

Soil Moisture Monitoring of the Plant Root Zone by Using Phenology as Context in Remote Sensing

Ayda Aktas  and Burak Berk Üstündağ , *Member, IEEE*

Abstract—In this study, the phenological behavior and energy balance of plants are used as a sensory mechanism for root-zone soil moisture monitoring using both *in-situ* and satellite remote sensing data. The commonly used *in-situ* measurements are not feasible for mapping soil moisture at large-scale agricultural areas. Local direct root-zone soil moisture measurements cannot be reliably interpolated owing to the high spatial variability of soil structure and the vegetative content. Remote sensing methods are negatively affected by vegetation coverage and density regarding penetration and backscattering characteristics. In order to overcome these limitations, we propose a root-zone soil moisture estimation method utilizing a context-aware data clustering process, which can be applied prior to any statistical analysis, for empirical evaluation of data. In this aspect, the crops' phenological stages and soil–air temperature differences are defined as the two contexts for data clustering. Parameters such as canopy–air temperature difference, land surface temperature, and solar radiation with respect to plant energy and water processes are used for the analysis. The proposed model is utilized using piecewise linear regression of data obtained from 16 rainfed wheat parcels distributed across Turkey, under different climatic and topographic conditions. It is shown that the proposed context-aware data clustering process enables the nonlinear plant behavior to be analyzed linearly. The correlation value of the whole season increased from 21% to a range between 78% and 95% for different clusters. The outliers became relevant and the parameters became significant after the proposed context-aware data clustering.

Index Terms—Land surface temperature (LST), piecewise linear regression, plant phenology, remote sensing, soil moisture.

I. INTRODUCTION

THE retrieval of reliable root-zone soil moisture data has been challenging because of the nonlinear relationship between surface moisture and root-zone soil moisture and its spatial variability [1]. Although yield efficiency and water management in agricultural areas are directly related to root-zone soil moisture on the basis of phenological stage (PS), remote sensing methods are generally used for measuring near-surface soil moisture. The aim of this study was to demonstrate the

efficiency of PS and soil–air temperature difference as a context for estimating the soil moisture based on the remote sensing of plants with known characteristics.

Soil moisture plays a vital role in the surface water cycle. It is the link in the water exchange process between the land surface and atmosphere and is a reflection of the groundwater status [2]. Because water is crucial for crop development, the ability to estimate and map soil moisture is important for agricultural applications of precision agriculture and sustainability, such as irrigation scheduling and basin water management. Moreover, soil moisture has a significant effect on crop yield estimation [3].

Soil moisture is affected by soil texture, topography, land cover, and climate. Therefore, soil moisture content varies both in space and time. There is a physical relationship based on diffusion processes between the surface and root-zone soil moisture [4]. Conventional *in-situ* soil moisture estimation methods are based on single-point observations of a specific location, representing a small area around the sensor [5], [6]. This point-based approach is poor for estimating soil moisture of large-scale areas owing to heterogeneity [7]. Moreover, it is not economically feasible to establish dense networks of agrometeorological stations. The extrapolation of single-point observations to larger scales is also time-consuming, complex, and expensive, especially for heterogeneous regions [8].

Advanced remote sensing technology can be used to estimate soil moisture at large and regional scales. Remote sensing methodologies that are used for estimating soil moisture are mainly grouped based on the focused electromagnetic spectrum regions; these are mainly optical and thermal methodologies, microwave methodologies, and synergistic approaches [9]. Optical and thermal methodologies are based on the relationship between soil moisture and soil reflectivity or surface temperature and soil thermal properties [10], [11]. Studies of soil moisture and reflection in different soil types have shown that increased water content reduces reflection in a nonlinear relationship [13], [14]. Vegetation growth is sensitive to water stress because drought or dry soil conditions affect the growth.

Therefore, another approach to retrieve soil moisture data is the use of vegetation indexes, such as the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI) [15]. A wide variety of drought indices have been developed based on vegetation indices with different contexts. Kogan [16] proposed the vegetation condition index to remove the weather and spatial differences from NDVI by using statistical NDVI time-series data. The normalized difference water index

Manuscript received February 28, 2020; revised July 27, 2020; accepted August 19, 2020. Date of publication September 4, 2020; date of current version October 15, 2020. This work was supported by Turkish Ministry of Development, in part by Agricultural Monitoring and Information Systems Project (2011A020100) and the relevant joint project funded by Ministry of Food, Agriculture and Livestock, Turkey. (Corresponding author: Ayda Aktas.)

Ayda Aktas is with the Informatics Institute, Istanbul Technical University, Istanbul 34469, Turkey (e-mail: akkartala@itu.edu.tr).

Burak Berk Üstündağ is with the Faculty of Informatics and Computer Engineering, Istanbul Technical University, Istanbul 34469, Turkey (e-mail: bustundag@itu.edu.tr).

Digital Object Identifier 10.1109/JSTARS.2020.3021990

(NDWI), which is not sensitive to atmospheric conditions, has been proposed based on the relationship between different water absorption bands [17]. Lie *et al.* improved NDWI by increasing the sensitivity and achieving a quick response to changes by utilizing a normalized multiband drought index (NMDI) [18]. Similarly, it has been shown that soil thermal properties and surface temperature variations show strong correlation with soil moisture. Different thermal infrared methodologies based on thermal inertia and temperature indexes have been proposed based on the thermal properties of land cover in the thermal infrared band [19], [20]. Based on the soil moisture and its effects on evapotranspiration, the crop water stress index (CWSI) was proposed by Idso *et al.* [21]. The CWSI method has higher precision over vegetated surfaces because it is based on a single canopy energy balance model.

Passive and active microwave methodologies are based on the backscatter and emission properties of soil influenced by soil texture, surface roughness, and vegetation [22], [23]. Changes in the amount of water content affects the scattering and absorption behaviors of the target surface owing to the changes in the dielectric properties. Active sensors have the advantage of high spatial resolution with lower soil moisture sensitivity, whereas passive sensors have higher soil moisture accuracy with coarse spatial resolution. Synergistic methods have been developed to decrease the deficiencies of different methods by combining different approaches, resulting in increased sensitivity to soil moisture. Optical and thermal data, active and passive microwave data, and active microwave and optical data are predominantly used as synergistic methods [9].

Moreover, global satellite-based soil moisture datasets have been collated, which started with the publication of the first global multiannual dataset derived from the European Remote Sensing (ERS) Satellites ERS-1 and ERS-2 scatterometer (SCAT) observations in 2003 [25]. Among these are the Advanced SCATterometer (ASCAT) by EUMETSAT [26], the Advanced Microwave Scanning Radiometer-2 (AMSR-2) by JAXA [27], Soil Moisture Ocean Salinity (SMOS) by ESA [28], and Soil Moisture Active Passive (SMAP) by NASA [29]. However, there are several limitations to these methods and the use of these datasets [29]. Using methodologies or products based on the passive microwave region of the spectrum has low spatial resolution for small catchments and field-based applications [29], [30]. On the other hand, optical remote sensing methodologies have the disadvantages of poor temporal resolution, weather dependencies, and night-time limitation. Furthermore, as the plants grow, the variations in the crops' vegetative status affect the performance of remote-sensing-based soil moisture monitoring [23]. As the level of vegetation cover becomes denser, it rapidly becomes opaque. Therefore, the reflecting energy no longer represents the soil but a mixture of soil and vegetation or entirely vegetation. Even though microwave bands can penetrate vegetation, they are also affected by surface roughness and vegetation cover density [23], [31]. Vegetation reduces the backscattering coming from the underlying soil and generates volume scattering [23].

To overcome spatial resolution limitations, disaggregation of remote sensing data has been developed for estimating root-zone soil moisture [32], [33]. Most of the root-zone soil moisture

estimation methodologies are based on extending surface soil moisture estimations. Basically, data assimilation algorithms have been established by directly using satellite-driven surface soil moisture data or global soil moisture datasets [34], [35]. Another research area has focused on the assimilation of disaggregated remote sensing data with land surface models or hydrological models to enhance root-zone soil moisture estimation [3], [36]. However, the accuracy of the soil moisture retrieval methods is highly influenced by the complex input parameters of both empirical equations and physically-based models.

Even though the access to a wide range of datasets has become easier, the datasets are usually not adequate for direct analysis. Therefore, a data refinement process has become a critical step before any statistical modeling. To the best of our knowledge, even though clustering has been used within different soil-moisture-related applications, it has not been used as a data-refinement technique. Van Arkel and Kaleita used *K*-means clustering to identify critical sampling locations for the field-scale near-surface soil moisture determination [37]. To characterize soil moisture variation on a steep hill slope, Lee and Kim used cluster analysis based on Euclidean similarity between soil moisture time series [38]. Matei *et al.* proposed the use of context-related data from easy to access web portals in addition to the local data provided from weather stations in real time for soil moisture estimation with an average correlation of 68% [39]. Avram *et al.* demonstrated the advantages of context-aware data mining over classical data mining with a case study on soil moisture prediction [40]. In this study, we proposed a context-aware clustering process, reflecting the contextual information within the data itself, to overcome the limitations mentioned above. Remote sensing data and *in-situ* measurements from agrometeorological stations that are easy to access and can be interpolated for the parcels without stations were fused based on the plant energy and water processes.

II. PLANT ENERGY AND WATER PROCESSES

There is a constant energy, water, and momentum exchange between the land surface and the atmosphere. This exchange is regulated by latent heat (λE) and sensible heat (H). λE is the heat flux from the surface to the atmosphere and associated with evapotranspiration. H is the heat energy between the surface and atmosphere and associated with conduction and convection. In both cases, the rate of exchange is explained by the driving force and a transfer coefficient corresponding to conductance. The efficiency of the transfer is described by the resistances that are reciprocal to conductance. The driving force of the exchange is the leaf-air temperature difference for H , whereas for λE , it is the leaf-air water vapor fraction difference

$$H = g_{aH} C_p (T_c - T_a) = \frac{(T_c - T_a)}{r_{aH}} \rho C_p \quad (1)$$

where

- T_c crop canopy surface temperature (°C);
- T_a air temperature (°C);
- ρ density of air (kg m^{-3});
- C_p molar specific heat of air ($29.3 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$);
- g_{aH} boundary layer conductance ($\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$);

r_{aH} corresponding resistance ($\text{m}^2 \cdot \text{s} \cdot \text{mol}^{-1}$);

$$\lambda E = \frac{(e_c(T_c) - e_a)}{r_{aW} + r_{cW}} \frac{\rho C_p}{\gamma} \quad (2)$$

in which

$e_c(T_c)$ saturated water pressure at given crop surface temperature (kPa);

e_a air vapor pressure (kPa);

r_{aW} aerodynamic resistance to vapor transport ($\text{s} \cdot \text{m}^{-1}$);

r_{cW} crop canopy resistance to vapor transport ($\text{s} \cdot \text{m}^{-1}$);

γ psychrometer constant ($\text{kPa}^\circ\text{C}^{-1}$).

Penman approximation can be used to replace the surface–air vapor pressure difference by vapor deficit of the ambient air and a term of surface temperature difference within the latent heat equation [41]

$$\begin{aligned} (e_c(T_c) - e_a) &= (e_c(T_a) - e_a) - S (T_a - T_c) \\ &= D + S (T_c - T_a) \end{aligned} \quad (3)$$

where $(e_c(T_a) - e_a) = D$ shows the vapor pressure deficit of the ambient air and S is the slope of the saturated vapor pressure temperature relation.

These fluxes must be in balance with the incoming radiation (R_n), both in short-wave and long-wave ranges, where the excess energy gives the rate of energy storage. M is the rate of metabolic storage representing the metabolic reactions, such as photosynthesis and respiration, and S is the net physical storage. When the whole canopy is considered, rather than a leaf, M and S can be neglected

$$R_n - H - \lambda E = M + S. \quad (4)$$

Using the energy balance equation, the relation between the crop canopy surface temperature and the air temperature can be written as follows [42]

$$\begin{aligned} T_c - T_a &= \frac{r_{aH} (r_{aW} + r_c) \gamma R_n}{\rho C_p [\gamma (r_{aW} + r_{cW}) + S r_{aH}]} \\ &- \frac{r_{aH} D}{\gamma (r_{aW} + r_{cW}) + S r_{aH}}. \end{aligned} \quad (5)$$

The following assumptions were made in (5). S was calculated at the temperature $(T_c + T_a)/2$ and it was assumed that S was constant within each $(T_c - T_a)$ temperature interval. The second assumption was that the net incoming radiation is not affected by leaf conditions even though it is a function of leaf temperature. It is shown that the canopy surface temperature is a function of air temperature (T_a), humidity, wind speed, and absorbed net radiation (R_n). Therefore, excess leaf temperature can be expressed with two terms based on net radiation and vapor pressure deficit, a parameter of evapotranspiration. The energy balance relation showing these parameters using a short-circuit analogy is used to establish the contextual model for root-zone soil moisture monitoring (see Fig. 1).

Crop temperature (T_c) is an important indicator of the health of plants. The main determinant of leaf temperature is the rate of evapotranspiration from the leaf [43]. Plants need root-zone soil moisture to meet their evapotranspiration needs. If soil

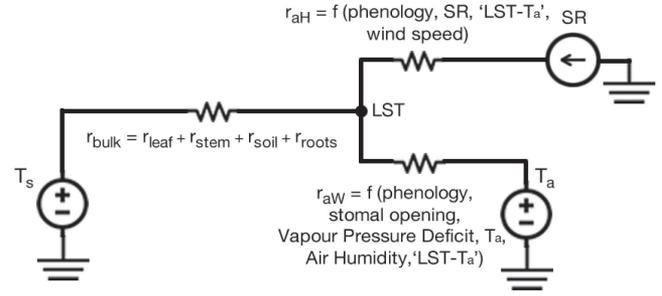


Fig. 1. Plants energy balance circuit with the effecting parameters used for simulation.

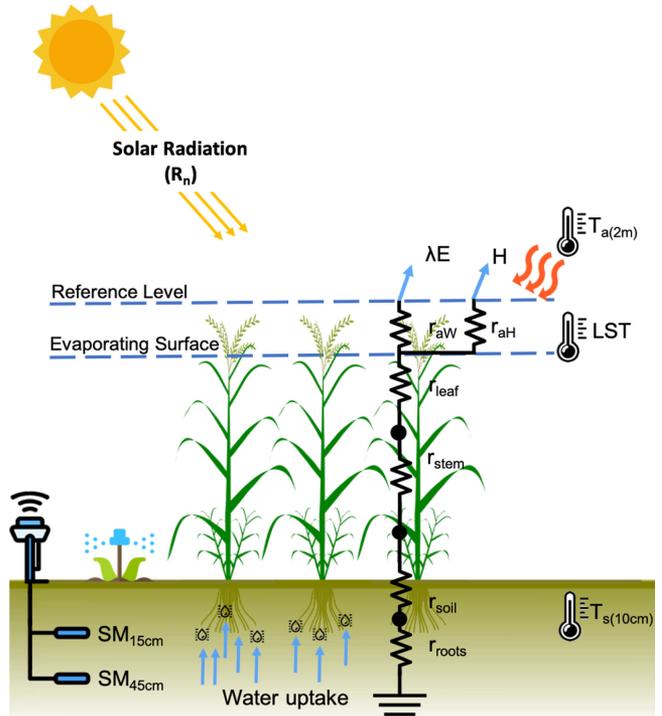


Fig. 2. Simplified representation of crop surface and aerodynamic resistances for hydraulic flow and heat transfer together with the parameters of reference soil moisture at 15- and 45-cm depth, atmospheric temperature at 2 m, LST, soil temperature at 10-cm depth, which are used for contextual model building.

moisture is available, plants can balance their temperature by evapotranspiration. The cooling mechanism uses the energy of latent heat to convert liquid water to water vapor, hence soil moisture is taken away from the plant in the form of evaporating water. If there is deficiency in soil moisture, the plant produces chemical and hydraulic signals that trigger physiological responses, such as stomatal closure, reduction in photosynthesis rate, reduction in evapotranspiration causing leaf area reduction, stunted growth, and wilting [44]. Soil moisture deficiency affects the crops' development because of the resulting reduced biomass and quality. Therefore, the contexts for the data clustering and the input parameters of the estimation model are defined with regard to these water and energy transfers. A simplified representation of the setup used in this study is given in Fig 2.

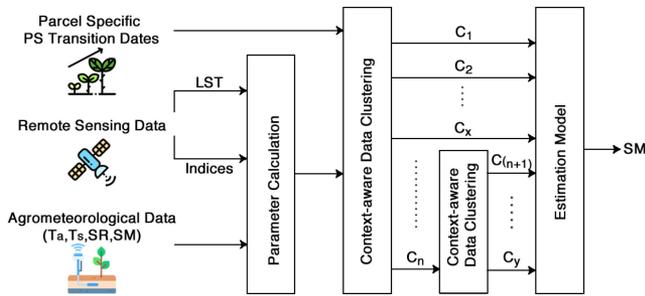


Fig. 3. Context-aware clustering approach for root-zone soil moisture estimation.

III. CONTEXT-AWARE CLUSTERS

The advancement of technology and the large number of datasets collected by the monitoring of various parameters all over the world allows access to a higher amount of *in-situ* measurements with higher spatial and temporal resolution. The main idea of context-aware data clustering is to fuse this data from various sources with remote sensing data and produce valuable information that cannot be obtained from only one data source. A context-aware data clustering process enables the data to be significant within the defined contexts within the aim of the application. To establish an efficient context-aware estimation model, a specified plant with known PS transition dates, whether derived from *in-situ* or satellite measurements, should be determined with regard to the aim of the agricultural application. In this study, we fused *in-situ* measurements of selected winter wheat parcels and their PS transition dates with satellite-imagery-driven parameters to monitor root-zone soil moisture (see Fig. 3).

A. Phenology

Phenology is the timing of the biological events in plants, from sowing to maturity. Plants' climatic and chemical needs and behavior differ at each PS. These variations are the main reason for nonlinearity in plant development and life cycle. Nonetheless, each PS represents a different logical state of this nonlinear system that switches between linear models [45]. In other words, a plant's behavior within each separate PS can be considered as linear.

Leaf area index (LAI), the percentage of soil covered with green leaf, increases as the plant grows. Increased LAI is one of the main problems of remote-sensing-based soil moisture estimation. The reflecting energy begins to represent not just soil, but a mixture of soil and vegetation or just vegetation itself. To overcome this issue of increasing vegetation, the PSs of winter wheat, based on FAO guidelines, were defined as the primary context (see Fig. 4) [46]. Up to PS3, when there is more bare soil than green vegetation, soil moisture can be directly associated with the input data. However, in PS4, PS5, and PS6, the context-aware data clustering approach for root-zone soil moisture estimation is expected to be more efficient. On the other hand, at the maturity stage until harvest, it will be negatively

affected by wilting. However, soil moisture at this stage is not essential for agricultural water management.

B. Soil–Air Temperature Difference

The soil–air temperature difference is evaluated using reference *in-situ* measurements and defined as the secondary context. To examine the climatic differences and their effects on the crop PS, the selected parcels were grouped into two, based on their latitudes: 37.0°N and 40.0°N. The daily variations in the 10-cm depth soil (T_s), 2-m air temperatures (T_a), and ($T_s - T_a$) were analyzed for a year covering the 2016–2017 winter wheat production season (see Fig. 5).

In summertime, the air temperature is expected to be higher than the near-surface soil temperature, whereas the opposite is expected in wintertime. As expected, the soil and air temperature difference ($T_s - T_a$) was positive for winter and negative for summer, with transitional periods corresponding to fall and spring. In addition, the average of each parcel's daily ($T_s - T_a$) values at the time of Landsat satellite passes was calculated for each latitude. To analyze the influence of different climatic conditions on crop growth, the average sowing and PS transition dates of the related parcels were used for interpretation (see Fig. 6). For visualization purposes, the five-day moving average values of the average ($T_s - T_a$) were used in the graph.

The graphs revealed latitude, and hence climate-related PS and seasonality differences. The average sowing times were relatively earlier with a longer seasonal period for the parcels at 40.0°N latitude. A notable difference between the phenological phase durations was observed. The dormancy phase was longer at 40.0°N latitude owing to relatively lower temperatures and snowfall in the wintertime. At a higher latitude, significantly shorter PS durations follow the longer dormancy phase. Also, the average harvesting time was earlier for the parcels closer to 37.0°N latitude. It is apparent from the graphs that there was a transition period especially corresponding to PS4 and PS5. Moreover, the transition from winter to summer corresponded to the anthesis stage (PS6).

IV. STUDY AREA

Turkey is located on both the Europe and Asia continents in the Mediterranean climate zone (see Fig. 7). The climatic conditions throughout the country changes from one region to another owing to diverse landscapes. In general, coastal regions have a milder climate, where the summers are hot and dry and the winters are mild and wet. On the Mediterranean and Aegean coasts, the annual precipitation varies from 580 to 1300 mm [47]. The coastal Black Sea region has a temperate climate with the highest annual precipitation rate reaching 2200 mm in the eastern parts. On the other hand, the inland regions have a mostly continental climate with extremes of high temperatures in summer and harsh winters with limited rainfall. In total, 65% of the land is arid and semiarid.

Wheat is the most widely grown cereal in Turkey. Turkey is home to 23 species of wild wheat and more than 400 cultivated wheat varieties [48]. Regardless of the irregular topography and various climates, it is possible to produce winter wheat in almost



Fig. 4. Series of winter wheat images showing PSs acquired from a TARBIL station.

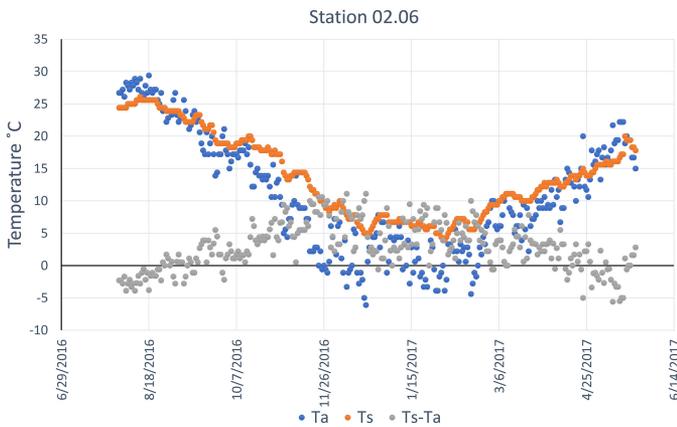


Fig. 5. T_a , T_s , and $T_s - T_a$ graph of station 02.06.

every region. Taking into consideration its economic importance, we chose winter wheat for the root-zone soil moisture monitoring. Throughout Turkey, 16 rainfed winter wheat parcels that reflect the versatility of climatic conditions and soil types are used as reference parcels (see Fig. 7).

V. DATA USED AND PARAMETERS

In-situ measurements of the selected 16 parcels were acquired from agrometeorological stations that are part of Istanbul Technical University—Agricultural and Environmental Informatics

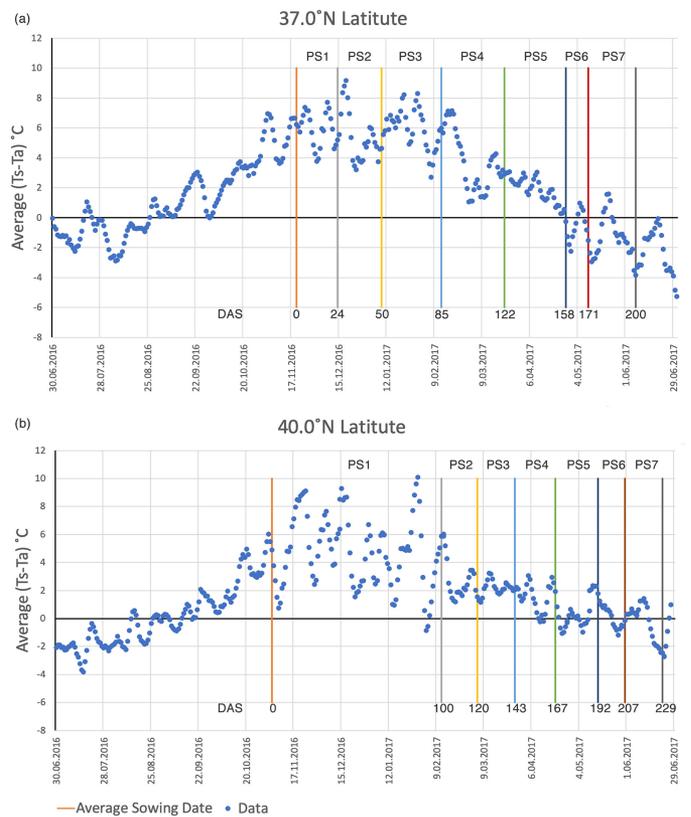


Fig. 6. Daily $(T_s - T_a)$ average of the selected winter wheat parcels grouped by their latitude (a) 37.0°N and (b) 40.0°N along with the average PS transition days and the corresponding day after sowing values.

TABLE I
IN-SITU MEASURED AND CALCULATED PARAMETERS

Acronym	Description	Unit	Use
SM ₄₅	Measured Soil moisture at 45cm depth	cbar	Reference
SM ₁₅	Measured Soil moisture at 15cm depth	cbar	Reference
SM ₃₀	(SM ₁₅ +SM ₄₅)/2	cbar	Reference
T _a	Measured Atmospheric Temperature at 2m	°C	Parameter
T _{s-10}	Measured Soil temperature at 10cm depth	°C	Reference
SR	Measured Solar radiation	W/m ²	Parameter
T _{a-24h}	Mean T _a of the last 24 hours	°C	Parameter
T _{a-7d}	Mean T _a of the last 7 days	°C	Parameter
1 / SR		W ⁻¹ m ²	Parameter
SR _{2h}	Cumulative SR of the last 2 hours	W/m ²	Parameter
T _a · SR		°C W/m ²	Parameter
GDD _{24h}	Growing Degree Days of the last 24 hours	°C	Parameter
GDD _{7d}	Growing Degree Days of the last 7 days	°C	Parameter

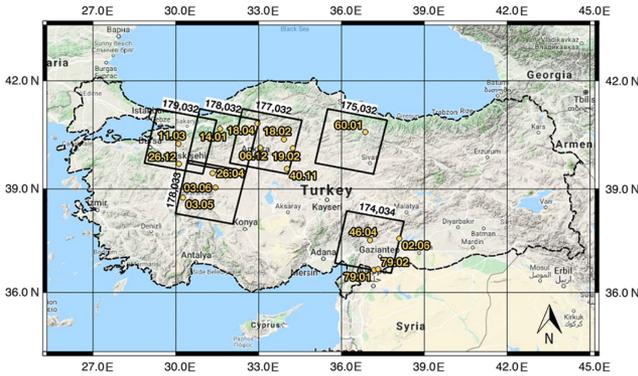


Fig. 7. Projection of 16 TARBIl agrometeorological stations covered by 6 Landsat tiles.

and Applied Research Center's (TARBIl's) observation network (see Fig. 7). Every TARBIl monitoring station is equipped with 35 sensors capable of atmospheric, soil, and phenological measurements, with a sampling rate of 10–30 min. All the *in-situ* measurements that were sampled at the time of satellite passes were used in the analysis unless otherwise stated.

Atmospheric temperature (°C) at 2 m and solar radiation (W/m²) was used within the model, whereas soil moisture (cbar) at depths of 15 and 45 cm and soil temperature (°C) at 10 cm depth were used as reference parameters (see Table I). Soil moisture at 30 cm depth (SM₃₀), defined as the corresponding root zone of winter wheat, was calculated as the mean value of SM₁₅ and SM₄₅. SM₃₀ was used as a reference parameter for the estimation model.

Tensiometric soil moisture measurements were used instead of volumetric measurements because the physical force that actually holds the water in the soil is more important than the percentage of water in a given amount of soil for agricultural applications, especially for irrigation purposes. Tensiometric soil moisture values varied from 0 to 200 cbar (see Table II).

TABLE II
INTERPRETATION OF TENSIMETRIC SOIL MOISTURE READINGS

Soil Moisture (cbars)	Interpretation
0-10	Saturated soil
10-30	Soil is adequately wet
30-60	Usual range for irrigation
60-100	Usual range for irrigation in heavy clay
100-200	Soil is becoming dangerously dry for maximum production.

Lower values indicate saturated wet water, whereas higher values indicate the need for irrigation based on soil type.

The thermal regime of the growing season is one factor that determines the development of plants [49]. Days warmer than the normal advance the growth, whereas days cooler than normal slow the growth. Therefore, besides measuring the *in-situ* parameters, supplementary input parameters devoted to the thermal regime of the plant were calculated using T_a and solar radiance (SR) (see Table I). Among them, the growing daily degree day (GDD) represents the heat value assigned to each day and is calculated as

$$\frac{T_{\max} + T_{\min}}{2} - T_{b\min} \quad (6)$$

where T_{min} and T_{max} represent the daily minimum and maximum temperatures, and T_{bmin} is the base temperature below which plant development stops for a selected plant. In this study, the base temperature was set to 5 °C and the minimum and maximum values of the last 24 h were used for the calculations. This approach allowed GDD to reflect the thermal time of the plant over the last 24 h instead of a calendar day. To examine the weekly thermal time influence on the root-zone soil moisture, 7-day GDD was also calculated for the statistical analysis using the same approach.

TABLE III
LANDSAT DATA AND CORRESPONDING STATIONS

Station No	Latitude (N)	Longitude (E)	City	Path	Row	Approx. Time (AM)
02-06	37,570306	38,088942	Adiyaman	174	34	08:10
03-05	38,757222	30,257972	Afyon	178	33	08:40
03-06	39,039972	31,435417	Afyon	178	33	08:40
06-12	40,165861	33,054306	Ankara	177	32	08:30
11-03	40,274167	30,093611	Bilecik	179	32	08:40
14-01	40,692806	31,595778	Bolu	178	32	08:30
18-02	40,395194	33,915389	Cankiri	177	32	08:30
18-04	40,834944	32,950611	Cankiri	177	32	08:30
				178	32	08:30
19-02	40,148389	34,234083	Corum	177	32	08:30
26-04	39,450056	31,329278	Eskisehir	178	33	08:40
26-12	39,713222	30,110750	Eskisehir	179	32	08:40
40-11	39,572750	34,024722	Kirsehir	177	32	08:30
46-04	37,534556	37,031556	Kahramanmaras	174	34	08:10
60-01	40,598833	36,862167	Tokat	175	32	08:20
79-01	36,668806	37,183611	Kilis	174	34	08:10
79-02	36,694667	37,327111	Kilis	174	34	08:10

TABLE IV
REMOTE SENSING DRIVEN MODEL INPUT PARAMETERS

Acronym	Definition	Formula	Source
NDVI	Normalized Difference Vegetation Index	$(\text{NIR-Red}) / (\text{NIR+Red})$	Downloaded
$\Delta\text{-NDVI}_1$	8-day interval NDVI derivative where possible	$(\text{NDVI}_{d1} - \text{NDVI}_{d2}) / 8$	Computed
$\Delta\text{-NDVI}_2$	16-day interval NDVI derivative where possible	$(\text{NDVI}_{d1} - \text{NDVI}_{d2}) / 16$	Computed
NDMI	Normalized Difference Moisture Index	$(\text{NIRSWIR}) / (\text{NIR+SWIR})$	Downloaded
LST	Land Surface Temperature	Section V-a	Computed
LST ²	Square of Land Surface Temperature	LST ²	Computed

Parcel-based PS transition dates from sowing to harvesting were also provided within the TARBIL project. These dates were computed by a semiautomated image processing software using the crop and geographical-location-based averages; suspicious data were visually controlled by operators [50].

In-situ parameters were used together with the satellite imagery-driven parameters. The evaluation of remote-sensing-based soil moisture retrieval has shown that high-resolution satellite imagery does not correlate with a significant improvement in the predicting power compared with moderate-resolution satellite imagery [51]. The Landsat archive is the longest continuously acquired, moderate-resolution satellite imagery archive, which is also free of charge since 2008. Landsat 7 and Landsat 8 satellites Collection-1 Level-1 images were used together, enabling an 8-day temporal resolution. The images used in the analysis were tier-1 images, which assured the highest data quality and suitability for time-series analysis. In total, 16 reference parcels were covered with 6 Landsat tiles (see Fig. 7 and Table III).

The satellite images, including the brightness temperature, pixel quality, and spectral indices of NDVI, and NDMI bands were downloaded using the ESPA (EROS Science Processing Architecture) interface. Taking into consideration the soil moisture effect on the crops' biomass, 8-day and 16-day interval NDVI changes were computed (see Table IV). Furthermore, as a representation of crop canopy surface temperature (T_c), the 8-day interval land surface temperature (LST) was calculated

using satellite imagery, where possible owing to weather conditions.

A specified pixel was assigned to each reference parcel. The selection was based on the images taken by the cameras mounted

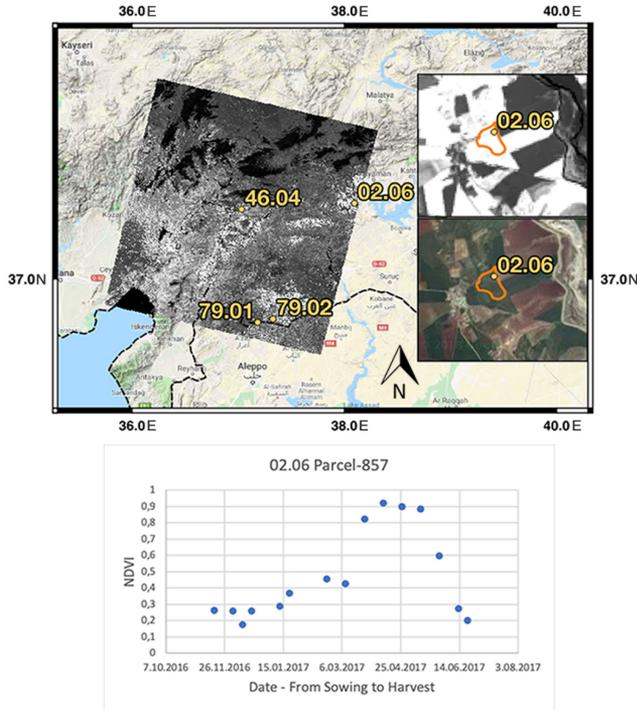


Fig. 8. Landsat-8 NDVI image of March 25, 2017 showing the general location of TARBIL stations around latitude 37.0°N and NDVI graph of Parcel-857, observed from station 02.06, from sowing to harvest.

TABLE V
PARAMETERS ANALYZED FOR MODEL DEVELOPMENT WITH
RESPECT TO THEIR ORIGIN

Remote Sensing Parameters	In-situ Parameters	Hybrid Parameters
NDVI	T_a	LST- T_a
Δ -NDVI ₁	SR	$(LST \cdot T_{a-24h}) / SR$
Δ -NDVI ₂	T_{a-24h}	$(LST \cdot T_{a-24h}) / SR^2$
NDMI	T_{a-7d}	$LST / (SR \cdot T_{a-24h})$
LST	$1 / SR$	LST / SR
LST ²	SR _{2h}	LST / SR ²
	$T_a \cdot SR$	LST \cdot SR
	GDD _{24h}	$(LST - T_a) / SR$
	GDD _{7d}	$(LST - T_a) / SR^2$
		LST / T_{a-24h}
		LST / $(T_a \cdot SR)$

on TARBIL stations to maintain homogeneity. Pixel-based remote sensing parameters were calculated and extracted, and a dataset was constructed by only using clear pixels based on the pixel quality band (see Fig. 8).

To evaluate the relationships and construct the proposed model, hybrid parameters were generated by combining remote sensing and *in-situ* parameters considering the plant energy balance and water processes. As a result, the final dataset was a combination of all the parameters given in Table V, based on the measurements on the days of clear pixel satellite passes.

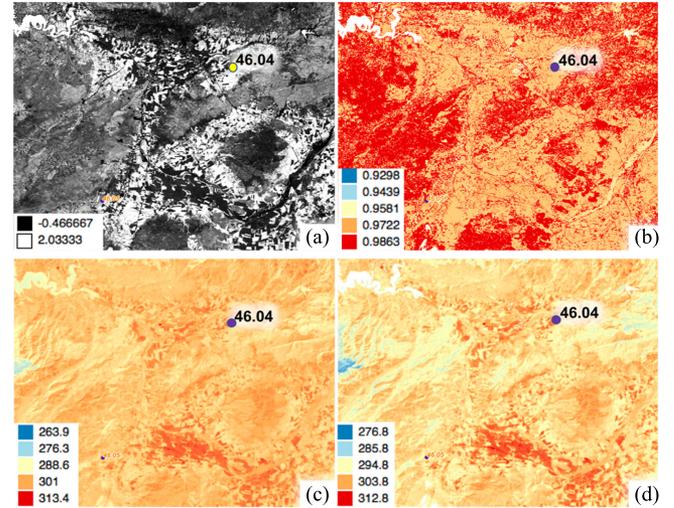


Fig. 9. Landsat 8 imagery showing (a) vegetation proportion, (b) emissivity, (c) brightness temperature, and (d) LST of station 46.04 and the surrounding area.

A. LST Retrieval

Landsat 7 ETM+ Thermal Infrared band (10.4–12.42 μm) and Landsat 8 TIRS band 10 (10.6–11.2 μm) were used for the computation of LST from Landsat images. Emissivity-corrected LST was computed using the brightness temperature downloaded from the ESPA website (see Fig. 9) [52]

$$\frac{T_B}{1 + \left(\lambda \times T_B / p \right) \times \ln(\varepsilon)} \quad (7)$$

where

- T_B brightness temperature (K);
- λ wavelength of the emitted radiance;
- $p = h \times c / s$ ($p = 1.438 \times 10^{-2}$ mK);
- h Planck's constant (6.626×10^{-34} J-s);
- s Boltzmann constant (1.38×10^{-23} J/K);
- c velocity of light (2.998×10^8 m/s);
- ε emissivity.

Emissivity values were computed based on the NDVI method considering three different conditions, which were only soil, full vegetation cover, and a mixture of both. The proportion of vegetation (P_v) was used for stating these three conditions [53]. $NDVI_{\min}$ corresponds to the soil that is in a nonvegetative state, which is set as NDVI value of 0.2, and $NDVI_{\max}$ represents the fully vegetated state, with an NDVI value set as 0.5

$$P_v = \left(\frac{NDVI - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}} \right)^2 \quad (8)$$

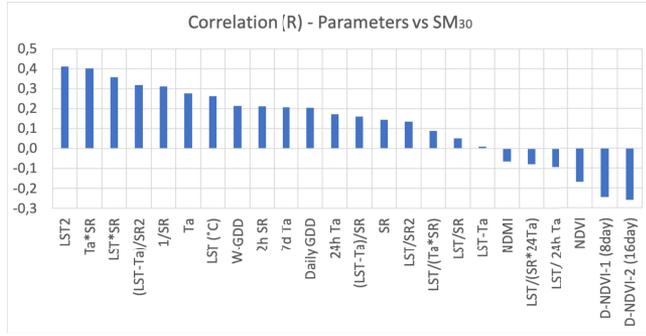
Emissivities for Landsat 7 and Landsat 8 images were calculated separately based on P_v (see Table VI and Fig. 9).

VI. CONTEXT-AWARE DATA CLUSTERING AND ANALYSIS

We analyzed the descriptive statistics of each parameter for the timespan from sowing to harvest as one dataset. The parameters'

TABLE VI
 EMISSIVITY CALCULATION FOR LANDSAT 7 AND LANDSAT 8 IMAGES

Landsat 7	Landsat 8	Vegetative State
0.979 – 0.035 · (Red)	0.973 – 0.047 · (Red)	$NDVI < 0.2$
$0.986 + 0.0004 \cdot P_v$	$0.9848 + 0.0015 \cdot P_v$	$0.2 \leq NDVI \leq 0.5$
0.990	0.9668	$NDVI > 0.5$


 Fig. 10. Correlation of parameters with SM_{30} for the whole growing season.

correlations with SM_{30} were evaluated within the descriptive analysis. None of the parameters had a standalone significant correlation with SM_{30} (see Fig. 10). Piecewise linear regression was utilized to analyze the relationship between the parameters and SM_{30} . The parameter selection procedure, which is carried out in all piecewise linear regression analysis, was demonstrated using the whole growing season dataset. Regarding the correlation graph given in Fig. 10, LST^2 had the highest absolute correlation value of 0.41. LST^2 was chosen as the primary parameter of regression. The selection of the second parameter depends on both the correlation rate to SM_{30} and its weakness of covariance with the primary parameter, LST^2 . A statistical significance value of S is defined as [45]

$$S = \max \left\{ \text{ABS} \left[\frac{R^2(EP, x)}{\text{cor}(PP, x)} \right] \right\} \quad (9)$$

where EP and PP represent the estimation parameter and primary parameter, respectively. The significance values of the parameters with respect to SM_{30} as the estimation parameter and LST^2 as the primary parameter are given in Fig. 11.

$\Delta\text{-NDVI}_1$ and $\Delta\text{-NDVI}_2$ had the highest values and were selected as the second and third parameters, respectively, based on the maximum S values. When the regression analysis of the selected two and three parameters were examined, R^2 was still not significant enough for the estimation of SM_{30} and adding more parameters to the regression analysis did not increase the performance (see Fig 12 and Table VII).

The absence of a nonlinearity context, even though the parameters were chosen to reflect the nonlinearity, affect the performance of the estimation model. Therefore, we used context-aware data clustering, with the phenology context, to represent the overall nonlinear model with piecewise linear models.

The same dataset was clustered into seven clusters using the PS of each reference parcel as context. For every cluster, the relation of the parameters to SM_{30} was analyzed. Each parameter

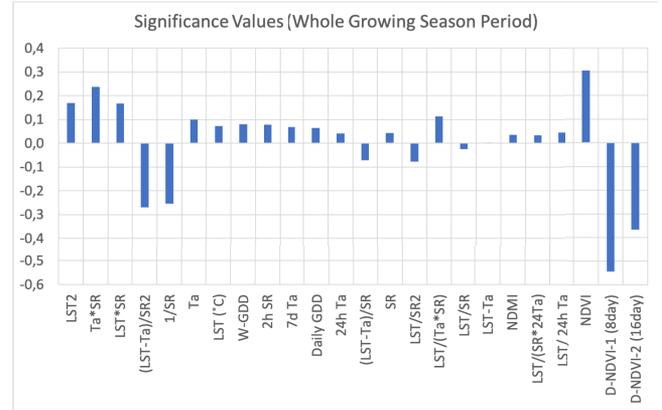
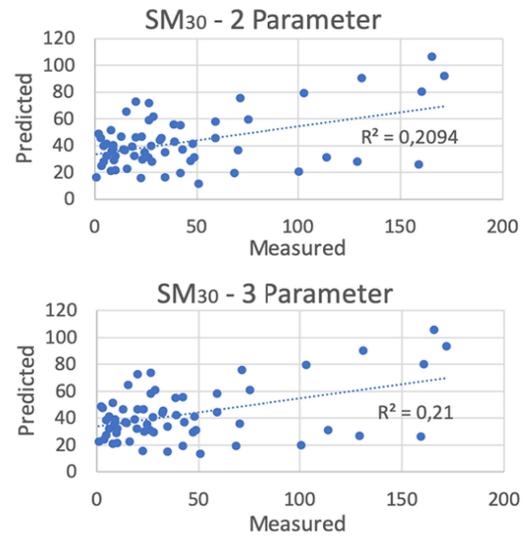

 Fig. 11. Significance values of parameters with respect to SM_{30} .


Fig. 12. Trendline of linear regression when (a) two and (b) three parameters with the highest correlation and minimum covariance were used.

 TABLE VII
 R^2 VALUES FOR THE WHOLE SEASON ANALYSIS BEFORE CONTEXT-AWARE DATA CLUSTERING

Parameter Quantity	Soil Moisture (SM_{30})	R^2
2	$17.05 + 0.04 \cdot (LST^2) - 772.66 \cdot (\Delta\text{-NDVI}_1)$	0.21
3	$17.24 + 0.04 \cdot (LST^2) - 608.73 \cdot (\Delta\text{-NDVI}_1) - 206.87 \cdot (\Delta\text{-NDVI}_2)$	0.21

had different characteristics based on the PS analyzed. PS-based correlation variations of some selected parameters are given in Fig 13; note that the soil moisture values in the graphs are tensiometric values. The lower soil moisture values indicate adequate wetness and saturation, whereas the higher values imply dry soil conditions with possible plant stress.

Clearly, a parameter can have minimum correlation in one PS and maximum correlation in another PS. In the first phenological phases, the soil moisture can be directly estimated from the soil-based variabilities. For example, $1/SR$ has one of the highest

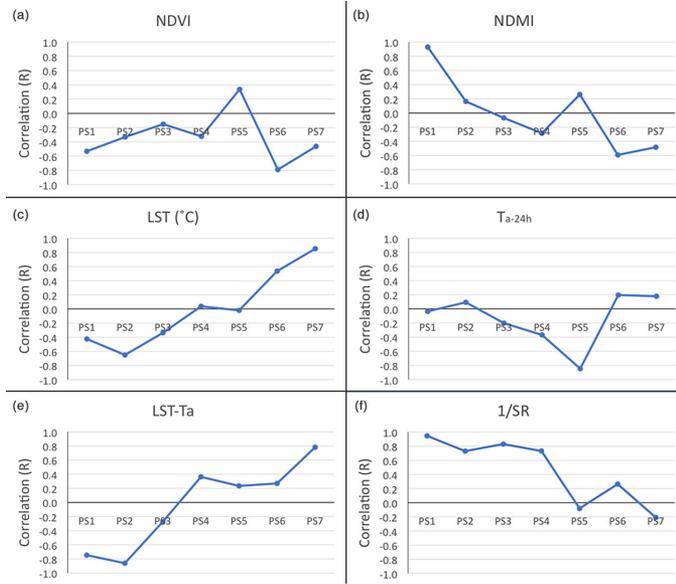


Fig. 13. Correlation of selected parameters with SM30. (a) NDVI. (b) NDMI. (c) LST ($^{\circ}\text{C}$). (d) $T_{a-24\text{ h}}$. (e) $(\text{LST}-T_a)$. (f) $(1/\text{SR})$.

correlation values for the first PS. In these stages, SR is directly interacting with the soil. As the vegetative part of the plant grows and prevents the direct estimation of the soil moisture, the correlation decreases. This implies that a higher SR decreases soil moisture content, as expected.

As the plant grows, the vegetative coverage of the soil surface increases, and it becomes impossible to use only soil-based parameters. However, other parameters, such as $T_{a-24\text{ h}}$, become dominant over soil-based ones for the estimation of root-zone soil moisture. In the last stages, the vegetative characteristics of the crop changes and the dominance of the parameters change accordingly.

As discussed in Section III-B, PS4 and PS5 correspond to transition periods showing both winter and summer characteristics at the higher latitudes because of climate variances. Therefore, PS3, PS4, and PS5 were clustered again using the soil–air temperature context for the $[(T_s - T_a) > +1\text{ }^{\circ}\text{C}]$ values for further analysis.

VII. RESULTS

As highlighted in the analysis section, the nonlinearity of a plant's growing period is one of the main difficulties associated with estimating soil moisture. Therefore, the dataset was clustered using the phenology context. When PS clusters were evaluated, the R^2 values became significant enough for soil moisture analysis, with an average R^2 value of 0.84 (see Table VIII). Although the parameters did not have significant correlation values considering the whole growing season, they may become significant within the context. This enabled the linear evaluation of a nonlinear system. Furthermore, these findings confirmed that the heat and mass transfer by convection or conduction was predominantly driven by sensible heat components of SR and LST, especially in the first and the last phenological phases.

TABLE VIII
 R^2 VALUES WITH RESPECT TO SM_{30} AFTER DATA CLUSTERING BASED ON PHENOLOGICAL CONTEXT

PS Stage	Soil Moisture (SM_{30})	R^2
PS1	$48.64 + 711.98 \cdot (1/\text{SR}) + 282.75 \cdot \text{NDMI}$	0.93
PS2	$106.84 - 3.14 \cdot (\text{LST}-T_a) - 75.62 \cdot \text{NDVI}$	0.82
PS3	$11.98 + 2307.74 \cdot (\text{LST}/\text{SR}^2)$	0.87
PS4	$11.37 + 31776.518 \cdot [(\text{LST}-T_a)/\text{SR}^2]$	0.82
PS5	$165.51 - 11.62 \cdot (T_{a-24\text{h}}) + 970.87 \cdot (\Delta\text{-NDVI}_2)$	0.80
PS6	$84.40 - 178.03 \cdot 2307\text{NDVI} + 3.79 \cdot \text{LST}$	0.83
PS7	$-286.17 + 8.96 \cdot \text{LST} + 0.008 \cdot (T_a \cdot \text{SR})$	0.81

TABLE IX
 R^2 VALUES OF GROUPED PSS WITH RESPECT TO SM_{30}

PS Stages	Soil Moisture (SM_{30})	R^2
PS1-2	$23.05 + 1110 \cdot (1/\text{SR})$	0.91
PS6-7	$-113.86 + 6.53 \cdot \text{LST}$	0.80

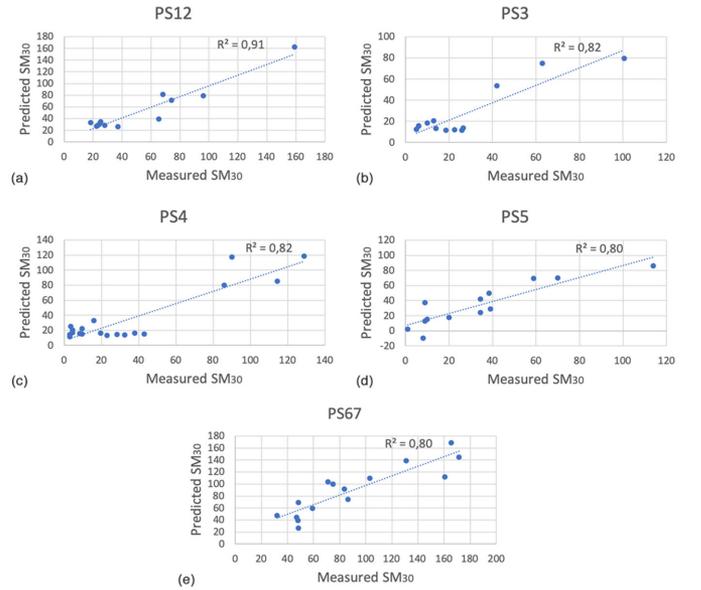


Fig. 14. Linear regression trendline of clusters where phenological stage is the context: Phenological Stage (a) 1 and 2 (PS1–2), (b) 3 (PS3), (c) 4 (PS4), (d) 5 (PS5), and (e) 6 and 7 (PS6–7).

In the first phenological phases, the parameters that correspond with the direct interaction with soil were more adequate. As plants grow, evapotranspiration and the cooling effect demonstrated a significant influence, as represented with the $(\text{LST}-T_a)$ parameter.

The dataset was clustered once more using the PS context. This time, PSs with similar characteristic responses to soil moisture content were clustered together. Two new clusters were defined: PS1 and PS2 (PS1–2) as one cluster and PS6 and PS7 (PS6–7) as another cluster.

Less dependent parameters were needed when combinations of PSs were used as clusters. Also, the R^2 values of 0.91% and 80% were still significant enough for soil moisture estimation considering different climatic conditions (see Table IX and Fig. 14).

TABLE X
 R^2 VALUES OF STAGES AFTER $(T_s - T_a)$ CONTEXT CLUSTERING WITH
 RESPECT TO SM_{30}

PS Stages	Soil Moisture (SM_{30})	R^2
PS3	$5.60 + 2520.89 \cdot (LST/SR^2)$	0.95
PS4	$5.85 + 31388.32 \cdot (LST/SR^2)$	0.78
PS5	$-273.16 + 420.93 \cdot NDVI$	0.84

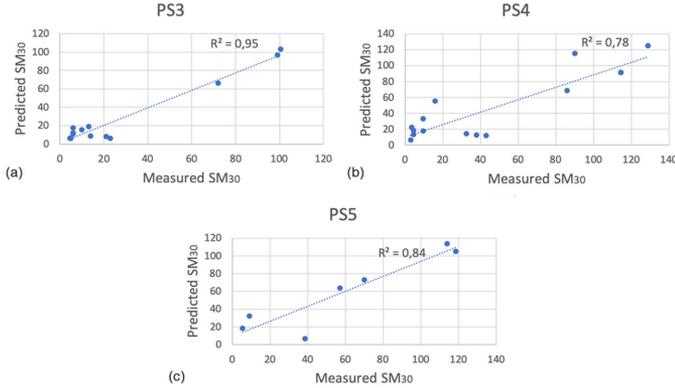


Fig. 15. Trendline of linear regression for $(T_s - T_a)$ context clustering of PSs (a) 3 (PS3), (b) 4 (PS4), and (c) 5 (PS5).

The $(T_s - T_a)$ analysis in Section III-B showed that the most complex PSs were PS3, PS4, and PS5. This complexity can be explained by the fact that the reflecting energy is a representation of both soil and vegetation with different percentages and increasing density of vegetation. Furthermore, the difference in climatic and topographic conditions increases this complexity. Therefore, the $(T_s - T_a)$ context is used as a second context to cluster the PS3, PS4, and PS5 clusters once more. Further statistical tests revealed a performance improvement for PS3 and PS5. Although the findings show a decrease in the performance of PS4, with 78% R^2 , it was still significant under the variability of the topographic and climatic conditions (see Table X and Fig. 15). These R^2 values were also shown to be possible using only one dependent parameter with the application of a second context-aware data clustering process. As shown, NDVI becomes significant as the LAI increases and the reflectance only represent the vegetation instead of soil. Moreover, the influence of LST and SR-based parameters on soil moisture based on the plant energy and water processes can be investigated from the equations.

The overall R^2 values of the clusters varied between 78% and 95%. When each cluster's regression statistics were examined, the highest "significance F " value was 0.003, indicating that the results were statistically significant. The p -values of each parameter together with the intercepts were less than 0.07. Matei *et al.* [39] proposed a data mining system using agrometeorological data from weather stations to estimate real-time soil moisture. In their proposed system, the average accuracies of various machine learning algorithms were 68%. Pradhan [54] proposed a method for growing season root-zone soil moisture using satellite-driven vegetation indices and physical properties

of soil moisture without the need for *in-situ* measurements with an average R^2 of 72%.

The context-aware data clustering approach can be used to explain a nonlinear system with piecewise linearity. While performing statistical analysis, the outliers of the nonlinear system became relevant within the context. As a result, even though the performance of a system is expected to decline with a decreasing number of training samples, the context-aware data clustering was shown to boost performance.

VIII. CONCLUSION

In this study, plants were used as a sensory mechanism for defining data clustering contexts based on phenology and soil-air temperature difference for reliable root-zone soil moisture retrieval in large-scale areas. Reference soil moisture and soil temperature data were only used for the model development, whereas the constituted model was formed by using *in-situ* measurements of solar radiation and atmospheric temperature with remote sensing parameters of NDVI, NDMI, and LST. Importantly, root-zone soil moisture at 30-cm depth was modeled without using soil structure and characteristics data for large-scale agricultural areas.

The performance of the proposed model was evaluated by using 16 rainfed winter wheat parcels distributed at different locations in Turkey. Wide-spread distribution of the parcels highlighted the convenience of the model under different climatic and soil conditions.

No significant correlation was found between the parameters and root-zone soil moisture before context-aware data clustering. Hence, the correlation of linear regression was only 21%. Thereafter, phenological-stage-based clusters demonstrated a significant performance improvement with 80% to 93% coefficient of determination. In light of these observations, we confirmed that clustering using PSs enables linear analysis and can be used as a solution for modeling the nonlinear character of a plant's life cycle. The study has also shown that it is possible to group PSs based on their characteristics without any performance loss.

A secondary data clustering context, the soil-air temperature difference, was proposed for phenological clusters that are affected by different vegetation density percentages and winter-summer transition temperatures. Further analysis performed on the PS3, PS4, and PS5 demonstrated a general improvement up to 95% coefficient of determination. Chronological timing of the biological systems is not the same every year. As a result, the PSs of agricultural products shift. This is mostly owing to the change in the seasons and the climatic conditions resulting in variations in the atmosphere and soil temperatures. The analysis provided an insight into this potential change from year to year by defining a context out of it.

The calculation of both contexts was possible using remote sensing techniques. After analyzing the time-series dataset as a whole, each cluster was analyzed individually. Usually, the reduced number of training samples is expected to have a negative impact on the performance of the estimation model as a result of clustering. In this case, including a context-aware data

clustering step before an estimation model enhanced the model's performance. The context-aware data clustering allows the use of piecewise linear regression analysis within clusters, as the defined contexts explain the time variance. Therefore, the model was obtained using the piecewise linear regression analysis. The proposed method is applicable for different agricultural applications using various estimation techniques, including machine learning, owing to the nature of context awareness.

Cloud cover can negatively affect passive remote sensing systems in springtime, which can also affect the number of sampling days and the continuity of the data. Another drawback of the proposed model is the time delay between the root-zone soil moisture level and vegetation moisture level and the time delay of observing these effects with remote sensing technologies. We plan to compensate for this by performing a time-series analysis of wavelet coefficients of the identified parameters in the next phase of this study. We also intend to perform a performance comparison by implementing a machine-learning-based estimation model in a future study. The data used in this study only cover one growing season of a particular crop type. The effect of using data from more than one season and the performance of the model when more crop types are introduced to the model are also planned for future studies.

REFERENCES

- [1] G. Dumedah, J. P. Walker, and O. Merlin, "Root-zone soil moisture estimation from assimilation of downscaled soil moisture and ocean salinity data," *Adv. Water Resour.*, vol. 84, pp. 14–22, 2015.
- [2] D. A. Robinson *et al.*, "Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review," *Vadose Zone J.*, vol. 7, no. 1, pp. 358–389, 2008.
- [3] G. C. Heathman, P. J. Starks, L. R. Ahuja, and T. J. Jackson, "Assimilation of surface soil moisture to estimate profile soil water content," *J. Hydrol.*, vol. 279, no. 1, pp. 1–17, 2003.
- [4] G. Petropoulos, *Remote Sensing of Energy Fluxes and Soil Moisture Content*. Boca Raton, FL, USA: CRC Press, 2013.
- [5] G. C. Topp, "State-of-the-art of measuring soil water content," *Hydrol. Process.*, vol. 17, no. 14, pp. 2993–2996, 2003.
- [6] H. R. Bogaen, J. A. Huisman, C. Oberdörster, and H. Vereecken, "Evaluation of a low-cost soil water content sensor for wireless network applications," *J. Hydrol.*, vol. 344, no. 1, pp. 32–42, 2007.
- [7] P. K. Srivastava, D. Han, M. A. R. Ramirez, and T. Islam, "Appraisal of SMOS soil moisture at a catchment scale in a temperate maritime climate," *J. Hydrol.*, vol. 498, pp. 292–304, 2013.
- [8] A. de Tomás, H. Nieto, R. Guzinski, J. Salas, I. Sandholt, and P. Berliner, "Validation and scale dependencies of the triangle method for the evaporative fraction estimation over heterogeneous areas," *Remote Sens. Environ.*, vol. 152, pp. 493–511, 2014.
- [9] G. P. Petropoulos, G. Ireland, and B. Barrett, "Surface soil moisture retrievals from remote sensing: Current status, products & future trends," *Phys. Chem. Earth, A/B/C*, vol. 83–84, pp. 36–56, 2015.
- [10] M. Minacapilli *et al.*, "Estimation of actual evapotranspiration of Mediterranean perennial crops by means of remote-sensing based surface energy balance models," *Hydrol. Earth Syst. Sci.*, vol. 13, no. 7, pp. 1061–1074, 2009.
- [11] R. Filion *et al.*, "Remote sensing for mapping soil moisture and drainage potential in semi-arid regions: Applications to the Campidano Plain of Sardinia, Italy," *Sci. Total Environ.*, vol. 543, pp. 862–876, 2016.
- [12] M. Nocita, A. Stevens, C. Noon, and B. van Wesemael, "Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy," *Geoderma*, vol. 199, pp. 37–42, 2013.
- [13] L. Weidong, F. Baret, G. Xingfa, T. Qingxi, Z. Lanfen, and Z. Bing, "Relating soil surface moisture to reflectance," *Remote Sens. Environ.*, vol. 81, no. 2, pp. 238–246, 2002.
- [14] W. T. H. Liu, O. Massambani, and C. A. Nobre, "Satellite recorded vegetation response to drought in Brazil," *Int. J. Climatol.*, vol. 14, no. 3, pp. 343–354, 1994.
- [15] F. N. Kogan, "Application of vegetation index and brightness temperature for drought detection," *Adv. Sp. Res.*, vol. 15, no. 11, pp. 91–100, 1995.
- [16] B. Gao, "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, 1996.
- [17] L. Wang and J. J. Qu, "NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing," *Geophys. Res. Lett.*, vol. 34, no. 20, 2007, Art. no. L20405.
- [18] S. Lei, Z. Bian, J. L. Daniels, and D. Liu, "Improved spatial resolution in soil moisture retrieval at arid mining area using apparent thermal inertia," *Trans. Nonferrous Metals. Soc. China.*, vol. 24, no. 6, pp. 1866–1873, 2014.
- [19] J. Qin, K. Yang, N. Lu, Y. Chen, L. Zhao, and M. Han, "Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia," *Remote Sens. Environ.*, vol. 138, pp. 1–9, 2013.
- [20] S. B. Idso, R. D. Jackson, P. J. Pinter, R. J. Reginato, and J. L. Hatfield, "Normalizing the stress-degree-day parameter for environmental variability," *Agric. Meteorol.*, vol. 24, pp. 45–55, 1981.
- [21] R. Panciera *et al.*, "The soil moisture active passive experiments (SMAPEx): Toward soil moisture retrieval from the SMAP mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 490–507, Jan. 2014.
- [22] M. Pan, A. K. Sahoo, and E. F. Wood, "Improving soil moisture retrievals from a physically-based radiative transfer model," *Remote Sens. Environ.*, vol. 140, pp. 130–140, 2014.
- [23] H. Wang, "Soil moisture retrieval from microwave remote sensing observations," *Soil Moisture*. London, U.K.: IntechOpen, 2018.
- [24] W. Wagner, K. Scipal, C. Pathe, D. Gerten, W. Lucht, and B. Rudolf, "Evaluation of the agreement between the first global remotely sensed soil moisture data with model and precipitation data," *J. Geophys. Res.*, vol. 108, no. D19, 2003, Art. no. 4611.
- [25] W. Wagner *et al.*, "The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications," *Meteorol. Zeitschrift*, vol. 22, no. 1, pp. 5–33, 2013.
- [26] K. Imaoka, M. Kachi, M. Kasahara, N. Ito, K. Nakagawa, and T. Oki, "Instrument performance and calibration of AMSR-E and AMSR2," *ISPRS Arch.*, vol. XXXVIII, pp. 13–16, 2010.
- [27] E. Jacquette *et al.*, "SMOS CATDS level 3 global products over land," *Proc. SPIE*, vol. 7824, 2010, Art. no. 78240K.
- [28] D. Entekhabi *et al.*, "The soil moisture active passive (SMAP) mission," *Proc. IEEE*, vol. 98, no. 5, pp. 704–716, May 2010.
- [29] D. Zhang and G. Zhou, "Estimation of soil moisture from optical and thermal remote sensing: A review," *Sensors*, vol. 16, no. 8, 2016, Art. no. 1308.
- [30] S. Louvet *et al.*, "SMOS soil moisture product evaluation over West-Africa from local to regional scale," *Remote Sens. Environ.*, vol. 156, pp. 383–394, 2015.
- [31] M. Owe, R. de Jeu, and T. Holmes, "Multisensor historical climatology of satellite-derived global land surface moisture," *J. Geophys. Res. Earth Surf.*, vol. 113, no. F1, 2008, Art. no. F01002.
- [32] M. Piles *et al.*, "A downscaling approach for SMOS land observations: Evaluation of high-resolution soil moisture maps over the Iberian Peninsula," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 7, no. 9, pp. 3845–3857, Sep. 2014.
- [33] N. N. Das, D. Entekhabi, R. S. Dunbar, E. G. Njoku, and S. H. Yueh, "Uncertainty estimates in the SMAP combined active & passive downscaled brightness temperature," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 640–650, Feb. 2016.
- [34] S. A. Margulis, D. McLaughlin, D. Entekhabi, and S. Dunne, "Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment," *Water Resour. Res.*, vol. 38, pp. 35:1–35:18, 2002.
- [35] R. H. Reichle, W. T. Crow, R. D. Koster, H. O. Sharif, and S. P. P. Mahanama, "Contribution of soil moisture retrievals to land data assimilation products," *Geophys. Res. Lett.*, vol. 35, no. 1, 2008, Art. no. L01404.
- [36] O. Merlin, A. Chehbouni, G. Boulet, and Y. Kerr, "Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data," *J. Hydrometeorol.*, vol. 7, no. 6, pp. 1308–1322, 2006.
- [37] Z. Van Arkel and A. L. Kaleita, "Identifying sampling locations for field-scale soil moisture estimation using K-means clustering," *Water Resour. Res.*, vol. 50, no. 8, pp. 7050–7057, 2014.

- [38] E. Lee and S. Kim, "Pattern similarity based soil moisture analysis for three seasons on a steep hillslope," *J. Hydrol.*, vol. 551, pp. 484–494, 2017.
- [39] O. Matei, T. Rusu, A. Petrovan, and G. Mihaş, "A data mining system for real time soil moisture prediction," *Procedia Eng.*, vol. 181, pp. 837–844, 2017.
- [40] A. Avram, O. Matei, C.-M. Pinte, P. C. Pop, and C. A. Anton, "Context-aware data mining vs. classical data mining: Case study on predicting soil moisture," in *Proc. 14th Int. Conf. Soft Comput. Models Ind. Environ. Appl.*, 2020, pp. 199–208.
- [41] H. L. Penman, "Natural evaporation from open water, bare soil and grass," *Proc. Roy. Soc. London. A, Math. Phys. Sci.*, vol. 193, no. 1032, pp. 120–145, 1948.
- [42] H. G. Jones, *Plants and Microclimate*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
- [43] H. G. Jones and P. Schofield, "Thermal and other remote sensing of plant stress," *General Appl. Plant Physiol.*, vol. 34, no. 1/2, pp. 19–32, 2008.
- [44] S. O. Ihuoma and C. A. Madramootoo, "Recent advances in crop water stress detection," *Comput. Electron. Agric.*, vol. 141, pp. 267–275, 2017.
- [45] B. Üstündağ, "An adaptive Mealy machine model for monitoring crop status," *J. Integr. Agric.*, vol. 16, no. 2, pp. 252–265, 2017.
- [46] E. Acevedo, P. Silva, and H. Silva, "Wheat growth and physiology," in *Bread Wheat*, B. C. Curtis, S. Rajaram, and H. Gómez Macpherson, Eds. Rome, Italy: FAO, 2002.
- [47] S. Sensoy, M. Demircan, and Y. Ulupinar, "Climate of Turkey," Turkish State Meteorol. Serv., Ankara, Turkey, 2008.
- [48] İ. Özberk, S. Atay, F. Altay, E. Cabi, H. Özkan, and A. Atlı, "Türkiye'nin Buğday Atlası," *İstanbul: WWF-Türkiye*, Sep. 2016.
- [49] P. Miller, W. Lanier, and S. Brandt, "Using growing degree days to predict plant stages," *Montana State Univ. Ext. Serv. MT00103 AG 7/2001*, vol. 9, 2001.
- [50] S. Bağış and B. B. Üstündağ, "Image based automated phenological stage detection of cereal plants," in *Proc. 1st Int. Conf. Agro-Geoinform.*, Shanghai, 2012, pp. 1–6, doi: [10.1109/Agro-Geoinformatics.2012.6311643](https://doi.org/10.1109/Agro-Geoinformatics.2012.6311643).
- [51] N. Rodríguez-Fernández, A. Al Bitar, A. Colliander, and T. Zhao, "Soil moisture remote sensing across scales," *Remote Sens.*, vol. 11, no. 2, 2019, Art. no. 190.
- [52] Q. Weng, D. Lu, and J. Schubring, "Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies," *Remote Sens. Environ.*, vol. 89, no. 4, pp. 467–483, 2004.
- [53] F. Wang, Z. Qin, C. Song, L. Tu, A. Karnieli, and S. Zhao, "An improved mono-window algorithm for land surface temperature retrieval from Landsat 8 thermal infrared sensor data," *Remote Sens.*, vol. 7, no. 4, pp. 4268–4289, 2015.
- [54] N. R. Pradhan, "Estimating growing-season root zone soil moisture from vegetation index-based evapotranspiration fraction and soil properties in the Northwest Mountain region, USA," *Hydrol. Sci. J.*, vol. 64, no. 7, pp. 771–788, 2019.



Ayda Aktas was born in Istanbul, Turkey, in 1982. She received the B.Sc. and M.Sc. degrees in geomatics engineering from Istanbul Technical University (ITU), Istanbul, Turkey, in 2004 and 2007, respectively. She also graduated from the Satellite Communication and Remote Sensing M.Sc. Program from ITU, in 2008. She is currently working toward the Ph.D. degree in satellite communication and remote sensing program with ITU.

From 2004 to 2006, she was a Student Assistant, and then from 2006 to 2009, she was a Specialist with the ITU Research and Application Center for Satellite Communications and Remote Sensing. She is currently a Lecturer with the Informatics Institute, ITU. Her research interests include change detection, agricultural spatial analysis, vegetation phenology, and vegetation mapping.

Ms. Aktas was the recipient of the Best Poster Award at the XXth ISPRS Congress.



Burak Berk Üstündağ (Member, IEEE) was born in Istanbul, Turkey, in 1970. He received the B.Sc. degree in electrical engineering, in 1991, and the M.Sc., and Ph.D. degrees in control systems and computer engineering from Istanbul Technical University (ITU), Istanbul, Turkey, in 1994 and 2000, respectively.

He is a Professor with the Computer Engineering Department with the Faculty of Computer and Informatics, ITU. He was a Science and Technology Advisor to Ministers and Governmental Institutions for more than 15 years. He has more than 100 scientific publications and patents. His research interests involve data fusion, artificial intelligence, signal and data processing, global optimization, cognitive communication, biosensors, bioprocessors, control systems, agricultural information systems, and precision farming.

Dr. Üstündağ is a member of the IEEE Communication Society and Computational Intelligence.