



OPEN Evapotranspiration estimation using Surface Energy Balance Model and medium resolution satellite data: An operational approach for continuous monitoring

S. Pareeth[✉] & P. Karimi

Monitoring spatial and temporal trends of water use is of utmost importance to ensure water and food security in river basins that are challenged by water scarcity and climate change induced abnormal weather patterns. To quantify water consumption by the agriculture sector, continuous monitoring is required over different spatial scales ranging from field (<1 ha) to basin. The demand driven requirement of covering large areas yet providing spatially distributed information makes the use of in-situ measurement devices unfeasible. Earth observation satellites and remote sensing techniques offer an effective alternative in estimating the consumptive use of water (Actual EvapoTranspiration (ET_a) fluxes) by using periodic observations from the visible and infrared spectral region. Optical satellite data, however, is often hindered by noises due to cloud cover, cloud shadow, aerosols and other satellite related issues such as Scan Line Corrector (SLC) failure in Landsat 7 breaking the continuity of temporal observations. These gaps have to be statistically filled in order to compute aggregated seasonal and annual estimates of ET_a. In this paper, we introduce an approach to develop a gap-filled multi-year monthly ET_a maps at medium spatial resolution of 30 m. The method includes two major steps: (i) estimation of ET_a using the python based implementation of surface energy balance model called PySEBAL and (ii) temporal interpolation using Locally Weighted Regression (LWR) model followed by spline based spatial interpolation to fill the gaps over time and space. The approach is applied to a large endorheic Lake Urmia Basin (LUB) basin with a surface area of ~52,970 km² in Iran for the years 2013–2015 using Landsat 7 and 8 satellite data. The results show that the implemented gap filling approach could reconstruct the monthly ET_a dynamics over different agriculture land use types, while retaining the high spatial variability. A comparison with a similar dataset from FAO WaPOR reported a very high correlation with R² of 0.93. The study demonstrates the applicability of this approach to a larger basin which is extendible and reproducible to other geographical areas.

Water scarcity is one of the most critical global issues of the twenty-first century which has affected the daily livelihood of millions. An estimated forty percent of the global population are experiencing water scarcity at different magnitudes^{1,2}. In Asia and the Middle East region, water scarcity is mainly driven by climate change, improper water use, and growing population³. Many studies have reported a rapid decrease in the level of groundwater table in the semi-arid regions of Asia and the Middle East mainly due to the overuse of water to meet irrigation needs^{4–6}. Therefore, there is a great need to establish monitoring systems to facilitate timely interventions ensuring sustainable management of land and water resources. One of the Sustainable Development Goal (SDG)—SDG 6 targets to substantially reduce the number of people suffering from water scarcity by improving water use efficiency by the year 2030⁷. The biggest share of water (around 70%) in Asia is allocated

Department of Land and Water Management, IHE Delft Institute for Water Education, 2611 AX Delft, The Netherlands. ✉email: s.pareeth@un-ihe.org

for irrigation⁸. Hence, to improve water use efficiency it is crucial to monitor the water use from the basin scale where water allocation to different sectors takes place^{9,10} to the field scale where the water saving interventions are often deployed¹¹. The requirement of covering large geographical areas, yet providing very fine resolution information makes the use of in-situ measurement devices unfeasible due to technological limitations as well as the large commitment required in terms of equipment costs and maintenance.

Remotely sensed information from satellites plays a crucial role in hydrology and water resources management, allowing for monitoring and assessment of surface water bodies, land and water use, precipitation patterns, and soil moisture levels over large spatial scales. It is considered as a viable alternative to in-situ data for monitoring land and water use in agriculture¹². Frequent observations of satellite data in visible and infrared spectral region has been developed to be a cost-effective, scalable, and reliable data source for measuring ET_a ¹³. In recent decades, the utilization of remote sensing has significantly expanded in terms of spectral and temporal resolution. This growth can be attributed to the launch of new satellite/sensor missions, the adoption of open data policies by agencies, and the advancement of algorithms used to accurately extract land use and geophysical parameters¹⁴. In agriculture, there is an increasing demand to use remotely sensed information to monitor and assess water use and irrigation performances at different scales from basin to field level⁹. Remote sensing can be used as a valuable source of information for water resource planners to conduct water accounting¹⁵ and for irrigation managers to estimate the spatial distribution of water use and productivity¹⁶. Satellite data offers unique spatial and temporal coverage with sensors like Moderate Resolution Imaging Spectroradiometer (MODIS) (with ~250–1000 m spatial and daily temporal resolution) which is ideal for basin scale studies and multiple Landsat missions (30–100 m spatial and 16 days temporal) suitable for field level studies^{14,17}. It is feasible now to setup operational monitoring systems across various scales from field to the basin and from districts to national and continental levels with daily, weekly, and biweekly temporal coverage using products derived from satellite data^{10,16}.

Many approaches have been developed over the past four decades to estimate ET_a from remotely sensed satellite data¹⁸. One of the most commonly used, albeit a complex approach, is to use physically based Surface Energy Balance (SEB) models which require a combination of meteorological and satellite data as inputs^{19–21}. There are many variants of SEB models including Atmosphere-Land Exchange Inverse (ALEXI)²², Mapping EvapoTranspiration at high Resolution with Internalized Calibration (METRIC)²³, Surface Energy Balance Model for land (SEBAL)¹⁹, operational Simplified Surface Energy Balance (SSEBop)²⁴, Simplified Surface Energy Balance Index (S-SEBI)²⁵ etc, which have been successfully implemented at various scales from basin to global to encompass all scales in which SEB type models have been used. METRIC and SSEBop are successfully implemented in the United States and other countries reporting high accuracies with R^2 reported ranging from 85 to 95% against field observations obtained from Lysimeter²⁶. The SEBAL model used in this study is implemented in various studies that reported higher accuracies ranging from 85 to 96% respectively on daily to seasonal time scales²⁷.

Implementing SEBAL model over a large area is often challenging due to complexity in data requirements, the need for internalized calibration, and geospatial big data processing²⁸. Hence SEBAL applications have been mostly spatially limited to smaller geographical areas, e.g. Irrigation districts or sub-basins, and temporally to processing a handful of images in one growing season or a year²⁹. However, with the development of cloud computing infrastructures there are community driven efforts such as OpenET aiming at increasing access to remote sensing based ET_a estimations irrespective of the models³⁰. SEBAL implementation in Google Earth Engine called geeSEBAL automates the ET_a estimation using Landsat data³¹. The most important input data required from remote sensing is thermal radiation data from satellite sensors. Currently, thermal data acquired from Landsat and MODIS sensors are being used extensively for ET_a estimation at 30 m and 1 km respectively²⁶. Note that the Landsat thermal bands are acquired at 100 m spatial resolution and resampled to 30 m by the data provider using cubic convolution method. However, the data obtained in the optical spectrum is often contaminated due to clouds leaving large gaps which makes it difficult to do monthly/seasonal aggregations. Statistical interpolation over spatial domain while applied to fill larger gaps results in over smoothing. Hence biophysical parameters following annual unimodal or multimodal cyclic patterns have been successfully gap-filled using a combination of temporal and spatial interpolations. Time series reconstruction and gap filling have been successfully implemented on daily NDVI using Fourier transform³², daily lake surface water temperature using the Harmonic ANalysis of Time Series (HANTS)³³, daily Land Surface Temperature (LST) using weighted temporal averaging³⁴ and Locally Weighted Regression (LWR) followed by spline based spatial interpolation³⁵. Data fusion approach Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM)^{36,37} and time series smoothing using Savitzky-Golay filter³⁸, were used to reconstruct daily ET_a from combining Landsat and MODIS spectral datasets. In this study, 30 m gap-filled monthly ET_a maps for two crop years from October 2013 to September 2015 over the large basin LUB in Iran was developed. An open source python based SEBAL implementation called PySEBAL library was used to compute ET_a from all the Landsat 7 and 8 images in the study years. A new approach based on temporal and spatial interpolation was developed to derive gap-filled monthly ET_a maps at high resolution over LUB in Iran.

Study area

Lake Urmia Basin (LUB) is located in the North-West region of Iran with a total surface area of ~52,970 km². Mean annual precipitation in the semi-arid basin is 350 mm and the mean annual temperature is 11.2 °C. The basin is an endorheic basin where no surface water outflow occurs. However, groundwater seepage and evaporation are the primary mechanisms of water loss from the basin. Lake Urmia is a terminal lake that is situated in this basin. All the runoff generated in the basin ends in this terminal lake through a network of rivers.

The past two decades have seen a staggering decline in the surface area of Lake Urmia. Once the world's second-largest hypersaline lake with a surface area of nearly 5000 km², the lake is now reduced to below 2000 km². This phenomenon had significant impact on the livelihood of millions of people living in the basin and causing

far reaching regional ecosystem and health consequences due to issues such as loss of biodiversity and salt storms. The large magnitude of the issues caused by the desiccation of Lake Urmia has triggered the scientific community to study the long-term land and water use in the basin thereby understanding the reasons behind such a phenomenon. Several terminal lakes such as Lake Chad in Africa³⁹, the Aral Sea in central Asia⁴⁰, and Lake Poopo in South America⁴¹, have been affected in the recent past and have witnessed significant shrinkages.

While the main reason behind the decline of the Lake Urmia remains a matter of dispute between scientists^{42,43}, it is plausible that a combination of natural factors e.g. reduced precipitation⁴⁴ and anthropogenic changes e.g. expansion of irrigated agriculture⁴⁵ are the drivers behind the lake shrinkage. Whatever the causes may be, the only way out is to significantly curb water consumption in the basin. The main burden is on agriculture that at the moment claims more than 90% of renewable water resources and reduces the groundwater storage causing a significant drop in the water table in the aquifers⁴⁶. Recognizing this, top-down strategies were drawn up by the government⁴⁷ including targeting a 40% reduction of agricultural water use centered around reducing surface water allocation to agriculture accompanied by investing in a shift to drip irrigation and proposing a cropping pattern change. These interventions have been successful in bringing fragile stability in an otherwise catastrophic situation⁴⁸. An important missing piece is, and has been, a detailed understanding of spatial and temporal variations in consumptive water use, i.e. ET_a , in the basin especially in irrigated areas. Such information is crucial for pinpointing hotspots for water savings and making targeted investments in scaling solutions that are proven useful by means of measurements. Furthermore, given the fragility of the gains, continuous monitoring of the water use in the basin remains pivotal for any prosperous future outlook.

Data

All the Landsat 7 and 8 images acquired between 1 October 2013 and 30 September 2015, were processed to estimate ET_a for the days of acquisition. This includes two tiles from path 167 and 169, and three tiles from path 168 (Fig. 1), and a total of 626 Landsat scenes were processed. Figure 2 shows the Landsat acquisitions in the study period used for applying SEBAL algorithm. The Landsat Collection 1 Level-1 data in Tier 1 (T1) inventory were used for this study, acquired from Google cloud public storage using *gsutil* python library. The Landsat data provided in the T1 inventory are terrain corrected with well characterized radiometry and are inter-calibrated across different Landsat sensors^{49,50}. For topography and elevation, Shuttle Radar Topography Mission (SRTM) data at 30 m spatial resolution obtained from the United States Geological Survey (USGS) repository were used.

Further, meteorological data at the time of satellite data acquisition (instantaneous) and 24 h average of the day of acquisition were required. The meteorological data required for SEBAL implementation are listed in Table 1. These data were extracted for the LUB from NASA Global Land Data Assimilation System (GLDAS v2.1), a global product developed by assimilating data from satellite and ground based observations⁵¹. GLDAS data is offered at 0.25 degree spatial resolution at 3 h interval. For comparison, Actual EvapoTranspiration and Interception (AETI—hereafter “ ET_a ” for consistency) data at Level 1 (250 m resolution) obtained from Food and Agriculture Organisation (FAO) portal to monitor WAtER Productivity through Open-access of Remotely sensed derived data (WaPOR) was used (url: <https://wapor.apps.fao.org>). FAO WaPOR is a remote sensing based spatial database on water productivity related indicators including ET_a which offer open data for African and near East countries. The ET_a data in WaPOR is computed using ETLook model based on the Penman–Monteith equation, which is modified to incorporate remote sensing input data and computes the combined rate of evaporation

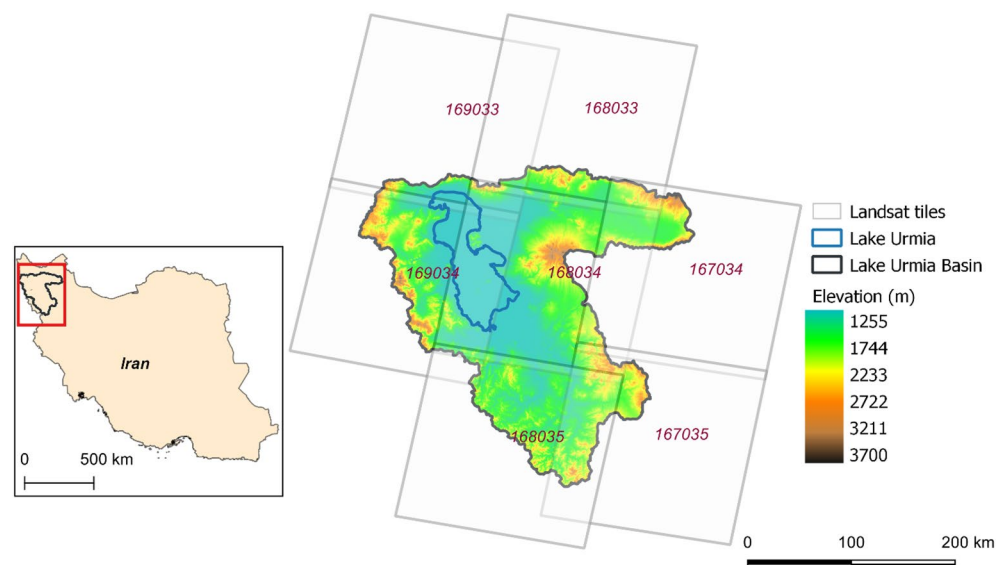


Figure 1. Study area—LUB boundary and Landsat tiles used for this study. The red box in the inset map shows the location of basin in Iran. Map created in QGIS 3.28.8 LTR software (<https://www.qgis.org/>).

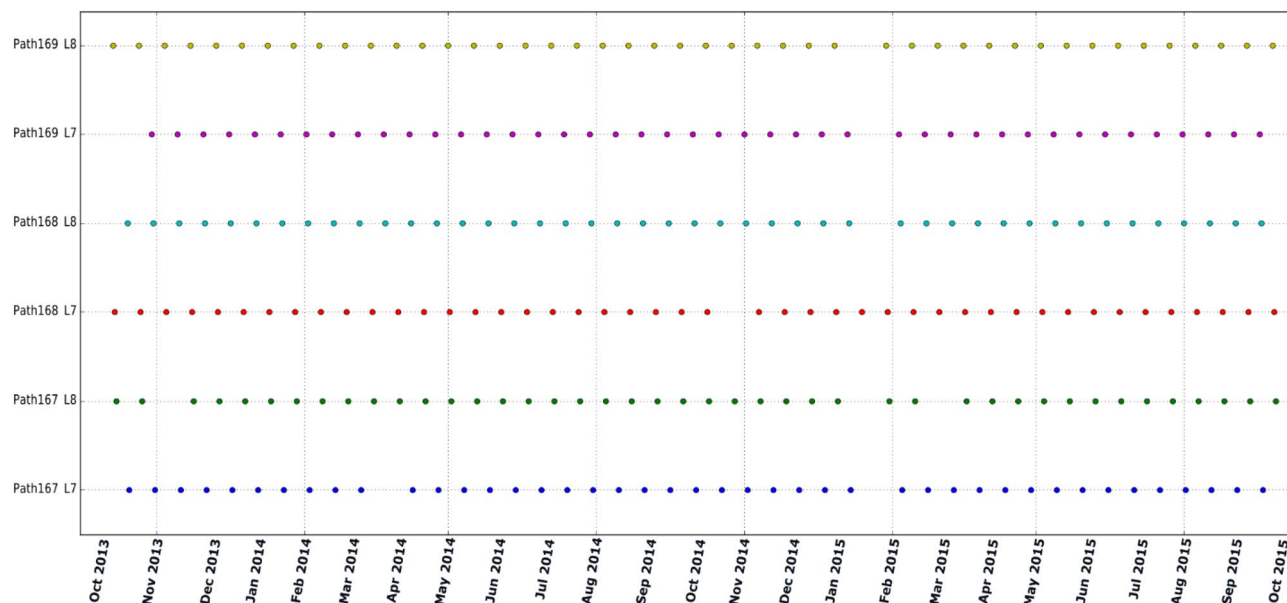


Figure 2. Available Landsat 7 and 8 over the study period from October 2013 to September 2015 separated by different paths covering LUB.

Parameter	Symbols	Unit
Downward shortwave radiation	SWdown	W/m ²
Wind speed	Ws	m/s
Air temperature	Tair	°C
Pressure	P	Mb
Relative humidity	Rh	%

Table 1. Input meteorological data required for PySEBAL model.

and transpiration⁵². It provides the highest available spatial resolution of 250 m for an operational open-access ET_a product at the continental scale⁹.

Methods

In this section, the implemented methodology grouped into (i) pre-processing of data, (ii) PySEBAL theory and (iii) the gapfilling approach are explained in detail.

Pre-processing satellite and meteorological data. The acquired Landsat 7 and 8 data were pre-processed to compute cloud masked Top Of Atmosphere (TOA) reflectance bands at 30 m spatial resolution. The pre-processing included conversion from Digital Number (DN) to TOA reflectance, cloud removal using the Quality Assessment (QA) band and mosaicking the images acquired on the same day over a single path. The QA band contains 16-bit integers representing certain atmospheric conditions. All the bit combinations showing medium or high cloud confidence were used to create cloud masks. The pre-processing of meteorological data included the following steps (i) extracting the required variables as listed in Table 1 from the three hourly data, (ii) clipping the global data to the extent of LUB, (iii) converting the units of air temperature from Kelvin to °C, pressure from Pascal (Pa) to Millibar (mb), (iv) specific humidity in kg/kg to relative humidity in % and (v) extracting instantaneous and daily average meteorological variables. The instantaneous data corresponding to the Landsat acquisition time (7:30 A.M. UTC) was estimated by averaging the 6H and 9H UTC data, while all the 8 three hourly data in a day were averaged to estimate daily averages.

To efficiently process the Landsat data for a large basin such as LUB entire processing chain was implemented in a High Performance Computing (HPC) infrastructure. The entire processing chain was implemented by using multiple open source libraries, to deploy multi-core jobs like processing satellite data and to implement SEBAL model. All the spatial and temporal processing were performed in GRASS GIS 7.4.0 software⁵³ which is open source under GNU General Public License (GPL).

The Landsat data processing for each path (167, 168 and 169) were performed in different nodes in parallel. This approach resulted in substantial reduction of the processing time. The data processing was divided into three Processing Units (PU) based on the Landsat path and row. Since the adjacent tiles in the same path have same acquisition date, we chose each path as a processing unit. Thus, data for each path were processed in parallel

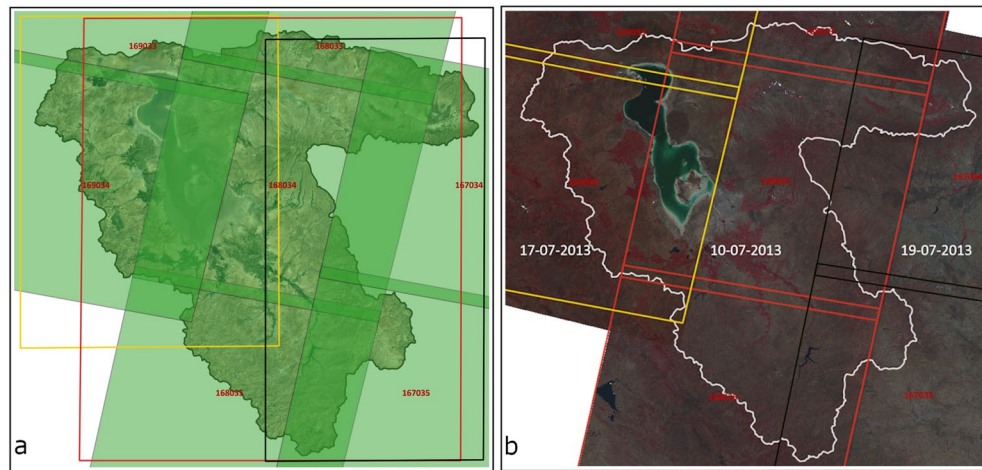


Figure 3. (a) Processing Units (PU) covering LUB; black rectangle—PU1 covering path 167; red rectangle—PU2 covering path 168; yellow rectangle—PU3 covering path 169, background satellite image source: Google Satellite API. (b) Mosaicked false colour composite from three paths acquired on three different dates. black rectangle—path 167; red rectangle—path 168; yellow rectangle—path 169, background satellite image source: Landsat 8. Maps created in QGIS 3.28.8 LTR software (<https://www.qgis.org/>).

in the HPC. Figure 3 shows the bounding box of the three defined PUs for LUB (see Fig. 1 for complete view of Landsat tiles covering LUB). The PU's are designed to incorporate minimum overlapping “rectangular” area between the respective Landsat path and the LUB. While processing in each PU, only those pixels inside the LUB boundary were considered further optimizing the efficiency in computing over larger area.

SEBAL implementation in Python—PySEBAL. SEBAL work on the basic principle of solving the energy balance model as shown in Eq. (1) where latent heat flux (LE), the energy consumed through evapotranspiration, is computed as residual energy from net radiation (Rn), ground heat flux (G) and sensible heat flux (H).

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

Instantaneous ET_a (mm/day) is then calculated using Eq. (2) where LE is divided by the latent heat of vaporization (L) and the water density (ρ_w).

$$ET_a = \frac{LE}{L * \rho_w} \quad (2)$$

PySEBAL is a python library to implement SEBAL using spectral reflectance values from satellite data, climatic parameters and topography as input to estimate the surface energy balance components. The outputs include parameters related to vegetation, energy balance, biomass, evapotranspiration, and water productivity. Currently, PySEBAL supports data from MODIS, Landsat, and Proba-V satellite sensors which facilitate the production of daily and seasonal ET_a maps. PySEBAL have been used for estimating water use in agriculture by^{54–56}.

Figure 4 shows the workflow of computing seasonal ET_a using PySEBAL. The steps 1 and 2 in the workflow diagram cover the pre-processing of Landsat data and cloud masking as explained in the previous section. In step 3, each spectral band of Landsat are patched/mosaicked to respective PU's. Step 4 is to create intermediate input layers that are required for closing energy balance equation which includes Normalized Difference Vegetation Index (NDVI), Soil-Adjusted Vegetation Index (SAVI), Leaf Area Index (LAI) and Surface Albedo (α) from Landsat spectral data. Step 5 computes the energy balance components Rn, G, and H as shown in Eq. (1).

The equations used behind computing the major energy balance components in SEBAL is well explained in^{19,20,57}. The computation of these components in PySEBAL is demonstrated in^{54,55}. The net radiation flux (Rn) represents the actual radiant energy available at the surface. Rn (W/m^2) is the difference between all incoming radiant fluxes and outgoing radiant fluxes using Eq. (3)²³.

$$R_n = [R_{S\downarrow} + R_{L\downarrow}] - [\alpha R_{S\downarrow} + R_{L\uparrow} + (1 - \epsilon_0)R_{L\downarrow}] \quad (3)$$

where $R_{S\downarrow}$ is the incoming solar radiation in W/m^2 , $R_{L\downarrow}$ and $R_{L\uparrow}$ is the incoming and outgoing longwave radiation respectively in W/m^2 , α is surface albedo (dimensionless) and ϵ_0 is the thermal emissivity in W/m^2 .

Ground heat flux (G) is the rate of heat storage into the soil and vegetation due to conduction. In PySEBAL, G in W/m^2 is computed using Eq. (4) as function of surface temperature (T_s), surface albedo (α), and NDVI¹⁹ after applying a water mask.

$$G = T_s(0.0038 + 0.0074\alpha)(1 - 0.98NDVI^4) \quad (4)$$

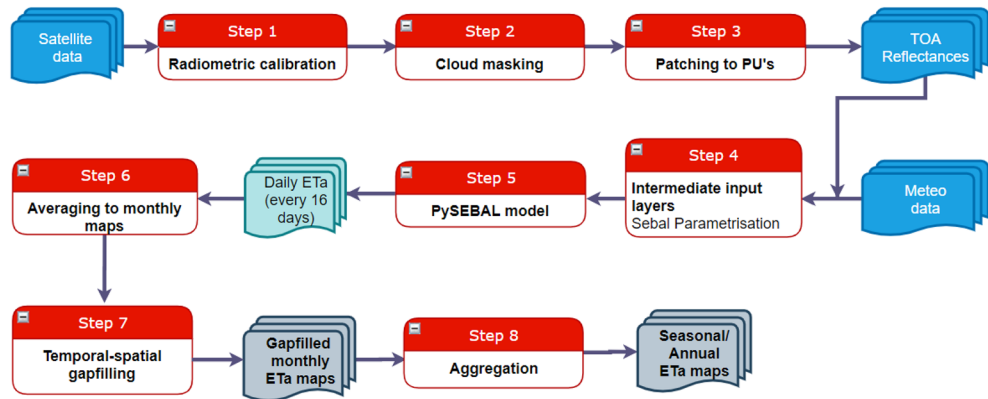


Figure 4. Work flow diagram showing all the steps included in the method including input data, PySEBAL model and gap filling process.

Sensible heat flux (H) is the rate of heat loss to the air by convection and conduction, due to a temperature difference and it is computed using the Eq. (5).

$$H = \frac{\rho_{air} C_p dT}{r_{ah}} \quad (5)$$

where ρ_{air} is air density, C_p is the specific heat of air at constant pressure, and dT is near-surface air temperature difference for each pixel and r_{ah} is the aerodynamic resistance. As the actual absolute values for air temperatures above each pixel are unknown, dT is computed by assuming a linear relationship between dT and altitude corrected surface temperature. To define the slope and offset of this linear model, SEBAL uses the two “anchor” pixels where a value for H can be reliably estimated. For each satellite image, these “anchor” pixels are selected by picking hot and cold pixels over the driest surface in the satellite scene and wettest surface such as water body or irrigated area respectively. The PySEBAL library facilitates automation of the entire process including the estimation of the hot and cold pixels from a satellite image. The cold pixels are usually selected from water bodies or crop areas with well-developed vegetation, whereas hot pixels are selected from completely dry soil surfaces. In PySEBAL, cold pixels are automatically identified based on thresholds (2nd and 5th percentiles) applied to distribution of surface temperature estimated from the Landsat thermal data, while hot pixels are identified from the distribution of NDVI, where 1st and 3rd percentiles are used as thresholds. The instantaneous latent heat flux (LE) and Evaporative Fraction (EF) at the time of satellite data acquisition is then computed from the energy balance components. The instantaneous EF is converted to the daily evaporative fraction (EF_{24}) by incorporating an advection factor, which takes into consideration the vapor pressure deficit and accounts for the rise in ET during the afternoon period⁵⁵. The daily ET_a is then computed by multiplying EF_{24} with 24-h net radiation where negligible ground heat flux over 24 h is avoided. The PySEBAL is capable of processing multiple satellite scenes in a single run. The optimal way is to set up and run PySEBAL for an entire season. Steps 6 to 8 in Fig. 4 represents the gap filling procedure which is explained in detail in the next section.

Gap filling of ET_a maps. For every month, Landsat has 4 observations, two from Landsat 7 and two from Landsat 8, 16 days being re-visit time of both satellites. It is possible to have up to 4 observations per month per pixel, but often this is not the case due to cloud cover in the region especially during the winter season. All available scenes from Landsat 7 and 8 are incorporated in the processing, in order to increase the probability of having maximum valid pixels over a year. First step in the gap-filling procedure (step 6 in Fig. 4) is to patch all the ET_a maps by averaging them per month. The monthly ET_a maps were then converted from mm/day to mm/month by multiplying each map by number of days in the respective month. The remaining gaps were then filled using a temporal and spatial interpolation.

A temporal interpolation based on LWR was applied to the monthly ET_a maps to reconstruct missing values and identify outliers^{35,58}. For each time series observation (pixel) in the map, a polynomial model of second order is computed using a set of neighboring pixels in the temporal dimension. Distance based weight is applied to the values in such a way that the observation farther away in time gets lower weights. All the observations in the time series were interpolated, as long as there were enough non-null observations. To keep the interpolated values within the seasonal limits, a maximum gap of 3 observations in the time series were interpolated, otherwise retained as NULL. The weight function used for LWR was Tricube which determines the influence of neighboring values in time to the current observation. High and low outliers in model fitting were ignored and extrapolation was avoided. The LWR add-on in GRASS GIS 7.4.0 (*r.series.lwr*) was used to implement temporal interpolation.

Due to the insufficient valid observations in the time series meeting the LWR conditions there were remaining gaps in the monthly ET_a maps (less than 10% of the surface area). These gaps were then reconstructed using bicubic spline spatial interpolation. This step is applied only to the NULL pixels using the neighboring valid pixels, which means the observations and the temporally interpolated values were kept unchanged. Bicubic spline

interpolation is a 2-dimensional approach to the linear spline. In this case, a minimum of neighboring 16 valid pixels are used to interpolate the null pixels using a cubic function. For each pixel, the interpolation takes into consideration the function itself, the gradients determined by one-dimensional splines and the cross derivatives. The values of the function and the derivatives are reproduced exactly at the pixels and they change continuously with the moving window crossing one pixel to another. The bicubic approach ensures the continuation of derivatives to the adjacent grids thereby reducing the artefacts. The interpolation module in GRASS GIS 7.4.0 (*r.fillnulls*) was used to implement spatial interpolation in each monthly map.

After the gap-filling process applied to all the monthly maps from October 2013 to September 2015, the monthly maps were aggregated to create annual maps (mm/year) for the years 2013/2014 and 2014/2015. To assess the performance of this gap-filling approach, the gap-filled ET_a maps were compared with data obtained from FAO WaPOR over the LUB. The data from WaPOR were resampled from 250 to 30 m using nearest neighbor algorithm to match the gap-filled ET_a maps. Further to check the spatial and temporal consistency of the gap-filled maps, ET_a dynamics over different agricultural land use types and its response to bio-physical parameters like NDVI were investigated. The monthly ET_a estimates were compared to aggregated monthly precipitation obtained from rainfall stations in the LUB to understand the influence of irrigation in driving the water use in the basin.

Results

Gap filled monthly ET_a maps (24 maps) at 30 m resolution for LUB were developed and analyzed. To demonstrate the different steps of gap-filling process, results obtained after each step is explained over the month of October 2014 as an example. Table 2 lists the available Landsat data for each PU (path) for October 2014 (see Fig. 2 for availability of Landsat 7 and 8 over the study period). There were total 10 Landsat scenes available in October 2014 (Four Landsat 7 and Six Landsat 8).

Figure 5 shows the individual 10 ET_a maps computed using PySEBAL from Landsat acquisitions of October 2014 for the three PU's. These 10 maps were then aggregated by averaging and clipped to LUB (see the last map of Fig. 5).

Figure 6 shows the gap filled ET_a map of October 2014 (mm/month) before and after the gap filling. Similarly, for each month from October 2013 to September 2015, there were a maximum of 12 observations per month for the basin and computed gap-filled monthly ET_a maps at 30 m spatial resolution. The variation of gaps in the ET_a maps over LUB after each step in the gap filling procedure was analyzed to understand the pattern of reconstructing the maps. The results show that, by combining Landsat 7 and 8 data substantially reduced the cloud coverage when aggregated to monthly maps. After monthly aggregation by averaging of all the individual ET_a maps the maximum cloud cover was 70% over entire LUB in January 2014 followed by 42% cloud cover in January 2015. After LWR interpolation majority of the gaps in the winter months were statistically filled to 90%, while all the summer months were completely gap-filled. Figure 7 shows the variation in gaps due to cloud cover over 24 months before and after the gap filling steps.

Figure 8 shows the final gap-filled annual ET_a maps for two crop years over LUB. The annual average ET_a estimated from the gap-filled monthly maps was 302 mm and 346 mm in the crop years of 2013/2014 and 2014/2015 respectively. The annual rainfall reported was 306 mm and 368 mm in the crop years 2013/2014 and 2014/2015 respectively. Correlation analysis between monthly gap filled ET_a and FAO WaPOR reported a coefficient of determination (R^2) of 0.93, Root Mean Square Deviation (RMSD) of 9 mm/month, Mean Absolute Deviation (MAD) of 7 mm/month and Mean Absolute Percentage Deviation (MAPD) of 36% over agriculture area. The land use map used to select agriculture area for correlation analysis is shown in Fig. 9.

Comparison statistics between gap filled ET_a and FAO WaPOR data over different agriculture land use types—irrigated wheat and barley, irrigated summer crops, orchards and rainfed are listed in Table 3. The scatter plot between gap filled ET_a and FAO WaPOR data is shown in Fig. 10 fitting the linear models over four agriculture land use types. Correlation of these two datasets over orchards was reported to be the highest with R^2 —0.97. RMSD and MAD over orchards were 6 and 4 mm/month respectively. For irrigated wheat and barley and irrigated summer cropped area the R^2 reported were 0.91 and 0.95 respectively showing very high correlation between the datasets. While the RMSD reported over these land use types were 7 and 8 mm/month, estimated MAD was 5 mm/month. Correlation over rainfed area was reported to be lowest with R^2 —0.84, RMSD—7 mm/month and MAD—5 mm/month. MAPD ranged from 30 to 46% for irrigated wheat and barley and rainfed area respectively. Table 3 also lists the comparison statistics between ET_a before gap filling and WaPOR data. The results show that gap filling didn't improve the ET values over irrigated wheat and Barley (winter season) and rainfed, while it improved substantially over the summer irrigated crops and orchards. In case of winter irrigated crops and rainfed, the RMSD increased from 4 to 6 mm/month. For orchards and irrigated summer crops the deviation

	Landsat 7	Landsat 8
PU1 (path 167)	02-10-2014	10-10-2014
	18-10-2014	26-10-2014
PU2 (path 168)	09-10-2014	01-10-2014
		17-10-2014
PU3 (path 169)	16-10-2014	08-10-2014
		24-10-2014

Table 2. List of Landsat 7 and 8 data available and processed for the month of October 2014.

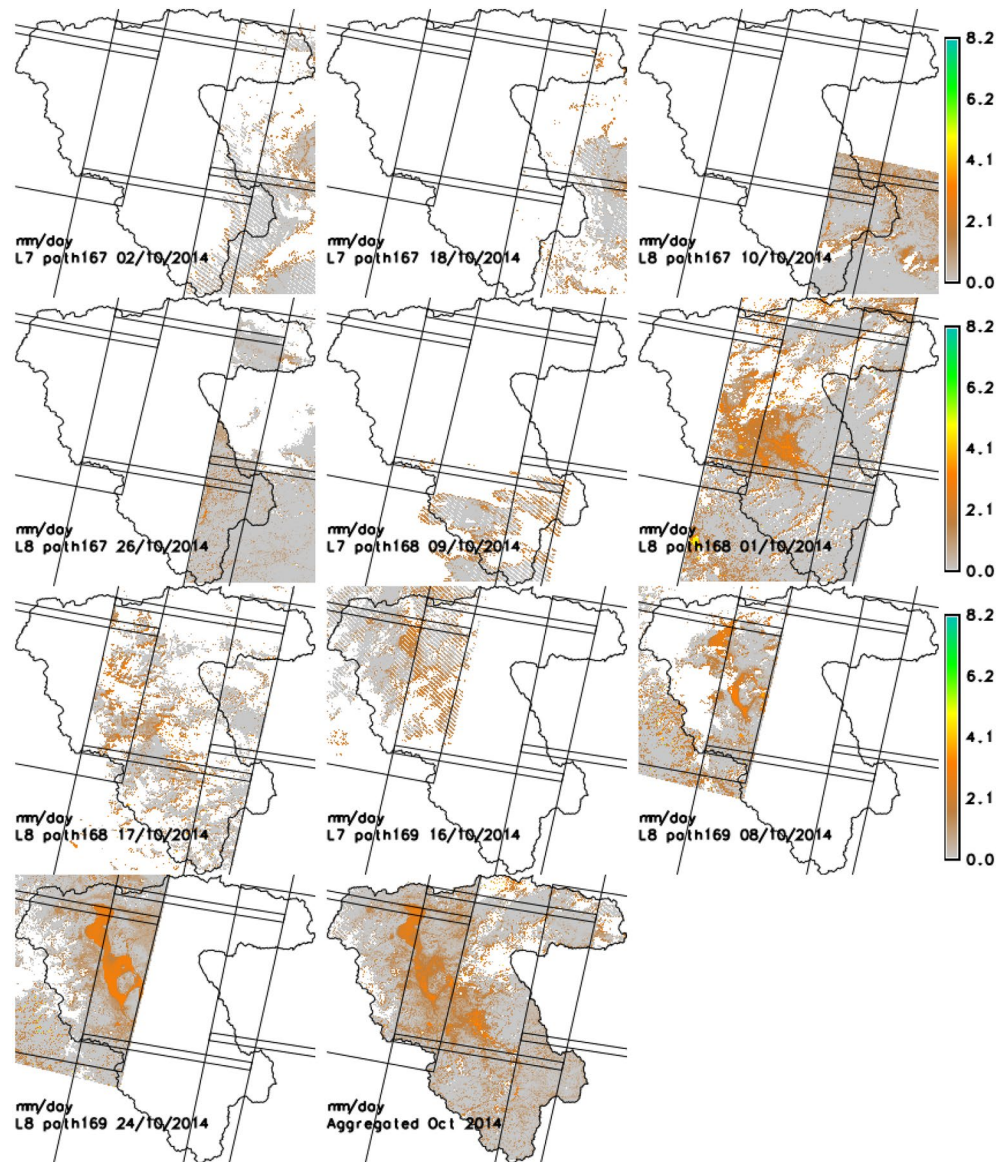


Figure 5. Individual ET_a maps from October 2014; the last block shows the aggregated average ET_a map for the same month. Unit is mm/day. Map created in GRASS GIS 7.8.4 software (<https://grass.osgeo.org/>).

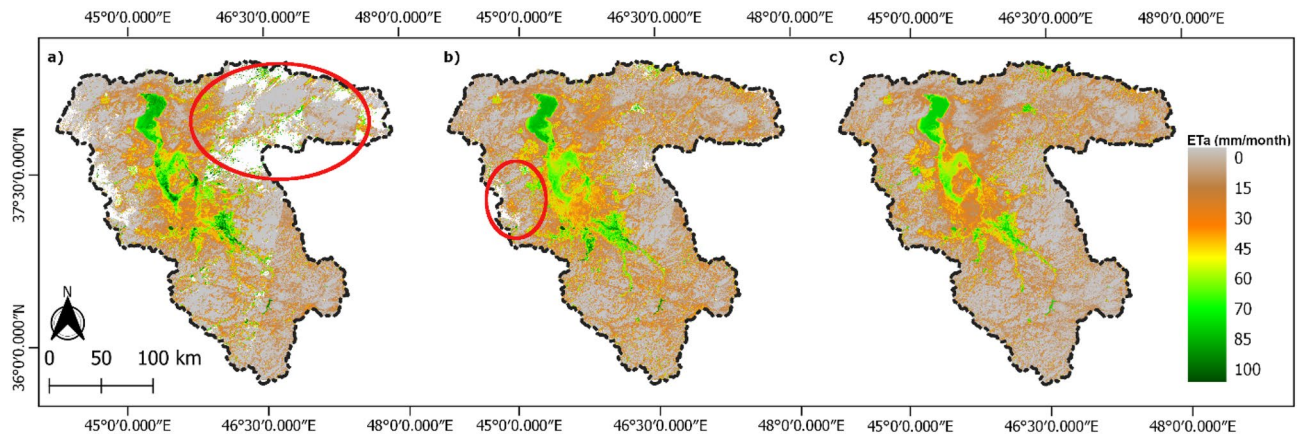


Figure 6. Monthly ET_a map for the month October 2014 (a) after aggregation of individual ET_a maps, (b) after LWR interpolation, (c) gap-filled ET_a monthly map after spatial interpolation; red circles shows the gaps due to clouds. Maps created in QGIS 3.28.8 LTR software (<https://www.qgis.org/>).

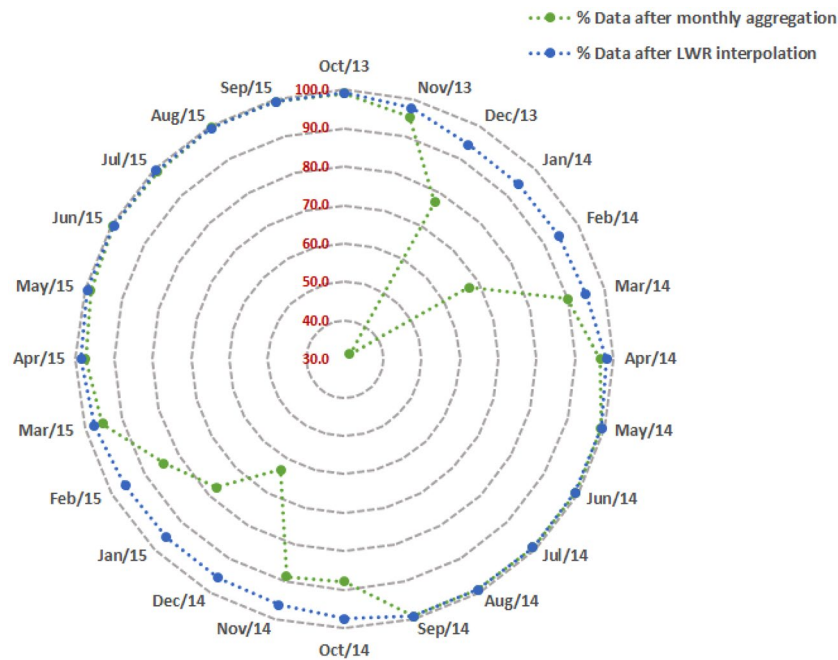


Figure 7. Percentage of valid monthly observations after monthly mean aggregation and LWR interpolation.

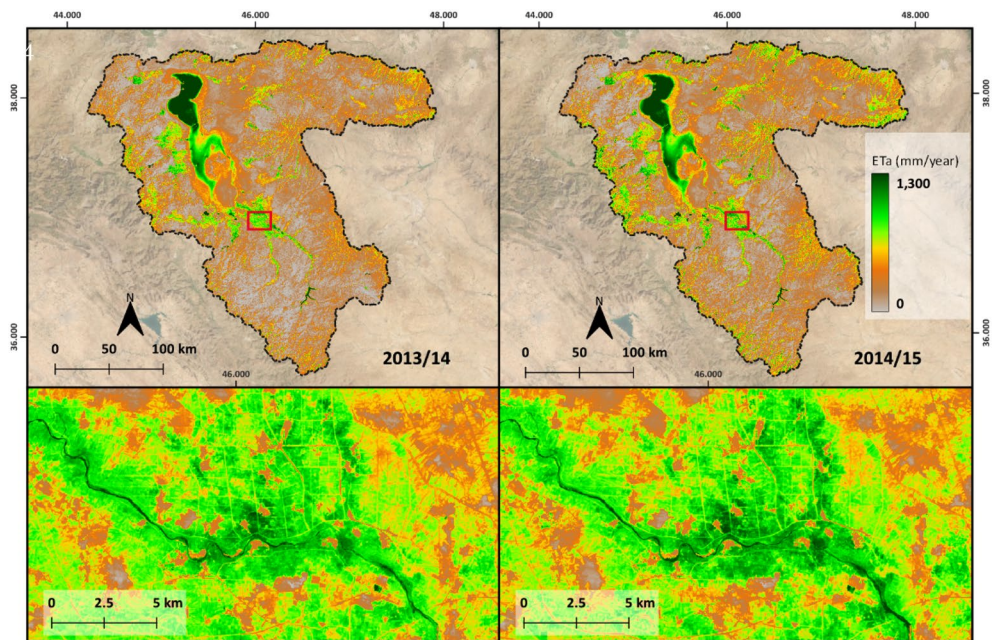


Figure 8. Top: Gap-filled annual ET_a maps of LUB for the years 2013/14 and 2014/15, background satellite image source: Sentinel-2 cloudless 2016 by EOX IT Services GmbH; Bottom: zoomed into an irrigated area in Miandoab irrigation scheme, this area is indicated in red box above. Map created in QGIS 3.28.8 LTR software (<https://www.qgis.org/>).

between gapfilled ET_a and WaPOR reduced substantially. RMSD improved from 25 to 8 mm/month for irrigated summer crops, while it improved from 19 to 6 over orchards. Further the response of gap-filled ET_a to NDVI showed a linear correlation with R^2 of 0.88 and an exponential fit with R^2 0.93 over irrigated wheat and barley, irrigated summer crops and orchards. Figure 11 shows the scatter plot between gap-filled ET_a and NDVI with fitted regression models.

The monthly ET_a and Precipitation (P) were compared to analyze the underlying drivers of the seasonal trend of water use over different land use types. Figure 12 shows the temporal variation of monthly ET_a (gapfilled and

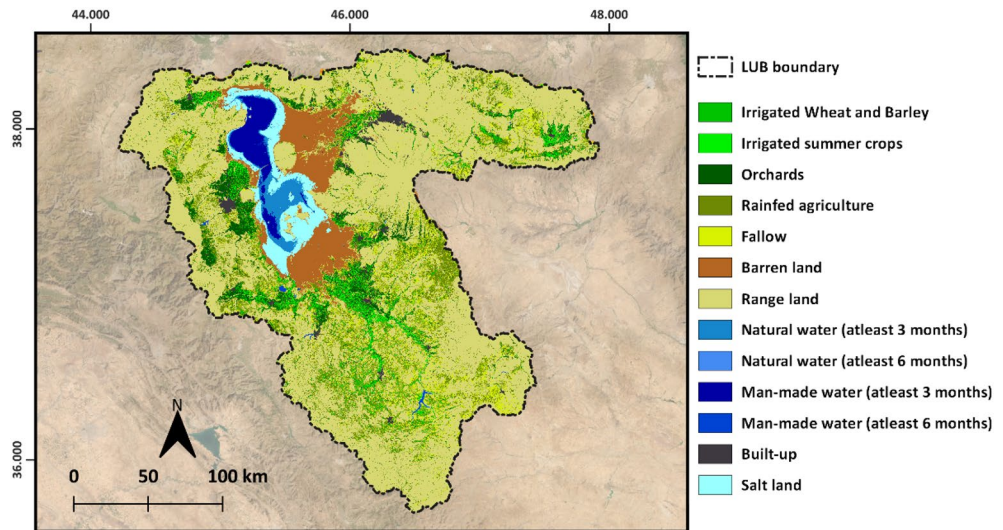


Figure 9. Land use map of Lake Urmia Basin (source: Tarbiat Modares University, Iran). Background satellite image source: Sentinel-2 cloudless 2016 by EOX IT Services GmbH. Map created in QGIS 3.28.8 LTR software (<https://www.qgis.org/>).

	R ²		RMSD (mm/month)		MAD (mm/month)		MAPD (%)	
	Before gapfilling	After gapfilling	Before gapfilling	After gapfilling	Before gapfilling	After gapfilling	Before gapfilling	After gapfilling
Irrigated Wheat and Barley	0.97	0.92	4	7	2	5	13	30
Irrigated Summer crops	0.72	0.96	25	8	18	5	82	35
Orchards	0.8	0.97	19	6	14	4	76	34
Rainfed	0.8	0.8	4	7	3	5	39	46

Table 3. Comparison statistics reported comparing monthly before and after gap filled ETa and FAO WaPOR data over two study years.

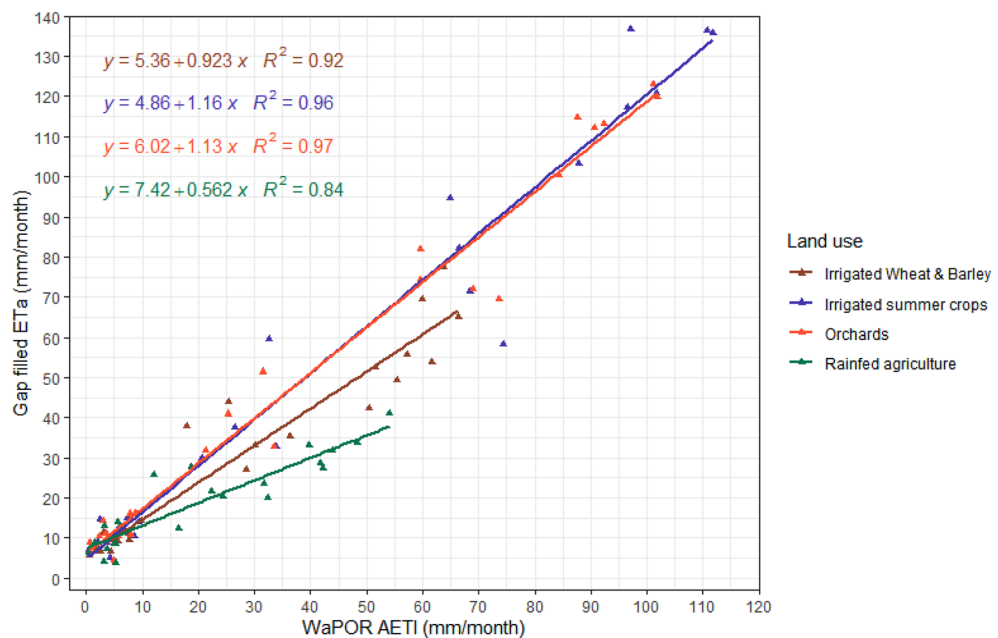


Figure 10. Scatter plot between monthly gap filled ETa and FAO WaPOR AETI with best fitting linear regression lines over four agricultural land use types.

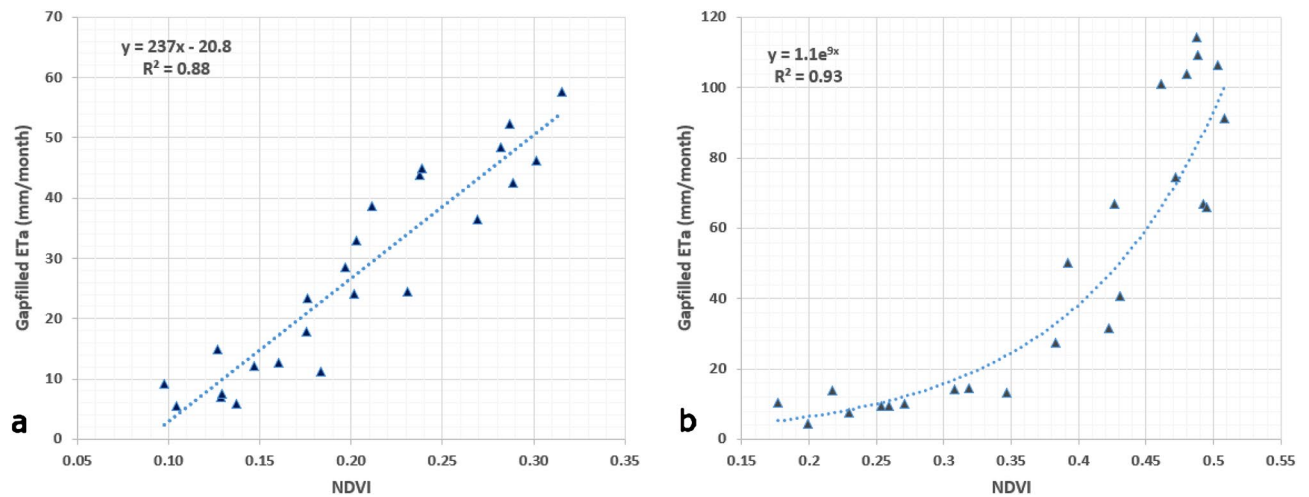


Figure 11. Scatterplot between monthly NDVI and ET_a (a) over the entire basin except water bodies; (b) over the irrigated wheat and barley, irrigated summer crops and orchards.

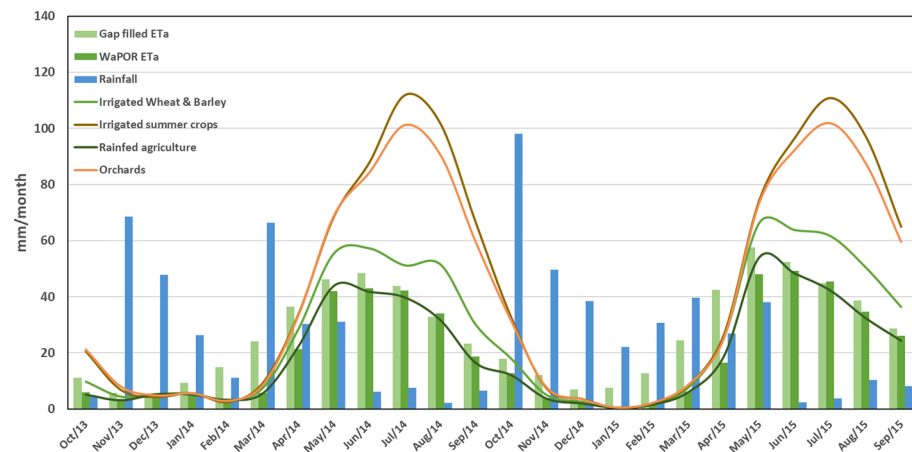


Figure 12. Monthly P and ET_a variations in LUB.

WaPOR) and P over the LUB and also the monthly gapfilled ET_a over the agricultural land use types for the two crop years from October 2013 to September 2015. The rainfall during the second cropping year (mean rainfall of 368 mm) was greater than the previous year (mean rainfall of 306 mm). It also reflects in the estimated ET_a with a higher average reported in the second year (346 mm) than first year (302 mm). There were extremely high rainfall events reported in the month of October 2014 (~95 mm) which also triggered the higher ET_a range in the second cropping year.

The land use based ET_a and P estimates over the study period are illustrated in Fig. 13. The higher ET_a values reported during the summer months from April to September in both cropping years were contributed by the irrigation events and evaporation from water bodies. To understand further the land use based dynamics of ET_a and P, mean ET_a based on land use types were extracted. Over the two crop years highest ET_a (except water bodies) of 800 mm was reported for the irrigated summer crops land use type followed by orchards with mean ET_a of 751 mm. The lowest average ET_a of 199 mm was reported for the barren land. Agriculture related land use types were reported to have highest ET_a over the study period as expected.

Discussion

Quantifying water use by agriculture land use types is of utmost importance to take actionable measures to avoid overuse of resources. The study demonstrates the use of publicly available earth observation satellite data acquired by Landsat 7 and 8 to monitor water use over a large basin at different temporal scales. The approach implemented over a large basin covered by 7 Landsat tiles could efficiently reconstruct the gaps in data due to clouds and stripes on Landsat 7 due to the scan line corrector failure. The parallelization of the process with the implementation of processing units covering tiles of same path substantially reduced the processing time.

The inclusion of data from both Landsat 7 and 8 increased the availability of observations in space-time domain thus reducing the gaps after step 6 of monthly aggregation in methodology (Fig. 4). This is reflected

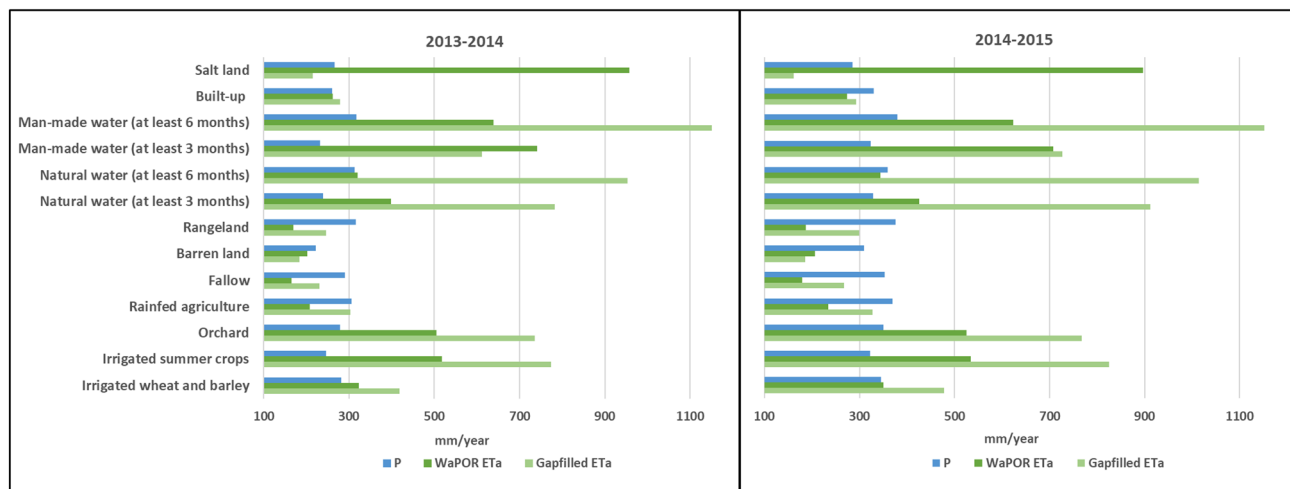


Figure 13. Annual P and ET_a over different land use classes for two crop years left—2013/14 and right—2014/15.

in the cloud dynamics after monthly aggregation (Fig. 7) where most of the months other than January have good observations for more than 60% of the basin. The possibility to use both Landsat 7 and 8 in PySEBAL and the inter-calibration of reflectance values by the Landsat data provider makes it possible to combine data from Landsat 7 and 8 (USGS, 2019).

The conversion of monthly ET_a maps in mm/day to mm/month by multiplying number of days in respective month assumes that the ET_a values are constant throughout that month. The ET_a can vary significantly within a day or a month depending on the land cover type. This might influence the accuracy of ET_a measured using sparse observations from Landsat. The LWR is able to capture the short-term fluctuations in the time series compared to approaches like harmonic analysis³⁵. This is useful while planning interventions based on short term water use of crops within a season. LWR in this case could fill the majority of the gaps even after applying stricter constraints especially in the number of samples per regression (limited to 3), thus avoiding overfitting of the predictions. After the time series smoothing using LWR, the gaps were reduced to less than 10% in all the months. For a diverse terrain like LUB, the time series reconstruction at high resolution allows the capturing of ET_a temporal dynamics for different land use types as demonstrated in this study.

The variation of monthly ET_a over agriculture land use types—irrigated wheat and barley, irrigated summer crops, orchards and rainfed showed distinctive seasonal variations in response to the rainfall and cropping patterns. For the irrigated wheat and barley, the ET_a is highest in April and declines gradually by late June. This match with the cropping calendar for winter wheat and barley in the region. For orchards, ET_a showed peak values during summer months (May to Aug) and lower values were observed from September onwards and in the winter season. As expected, the water consumption by orchards were much higher than irrigated wheat and barley and lower than Irrigated summer crops⁵⁹. This could be attributed to inclusion of young orchards in this land use class and also the fact that orchards besides irrigation schemes are present in the valleys in the vicinity of surface streams throughout the basin where irrigation is not as intensive as in the schemes.

The non-irrigated land use types showed much lower ET_a compared to the irrigated classes. The peak ET_a was observed in months of May at an average value of about 2 mm/day for rainfed crops and 1.5 mm/day for rangelands. Similar study reported an average ET_a of 736 ± 42 mm for over the irrigated agriculture in the basin in 2014, the average ET_a for irrigated agriculture in our assessment was 696 mm⁵⁹. It demonstrates the usability of such maps in comparison with rainfall to give us an overall account on critical water availability versus water consumption over managed landscapes and water deficit. The linear correlation between ET_a and NDVI over the entire basin as shown in Fig. 11a is influenced by the rangeland pixels which forms major land cover in the basin (64% of total area). While the scatterplot as shown in Fig. 11b over the irrigated area which is a managed land cover type shows an exponential relation representing the yield response to water applied for the irrigated crops.

The comparative analysis between newly developed monthly gap-filled ET_a and AETI from FAO WaPOR dataset reported R^2 greater than 0.9 for irrigated wheat and barley and summer crops, while a R^2 of 0.84 was reported in rainfed areas. The FAO WaPOR level 1 AETI data used in this study was developed using 8-day MODIS products at 250 m and the difference in spatial resolution could be a contributing factor for higher RMSD of 7–9 mm/month and MAPD greater than 30%. The correlation was worsened after gapfilling over the irrigated wheat and barley potentially due to multiple factors like higher percentage of gaps due to cloud cover during the winter months (see Fig. 7), undetected cloud and cloud shadow pixels, coarse spatial resolution and the difference in interpolation techniques used in WaPOR database. WaPOR level 1 data is reported to be underestimating the ET_a values due to the coarse resolution of input land surface temperature data (1 km) from MODIS sensor which is used to derive moisture stress and thus affecting the spatial variation⁶⁰. A recent study evaluating the consistency between different levels of WaPOR data found higher correlation between level 1 AETI and the field observations over Zankalon irrigated area in Egypt⁶¹.

Remote sensing based ET_a mapping in many studies is performed by establishing direct empirical relation between NDVI and crop coefficient (Kc) instead of solving surface energy balance equation^{62,63}. Hence, investigating relation between aggregated ET_a values derived from PySEBAL and the NDVI values directly acquired from satellites provide insights into the quality of derived ET_a . It is expected that ET_a and NDVI show a general correlation which could be stronger in land use classes that typically have higher vegetation growth such as orchards and summer crops. The linear relationship between ET_a and NDVI over all the land use types demonstrates vegetation growth with increasing availability of water. The vegetation indices from multi-spectral satellite data can well describe the vegetation growth and thereby can explain the water use trend. Hence the expected response of ET_a to the vegetation growth as indicated by the time series of NDVI would give us quality assurance of the developed monthly ET_a maps.

For irrigated landscapes such as orchards and irrigated summer crops, seasonal prevalence of evaporation and transpiration play key role in formulating the relationship between ET_a and NDVI. The models revealed that the vegetation growth in these land use types are not limited to availability of water. During the winter months and beginning of the growing season, a rapid response of vegetation growth with lower ET_a values with prevalent contribution from evaporation was observed. After the vegetation growth is matured (corresponding to NDVI value of 0.43 and ET_a of 60 mm—see Fig. 11), the response becomes slow which requires much more water to attain rest of the growth which are provided by irrigation.

The newly developed gap-filling approach can be used to monitor the water use estimation by different land use types in a larger basin like LUB. In LUB, periodic monitoring of land and water use will support the water management interventions to be taken in order to revive the Lake Urmia. This data can be further used for assessing water productivity, irrigation performance and computing water accounts at different scales from field to basin¹². Validating the remote sensing based ET_a data remains a challenge due to lack of wide network of in-situ flux data. Hence, the alternative is to perform inter-comparison with similar products and analyze the spatio-temporal trends of ET_a in different land use types. There are multiple projects being carried out to improve this and to develop robust protocols to assess remote sensing based ET_a products (see WaterPIP project; url: <http://waterpip.un-ihe.org/>).

Conclusion

The increasing availability of open access satellite data and new advances in remote sensing techniques are paving the way to systems which can monitor water use by different stakeholders near real time at various spatial scales. However, to implement an operational monitoring system based on earth observation data there needs to be established approaches with robust protocols to extract information at required spatio-temporal scale. In this study, a new approach was implemented to extract annual ET_a at high spatial resolution of 30 m over a large basin—LUB in Iran. The established approach demonstrates how to compute ET_a using a surface energy balance model over a large area covering multiple Landsat tiles of different acquisition dates and introduced a novel gap-filling approach to fill the gaps due to clouds followed by aggregate to monthly and annual maps. The monthly maps thus developed for two crop years 2013/2014 and 2014/2015 were compared with AETI data from FAO WaPOR over agriculture land use types reporting R^2 greater than 0.9.

To logically fill the gaps in satellite derived ET_a maps due to clouds and stripes on Landsat 7 due to the scan line corrector failure, a combined approach of temporal interpolation followed by a spatial interpolation is recommended. The approach should also work for gapfilling other bio-physical parameters like surface temperature, vegetation indices etc., which follows a cyclic change pattern over seasons. The LWR approach to fill the gaps over time captures the temporal dynamics of ET_a over different agricultural land use types. This is validated by comparing with AETI data from FAO WaPOR. The approach can be extended to any other geographical area with Landsat coverage, but recommend to perform a validation analysis as demonstrated in this study before using the derived information for interventions. The open source code and documentation developed to implement this approach further facilitate the uptake by the community.

Data availability

The PySEBAL and the gap filling procedure developed in this study are available as an open “github” repository (url: https://github.com/wateraccounting/PySEBAL_dev) with a detailed documentation provided in “Read the Docs” technical documentation repository (url: <https://pysebal.readthedocs.io/en/latest/>). All datasets are disseminated as Open Geospatial Consortium (OGC) services through the project web mapping link: <http://www.wa-urmia.org/>.

Received: 8 May 2023; Accepted: 11 July 2023

Published online: 25 July 2023

References

- Gosling, S. N. & Arnell, N. W. A global assessment of the impact of climate change on water scarcity. *Clim. Change* **134**, 371–385 (2016).
- Rijsberman, F. R. Water scarcity: Fact or fiction?. *Agric. Water Manag.* **80**, 5–22 (2006).
- Vörösmarty, C. J., Green, P., Salisbury, J. & Lammers, R. B. Global water resources: Vulnerability from climate change and population growth. *Science* **289**, 284–288 (2000).
- Khaki, M. & Hoteit, I. Monitoring water storage decline over the Middle East. *J. Hydrol.* **603**, 127166 (2021).
- Döll, P., Müller Schmied, H., Schuh, C., Portmann, F. T. & Eicker, A. Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. *Water Resour. Res.* **50**, 5698–5720 (2014).
- Seckler, D., Barker, R. & Amarasinghe, U. Water scarcity in the twenty-first century. *Int. J. Water Resour. Dev.* **15**, 29–42 (1999).
- Bhaduri, A. *et al.* Achieving sustainable development goals from a water perspective. *Front. Environ. Sci.* **4**, 64 (2016).

8. Ajaz, A., Karimi, P., Cai, X., Fraiture, C. D. & Akhter, M. S. Statistical data collection methodologies of irrigated areas and their limitations: A review. *Irrig. Drain.* **68**, 702–713 (2019).
9. Blatchford, M. L., Karimi, P., Bastiaanssen, W. G. M. & Nouri, H. From global goals to local gains—A framework for crop water productivity. *ISPRS Int. J. Geo-Inf.* **7**, 414 (2018).
10. Fritz, S. *et al.* A comparison of global agricultural monitoring systems and current gaps. *Agric. Syst.* **168**, 258–272 (2019).
11. Molden, D. & Saktivadivel, R. Water accounting to assess use and productivity of water. *Int. J. Water Resour. Dev.* **15**, 55–71 (1999).
12. Blatchford, M. L., Mannaerts, C. M., Zeng, Y., Nouri, H. & Karimi, P. Status of accuracy in remotely sensed and in-situ agricultural water productivity estimates: A review. *Remote Sens. Environ.* **234**, 111413 (2019).
13. Weerasinghe, I., Bastiaanssen, W., Mul, M., Jia, L. & van Griensven, A. Can we trust remote sensing evapotranspiration products over Africa?. *Hydrol. Earth Syst. Sci.* **24**, 1565–1586 (2020).
14. Toth, C. & Józkó, G. Remote sensing platforms and sensors: A survey. *ISPRS J. Photogramm. Remote Sens.* **115**, 22–36 (2016).
15. Karimi, P., Bastiaanssen, W. G. M. & Molden, D. Water Accounting Plus (WA+)—A water accounting procedure for complex river basins based on satellite measurements. *Hydrol. Earth Syst. Sci.* **17**, 2459–2472 (2013).
16. Blatchford, M., Mannaerts, C., Zeng, Y., Nouri, H. & Karimi, P. Influence of spatial resolution on remote sensing-based irrigation performance assessment using WaPOR data. *Remote Sens.* **12**, 2949 (2020).
17. Pareeth, S., Karimi, P., Shafiei, M. & De Fraiture, C. Mapping agricultural landuse patterns from time series of Landsat 8 using random forest based hierarchical approach. *Remote Sens.* **11**, 601 (2019).
18. Zhang, K., Kimball, J. S. & Running, S. W. A review of remote sensing based actual evapotranspiration estimation. *Wiley Interdiscip. Rev. Water* **3**, 834–853 (2016).
19. Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A. & Holtslag, A. A. M. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *J. Hydrol.* **212–213**, 198–212 (1998).
20. Bastiaanssen, W. G. M. *et al.* A remote sensing surface energy balance algorithm for land (SEBAL): Part 2: Validation. *J. Hydrol.* **212**, 198–212 (1998).
21. Bezerra, B. G., da Silva, B. B., dos Santos, C. A. C. & Bezerra, J. R. C. Actual evapotranspiration estimation using remote sensing: Comparison of SEBAL and SSEB approaches. *Adv. Remote Sens.* **04**, 234 (2015).
22. Anderson, M. C., Norman, J. M., Diak, G. R., Kustas, W. P. & Mecikalski, J. R. A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sens. Environ.* **60**, 195–216 (1997).
23. Allen, R. G. *et al.* Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Applications. *J. Irrig. Drain. Eng.* **133**, 395–406 (2007).
24. Senay, G. B. *et al.* Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach. *JAWRA J. Am. Water Resour. Assoc.* **49**, 577–591 (2013).
25. Roerink, G. J., Su, Z. & Menenti, M. S-SEBI: A simple remote sensing algorithm to estimate the surface energy balance. *Phys. Chem. Earth Part B Hydrol. Oceans Atmos.* **25**, 147–157 (2000).
26. McShane, R. R., Driscoll, K. P. & Sando, R. A review of surface energy balance models for estimating actual evapotranspiration with remote sensing at high spatiotemporal resolution over large extents. *Sci. Investig. Rep.* <https://doi.org/10.3133/sir20175087> (2017).
27. Bastiaanssen, W. G. M. *et al.* SEBAL model with remotely sensed data to improve water-resources management under actual field conditions. *J. Irrig. Drain. Eng.* **131**, 85–93 (2005).
28. Tang, R. *et al.* Spatial-scale effect on the SEBAL model for evapotranspiration estimation using remote sensing data. *Agric. For. Meteorol.* **174–175**, 28–42 (2013).
29. Karimi, P. & Bastiaanssen, W. G. M. Spatial evapotranspiration, rainfall and land use data in water accounting—Part 1: Review of the accuracy of the remote sensing data. *Hydrol. Earth Syst. Sci.* **19**, 507–532 (2015).
30. Melton, F. S. *et al.* OpenET: Filling a critical data gap in water management for the western United States. *JAWRA J. Am. Water Resour. Assoc.* **58**, 971–994 (2022).
31. Laipelt, L. *et al.* Long-term monitoring of evapotranspiration using the SEBAL algorithm and Google Earth Engine cloud computing. *ISPRS J. Photogramm. Remote Sens.* **178**, 81–96 (2021).
32. Roerink, G. J., Menenti, M. & Verhoef, W. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *Int. J. Remote Sens.* **21**, 1911–1917 (2000).
33. Pareeth, S., Salmaso, N., Adrian, R. & Neteler, M. Homogenised daily lake surface water temperature data generated from multiple satellite sensors: A long-term case study of a large sub-Alpine lake. *Sci. Rep.* **6**, 31251 (2016).
34. Neteler, M. Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data. *Remote Sens.* **2**, 333–351 (2010).
35. Metz, M., Andreo, V. & Neteler, M. A new fully gap-free time series of land surface temperature from MODIS LST data. *Remote Sens.* **9**, 1333 (2017).
36. Yang, Y. *et al.* Daily Landsat-scale evapotranspiration estimation over a forested landscape in North Carolina, USA, using multi-satellite data fusion. *Hydrol. Earth Syst. Sci.* **21**, 1017–1037 (2017).
37. Zhuang, Q., Shao, H. & Guan, D. Operational daily evapotranspiration mapping at field scale based on SSEBop model and spatiotemporal fusion of multi-source remote sensing data. *PLoS One* **17**, e0264133 (2022).
38. Anderson, M. C. *et al.* Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrol. Earth Syst. Sci.* **15**, 223–239 (2011).
39. Birkett, C. M. Synergistic remote sensing of Lake Chad: Variability of basin inundation. *Remote Sens. Environ.* **72**, 218–236 (2000).
40. Micklin, P. Desiccation of the Aral Sea: A water management disaster in the Soviet Union. *Science* **241**, 1170–1176 (1988).
41. Satgé, F. *et al.* Role of climate variability and human activity on Poopó Lake droughts between 1990 and 2015 assessed using remote sensing data. *Remote Sens.* **9**, 218 (2017).
42. AghaKouchak, A. *et al.* Aral Sea syndrome desiccates Lake Urmia: Call for action. *J. Gt. Lakes Res.* **41**, 307–311 (2015).
43. Bakhshianlamouki, E., Masia, S., Karimi, P., van der Zaag, P. & Sušnik, J. A system dynamics model to quantify the impacts of restoration measures on the water-energy-food nexus in the Urmia lake Basin, Iran. *Sci. Total Environ.* **708**, 134874 (2020).
44. Shadkam, S., Ludwig, F., van Oel, P., Kirmit, Ç. & Kabat, P. Impacts of climate change and water resources development on the declining inflow into Iran's Urmia Lake. *J. Gt. Lakes Res.* **42**, 942–952 (2016).
45. Hesami, A. & Amini, A. Changes in irrigated land and agricultural water use in the Lake Urmia basin. *Lake Reserv. Manag.* **32**, 288–296 (2016).
46. Javadzadeh, H., Ataie-Ashtiani, B., Hosseini, S. M. & Simmons, C. T. Interaction of lake-groundwater levels using cross-correlation analysis: A case study of Lake Urmia Basin, Iran. *Sci. Total Environ.* **729**, 138822 (2020).
47. ULRP. *The necessity for Restoration of Urumia Lake—The Cause and Consequence of Drought* (2015).
48. Saemian, P., Elmi, O., Vishwakarma, B. D., Tourian, M. J. & Sneeuw, N. Analyzing the Lake Urmia restoration progress using ground-based and spaceborne observations. *Sci. Total Environ.* **739**, 139857 (2020).
49. Helder, D. *et al.* Absolute radiometric calibration of Landsat using a pseudo invariant calibration site. *IEEE Trans. Geosci. Remote Sens.* **51**, 1360–1369 (2013).
50. USGS. Landsat collection 1 level 1 product definition. https://prd-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/atoms/files/LSDS-1656_%20Landsat_Collection1_L1_Product_Definition-v2.pdf (2019).
51. Rodell, M. *et al.* The global land data assimilation system. *Bull. Am. Meteorol. Soc.* **85**, 381–394 (2004).

52. Bastiaanssen, W. G. M., Cheema, M. J. M., Immerzeel, W. W., Miltenburg, I. J. & Pelgrum, H. Surface energy balance and actual evapotranspiration of the transboundary Indus Basin estimated from satellite measurements and the ETLook model. *Water Resour. Res.* **48** (2012).
53. Neteler, M., Bowman, M. H., Landa, M. & Metz, M. GRASS GIS: A multi-purpose open source GIS. *Environ. Model. Softw.* **31**, 124–130 (2012).
54. Caiserman, A., Amiraslani, F. & Dumas, D. Assessment of the agricultural water budget in southern Iran using Sentinel-2 to Landsat-8 datasets. *J. Arid Environ.* **188**, 104461 (2021).
55. Jaafar, H. H. & Ahmad, F. A. Time series trends of Landsat-based ET using automated calibration in METRIC and SEBAL: The Bekaa Valley, Lebanon. *Remote Sens. Environ.* **238**, 111034 (2020).
56. Xue, J., Bali, K. M., Light, S., Hessels, T. & Kisekka, I. Evaluation of remote sensing-based evapotranspiration models against surface renewal in almonds, tomatoes and maize. *Agric. Water Manag.* **238**, 106228 (2020).
57. Allen, R. *et al.* Satellite-based ET estimation in agriculture using SEBAL and METRIC. *Hydrol. Process.* **25**, 4011–4027 (2011).
58. Cleveland, W. S. & Devlin, S. J. Locally weighted regression: An approach to regression analysis by local fitting. *J. Am. Stat. Assoc.* **83**, 596–610 (1988).
59. Taheri, M. *et al.* Investigating the temporal and spatial variations of water consumption in Urmia Lake River Basin considering the climate and anthropogenic effects on the agriculture in the basin. *Agric. Water Manag.* **213**, 782–791 (2019).
60. Chukalla, A. D. *et al.* A framework for irrigation performance assessment using WaPOR data: The case of a sugarcane estate in Mozambique. *Hydrol. Earth Syst. Sci.* **26**, 2759–2778 (2022).
61. Blatchford, M. L. *et al.* Evaluation of WaPOR V2 evapotranspiration products across Africa. *Hydrol. Process.* **34**, 3200–3221 (2020).
62. Glenn, E. P., Neale, C. M. U., Hunsaker, D. J. & Nagler, P. L. Vegetation index-based crop coefficients to estimate evapotranspiration by remote sensing in agricultural and natural ecosystems. *Hydrol. Process.* **25**, 4050–4062 (2011).
63. Mateos, L., González-Dugo, M. P., Testi, L. & Villalobos, F. J. Monitoring evapotranspiration of irrigated crops using crop coefficients derived from time series of satellite images. I. Method validation. *Agric. Water Manag.* **125**, 81–91 (2013).

Acknowledgements

This work was funded by the United Nations Food and Agricultural Organization (FAO) as part of the “Rapid Assessment of the Water Accounts in Lake Urmia Basin, Iran” project (GCP/IRA/066/JPN), in which we undertook the ET_a mapping efforts. We acknowledge Tarbiat Modares University for providing the high-resolution land use map. We acknowledge NASA/USGS for providing Landsat data in public through various platforms including cloud storages. We also acknowledge the reviews by anonymous reviewers which helped in improving the manuscript substantially.

Author contributions

S.P. and P.K. developed the concept, S.P. conducted the analysis, S.P. and P.K. wrote the main manuscript text, S.P. prepared figures and tables. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023