uses actual WaPOR product, downloaded at level 2, the other sites uses either ETLook or TSEB-PT model run with WaPOR-like inputs

Site	Source	Model	$\overline{bias}$	MAE	RMSE	$fMSE_u$	$fMSE_s$	a	scale	r	d
All	Copernicus	TSEB-PT	0.29	0.75	0.96	0.56	0.44	0.67	1.35	0.90	0.92
		ETLook	1.39	1.44	1.69	0.89	0.11	0.60	1.51	0.91	0.78
	WaPOR-Like	TSEB-PT	0.62	0.94	1.23	0.63	0.37	0.62	1.37	0.85	0.87
	WaPOR-Like*	ETLook	1.88	1.90	2.28	0.96	0.04	0.39	2.25	0.87	0.61
Olive	Copernicus	TSEB-PT	0.09	0.43	0.59	0.25	0.75	0.58	1.04	0.60	0.75
		ETLook	0.46	0.72	0.95	0.88	0.12	-0.13	1.98	-0.26	0.30
	WaPOR-Like	TSEB-PT	-0.21	0.54	0.70	0.37	0.63	0.44	1.07	0.47	0.69
	WaPOR	ETLook	0.79	0.84	1.04	0.93	0.07	0.07	2.43	0.16	0.43
Festuca	Copernicus	TSEB-PT	1.05	1.14	1.38	0.81	0.19	0.55	1.47	0.81	0.75
		ETLook	2.22	2.22	2.30	0.97	0.03	0.67	1.40	0.94	0.62
	WaPOR-Like	TSEB-PT	1.55	1.55	1.78	0.89	0.11	0.62	1.40	0.87	0.73
		ETLook	3.23	3.23	3.35	0.98	0.02	0.52	1.69	0.89	0.51
Potato	Copernicus	TSEB-PT	0.54	0.82	0.99	0.46	0.54	0.79	1.14	0.91	0.92
		ETLook	1.79	1.79	2.00	0.85	0.15	0.78	1.14	0.89	0.76
	WaPOR-Like	TSEB-PT	1.47	1.47	1.75	0.74	0.26	0.82	1.07	0.88	0.82
		ETLook	2.66	2.66	3.07	0.96	0.04	0.29	2.38	0.68	0.54
Grapevine	Copernicus	TSEB-PT	-0.42	0.64	0.82	0.73	0.27	0.20	1.58	0.32	0.53
		ETLook	0.92	0.94	1.12	0.93	0.07	0.21	2.13	0.45	0.53
	WaPOR-Like	TSEB-PT	0.12	0.52	0.61	0.06	0.94	0.88	0.85	0.74	0.85
		ETLook	1.08	1.10	1.26	0.98	0.02	0.16	3.36	0.55	0.50
Almond	Copernicus	TSEB-PT	0.31	0.68	0.75	0.82	0.18	0.17	2.15	0.36	0.53
		ETLook	1.34	1.34	1.44	0.97	0.03	0.33	2.17	0.71	0.50
	WaPOR-Like	TSEB-PT	0.10	0.57	0.72	0.40	0.60	0.41	1.18	0.48	0.69
		ETLook	1.61	1.65	1.76	0.97	0.03	0.17	2.16	0.36	0.43
Wheat	Copernicus	TSEB-PT	-0.21	0.64	0.84	0.70	0.30	0.49	1.67	0.81	0.84
		ETLook	1.45	1.49	1.73	0.93	0.07	0.37	1.97	0.72	0.57
	WaPOR-Like	TSEB-PT	0.70	0.98	1.17	0.35	0.65	1.01	0.81	0.81	0.84
		ETLook	1.59	1.59	1.88	0.98	0.02	0.26	2.90	0.77	0.54

uses the same canopy height/roughness input, seems to overcome this limitation, as it computes ET from the relative position of actual LST within some theoretical wet and cold boundaries.

It is also noteworthy that for the grapevine and almond the lowest error metrics (mean bias, MAE and RMSE) are found for TSEB-PT run with WaPOR-like data. These two sites are drip-irrigated, with a large fraction of bare soil exposed and subject to some crop water stress (e.g. lower ET rates shown in Fig. 3.1), especially relevant in the case of the grapevine as applying a controlled stress is strategic in order to enhance the berry quality. In such cases, actual Landsat surface temperature data plays a significant role in modulating stress and improving ET estimates of TSEB-PT than using Copernicus sharpened LST.

The ability to track ET temporal changes as well as the error deviations from the *in situ* measurement are shown in Figures 3.6 to 3.11. In these figures, and for the sites with EC towers, an uncertainty band around the observed dekadal ET is displayed, due to energy imbalance closure. These figures also depict that ETLook yields errors frequently larger than 1 mm/day, caused by the systematic underestimation that the model produces. TSEB-PT on the contrary, produces fewer cases with errors above 1 mm/day, even in the sites where it showed a lower correlation as compared with ETLook.

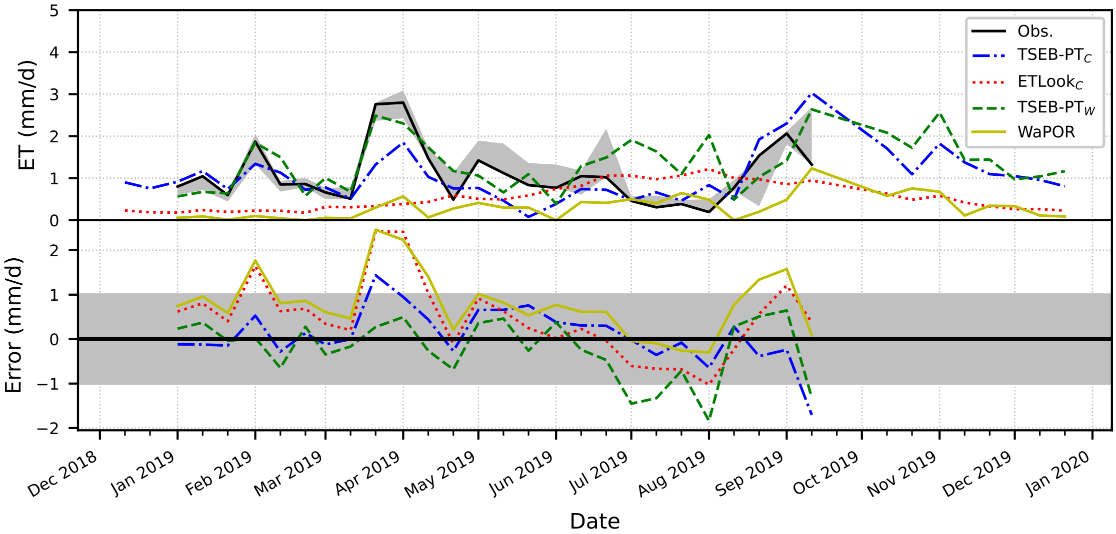


Figure 3.6: Timeseries at the olive EC site in Tunisia of a) *in situ* dekadal ET, with greyed area the uncertainty band due to residual correction, and estimated dekadal ET ; b) errors between observed and estimated dekadal ET, with greyed area corresponding to errors within ±1 mm/day.

Finally, Figure 3.12 shows the observed vs. predicted scatterplot only for the olive site, since it is located Tunisia and hence is the only one that is validated against the product downloaded from the WaPOR database.

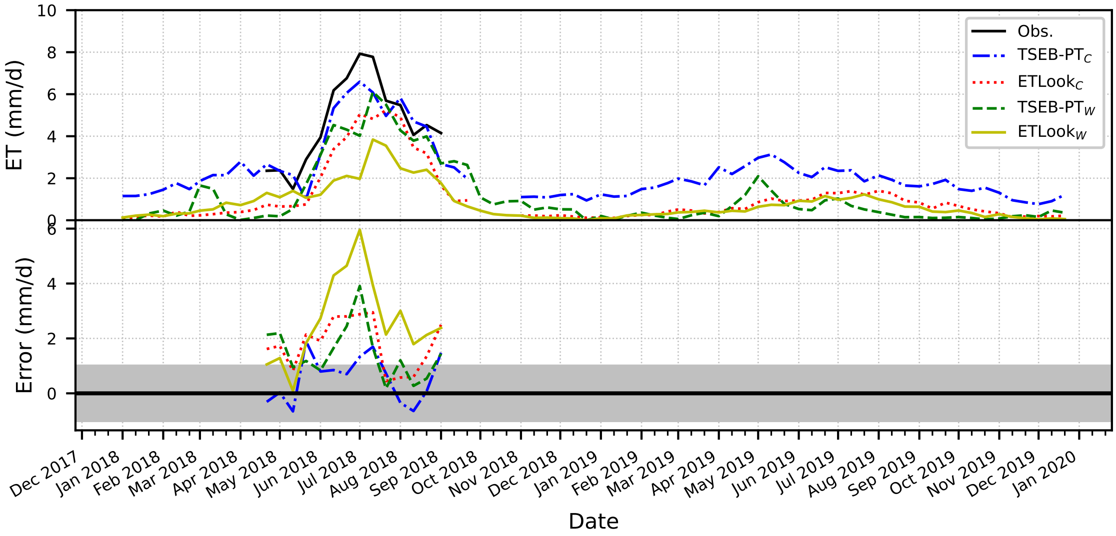


Figure 3.7: Timeseries at the potato lysimeter site at Spain of a) *in situ* dekadal ET and estimated dekadal ET ; b) errors between observed and estimated dekadal ET, with greyed area corresponding to errors within ±1 mm/day.

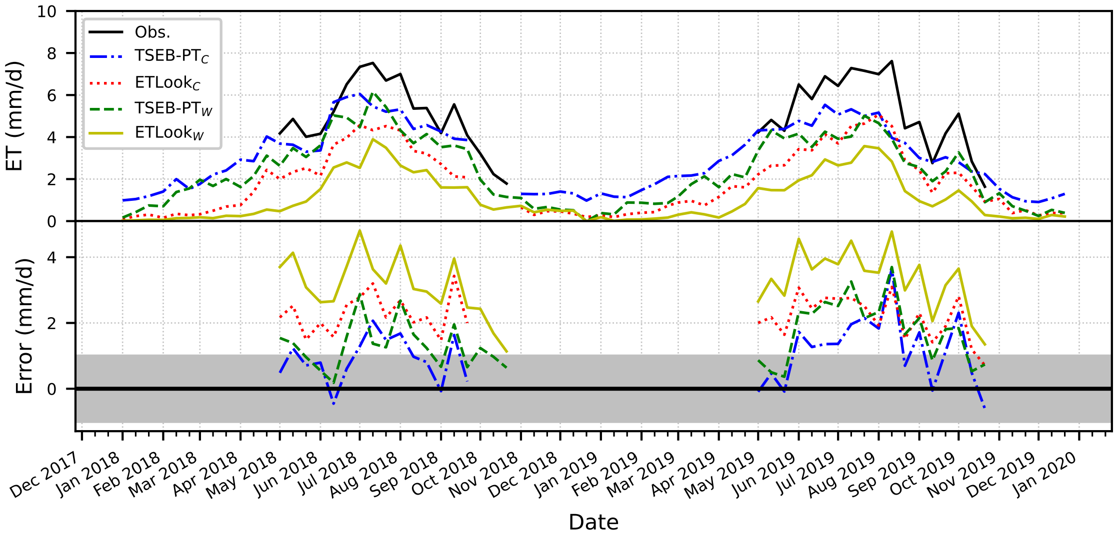


Figure 3.8: Timeseries at the grass lysimeter site in Spain of a) *in situ* dekadal ET and estimated dekadal ET ; b) errors between observed and estimated dekadal ET, with grayed area corresponding to errors within ±1 mm/day.

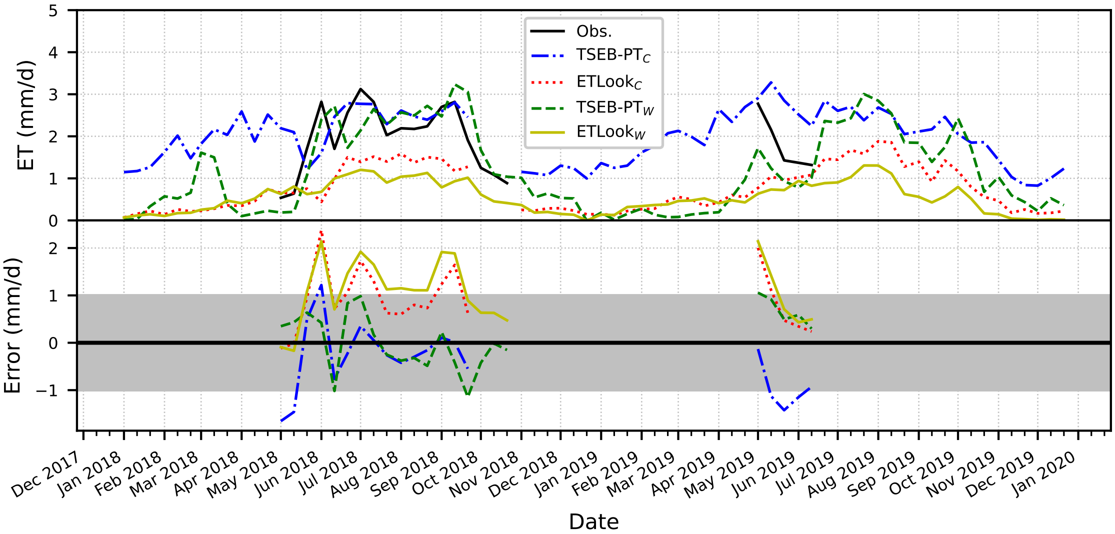


Figure 3.9: Timeseries at the grapevine lysimeter site in Spain of a) *in situ* dekadal ET and estimated dekadal ET ; b) errors between observed and estimated dekadal ET, with greyed area corresponding to errors within ±1 mm/day.

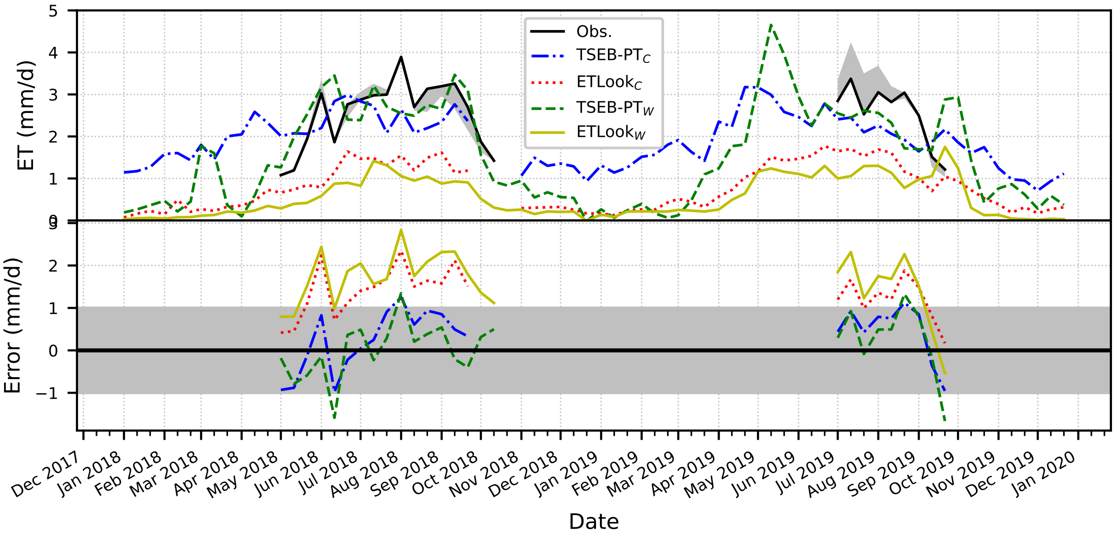


Figure 3.10: Timeseries at the almond EC site in Spain of a) *in situ* dekadal ET, with greyed area the uncertainty band due to residual correction, and estimated dekadal ET ; b) errors between observed and estimated dekadal ET, with greyed area corresponding to errors within ±1 mm/day.

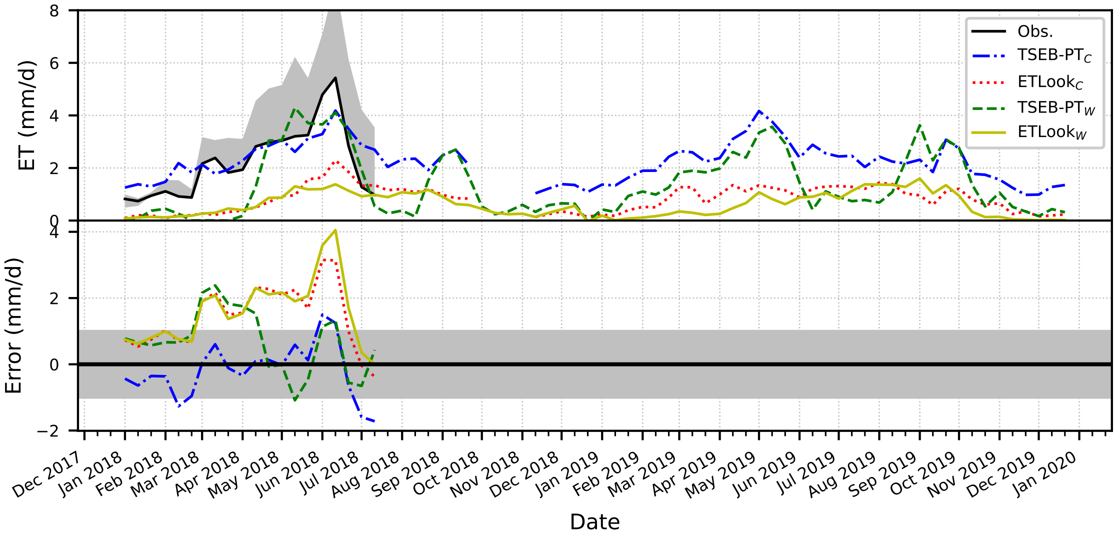


Figure 3.11: Timeseries at the wheat EC site in Spain of a) *in situ* dekadal ET, with greyed area the uncertainty band due to residual correction, and estimated dekadal ET ; b) errors between observed and estimated dekadal ET, with greyed area corresponding to errors within ±1 mm/day.

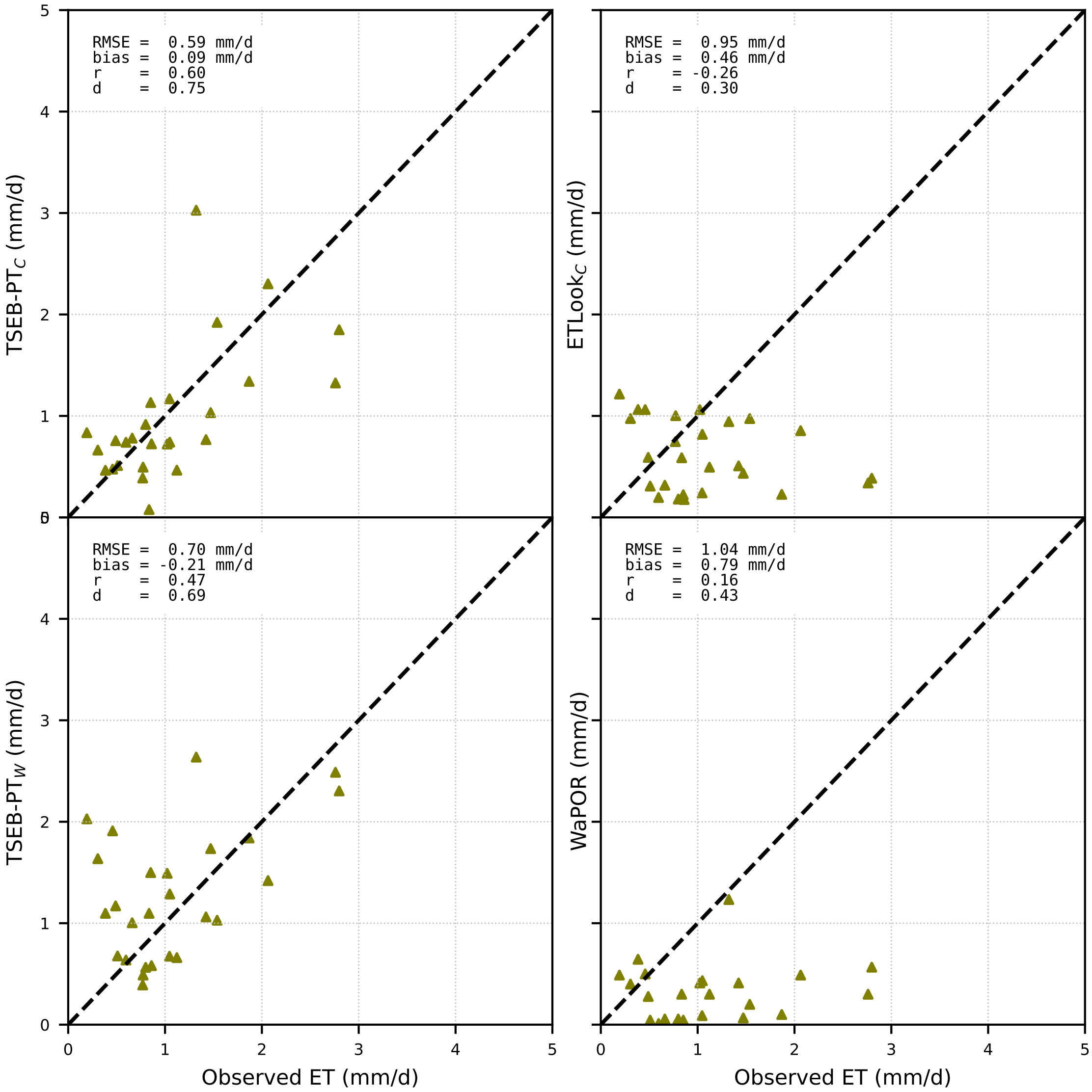


Figure 3.12: Scatterplot between the *in situ* dekadal ET and estimated dekadal ET for all sites in the olive site in Tunisia, where WaPOR ET product was downloaded from the FAO portal.

Chapter 4

Spatio-temporal assessment

4.1 Introduction

In this chapter, a spatio-temporal intercomparison will be performed between models, data sources and scales. For that purpose we will focus on the two main study regions of ET4FAO, Lebanon and Tunisia. For Lebanon, the three processing levels are available, as WaPOR portal provides Level 3 (20/30 m) data for the Bekaa Valley, whereas for Tunisia the intercomparison will be done using Levels 1 (250/300 m) and 2 (100 m), since no WaPOR data in Level 3 is available for that region.

The first part of this analysis will compare monthly and annual spatio-temporal variability between products at the same level, using the best level available for all datasets for a given region. Then, in a second step, the robustness between levels for a given product is evaluated, considering that energy and mass must be preserved between the different scales.

Data produced during the ET4FAO project and used in this assessment can be visualized online at `https://et4fao.dhigroup.com`.

4.2 Cumulative maps

4.2.1 Tunisia

In Tunisia WAPOR L3 data is not available. Therefore, this comparison is limited between the products generated at Level 2. Figure 4.1 show the cumulative monthly maps for TSEB-PT run with both Copernicus and WaPOR-like data, ETLook run with Copernicus data, and WaPOR product directly downloaded from the FAO portal.

Monthly maps show a higher spatial variability with TSEB-PT, where northern areas show overall larger ET rates in all months. It also shows some regional features of higher rates in April and September, as compared to ETLook. These differences are also summarized in annual cumulative ET maps of Fig. 4.2. It this case it is more clear that the differences between models show a S-N gradient. Regarding models produced with Copernicus data, ET rates show larger values in TSEB-PT than ETLook in the North, with also a higher temporal variability $\sigma_{TSEB-PT}/\sigma_{ETLook} > 1$, while in the South, ETLook yields higher ET annual rates and higher temporal variability $\sigma_{TSEB-PT}/\sigma_{ETLook} < 1$. While comparing WaPOR product with Copernicus-based models, it tends to provide overall

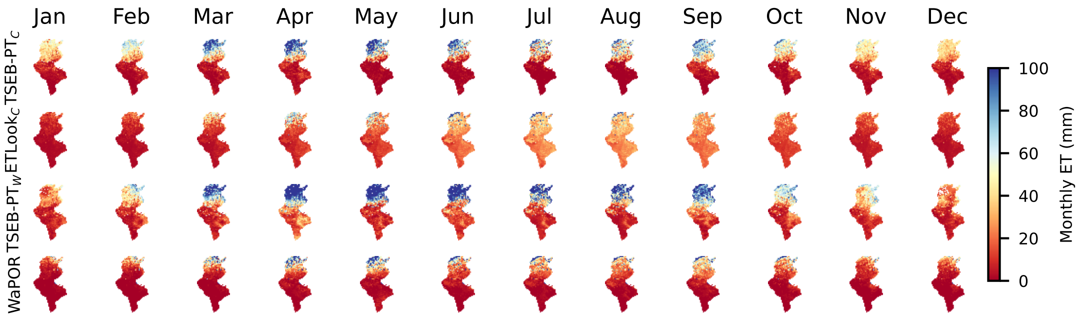


Figure 4.1: Monthly cumulative ET maps for ET products at Level 2 in Tunisia. All maps share the same scale and colorbar

lower ET rates, most notably in the northern areas and against TSEB-PT, and lower temporal variability, mostly in the southern areas.

4.2.2 Lebanon

The model intercomparison in Lebanon is done using Level 3 (20/30m) over the Bekaa Valley, as it is the area of interest with availability of WaPOR Level 3. Figure 4.3 shows the cumulative monthly maps for the four products evaluated. In addition, similar maps are shown for the whole country of Lebanon using products at Level 2 in Figure 4.4

In this case the monthly maps show similar spatio-temporal patterns in all products, with irrigated areas showing larger ET rates during the same area than the rainfed and natural vegetation surfaces. Nevertheless, Copernicus TSEB-PT tends to yield higher ET rates in winter compared to the other three products.

The annual cumulative ET maps and model intercomparison for this region are depicted in Figs. 4.6 and 4.5, for Level 3 in Bekaa Valley and Level 2 for the whole country respectively. Similarly as it has been previously observed, TSEB-PT tends to produce higher ET rates than ETLook. In addition, ETLook shows a larger temporal variability ($\sigma_{TSEB-PT}/\sigma_{WaPOR} < 1$) within irrigation perimeters while in rainfed areas it occurs the opposite.

4.3 Comparison with other ET products

To help interpret the observed spatial patterns we compared them against other available global ET products, both at monthly (Figures 4.7 and 4.8) and annual scales (Figures 4.9 and 4.10).

TerraClimate [22] dataset has a spatial resolution of ca. 4 km and temporal resolution of 1 month. It is based on a water balance model forced with monthly reanalysis meteorological data and using as ancillary data climatic normals of land cover and phenology. The actual ET product is estimated as reference ET constrained by available liquid and root-zone water with the root depth being time invariant and based on a map with resolution of 0.5°. Therefore, it does not take into account parameters such as irrigation water input, variable root depth, vegetation height or functional type, etc.

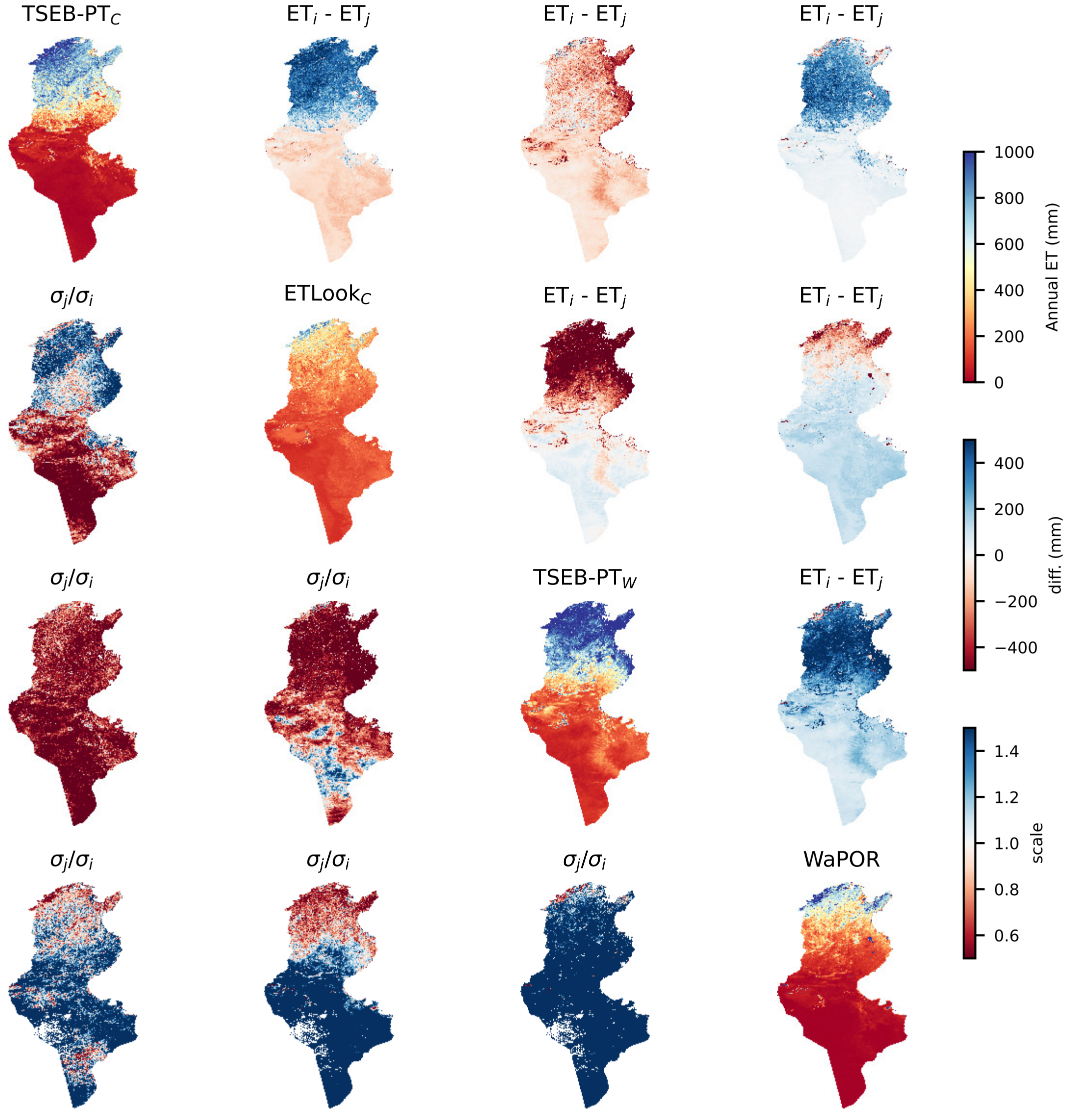


Figure 4.2: Annual comparison of ET for ET products at Level 2 in Tunisia. Maps in the diagonal show each model cumulative annual ET. The off-diagonal represent model intercomparisons, both in ET differences (upper-right) and in standard deviations scale (lower-left). "i" represents the model in row and "j" is the model in the column

The Global Land Evaporation Amsterdam Model (GLEAM) v3.5b [23] constrains the Priestley-Taylor potential evaporation equation with satellite microwave estimates for both topsoil moisture (ESA-CCI) and vegetation optical depth. GLEAM v3.5b is also driven by ancillary satellite data such as radiation and air temperature and produces ET globally at 0.25° spatial resolution.

EUMETSAT's Land Surface Analysis (LSA-SAF) daily ET product [24] combines ECMWF meteorological forcing with other LSA-SAF products, derived from Meteosat Second Generation MSG-SEVIRI, into a simplified land surface model to provide 30 minute

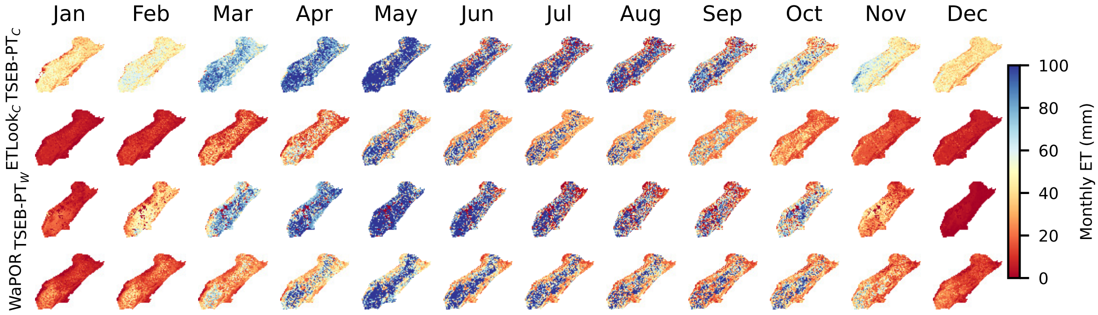


Figure 4.3: Monthly cumulative ET maps for ET products at Level 3 in the Bekaa Valley. All maps share the same scale and colorbar

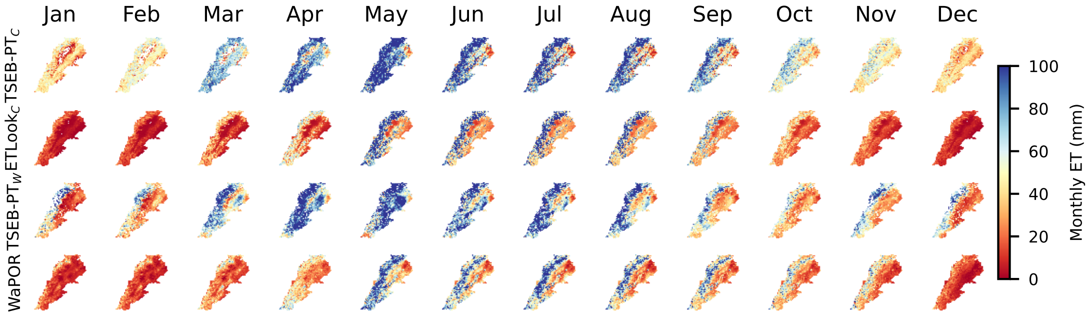


Figure 4.4: Monthly cumulative ET maps for ET products at Level 2 in Lebanon. All maps share the same scale and colorbar

and daily ET estimates at 3km sub-satellite point. LSA-SAF input products include both shortwave and longwave irradiance, albedo, LAI, and recently soil moisture derived from LST.

All these products offer an independent comparison dataset and indeed the three products significantly differ one each other, as TerraClimate is purely driven by meteorological forcing into a water balance model, GLEAMv3.5b is mainly driven by microwave satellite data, and LSA-SAF is driven by geostationary satellite shortwave and thermal products. In addition, we included the Copernicus Global Land Cover maps to help with the inter-comparison between products (Figure 4.11).

Considering the Tunisian annual ET (Figures 4.2 and 4.10), WaPOR produces the closest spatial patterns and absolute values compared to TerraClimate and GLEAMv3.5b, with ET of 600 - 800 mm/year on the north coast, around 500 mm/year in northern part of the country and 0 - 300 mm/year in the rest of Tunisia. Both TSEB-PT datasets produce generally higher ET in the north and lower ET in the barren and desert southern areas compared to the global products. WaPOR dataset has clearly higher ET values in areas classified as forests in CGLC, while TSEB-PT$_C$ has equally high ET in forested and agricultural areas of northern Tunisia. This could be due to influence of landcover map on ET models, in particular vegetation height, but also due to geography with forests being mostly located in mountainous areas and agriculture placed in areas classified as temper-

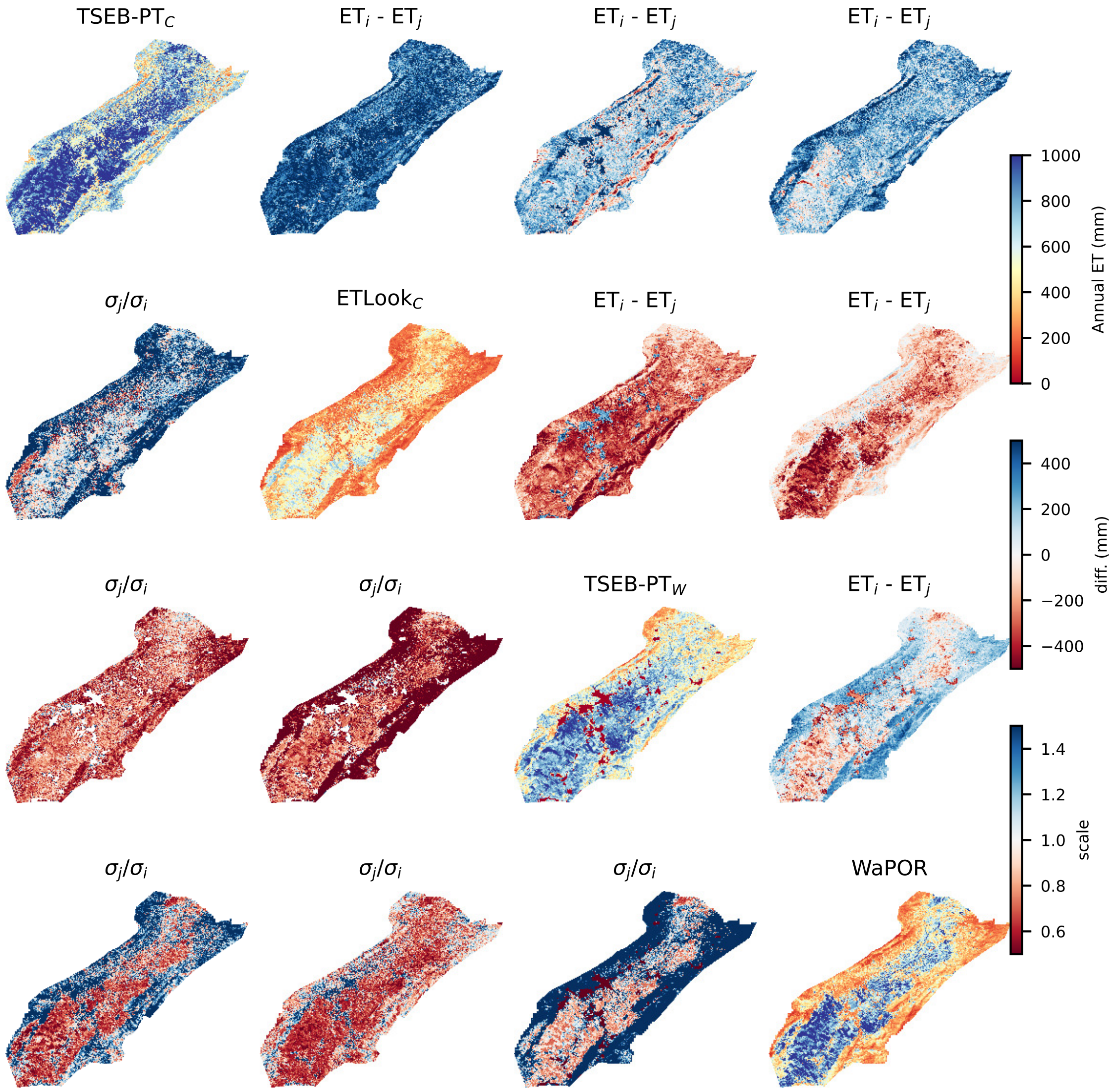


Figure 4.5: Annual comparison of ET for ET products at Level 3 in Bekaa Valley. Maps in the diagonal show each model cumulative annual ET. The off-diagonal represent model intercomparisons, both in ET differences (upper-right) and in standard deviations scale (lower-left). “i” represents the model in row and “j” is the model in the column

ate or arid-steppe by Köppen-Geiger classification and rest of Tunisia being classified as arid-desert. In Lebanon, and particularly in Bekaa valley, TerraClimate and GLEAMv3.5b annual ET rates are much lower in irrigated areas than the thermal-based datasets due to not accounting for irrigation (for Terraclimate) nor root-zone soil moisture (in GLEAM v3.5b), which is widely used in this area. Outside of agricultural areas, TSEB-PT$_C$ produces higher ET and both ETLook produce lower ET than these global models. All the models capture the agricultural areas, although with large differences in the absolute ET values.

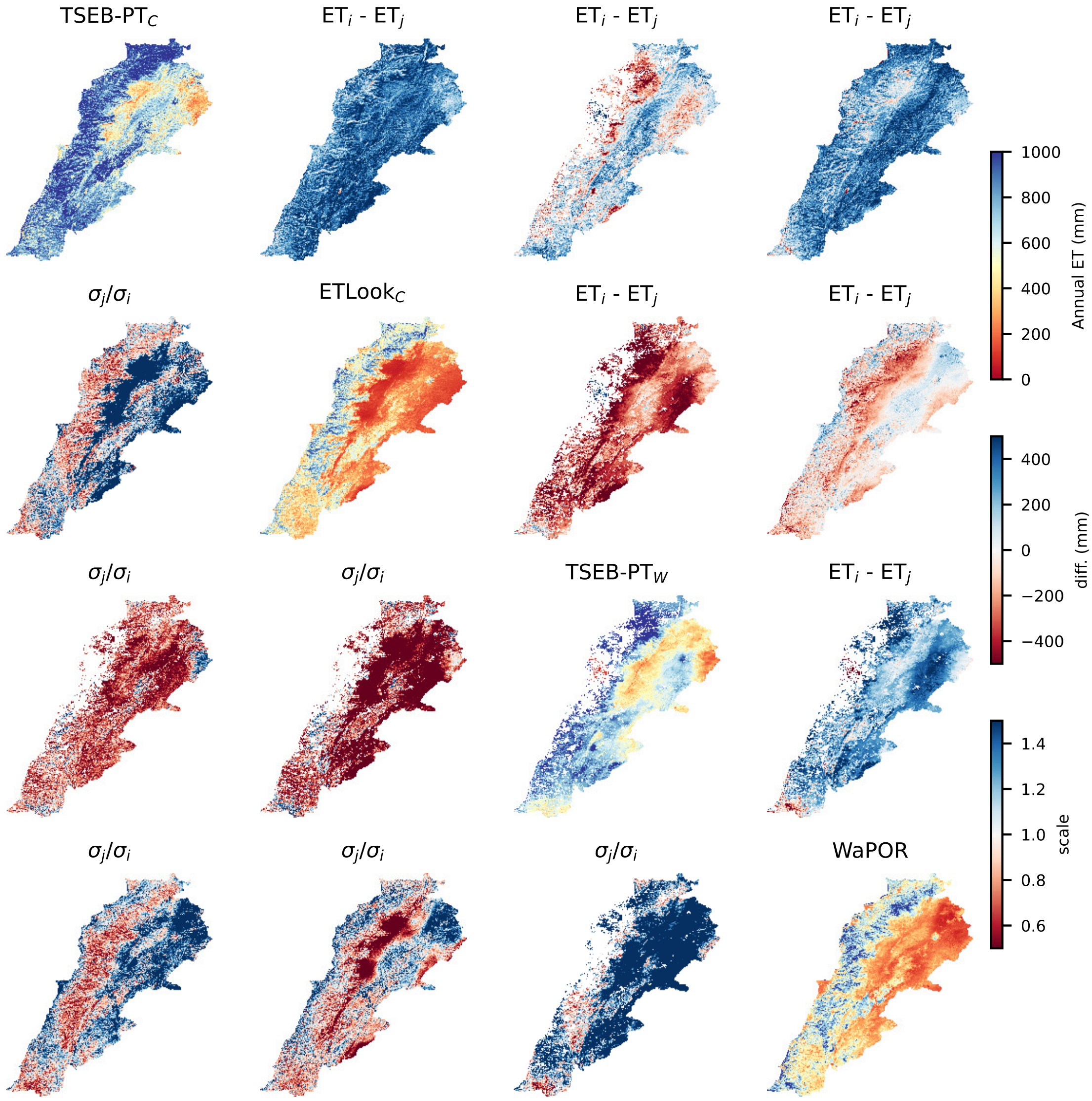


Figure 4.6: Annual comparison of ET for ET products at Level 2 in Lebanon. Maps in the diagonal show each model cumulative annual ET. The off-diagonal represent model intercomparisons, both in ET differences (upper-right) and in standard deviations scale (lower-left). "i" represents the model in row and "j" is the model in the column

In addition, the effect of irrigation and land cover/land use can be observed when comparing the monthly trends between TerraClimate and GLEAMv3.5b and the thermal-based models. Both Terraclimate and GLEAMv3.5b yields ET rates closer to zero during the late spring and summer months (i.e. May to September) in the Bekaa Valley as compared to the thermal-based models (Figures 4.3 and 4.7). Considering this area as intensively irrigated, thanks to the snowmelt and the existing reservoir infrastructure, one could assume that the thermal-based models provide a more realiable estimate of water use than soil-water-balance- and microwave-based models during this season.

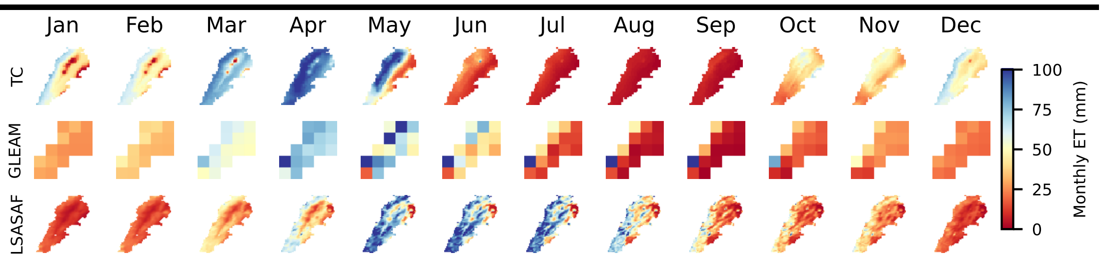


Figure 4.7: Monthly maps of TerraClimate Land Surface Model, GLEAMv3.5b and MSG LSASAF in Lebanon.

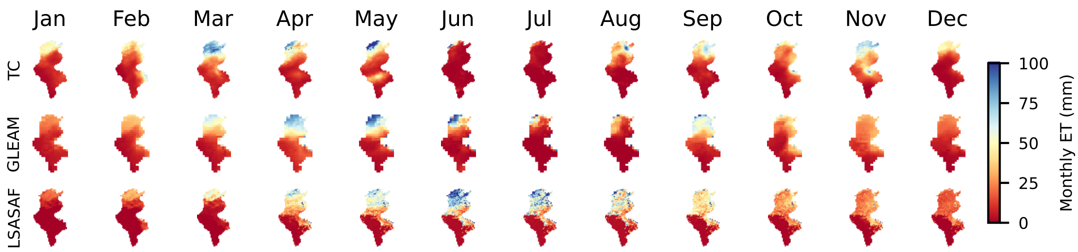


Figure 4.8: Monthly maps of TerraClimate Land Surface Model, GLEAMv3.5b and MSG LSA-SAF in Lebanon.

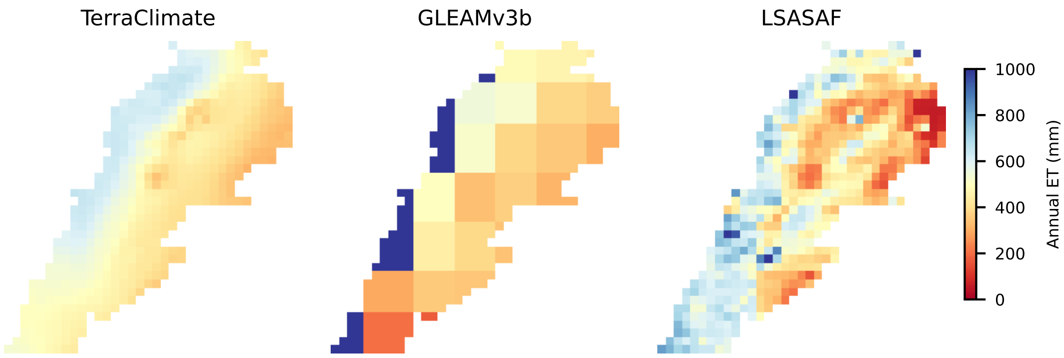


Figure 4.9: Annual maps of TerraClimate Land Surface Model, GLEAMv3.5b and MSG LSASAF in Lebanon.

Similar spatiotemporal disagreement is found in the monthly trends for Tunisia (Figures 4.1 and 4.8), which is evident in the northern part where most of the croplands are located (Figure 4.11c) during June and July. Based on the observations made before, we could hypothesize that these differences are also due to irrigation practices in the country during the dry season. Indeed, according to FAO, 455 070 ha were accounted as irrigated lands in Tunisia (`http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/irrigation-by-country/country/TUN`, last visited 11.06.2021), which approximately correspond (making use of the CGLC map of Figure 4.11c) to ca. 25% of total agricultural land. The location of these irrigated areas are estimated in FAO's AQUAMAPS portal (`https://data.apps.fao.org/aquamaps/`), and depicted for

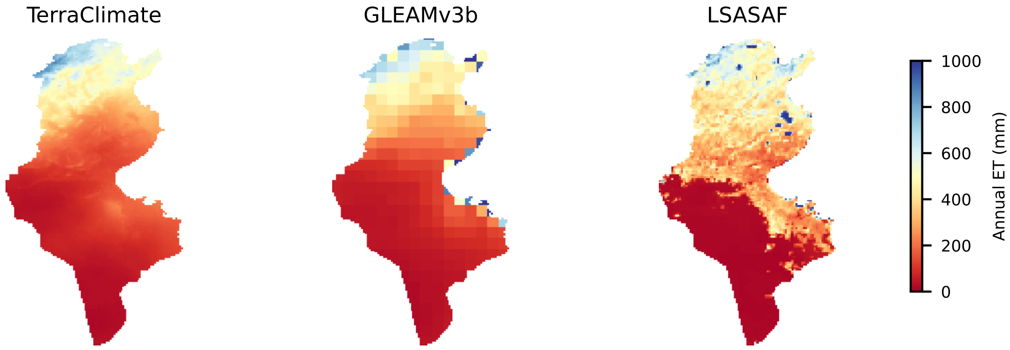


Figure 4.10: Annual maps of TerraClimate Land Surface Model, GLEAMv3.5b and MSG LSASAF in Tunisia.

Lebanon and Tunisia in Figure 4.12 , and these areas primarily coincide with the largest summer and annual ET rates in the thermal-based products (Figures 4.1 to 4.6).

Finally it is also worth noting that the largest monthly ET rates shown by the soil-water-based models (TerraClimate and GLEAMv3.5b) during the spring (Figures 4.7 and 4.8), which is the typical season in these regions when the largest water availability and vegetative growth occur, agree better with TSEB-PT models (either using Copernicus or WaPOR-like inputs) than with ETLook models, as the latter one tends to yield lower ET rates than the rest of the products (Figures 4.1 - 4.3), as it was already observed when validating this model against *in situ* measurements.

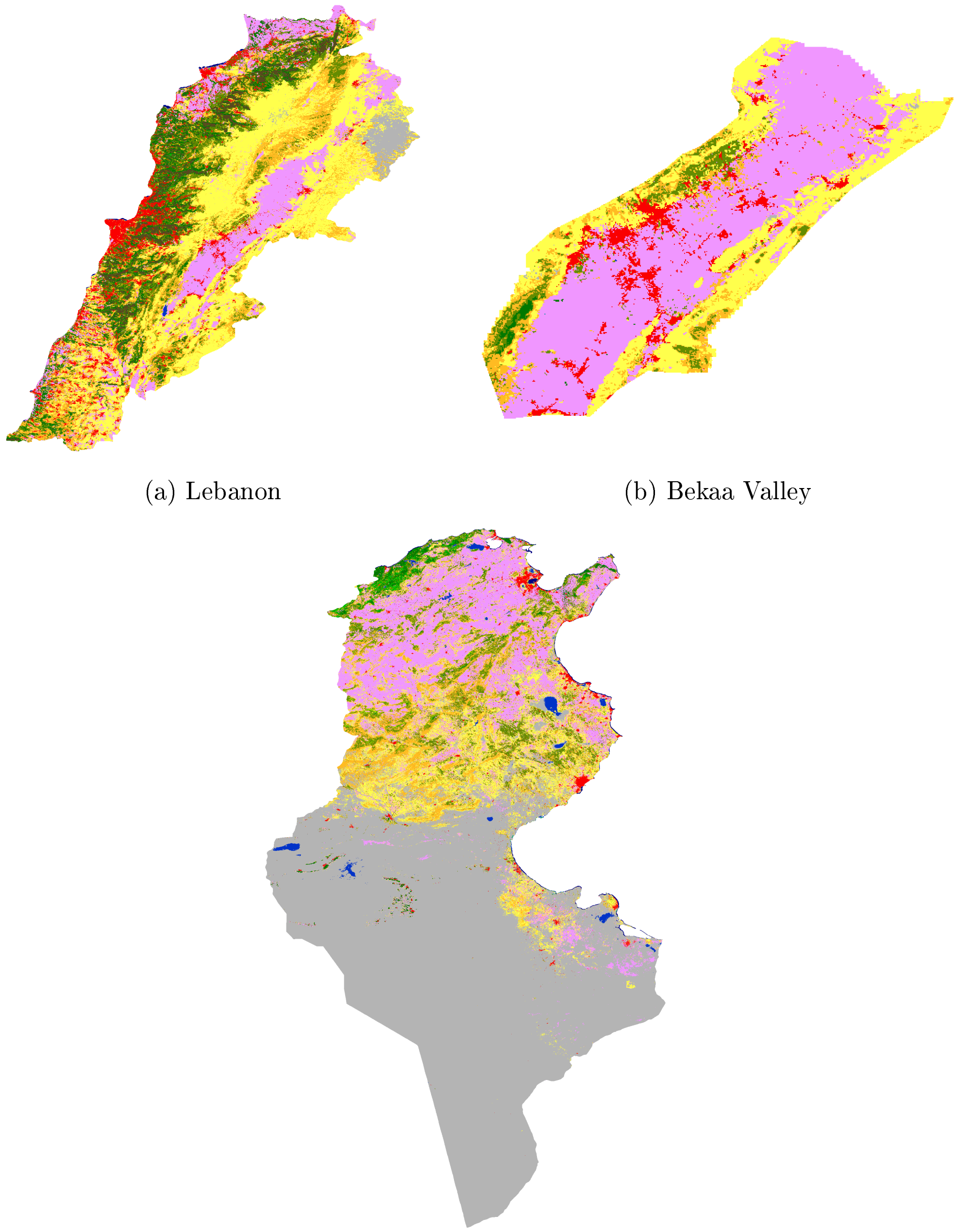
4.4 Robustness across scales

Considering that mass and energy must be preserved between scales, this analysis will evaluate the robustness of the models/data sources to produce sound estimates across all available levels. In a first step the dekadal ET (in mm/day) was first resampled by averaging, from the highest level to the lower one. With the resampled product at the lower scale, annual cumulative ET and annual standard deviation were compared against both levels (i.e. resampled high-resolution and original coarse-resolution levels).

Only Copernicus TSEB-PT, WaPOR-based TSEB-PT and Copernicus ETLook are available at the three levels, while no Level 3 is available for WaPOR in Tunisia. Therefore, Fig. 4.13 shows the differences in Tunisia between Levels 1 and 2 for the four products.

In the case of Lebanon, both Copernicus and WaPOR are available for Level 3 in the Bekaa Valley. As analogous to Tunisia, Fig. 4.14 shows the intercomparison between scales in the whole country, where only Levels 1 and 2 are available for WaPOR. On the other hand level 3 is compared against both Levels 1 and 2 in the Bekaa Valley (Figures 4.15 and 4.16).

Models using Copernicus data shows more robust estimates between Level 3 and 2, since the differences between scales are smaller and closer to 0 mm/day and the annual standard deviation are more similar ($\sigma_{L2}/\sigma_{L3} \approx 1$) On the other hand, WaPOR Level 2 yields significant lower ET estimates, with also lower temporal variability, likely due to the fact that WaPOR ET at 100m uses original MODIS LST at 1km by a simple resampling



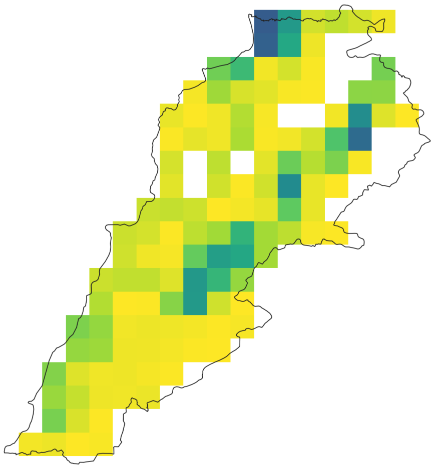
(a) Lebanon (b) Bekaa Valley

(c) Tunisia

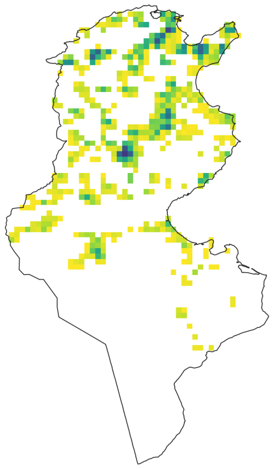
Figure 4.11: Copernicus Global Land Cover product in Lebanon, Bekaa Valley and Tunisia. Agricultural areas are pink, urban areas are red, forests are green, grasslands are yellow, shrublands are orange, and bare areas are gray.

at 100m scale. This issue causes that many irrigated field and other spatial features are not mapped properly in the LST input for ETLook.

In the datasets produced in this study using Copernicus data (TSEB-PT$_C$ and ETLook$_C$), we tried to achieve this consistency by utilizing the same data across all spatial scales (in-



(a) Lebanon



(b) Tunisia

Figure 4.12: Global Map of Irrigation Areas - Version 5 Grid with percentage of area equipped for irrigation with a spatial resolution of 5 arc minutes. Percentage of area closer to 100% is depicted in blue colour.

cluding thermal data), apart from shortwave optical which came from S3 observations at L1 and S2 observations at L2 and L3. In addition, we ensured conservation of thermal energy when sharpening LST between the different resolutions and developed a method to ensure consistency between biophysical parameters derived from S2 and S3 L2A products. Finally, since Copernicus-based L2 and L3 ET is based on exactly the same inputs, we resampled the output of L3 processing to obtain instead of resampling the model inputs. When compared to WaPOR ET maps, which rely on different satellite data at all three levels, the Copernicus-based ETLook ET maps do provide improved consistency across all levels (Figs. 4.13 - 4.16).

However, some differences remain between L1 and other levels even for ETLook_C and $\mathrm{TSEB\text{-}PT}_C$ datasets. Those differences can be attributed to two main factors. Firstly, both TSEB-PT and ETLook models depend on landcover map for setting ancillary parameters (see Table 2.1) and that landcover map is aggregated to the coarser resolution using the statistical mode of the discrete land cover classes. This means that, e.g. a patchwork of urban and agricultural pixels at L3 might become an urban pixel at L1. The TSEB-PT model is more sensitive to those ancillary parameters compared to ETLook (especially to canopy height) and thus the differences between L1 and other levels are larger for $\mathrm{TSEB\text{-}PT}_C$ ET outputs (Figs. 4.13, 4.14, and 4.15). A possible solution could be to first produce maps of ancillary parameters at the highest spatial resolution before aggregating them using averaging to the lower resolutions. The second factor, is the models' assumption of sub-pixel homogeneity which can produce increased output uncertainty in environments in which sub-pixel heterogeneity is present [25, 26]. This assumption becomes increasingly less valid as the pixel size increases. This is a more complex issue to solve and could require a change of paradigm of ET model assumptions.

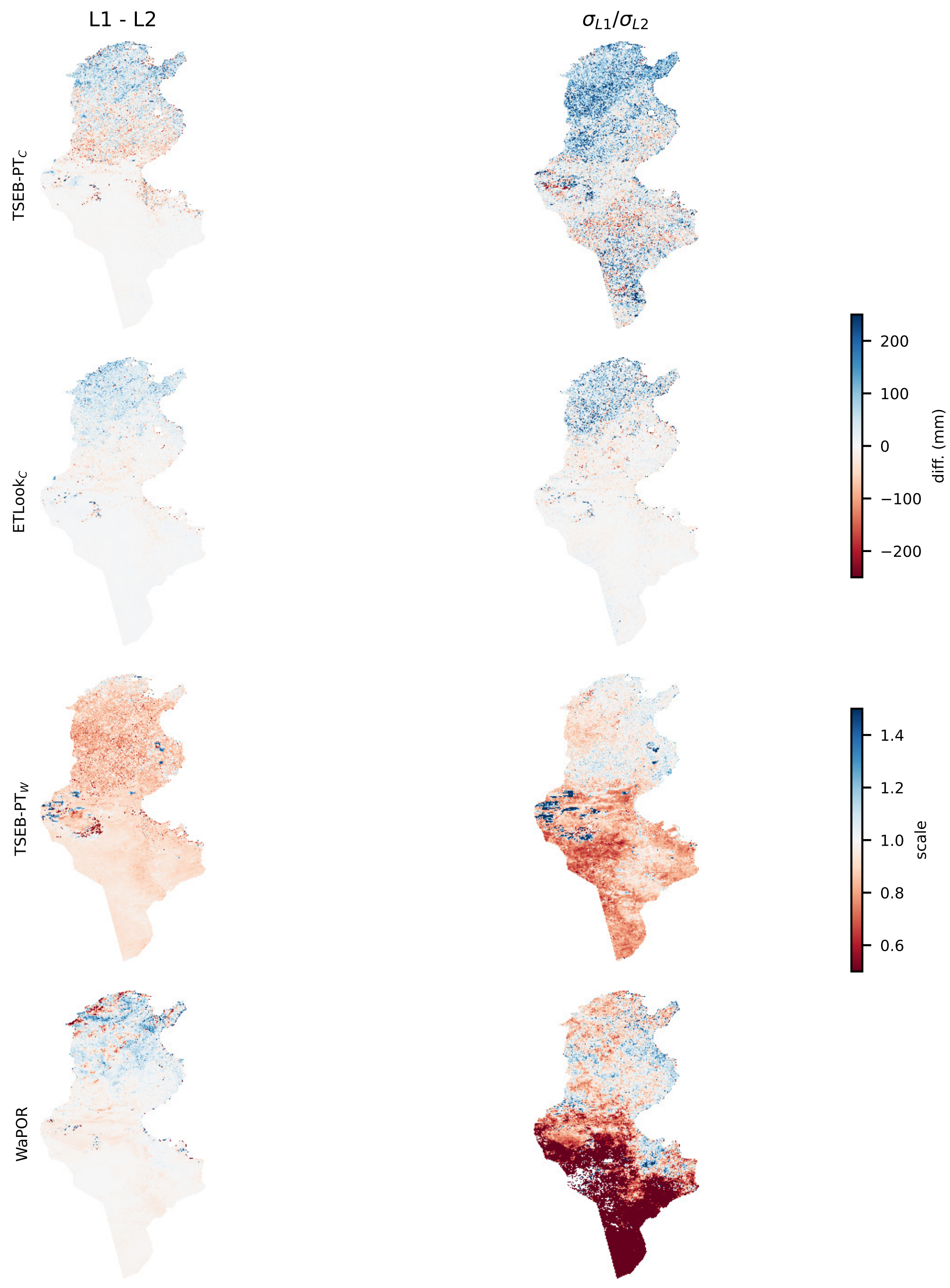


Figure 4.13: Comparison between Levels 1 and 2 for ET products in Tunisia, both in ET differences (left panels) and in standard deviations scale (fight panels)

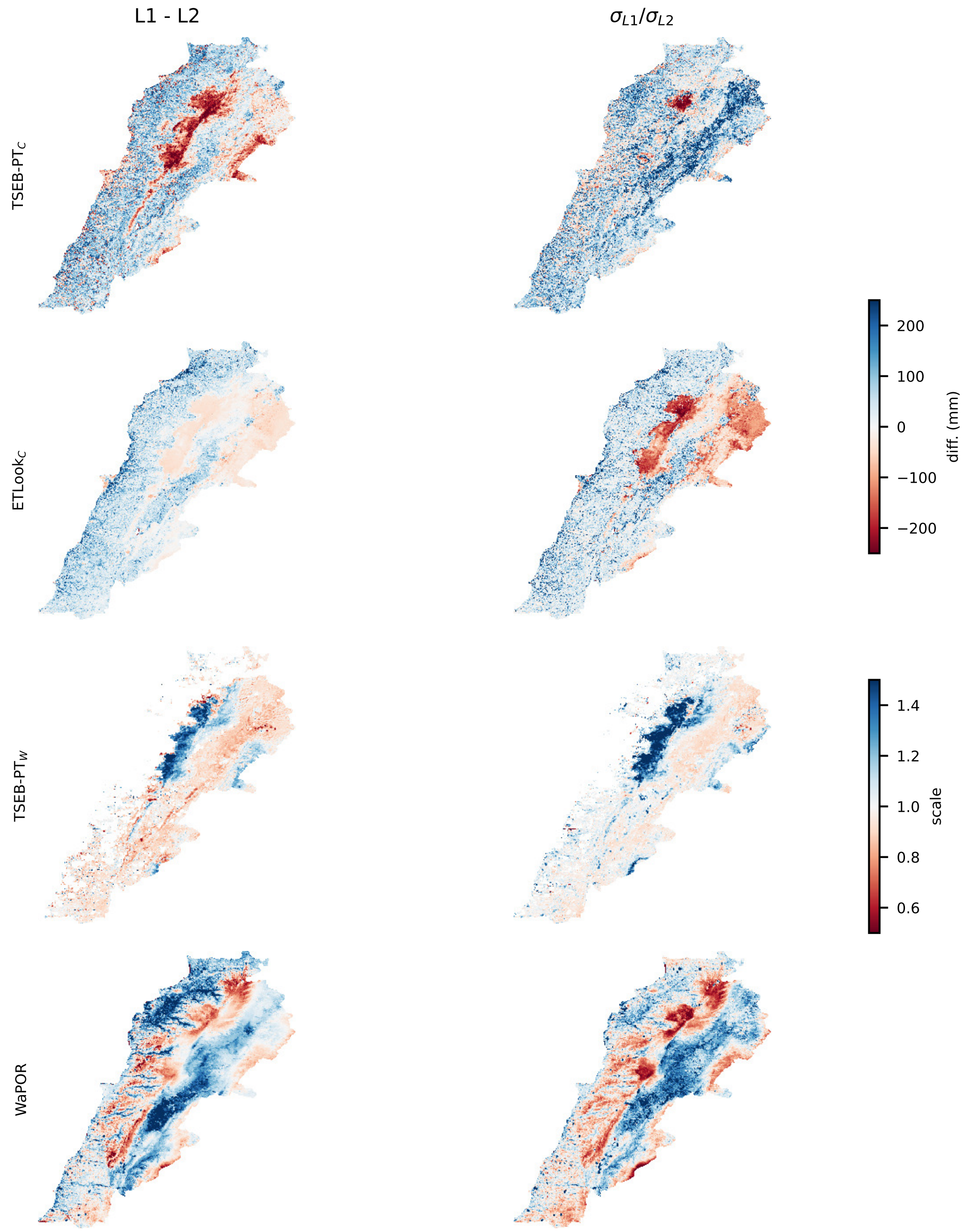


Figure 4.14: Comparison between Levels 1 (L1) and 2 (L2) for ET products in Lebanon, both in cummulative annual ET difference (left panels) and in intraannual standard deviations scale (right panels)

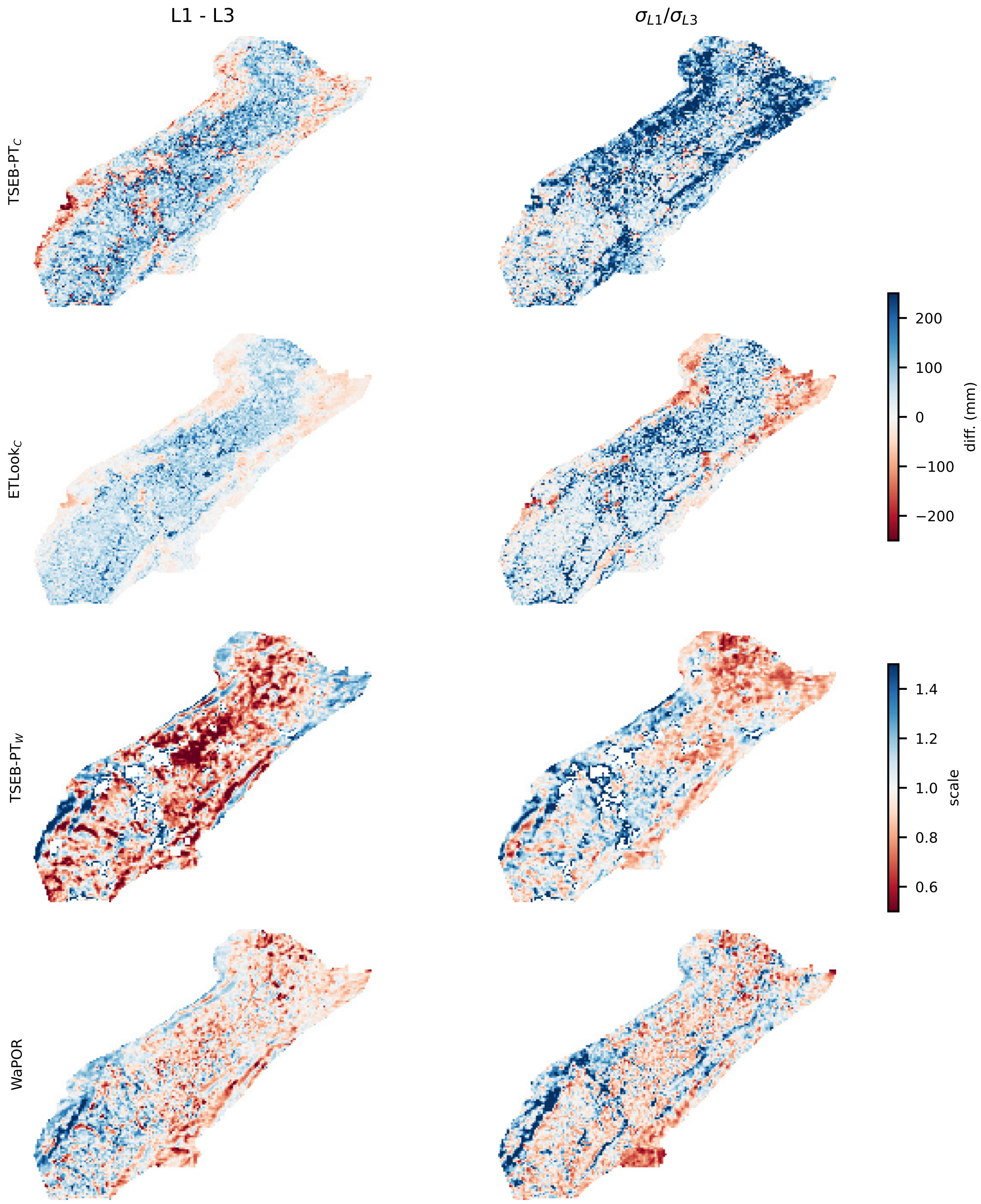


Figure 4.15: Comparison between Levels 1 (L1) and 3 (L3) for ET products in Bekaa Valley, both in cummulative annual ET difference (left panels) and in intraannual standard deviations scale (right panels)

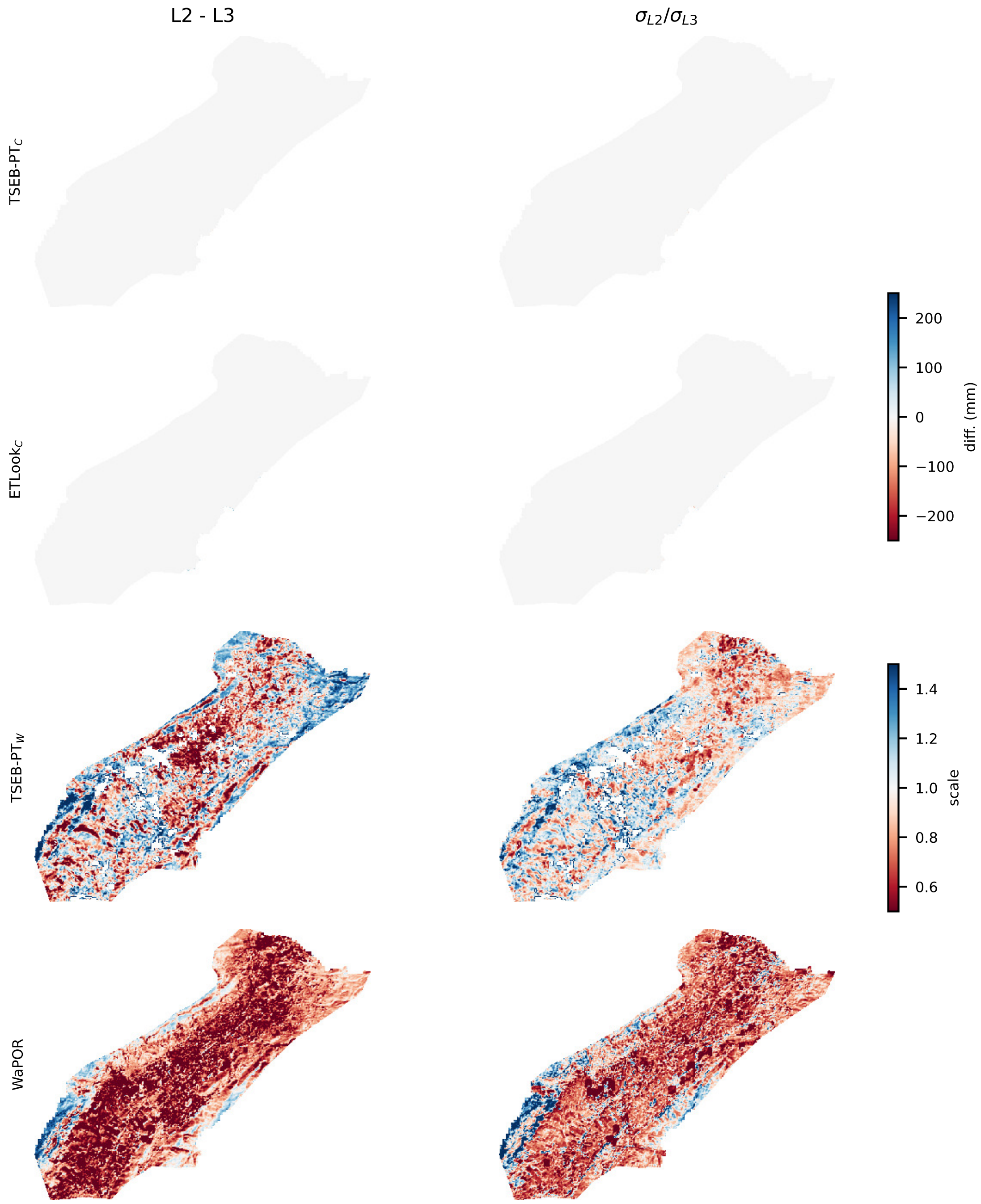


Figure 4.16: Comparison between levels 2 (L2) and 3 (L2) for ET products in Bekaa Valley, both in cummulative annual ET difference (left panels) and in intraannual standard deviations scale (right panels)

Chapter 5

Comparison with previous studies

Since ETLook is a rather recent model published in 2012 [14], fewer studies have been issued with the aim validating its performance as compared to TSEB-PT, which first version was published in 1995 [27] and hence applied and validated at a larger number of studies and different environments [28, 6]. Nevertheless, ETLook has shown its potential in tracking the spatio-temporal variability of ET, with correlation coefficients between the observed and the predicted of 0.9 in this study, which is confirmed from a previous study that showed a R^2 of 0.54 (equivalent to a correlation of 0.73) for a larger number of study sites [29]. However, ETLook still have room for improvement in capturing the ET rate magnitude. The *in situ* validation in this study showed that ETLook systematically underestimated dekadal ET in irrigated and rainfed semi-arid croplands, but also tend to yield lower cumulative monthly and annual ET rates as compared to TSEB-PT model using Copernicus data. [29] also found significant biases for the WaPOR product in semi-arid rainfed croplands, with relative errors higher than 50% of the measured dekadal ET rates. However, the same study also showed that WaPOR product tends to overestimate ET at irrigated croplands, which is the opposite as showed in our study in the Barrax sites. Furthermore, [30] also pointed out the trend of WaPOR of underestimating ET as compared to a lysimeter measurements in an alfalfa field in Iran.

Regarding the spatial consistency between levels, similar behaviour was found on WaPOR L3 yielding largest ET annual and monthly trends as compared to L1 and L2 by [29] and [31]. The authors of these two studies also pointed out the effect of using PROBA-V for L2 (100m) ET after 2014, which produced as loss of consistency against L1 (300m). In ET4FAO we proposed a method for ensuring spatial consistency of the different inputs across scales, which minimized such disagreements of water accounting at different spatial resolutions. However, using Copernicus data, due to the lack of high-spatial thermal infrared mission, might also lead in larger uncertainties in tracking either stress or wetting events in the shorter term. Until an operational high-spatial and high-revisit thermal infrared satellite mission is launched, future research should focus in exploring new sharpening and fusion methods that could integrate Landsat LST data in order to improve the dynamic range of surface temperature within the spatial domain.

Chapter 6

Recommendations and Conclusions

6.1 Utility of Copernicus data for ET modeling

The results indicate the suitability of Copernicus data as inputs for consistent ET modeling at various spatial resolutions from 20 m to 300 m. TSEB-PT$_C$ achieved correlation of 0.9 and bias of less than 0.3 mm/day when validating L3 ET against all the field measurements combined. The bias of both TSEB-PT$_C$ and ETLook$_C$ points to underestimation of dekadal ET and it could partially be caused by the choice of ET gap-filling method (see section 2.3.3 and [32]). At the same time, both TSEB-PT$_C$ and ETLook$_C$ achieved better accuracy than WaPOR at the one site where L2 ET was available from all three datasets (Tunisian rainfed olive grove). In addition, ETLook$_C$ performed the best when comparing consistency across the three spatial scale. While the validation sites represented a wide selection of irrigation practices and crop types, they are all located in semi-arid Mediterranean climate. It could be argued that this represents the regions where irrigation demand is highest and water shortages most pressing. However, irrigation is being developed in countries as diverse as Uganda and Denmark and therefore the validation effort should be extended to other climates. It should also be noted that WaPOR has to operate across all the African and Middle East climates and while no climate or site specific adjustments were included in the preprocessing of Copernicus data or in modeling ET, some trade-offs might be involved when the geographical area of interest is expanded.

The robustness of LST sharpening approach for producing high spatio-temporal resolution representation of the LST based on S3 observations is also demonstrated by the field validation results presented in section 3.2. ET derived with sharpened LST is well able to capture spatial and temporal patterns even of small fields with different irrigation and growing regimes compared to the neighboring parcels (e.g. potato and vineyard fields as shown in Figure 3.2)). However, those results also illustrate the limitation of sharpening low-resolution LST, namely the difficulty in capturing LST values which are outside of the range of the low-resolution LST, which by its nature is an aggregated value. This can be observed in the highest bias being present in the festuca site (well irrigated and surrounded by semi-arid landscape) and was also reported in previous studies [33, 28]. While progressive enhancements to various LST sharpening methods are being proposed [33, 34], this issue cannot really be resolved without the use of thermal sensors with high spatio-temporal resolution. In the context of Copernicus this will be addressed by the proposed Land Surface Temperature Monitoring (LSTM) mission [35], while in the meantime a fusion between S3 and Landsat (high spatial but low temporal resolution) thermal observations could be explored, based on existing methods [36].

The meteorological forcing derived from ERA5 has the lowest original spatial resolution of all the input datasets, of around 30 km. We performed topographic corrections of these data through which the resolution is increased. To assess how this might impact local, parcel-scale ET estimates we compared the orographic sharpened ERA5 fields against one meteorological station located in Bekaa valley, with results shown in Figure 6.1 for reference ET (ET_{ref}). ET_{ref} combines all the relevant meteorological parameters needed for ET modeling and shows very high correlation and very low bias. Similar results have been obtained with other agrometeorological stations placed in areas outside the regions which are the subject of this study (results therefore not shown), in which instantaneous air temperature and solar irradiance also corresponded well to ground measurements, while wind speed, which is the most difficult parameter to model at local scales, still showed acceptable results. This implies the suitability of ERA5 inputs even for high-resolution ET modeling. Copernicus Climate Data Store also provide access to ERA5-Land dataset which contains only surface meteorological outputs of ERA5 but with a 9 km resolution [37]. It uses a conservative land mask which makes the use of this dataset impractical in coastal areas. However, in areas away from the coast, ERA5 Land might lead to even better agreement with local measurements.

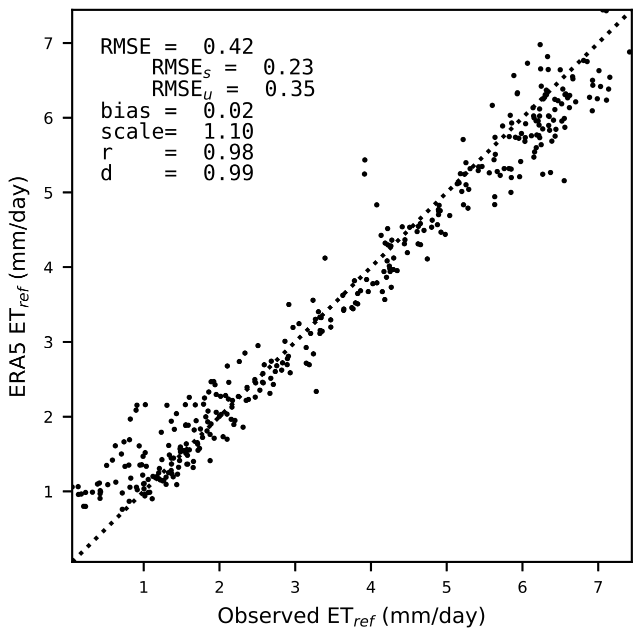


Figure 6.1: Comparison of daily reference ET modeled with topographically corrected ERA5 data against measurements in Tal Amara station in Bekaa Valley.

Finally, land-cover map can impact ET outputs through its influence of ancillary parameters such as vegetation height. CGLC has 23 land-cover classes, high spatial resolution of 100 m, is updated annually since 2015 and was extensively validated resulting in overall accuracy of 80% [38]. However, it still has some limitations when used in ET models. The first, particularly relevant for SDG indicator 6.4.1 reporting in agriculture, is the presence of a single agricultural class. This class contains such diverse types as orchards, vineyards and herbaceous annual crops and all of them had to be assigned the same ancillary canopy parameters despite being clearly different. For example, in the approach used in this study the vegetation height in agricultural pixels is scaled with LAI up to a maximum value of 1.5 m [6]. This results in underestimation of vegetation height in both the grapevine and young almond sites while overestimation is present in potato and reference festuca sites. Olive grove location was classified as unknown forest type and therefore had constant height of 1.5 m. In TSEB-PT model, underestimation of vegetation height leads to potential underestimation of sensible heat flux (through underestimation of surface roughness) and therefore overestimation of latent heat flux, and vice versa. In Table 3.2 it can be seen that the largest underestimation of TSEB-PT$_C$ ET happens in festuca and potato sites where vegetation height is overestimated and largest overestimation of ET happens in grapevine site where vegetation height is underestimated. The second issue, is the mismatch between CGLC spatial resolution and the 20 m S2 data used to set the output resolution of Level 3 ET product. This results in visible 100 m by 100 m blocks in the output ET maps, especially apparent at the borders between two different land-cover classes, such as agriculture and forests. Landcover maps with resolution of up to 10 m are being produced using Sentinel-2 data [39] and could potentially be used as inputs to ET modeling.

6.2 Ensuring consistency between spatial scales

An important aspect when estimating ET at multiple spatial scales is to ensure consistency across those scales. This is important from a theoretical point of view because mass and energy should be conserved, and from a practical point of view because regional or national estimates of water use should not change depending the spatial resolution of the map that is being used. In the new datasets produced in this study (TSEB-PT$_C$ and ETLook$_C$) we tried to achieve this consistency by utilizing the same data across all spatial scales (including thermal data), apart from shortwave optical which came from S3 observations at L1 and S2 observations at L2 and L3. In addition, we ensured conservation of thermal energy when sharpening LST between the different resolutions and developed a method to ensure mass consistency between biophysical parameters derived from S2 and S3 L2A products. Finally, since Copernicus-based L2 and L3 ET is based on exactly the same inputs, we resampled the output of L3 processing to obtain instead of resampling the model inputs. When compared to WaPOR ET maps, which rely on different satellite data at all three levels, the Copernicus-based ETLook ET maps do provide improved consistency across all levels (Figs. 4.13 - 4.16).

6.3 On WaPOR interception product

Besides of soil evaporation and canopy transpiration, WaPOR provides and additional dataset named as Interception and defined in the WaPOR methodology documents as "[...] the rainfall intercepted by the leaves of the plants that will be directly evaporated from their surface". Such evaporation from intercepted rainfall is estimated simply as a function of daily rainfall and Leaf Area Index. However, especial care should be taken in using and interpreting this product:

- This approach assumes that all intercepted water by the canopy after a rainfall is evaporated. This is not usually true as a significant fraction of this intercepted rainfall is removed from the leaves by the wind gusts after the rainfall happened.
- It is unable to quantify evaporation at the canopy caused by dew and/or intercepted overhead irrigation. Specially the latter would yield to a large underestimation of such interception in irrigated areas.
- Despite most energy balance models do not explicitly account for this rather complex process, when using thermal-based information the evaporation of intercepted water is reflected by a decrease of LST. Therefore this process is already implicitly accounted for when using any thermal-based energy balance model such as ETLook or TSEB-PT.

For all these reason we discourage using this product together with WaPOR Evaporation and Transpiration products when accounting for total water use (as provided in the WaPOR ETI product). In any case documentation should clarify these issues and hence let the end-users decide whether or not to include this Interception product in their water accounting activities.

6.4 Conclusions

Estimating spatial and temporal patterns of evapotranspiration is essential for accurate reporting of the agricultural component of SDG target 6.4. The use of earth observation based ET estimates can improve consistency of this reporting across administrative and natural boundaries, thus increasing transparency and trust. In this study, we evaluated whether Copernicus products are suitable as input datasets for ET models. The ET product available on the WaPOR portal, run by FAO with the aim of encouraging the use of satellite observations in SDG 6.4 reporting, was used as a benchmark. Therefore we assessed the accuracy, consistency and spatial patterns of Copernicus-based ET at 10-day timestep and three spatial resolutions (20 m, 100 m and 300 m).

The results from validating the estimated Copernicus-based ET against measurements from six field sites spread across irrigated and rainfed agriculture in semi-arid Mediterranean climate indicate an accuracy of less than 0.3 mm/day using TSEB-PT model. At the same time, when Copernicus inputs are used with the same ET model as used in WaPOR (ETLook) a better consistency across spatial scales is obtained compared to WaPOR. This is due to limiting the number of different satellite sensors when modeling at different spatial resolutions and the use of inputs' pre-processing methods designed to ensure consistency. Large scale spatio-temporal patterns resulting from monthly and annual aggregations of the different ET products are more difficult to interpret, although the models do show the same general outlines and spatio-temporal trends consistent with irrigation patterns, as oposed to model purely driven with meteorological forcing.

Although the results show high suitability of Copernicus-based ET for SDG reporting, a number of issues should be addressed to further increase the quality of the outputs. Some of them can be addressed in the shorter term while others require a long-term perspective. Among the former, is the fact that the ET model which produced the most accurate fluxes (TSEB-PT) was also less consistent across spatial scales compared to the model used in WAPOR (ETLook). This is mainly due to its sensitivity to land-cover based parameters and might require a modification to the pre-processing of model input data or including additional remote sensing data with information on canopy structure such as SAR and/or LiDAR. Secondly, the validation should be extended to other climatic zones. Even though the Mediterranean region has among highest proportion of fresh water withdrawals used in irrigated agriculture in the world, irrigation is widely used across the globe. An issue which requires a longer-term perspective is the lack of high spatio-temporal resolution thermal sensor within the Copernicus constellation. A Copernicus Land Surface Temperature Monitoring (LSTM) mission addressing this data gap is being planned but it will not be operational for a number of years yet. In the meantime, advanced data fusion methods between different satellite sensors can partially fill this gap.

In summary, this study has demonstrated that products based on satellite observations and meteorological models made freely and openly available by the Copernicus program are highly suitable for consistent and robust estimation of ET in the context of SDG reporting. By relying on predominantly Copernicus data it is possible to take advantage of its operational data quality and guaranteed long-term continuity, thus laying a robust baseline for monitoring of changes in the SDG 6.4 indicators.

Bibliography

[1] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, Andras Horanyi, Joaquin Munoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, and et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. ISSN 1477-870X. doi: https://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.

[2] Pierre Ineichen. A broadband simplified version of the Solis clear sky model. *Solar Energy*, 82(8):758–762, August 2008. ISSN 0038-092X. doi: 10.1016/j.solener.2008.02.009. URL `https://www.sciencedirect.com/science/article/pii/S0038092X08000406`.

[3] C. Aguilar, J. Herrero, and M. J. Polo. Topographic effects on solar radiation distribution in mountainous watersheds and their influence on reference evapotranspiration estimates at watershed scale. *Hydrology and Earth System Sciences*, 14(12): 2479–2494, December 2010. ISSN 1607-7938. doi: 10.5194/hess-14-2479-2010. URL `https://www.hydrol-earth-syst-sci.net/14/2479/2010/`.

[4] Magdalena Main-Knorn, Bringfried Pflug, Jerome Louis, Vincent Debaecker, Uwe Müller-Wilm, and Ferran Gascon. Sen2Cor for Sentinel-2. In *Image and Signal Processing for Remote Sensing XXIII*, volume 10427, page 1042704. International Society for Optics and Photonics, October 2017. doi: 10.1117/12.2278218. URL `https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10427/1042704/Sen2Cor-for-Sentinel-2/10.1117/12.2278218.short`.

[5] David Frantz, Erik Haß, Andreas Uhl, Johannes Stoffels, and Joachim Hill. Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sensing of Environment*, 215: 471–481, September 2018. ISSN 0034-4257. doi: 10.1016/j.rse.2018.04.046. URL `http://www.sciencedirect.com/science/article/pii/S0034425718302037`.

[6] Radoslaw Guzinski, Hector Nieto, Inge Sandholt, and Georgios Karamitilios. Modelling High-Resolution Actual Evapotranspiration through Sentinel-2 and Sentinel-3 Data Fusion. *Remote Sensing*, 12(9):1433, January 2020. doi: 10.3390/rs12091433. URL `https://www.mdpi.com/2072-4292/12/9/1433`. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

[7] G. S. Campbell. Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution. *Agricultural and Forest Meteorology*, 36

(4):317–321, April 1986. ISSN 0168-1923. doi: 10.1016/0168-1923(86)90010-9. URL https://www.sciencedirect.com/science/article/pii/0168192386900109.

[8] FRAME Consortium. WaPOR Data Manual - Evapotranspiration v2, June 2020. URL https://bitbucket.org/cioapps/wapor-et-look/downloads/FRAME_ET_v2_data_manual_finaldraft_v2.2.pdf.

[9] H. RAHMAN and G. DEDIEU. SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing*, 15(1):123–143, January 1994. ISSN 0143-1161. doi: 10.1080/01431169408954055. URL https://doi.org/10.1080/01431169408954055. Publisher: Taylor & Francis_eprint: https://doi.org/10.1080/01431169408954055.

[10] Jin Chen, Per. Jonsson, Masayuki Tamura, Zhihui Gu, Bunkei Matsushita, and Lars Eklundh. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzkyâ€“Golay filter. *Remote Sensing of Environment*, 91(3):332–344, June 2004. ISSN 0034-4257. doi: 10.1016/j.rse.2004.03.014. URL https://www.sciencedirect.com/science/article/pii/S003442570400080X.

[11] William P. Kustas, Hector Nieto, Laura Morillas, Martha C. Anderson, Joseph G. Alfieri, Lawrence E. Hipps, Luis Villagarcía, Francisco Domingo, and Monica Garcia. Revisiting the paper "Using radiometric surface temperature for surface energy flux estimation in Mediterranean drylands from a two-source perspective". *Remote Sensing of Environment*, 184:645–653, October 2016. ISSN 0034-4257. doi: 10.1016/j.rse.2016.07.024. URL http://www.sciencedirect.com/science/article/pii/S0034425716302814.

[12] John R. Mecikalski, George R. Diak, Martha C. Anderson, and John M. Norman. Estimating Fluxes on Continental Scales Using Remotely Sensed Data in an Atmosphericâ€“Land Exchange Model. *Journal of Applied Meteorology*, 38(9):1352–1369, September 1999. ISSN 0894-8763. doi: 10.1175/1520-0450(1999)038<1352:EFOCSU>2.0.CO;2. URL http://journals.ametsoc.org/doi/full/10.1175/1520-0450(1999)038%3C1352:EFOCSU%3E2.0.CO%3B2.

[13] C. Cammalleri, M. C. Anderson, and W. P. Kustas. Upscaling of evapotranspiration fluxes from instantaneous to daytime scales for thermal remote sensing applications. *Hydrology and Earth System Sciences*, 18(5):1885–1894, 2014. doi: 10.5194/hess-18-1885-2014. URL https://hess.copernicus.org/articles/18/1885/2014/.

[14] W. G. M. Bastiaanssen, M. J. M. Cheema, W. W. Immerzeel, I. J. Miltenburg, and H. Pelgrum. Surface energy balance and actual evapotranspiration of the transboundary Indus Basin estimated from satellite measurements and the ET-Look model. *Water Resources Research*, 48(11), 2012. ISSN 1944-7973. doi: 10.1029/2011WR010482. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011WR010482.

[15] Yuting Yang, Huade Guan, Di Long, Bing Liu, Guanghua Qin, Jun Qin, and Okke Batelaan. Estimation of Surface Soil Moisture from Thermal Infrared Remote Sensing Using an Improved Trapezoid Method. *Remote Sensing*, 7(7):8250–8270, July

2015. doi: 10.3390/rs70708250. URL `https://www.mdpi.com/2072-4292/7/7/8250`. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.

[16] W. Chebbi, G. Boulet, V. Le Dantec, Z. Lili Chabaane, P. Fanise, B. Mougenot, and H. Ayari. Analysis of evapotranspiration components of a rainfed olive orchard during three contrasting years in a semi-arid climate. *Agricultural and Forest Meteorology*, 256-257:159–178, 2018. doi: 10.1016/j.agrformet.2018.02.020.

[17] Domenico Vitale, Gerardo Fratini, Massimo Bilancia, Giacomo Nicolini, Simone Sabbatini, and Dario Papale. A robust data cleaning procedure for eddy covariance flux measurements. *Biogeosciences*, 17(6):1367–1391, 2020. doi: 10.5194/bg-17-1367-2020.

[18] R. López-Urrea, F. Martín de Santa Olalla, C. Fabeiro, and A. Moratalla. Testing evapotranspiration equations using lysimeter observations in a semiarid climate. *Agricultural Water Management*, 85(1):15–26, 2006. doi: 10.1016/j.agwat.2006.03.014.

[19] Juan M. Sánchez, Ramón López-Urrea, Francisco Valentín, Vicente Caselles, and Joan M. Galve. Lysimeter assessment of the Simplified Two-Source Energy Balance model and eddy covariance system to estimate vineyard evapotranspiration. *Agricultural and Forest Meteorology*, 274:172–183, 2019. doi: 10.1016/j.agrformet.2019.05.006.

[20] T.E. Twine, W.P. Kustas, J.M. Norman, D.R. Cook, P.R. Houser, T.P. Meyers, J.H. Prueger, P.J. Starks, and M.L. Wesely. Correcting eddy-covariance flux underestimates over a grassland. *Agricultural and Forest Meteorology*, 103(3):279–300, 2000. doi: 10.1016/S0168-1923(00)00123-4.

[21] Cort J. Willmott. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society*, 63(11):1309–1313, 1982. doi: 10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2.

[22] John T. Abatzoglou, Solomon Z. Dobrowski, Sean A. Parks, and Katherine C. Hegewisch. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. *Scientific Data*, 5(1):170191, January 2018. ISSN 2052-4463. doi: 10.1038/sdata.2017.191. URL `https://www.nature.com/articles/sdata2017191`. Number: 1 Publisher: Nature Publishing Group.

[23] B. Martens, D. G. Miralles, H. Lievens, R. van der Schalie, R. A. M. de Jeu, D. Fernández-Prieto, H. E. Beck, W. A. Dorigo, and N. E. C. Verhoest. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5):1903–1925, 2017. doi: 10.5194/gmd-10-1903-2017.

[24] N. Ghilain, A. Arboleda, and F. Gellens-Meulenberghs. Evapotranspiration modelling at large scale using near-real time MSG SEVIRI derived data. *Hydrology and Earth System Sciences*, 15(3):771–786, 2011. doi: 10.5194/hess-15-771-2011.

[25] Vicente Burchard-Levine, Héctor Nieto, David Riaño, Mirco Migliavacca, Tarek S. El-Madany, Radoslaw Guzinski, Arnaud Carrara, and M. Pilar Martín. The effect of pixel heterogeneity for remote sensing based retrievals of evapotranspiration in a semi-arid tree-grass ecosystem. *Remote Sensing of Environment*, 260:112440, July 2021.

ISSN 0034-4257. doi: 10.1016/j.rse.2021.112440. URL https://www.sciencedirect.com/science/article/pii/S0034425721001589.

[26] Ayman Nassar, Alfonso Torres-Rua, William Kustas, Hector Nieto, Mac McKee, Lawrence Hipps, David Stevens, Joseph Alfieri, John Prueger, Maria Mar Alsina, Lynn McKee, Calvin Coopmans, Luis Sanchez, and Nick Dokoozlian. Influence of Model Grid Size on the Estimation of Surface Fluxes Using the Two Source Energy Balance Model and sUAS Imagery in Vineyards. *Remote Sensing*, 12(3):342, January 2020. doi: 10.3390/rs12030342. URL https://www.mdpi.com/2072-4292/12/3/342.

[27] J.M. Norman, W.P. Kustas, and K.S. Humes. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agricultural and Forest Meteorology*, 77(3):263–293, 1995. ISSN 0168-1923. doi: https://doi.org/10.1016/0168-1923(95)02265-Y. URL https://www.sciencedirect.com/science/article/pii/016819239502265Y. Thermal Remote Sensing of the Energy and Water Balance over Vegetation.

[28] Radoslaw Guzinski and HÃ©ctor Nieto. Evaluating the feasibility of using Sentinel-2 and Sentinel-3 satellites for high-resolution evapotranspiration estimations. *Remote Sensing of Environment*, 221:157–172, February 2019. ISSN 0034-4257. doi: 10.1016/j.rse.2018.11.019. URL http://www.sciencedirect.com/science/article/pii/S0034425718305285.

[29] Megan L. Blatchford, Chris M. Mannaerts, Sammy M. Njuki, Hamideh Nouri, Yijian Zeng, Henk Pelgrum, Steven Wonink, and Poolad Karimi. Evaluation of wapor v2 evapotranspiration products across africa. *Hydrological Processes*, 34(15):3200–3221, 2020. doi: https://doi.org/10.1002/hyp.13791. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.13791.

[30] Mostafa Javadian, Ali Behrangi, Mohsen Gholizadeh, and Masoud Tajrishy. Metric and wapor estimates of evapotranspiration over the lake urmia basin: Comparative analysis and composite assessment. *Water*, 11(8), 2019. ISSN 2073-4441. doi: 10.3390/w11081647. URL https://www.mdpi.com/2073-4441/11/8/1647.

[31] Megan Blatchford, Chris M. Mannaerts, Yijian Zeng, Hamideh Nouri, and Poolad Karimi. Influence of spatial resolution on remote sensing-based irrigation performance assessment using wapor data. *Remote Sensing*, 12(18), 2020. ISSN 2072-4292. doi: 10.3390/rs12182949. URL https://www.mdpi.com/2072-4292/12/18/2949.

[32] Emilie Delogu, Albert Olioso, Aubin Alliès, Jérôme Demarty, and Gilles Boulet. Evaluation of Multiple Methods for the Production of Continuous Evapotranspiration Estimates from TIR Remote Sensing. *Remote Sensing*, 13(6):1086, January 2021. doi: 10.3390/rs13061086. URL https://www.mdpi.com/2072-4292/13/6/1086. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

[33] Juan M. Sánchez, Joan M. Galve, José González-Piqueras, Ramón López-Urrea, Raquel Niclòs, and Alfonso Calera. Monitoring 10-m LST from the Combination MODIS/Sentinel-2, Validation in a High Contrast Semi-Arid Agroecosystem. *Remote Sensing*, 12(9):1453, January 2020. doi: 10.3390/rs12091453. URL https://www.mdpi.com/2072-4292/12/9/1453.

[34] Sammy M. Njuki, Chris M. Mannaerts, and Zhongbo Su. An Improved Approach for Downscaling Coarse-Resolution Thermal Data by Minimizing the Spatial Averaging Biases in Random Forest. *Remote Sensing*, 12(21):3507, January 2020. doi: 10.3390/rs12213507. URL `https://www.mdpi.com/2072-4292/12/21/3507`. Number: 21 Publisher: Multidisciplinary Digital Publishing Institute.

[35] B. Koetz, W. Bastiaanssen, M. Berger, P. Defourney, U. Del Bello, M. Drusch, M. Drinkwater, R. Duca, V. Fernandez, D. Ghent, R. Guzinski, J. Hoogeveen, S. Hook, J. Lagouarde, G. Lemoine, I. Manolis, P. Martimort, J. Masek, M. Massart, C. Notarnicola, J. Sobrino, and T. Udelhoven. 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*, pages 8160–8162, July 2018. doi: 10.1109/IGARSS.2018.8517433.

[36] K. R. Knipper, W. P. Kustas, M. C. Anderson, H. Nieto, J. G. Alfieri, J. H. Prueger, C. R. Hain, F. Gao, L. G. McKee, M. Mar Alsina, and L. Sanchez. Using high-spatiotemporal thermal satellite ET retrievals to monitor water use over California vineyards of different climate, vine variety and trellis design. *Agricultural Water Management*, 241:106361, November 2020. ISSN 0378-3774. doi: 10.1016/j.agwat.2020.106361. URL `http://www.sciencedirect.com/science/article/pii/S0378377420305473`.

[37] Joaquín Muñoz Sabater, Emanuel Dutra, Anna Agustí-Panareda, Clément Albergel, Gabriele Arduini, Gianpaolo Balsamo, Souhail Boussetta, Margarita Choulga, Shaun Harrigan, Hans Hersbach, Brecht Martens, Diego G. Miralles, María Piles, Nemesio J. Rodríguez-Fernández, Ervin Zsoter, Carlo Buontempo, and Jean-Noël Thépaut. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data Discussions*, pages 1–50, March 2021. ISSN 1866-3508. doi: 10.5194/essd-2021-82. URL `https://essd.copernicus.org/preprints/essd-2021-82/`. Publisher: Copernicus GmbH.

[38] Nandin-Erdene Tsendbazar, Agnieszka Tarko, Linlin Li, Martin Herold, Myroslava Lesiv, Steffen Fritz, and Victor Maus. Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Validation Report. Technical report, Zenodo, Geneve, Switzerland, September 2020. URL `https://zenodo.org/record/4723975#.YLR1cKF8LmE`.

[39] Radek Malinowski, Stanislaw Lewinski, Marcin Rybicki, Ewa Gromny, Malgorzata Jenerowicz, Michal Krupinski, Artur Nowakowski, Cezary Wojtkowski, Marcin Krupinski, Elke Krätzschmar, and Peter Schauer. Automated Production of a Land Cover/Use Map of Europe Based on Sentinel-2 Imagery. *Remote Sensing*, 12(21):3523, January 2020. doi: 10.3390/rs12213523. URL `https://www.mdpi.com/2072-4292/12/21/3523`. Number: 21 Publisher: Multidisciplinary Digital Publishing Institute.