Image-To-Image Translation Networks for Estimating Evapotranspiration Variations

IMAGE-TO-IMAGE TRANSLATION NETWORKS FOR ESTIMATING EVAPOTRANSPIRATION VARIATIONS: SAR2ET

Samet Çetin[1], *Berk Ülker*[2], *Gökberk Cinbis*[1], *Esra Erten*[3*]

[1]Department of Computer Engineering, Middle East Technical University, TR-06800 Ankara, Turkey
[2]Department of Electronics Engineering, Eindhoven University of Technology, 6825MC, The Netherlands
[3]Department of Geomatics Engineering, Istanbul Technical University, TR-34469 Istanbul, Turkey

ABSTRACT

Evapotranspiration (ET) plays a significant role in understanding the water necessities of crops during their growing season, and hence, aids to make a decision in agriculture (planting time, applying fertilizer, irrigation, yield prediction and etc.). In this context, over the past few years, a wide range of research studies have been implemented for learning field-level ET from low-resolution ET products by downscaling and/or data fusion strategies. Unlike these previous studies, this research aims to leverage deep learning based models to learn ET from temporally and spatially dense imaging data; Sentinel-1 and climate data; ERA-5, both provided by Copernicus Climate Change Service. The model is formed by weak supervision from high spatial resolution Sentinel-1 coupled with climate data and analysis ready ET product as target. We evaluated the framework across two geographically distributed regions, namely; The Balkans and The Aegean in order to understand how well weak supervision estimates ET over croplands in different ecosystems.

The code for the SAR2ET model is publicly available at `https://github.com/Agcurate/SAR2ET`, where you can access all the details regarding the model.

***Index Terms*—** Sentinel-1, Evapotranspiration, deep learning, disaggregation, weak supervision

1. INTRODUCTION

The use of Earth Observation (EO) data for agriculture is in its golden age with increasing publicly available wide range of high spatial resolution data with dense-in-time and dense-in-spectral information. This recent advances pave the way for the technological advancement in agriculture. Firstly, supervised learning algorithms have become standard approaches for leveraging EO data in its agricultural research processes. Extensive research has been conducted in the field of crop type classification, biophysical parameter estimation, soil moisture estimation and etc., which aims to learn the complex relationship between the EO data (inputs) and agronomic variable (target). In perspective of supervised learning, one of the main challenges with EO data is the requirement of large volumes of ground truth or labelled data for training stage. Although there are successful models with limited ground truth data, those model performance are highly limited with small geographic area where the ground truth information is collected. Addressing this challenge, employing unsupervised domain adaptation and weakly supervised deep learning frameworks are getting popular in the field of EO satellite-based agriculture in order to capture the biophysical parameters' variability even if there is no direct ground truth [1–3].

The second great value, as a result of huge amount of publicly available accumulated EO data, is the availability of Analysis Ready Data (ARD) for agricultural purposes, such as Leaf Area Index, Enhanced Vegetation Index, Evapotranspiration (ET), soil moisture products and etc. [4]. These data are generally produced from optical sensors due to their spectral richness being likely to provide more chemical information about crops, but, their availability can be limited based on the meteorological conditions of the acquisition time and the requirement of cloud-cleared reflectance information [5]. Landsat's ARD ET products, obtained from the Earth Engine Evapotranspiration Flux (EEFlux) platform, have been widely used as a proxy of water amount information of croplands. These products are generally produced by surface-energy balance and water balance methods [6] and constrained by the availability of optical sensors providing information about the croplands (e.g., available NDVI data). Instead, synthetic aperture radar (SAR) sensors-based images are not restricted by weather conditions and their utilization has been increasing rapidly due to the recent sustainable open-access data policies with high temporal resolution, specifically with Sentinel-1 mission.

In this study, to address this issues, a weakly supervised deep learning framework is discussed with ARD ET product as target and high spatial resolution Sentinel-1 and climate variables as inputs for having cloud-free regular ET product.

*This project entitled "Improving Resiliency of Malian Farmers with Yield Estimation: IMPRESSYIELD" was funded by the Climate Change AI Innovation Grants program, hosted by Climate Change AI with the additional support of Canada Hub of Future Earth.

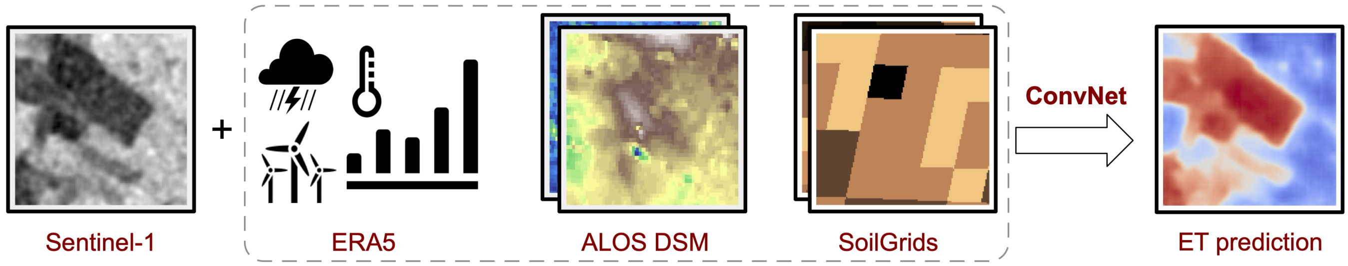

Fig. 1. The framework for dense-in-time high resolution ET product [7]. The weakly supervised framework used Sentinel-1 data having 6-days re-visit time from high spatial resolution ($\sim 10m$) and ET product having 16-days re-visit time from low spatial resolution ($\sim 30m$) as target. Agri-environmental indicators namely climate (dynamic) and soil & topographic (static) variables are used as auxiliary data in order to understand their impacts on ET estimation.

Our objective is to evaluate the potential of Sentinel-1 coupled with climate variables to produce high resolution ET product. Our analysis included cropland masks produced by Dynamic World (DW) Project; almost real-time land cover data set [8], representing different ecosystems from The Balkans to Aegean regions.

2. SAR2ET DATA SOURCE AND METHODS

The task of producing dense-in-time field-level ET product is discussed by deep learning framework as shown in Figure 1. U-net-like architecture is used to extract features from entire data, including SAR imaging data; Sentinel-1 and agri-environmental indicators such as soil texture variables, DEM-based topographic variables and climate variables. All variables are downloaded from a cloud-based geographic information computing database namely Google Earth Engine (GEE) [7]. Landsat based ARD ET images namely METRIC (Mapping Evapotranspiration at high Resolution with Internalized Calibration) EEFLUX were used as target variable. Its corresponding SAR imagery and auxiliary data such as climate (*temperature, dewpoint temperature, total precipitation, components of wind, surface net solar radiation sum and surface pressure*) variables, DEM (*elevation, slope and aspect*) and soil texture (*the amount of clay, sand, silt, bulk density*) were obtained by Sentinel-1, ERA5, ALOS DSM and SoilGrids data sources, respectively [7].

Firstly, a benchmark data set covering the croplands in The Balkans and Aegean Regions was constructed. Considering the agricultural activities in the region all the available target variable; ET, between March 1 and September 30, 2021 were downloaded. For each target ET, all the available Sentinel-1 data having an incidence angle between 30^o and 42^o from a week prior to and a week after the date of ET data were gathered. Sentinel-1 VV and VH bands and their ratio image were used to facilitate ET estimation task. Along with all the other auxiliary dynamic and static data, 128×128 pixel patches were extracted from all images resampled to 10 m (see Fig. 2). Once completing the pairing of the patches and splitting the data into train (70%), test (15%), and validation (15%) by applying the same split ratio to each individual ET raster, knowledge learned from ET product at $30m$ was transferred to retrieve from Sentinel-1 at $10m$.

To evaluate the performance of multiple experiments with varying input, ET estimation problem was formulated as a regression problem and the mean squared error (MSE) loss function was used for model training. The quantitative of each experiment is given by MSE and coefficient of correlation (R^2).

3. EXPERIMENTAL RESULTS

Optimizing the training of deep learning models hinges greatly on tuning the learning rate, a particularly sensitive hyper-parameter. In our approach, we employed the Adam optimizer with default settings, except for weight decay, and implemented the 1-cycle learning rate scheduling policy [9]. This policy offers a systematic and effective method for learning rate adjustment, with higher values applied during the middle of training to induce regularization effects. Our strategy entails a gradual increase in the learning rate from its initial value to a maximum rate, followed by a subsequent decrease to a minimum rate, significantly lower than the initial one. We set the maximum learning rate to 1e-4, with the initial rate initialized at 1/25 of this maximum. The learning rate then ascends, reaching its peak around 10 epochs before gradually declining to 1/1000 of the maximum value until the training process concludes. Additionally, across all experiments, we set the weight decay to 1e-6, the batch size to 128, and conducted training for 50 epochs.

The validation results of SAR2ET models run over two distinct geographical regions considering different combinations of static and dynamic variables are given in Table 1. Note that, the table gives the R^2 on the test data, however, the training data had almost the similar values in terms of the R^2, preceding the proper training stage. The accuracy was improved specifically when climate variables were considered. The added value of soil texture and DEM in the validation

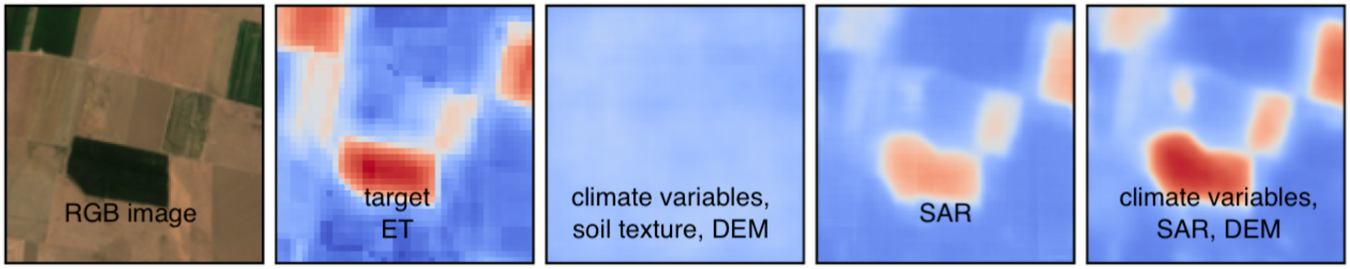

Fig. 2. An example of test patches; the colors are scaled from low (white) to high (red). Please note that RGB images (or any optical sensor-based images) are not used in the weakly supervised framework for the ET estimation.

Table 1. Comparison of multiple SAR2ET models trained with data from various static and dynamic data sources along with SAR for estimating ET by The Balkans and Aegean regions in terms of R^2 on the test data.

Model Inputs	The Balkans	Aegean R.
SAR	0.71	0.56
SAR + soil texture	0.75	0.59
SAR + DEM	0.75	0.63
SAR + climate variables	0.85	0.75
SAR + DEM + soil texture	0.64	0.78
SAR + DEM + climate variables	0.86	0.78

stage was not as evident as climate variables. This might be related to the fact that ET learning stage was conducted over only croplands, having similar topographic and texture properties. The best model for both regions was SAR with climate variables and DEM with mean MSE of 0.46, being much better than the ones with the only SAR having mean MSE of 0.94.

We also evaluated the best model's (SAR+DEM+climate variables) performance across different months and regions. Fig. 3 shows the R^2 values based on the clay content of soil and NDVI values for entire data set. The NDVI values, giving a hint about agricultural activity over the cropland mask, were calculated using the NIR and RED band of Sentinel-2. For both regions, the results showed that SAR2ET model worked perfectly when there is an agricultural activity in the test patches. This can be explained by the fact that the Sentinel-1 data is able to capture variations from both soil and vegetation layer, regardless of the soil texture information. Although, there is no crop type information for the cropland mask, generally the soil with high amount of clay is ideal for the non-rainfed agricultural activity due to the fact that it shows the water retention in the soil. In line with this outcome, the monthly average R^2 values for March to August depicted a sharp increase from 0.74 to 0.85. Similarly, if we examine the R^2 values by month, we observed slowly increasing patterns by months. The baseline SAR model started with relatively lower R^2 values in March (the beginning time of agricultural activities in the regions) but gradually improved its performance as the year progresses, achieving its highest one in July and August according to the region. It decreased in

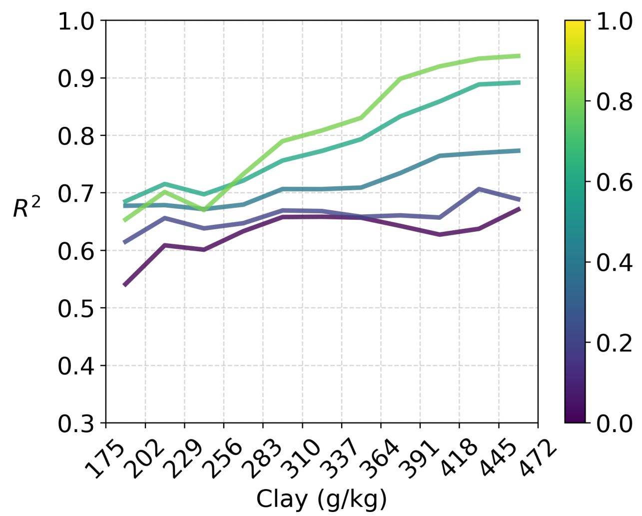

Fig. 3. Performance of the best-performing SAR2ET model (SAR+DEM+climate variables) including entire data set with varying soil texture parameter; namely clay and vegetation indicator; namely NDVI in terms of R^2 on the test data set.

the harvesting time again. This suggests that the SAR model becomes more accurate in capturing ET variations as vegetation and environmental conditions change during the warmer months and the crops' leaves are developed enough for transpiration. As an example, Fig. 4 illustrates the boxplot of response of Sentinel-1 VH backscattering due to the ERA-5

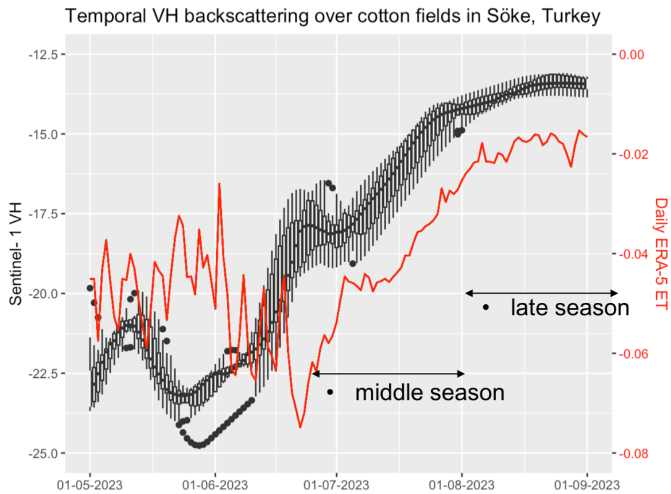

Fig. 4. Temporal boxplot of Sentinel-1 VH backscattering over cotton fields in a commune in the Aegean region with ERA-5 based ET values.

based ET variation over all cotton fields in Söke (a commune) in the Aegean region. The figure reveals that once the crops get volume (end of June), there is a clear relationship between backscattering and ET. This relationship gradually diminishes towards the end of the late season of cotton (begining of October). This trend is attributed to the radar signals' high penetration into dry crop volume. These physical relationships have been captured by our data-driven model, as emphasized in Fig. 3.

4. CONCLUSION

In this study, multi-source SAR2ET problem as a patch-level mapping task using weak supervision was discussed. The main motivation of the study was to discuss whether it is possible to learn ET from Sentinel-1 in order to have regular high resolution ET product. Additionally, the non-optical auxiliary (climate, soil texture and topographic variables) data were presented to increase the SAR2ET model performance. Specifically, climate variables boosted the performance of the SAR2ET model. The primary results indicate that introducing Sentinel-1 can generate dense in-time ET product and can be used for generating higher quality ARD ET product, specifically if there is a vegetation. The models' sensitivity to the vegetation condition of the crops was proven with temporal analysis.

5. REFERENCES

[1] Natalia Efremova, Mohamed El Amine Seddik, and Esra Erten, "Soil moisture estimation using Sentinel-1/-2 imagery coupled with CycleGAN for time-series gap filing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2022.

[2] Dilli Paudel, Diego Marcos, Allard de Wit, Hendrik Boogaard, and Ioannis N. Athanasiadis, "A weakly supervised framework for high-resolution crop yield forecasts," *Environmental Research Letters*, vol. 18, no. 9, 2023.

[3] Bulut Aygunes, Ramazan Gokberk Cinbis, and Selim Aksoy, "Weakly supervised instance attention for multisource fine-grained object recognition with an application to tree species classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 176, pp. 262–274, 2021.

[4] Mehmet Furkan Celik, Mustafa Serkan Isik, Gulsen Taskin, Esra Erten, and Gustau Camps-Valls, "Explainable artificial intelligence for cotton yield prediction with multisource data," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023.

[5] Yao Xiao, Wei Zhao, Mingguo Ma, Wenping Yu, Lei Fan, Yajun Huang, Xupeng Sun, and Qing Lang, "An integrated method for the generation of spatio-temporally continuous LST product with MODIS/Terra observations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–14, 2023.

[6] Maurizio Pieri, Marta Chiesi, Piero Battista, Luca Fibbi, Lorenzo Gardin, Bernardo Rapi, Maurizio Romani, Francesco Sabatini, Luca Angeli, Claudio Cantini, Alessio Giovannelli, and Fabio Maselli, "Estimation of actual evapotranspiration in fragmented Mediterranean areas by the spatio-temporal fusion of NDVI data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 12, pp. 5108–5117, 2019.

[7] Samet Çetin, Berk Ülker, Esra Erten, and Ramazan Gokberk Cinbis, "SAR2ET: End-to-end SAR-driven multisource ET imagery estimation over croplands," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, p. under review, 2024.

[8] Christopher F Brown, Steven P Brumby, Brookie Guzder-Williams, Tanya Birch, Samantha Brooks Hyde, Joseph Mazzariello, Wanda Czerwinski, Valerie J Pasquarella, Robert Haertel, Simon Ilyushchenko, Kurt Schwehr, Mikaela Weisse, Fred Stolle, Craig Hanson, Oliver Guinan, Rebecca Moore, and Alexander M Tait, "Dynamic World, Near real-time global 10m land use land cover mapping," *Scientific Data*, vol. 9, no. 1, pp. 251, 2022.

[9] Leslie N Smith and Nicholay Topin, "Super-convergence: Very fast training of neural networks using large learning rates," in *Artificial intelligence and machine learning for multi-domain operations applications*. SPIE, 2019, vol. 11006, pp. 369–386.