Remote Sensing Applications: Society and Environment 35 (2024) 101264

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

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

Added utility of temperature zone information in remote sensing-based large scale crop mapping

E. Donmez [a,b,*], M.T. Yilmaz [a], I. Yucel [a]

[a] Civil Engineering Department, Middle East Technical University, Ankara, Turkiye
[b] University of Bonn, The Institute of Food and Resource Economics, Nussallee 21 D-53115 Bonn, Germany

ARTICLE INFO

Keywords:
Remote sensing
Study area division
Crop cover mapping
Machine learning
Supervised classification
Temperature zones

ABSTRACT

Accurate crop cover maps are beneficial for various aspects like water resources management, crop yield prediction, regulation insurance policies, and investigation of the effects of climate change. When making large-scale crop classification, regional harvest time and phenological growth differences occur due to varying temperatures along the study area, and using regional temperature differences while performing crop cover classification may increase the map accuracy. Therefore, in this study, we investigated for the first time the contribution of temperature information over large areas to the classification of agricultural products. Agricultural crop mapping is performed over Türkiye using Sentinel-2 Level-2A images with 10-m spatial resolution acquired between March 15, 2019, and October 15, 2019. In addition to spectral bands, Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) are used as classification features. Twenty years of ERA5-Land 2-m temperature data is averaged to divide the study area into three temperature zones Low (LTZ), Medium (MTZ), and High-Temperature Zone (HTZ). Before the classification, feature selection using random forest importance is performed to select the most successful features. After that, a random forest classifier is created for each temperature zone. LTZ reached 89% overall accuracy (OA) with a 0.88 Kappa. MTZ reached 91% OA with 0.92 Kappa, and HTZ reached 94% OA with 0.94 Kappa, giving the best accuracy among the classifiers. Finally, test sets of all temperature zones are combined, and an OA of 92% with a Kappa of 0.93 is achieved with this combined test set. To test the advantage of temperature zoning, classification is also performed without the temperature zones, and it is observed that temperature zoning increases the OA and Kappa by 1%. A land cover classification map is then created using temperature zone classifiers with 34 crop classes and six non-agricultural classes.

1. Introduction

Accurate information on seasonal or yearly agricultural land cover benefits many aspects, like water resources management, agricultural insurance, crop yield prediction, and planning of domestic and foreign policies and actions for governments (Bauer et al., 1978). Agriculture is the biggest consumer of freshwater, with almost 75% of current human water consumption worldwide (Wallace, 2000). The increasing population of the world (UNDESA, 2017) and the increasing water use, which exceeds twice that of the population increase (Steduto et al., 2012), show how necessary it is to manage water resources with extreme caution. Creating accurate

* Corresponding author. Nussallee 21, D-53115, Bonn, Germany.
E-mail address: elif.donmez@ilr.uni-bonn.de (E. Donmez).

https://doi.org/10.1016/j.rsase.2024.101264
Received 16 August 2023; Received in revised form 18 May 2024; Accepted 31 May 2024
Available online 1 June 2024

agricultural maps and interpreting these maps can be steps toward the well-planned use of water resources. The creation of nationwide crop maps, in addition to its contribution to water resources planning, can assist policymakers in agricultural development plans, which could be a step towards more sustainable and efficient agricultural policies.

Remote sensing-based crop classification has been performed over different regions in Türkiye (e.g., Abdikan et al., 2018; Alganci et al., 2013; Turker and Arikan, 2005). However, country-level mapping studies remain missing. Ministry of Agriculture and Forestry has datasets showing ground-truth conditions, but these datasets depend on farmers' declarations which may not reflect the reality, particularly during disaster years. There are also large-scale landcover mapping projects like CORINE (Coordination of Information on The Environment), but these projects lack local information and features unique to the country (Özür and Ataol, 2018). So, there is a need for an objective methodology to obtain nationwide agricultural crop classification maps.

When constructing remote sensing-based, nationwide, or large-scale crop maps, it should be considered that the subject region may have different zones that can contain different crop types or growth stages of certain crops. Although Türkiye is located in the Mediterranean geography with relatively temperate climatic conditions, significant differences in climatic conditions from one region to another occur due to the variable nature of the landscape and especially the presence of mountains extending parallel to the coast (Sensoy et al., 2008). Previous studies show that phenology correlates highly with temperature (Siebert and Ewert, 2012; Zhang and Tao, 2013). Supporting the information in the literature, when ground truth data used in this study is inspected, major crop types like wheat, barley, and potatoes had different growing patterns in different regions. In Fig. 1 this phenology difference is illustrated by the Normalized Difference Vegetation Index (NDVI) (Bremer et al., 2011) time series data of two wheat pixels derived from Sentinel-2 (European Space Agency, 2018) data. One of the pixels is located in the east of Türkiye, while the other pixel is located in the west of Türkiye, where average temperatures are much higher than in the east of Türkiye, as shown in the same figure with the temperature profiles obtained from 2-m surface temperature band of ERA5. There is a visible difference between the growing periods of two pixels in different temperature conditions.

Studies in the literature have used a variety of characteristics affecting crop growth across regions to improve classification accuracy, including climatic regions (e.g., Inglada et al., 2017), landscape regions (e.g., Asam et al., 2022) and agro-climatic regions (e.g., Arias et al., 2020). However, there is currently a lack in the literature regarding how to divide the study area using computationally inexpensive preprocessing and publicly available and widely accessible ERA5-Land temperature data to produce a crop classification map with better accuracy. Therefore, the primary objective of this study is to investigate the additional value of dividing the study area according to temperature data. And the secondary objective is to utilize state-of-the-art methods and tools to ensure the accuracy of the resultant crop map.

This study aims to investigate the added utility of temperature data-based study area division for the first time to perform a more accurate country-level high-resolution crop map of 2019, including 40 land cover classes with 10 m spatial resolution over Türkiye. Apart from utilizing state-of-the-art methods and tools, this study pays attention to Türkiye's crop production and climatic factors shaping the growth stages of targeted crops to represent the land cover more accurately.

2. Methodology

In this part of the article, the methodology for crop classification is explained. A flowchart is given in Fig. 2 for easier follow-up of the methodology sequence. In the next subchapters of the methodology, the box number for each step is given in parentheses as they are explained in the text.

2.1. Study area

Türkiye is located in Anatolia in Western Asia, with a portion in Southeast Europe, between latitudes from 35.9 to 42.0 Northern and longitudes from 25.9 to 44.6 Eastern. Although Türkiye is situated in a Mediterranean region with relatively temperate climate, the varied terrain, particularly the presence of mountains running parallel to the coast, causes notable variations in climatic conditions from one region to another (Sensoy et al., 2008). Due to its favorable geography, fertile soil structure, and suitable climatic conditions, various crop products such as vegetables, grains, and cotton are grown in almost every part of the country (Cakirli Akyüz and

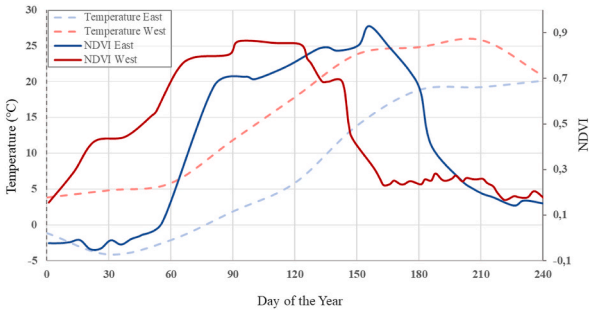

Fig. 1. NDVI and Temperature vs. Day of the Year for Two Wheat Pixels. Profiles shown in Blue is Located in Eastern Türkiye (37.473°E,39.182°N), While the Red Ones are in Western Türkiye (26.424°E,41.171°N). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

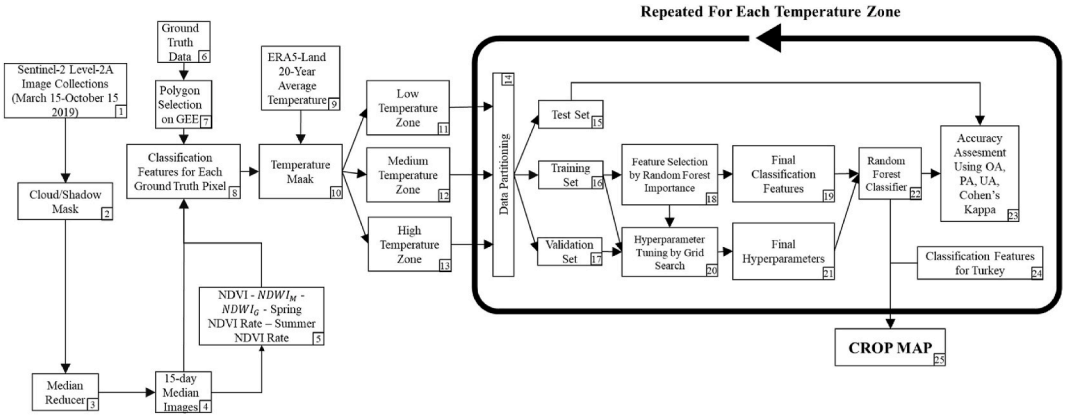

Fig. 2. Methodology flowchart.

Theuvsen, 2021). According to Plant Production Data provided by the Turkish Statistics Institute (2020), 29.5% of Türkiye's land was used for agricultural production in 2019.

2.2. Ground truth data

Ground truth data (Box 6 in Fig. 2) is received from the observations obtained from the firm named Tarla.io (https://tarla.io/eng/index.html). This company, which provides agricultural risk management services, does business with many farmer customers by obtaining information from the field where they plant crops. For this reason, the information they provide is considered to be reliable. This dataset is not publicly released online; however, it can be accessed through the company Tarla.io per request. In its raw form, data consist of parcel coordinate information of 105 types of crop fields all over Türkiye. To perform supervised classification, for which the classes should be determined prior to the classification, crop classes are selected with the help of agricultural production data for 2018 showing production of crops in tons of crops provided online by the Turkish Statistics Institute (2018). Crops are sorted according to the total ton of production, and the top 34 crops with an adequate number of data are selected as classes for the crop map. These selected crops take up to 86% of the total ton of production, and the remaining crops are not included in the study since they have lower production, thus assumed to cover less area than selected crops. In addition to the crop classes, six non-crop classes: greenhouse, bare soil, water, urban, steppe, and forest, are chosen to represent the study area more accurately. The final class names are shown in the legend of Fig. 3. To obtain homogeneous training and test data, provided ground truth data is visualized on the Google Earth Engine platform (GEE) (Gorelick et al., 2017), and rectangular polygons are selected manually from this dataset (Box 7 in Fig. 2). Geometry size and the pixel number in the polygons change according to the original field size. Data for non-agricultural classes are selected manually by visualizing selected scenes at the high-resolution satellite view of GEE. After the data selection, ground truth data is exported from GEE in shapefile format for classification.

2.3. Sentinel-2 for classification features

The European Space Agency's Multispectral Instrument on the Sentinel-2 satellite has been providing multispectral images with 10-m spatial resolution since 2015 on a global extent. The Sentinel-2 mission includes two satellites developed to monitor vegetation, land cover, and the environment. The Sentinel-2A and Sentinel-2B satellites operate in a sun-synchronous orbit with a 10-day temporal resolution, and they cover Earth's surface in 5-day intervals.

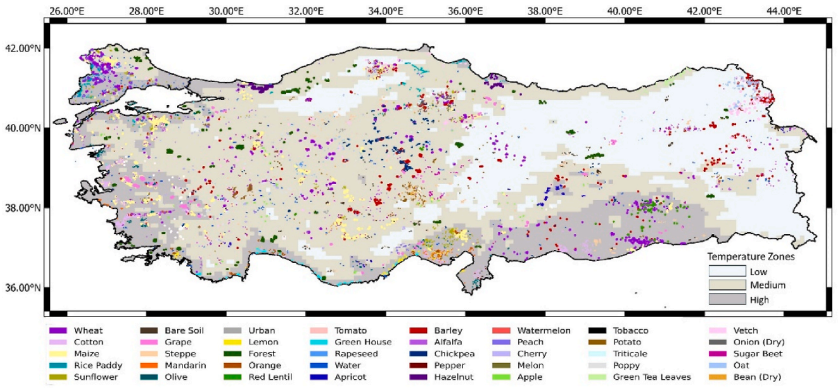

Fig. 3. Distribution of agricultural and non-agricultural class pixels on the resultant temperature map.

Predictor variables are usually called features in classification and pattern recognition literature (Hastie et al., 2009). To obtain the classification features for this study, from March 15, 2019, to October 15, 2019, for each 15-day interval, tiles of Sentinel-2 Level-2A are collected to create 15-day image collections over the whole study area (Box 1 in Fig. 2). After that, images in each image collection are cleared from cloud cover using the S2 cloud probability dataset (s2cloudless) on GEE with a maximum cloud probability of 20% (Box 2 in Fig. 2). Cloud shadow pixels are also masked from the images by considering cloud projection intersection with low-reflectance NIR pixels, as demonstrated by Braaten (n.d.) (Box 2 in Fig. 2). After the cloud masking process, cloud and cloud shadow pixels in the original satellite images have NA values. Since the Sentinel-2 Level-2 products have an image every five days, each part of the study area has an image acquired every five days. In order to reduce the computational cost, after obtaining image collections that are cleared from cloud and cloud shadow pixels, median of the images that cover the same area is taken for each 15-day interval using the median reducer tool on GEE (Box 3 in Fig. 2). This way, the computational cost is decreased by decreasing the number of images used in the study, and spectral changes are captured twice a month. As a result, 14 images are obtained for classification as one image for every 15-day interval between March 15 – October 15, 2019. All bands of obtained images are exported separately for each image from GEE except for Band 10 (short wave infrared – cirrus), which is used only for cloud probability calculation before image reduction. Finally, 12 bands of each 14 images (168 spectral bands in total) are obtained to be used as classification inputs.

2.4. Era5-land for temperature zones

This study utilizes crop phenology to identify different crops using multi-temporal satellite imagery. Previous studies show that phenology correlates highly with temperature (Siebert and Ewert, 2012; Zhang and Tao, 2013). To capture the phenology difference occurring due to temperature differences between regions, ERA5-Land is utilized to create a temperature mask.

ERA5-Land is produced by replaying the land component of the ECMWF ERA5 climate reanalysis (Muñoz-Sabater et al., 2021). It contains regular latitude-longitude gridded record with a temporal resolution of 1 h since 1950. The horizontal coverage of the ERA5-Land product is $0.1^\circ \times 0.1^\circ$; (Native resolution is 9 km), and its vertical coverage is from 2 m above the surface level to a soil depth of 289 cm. It provides information globally with an update frequency of monthly with a delay of 2–3 months from the actual date. Some of the main variables of ERA5-Land are 2m temperature, 2m dewpoint temperature, total precipitation, and total evaporation. The data used in the study is downloaded as post-processed by monthly averaging the full ERA5-Land dataset.

For the temperature mask, the 'temperature_2 m' band of ERA5-Land monthly averaged dataset is averaged for 20 years (1999–2019) (Box 9 in Fig. 2). The reason for averaging multiple years instead of selecting the year of the map is to capture the crop planting tendencies of farmers for a more realistic crop distribution over the study area. This implementation would also reduce the sampling uncertainty in selection of a single year rather than longer-term averages. Another reason for not using only 2019 data is to be able to apply these classifiers to other years when Sentinel-2 data is available. Using only 2019 temperatures would not give representative results and thus could not be used for crop mapping in other years. The average temperature map is visualized in GEE and temperature thresholds that determine the temperature regions are selected so that the major crop plains are not divided. Different thresholds were also tested, but they resulted in a greater number of crop plains that are divided to fall into two temperature zones. As a result, the Low-Temperature Zone (LTZ) is set to have a maximum temperature of 9 °C; the Medium-Temperature Zone (MTZ) is set to have a minimum temperature of 9 °C and a maximum temperature of 14 °C. Finally, the High-Temperature Zone (HTZ) is set to have a minimum temperature of 14 °C. The temperature mask raster (Box 10 in Fig. 2) is later exported from GEE and resampled to match the 10-m spatial resolutions of Sentinel-2 data on QGIS software (QGIS.org, 2024). Due to a lack of temperature information on coastal regions resulting from spatial resolution differences between the two products, the temperature mask is extended using the R programming language (R Core Team, 2020) so that each empty pixel is filled with the value of the nearest pixel with a non-NA value. After these pre-processing steps, Sentinel-2 images and the temperature mask are resized and reprojected to have the same Coordinate Reference Systems (CRS) and boundaries keeping the 10-m spatial resolution. A total of 169 rasters (12 spectral bands for each 14 date intervals and one temperature mask) is cropped with the extent of vector coordinate data on R (Box 8 in Fig. 2), resulting in a data frame with 168 spectral reflectance values, one temperature mask value, and the land cover class number for each pixel. The data frame is later divided into three separate data frames based on the temperature zone information (Boxes 11,12 and 13 in Fig. 2). The resultant temperature map is shown in Fig. 3, along with the distribution of agricultural and non-agricultural class pixels used in the study.

In addition to separate data frames for each temperature zone, one classifier is created without separating the dataset into temperature zones. The reason for this is to test whether or not the division of the study area into temperature zones increases the classification accuracy. The details for this dataset are given in the '2.7 Data Partitioning' subtitle of the Methodology chapter.

2.5. Gap filling

After applying cloud and cloud shadow masks, inevitable gaps occur in the data. The masked pixels that make up these gaps have not available (NA) values instead of spectral reflectance values. So, these pixels become unusable since each pixel should have the same number of classification features for the classifier to work. To use as many pixels as possible, a computationally efficient temporal gap-filling procedure based on linear interpolation is followed on a temporal scale similar to the method used by Griffiths et al. (2019). Bands with empty (NA) values are filled with the average of the next and previous months' same band. Later, cloud gaps could not be filled with this method in the first 15-day image of April due to consecutive cloud pixels, and these gaps are filled with the values of the second 15-day of March. After the gap-filling process, the remaining NA pixels that could not be filled due to consecutive missing data are removed from the dataset. At the end of the gap-filling step, 28% of all ground truth data is removed while 15% is restored, allowing further use.

2.6. Spectral indices

Different land cover types have different spectral curves, meaning they reflect different portions of the incoming energy in each wavelength. Spectral indices are calculated using combinations of spectral bands that enhance specific spectral properties (Palacios-Orueta et al., 2006) (Box 5 in Fig. 2). NDVI is a commonly used spectral index that can be used to benefit from the spectral reflectance difference between near-infrared (NIR) and red wavelengths (Bremer et al., 2011). Maximum chlorophyll absorption of a green leaf occurs at about 690 nm, which corresponds to red wavelength, while at the NIR wavelength interval (650–850 nm), absorption shows a significant decrease (Myneni et al., 1995). This spectral difference is suitable for differentiation between vegetation and other classes. Also, the magnitude or time interval of maximum NDVI can be used to differentiate vegetation classes in land cover classification. The NDVI is calculated using Equation (1).

$$NDVI = (NIR - RED) / (NIR + RED) \tag{1}$$

The Normalized Difference Water Index (NDWI), proposed by McFeeters (1996) ($NDWI_M$), is a spectral index that increases the visibility of open water features in remote sensing products. $NDWI_M$ enhances the reflectance of water bodies utilizing green wavelengths and exploits the low NIR reflectance of water bodies compared to vegetation and soil. This way, water bodies that have high positive $NDWI_M$ can be identified very easily (McFeeters, 1996). The $NDWI_M$ is calculated using Equation (2).

$$NDWI_M = (GREEN - NIR) / (GREEN + NIR) \tag{2}$$

The Normalized Difference Water Index proposed by Gao (1996) ($NDWI_G$) aims to assess vegetation water content from remotely sensed data using the combination of NIR and short wave infrared (SWIR) channels (Zhang et al., 2017). Since dry soil, dry vegetation, and green vegetation have different ranges of NDWI values (Gao, 1996), $NDWI_G$ can be used to differentiate these land cover classes. The $NDWI_G$ is calculated using Equation (3).

$$NDWI_G = (NIR - SWIR) / (NIR + SWIR) \tag{3}$$

NDVI, $NDWI_G$, and $NDWI_M$ are calculated for each 15-day period and added to the classification features. In addition to vegetation indices (VIs), two features, "spring rate" and "summer rate", represent NDVI increase in spring (from the first half of May to the second half of June) and summer (from the first half of July to the second half of August) are calculated and added to classification features. With this addition, the classification algorithm has 212 input features (168 bands + 3 indices for 14 date intervals + 2 NDVI rates).

2.7. Data partitioning

To perform an unbiased classification, the dataset is divided into three subsets as the training set (~70%), the validation set (~15%), and the test set (~15%) for each crop type (Boxes 15, 16, and 17 in Fig. 2). Since ground truth polygons contain more than one pixel, dividing a polygon into more than two subsets may result in overfitting, which occurs when the classification algorithm fits the training data too well and fails to predict when classification is performed on another set of data. To avoid that, polygons are distributed to subsets so that no polygon has pixels in more than one subset. After data is separated into three independent sets, the training set is used to train the classification algorithm, the validation set is used for hyperparameter tuning, and the test set is used for the final classification accuracy assessment. If a crop class does not have sufficient ground truth data in a temperature zone to satisfy the needs for data partitioning, that crop is removed from the dataset for that temperature zone.

A dataset without temperature zones is created to test whether dividing the study area into temperature zones improves classification accuracy. For this dataset, the validation and test subsets, as well as the training subset of each temperature zone, are combined.

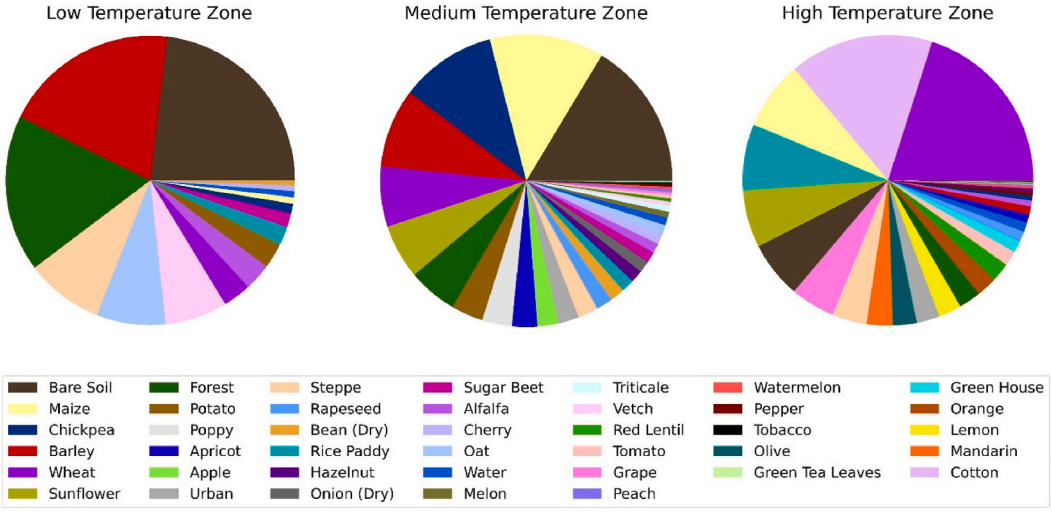

Fig. 4. Proportion of Crop classes in the ground truth dataset for each temperature zone.

Therefore, the classifier created with this data set has the number of training pixels equal to the sum of the training sets of each temperature zone. Pie charts for each temperature zone showing the proportion of each class in the ground truth set are given in Fig. 4. Table A1 contains more extensive information on the number of class samples over each temperature zone with numbers of training, test, and validation datasets.

2.8. Random forests

Random forests (RF; Breiman, 2001) are machine learning algorithms that rely on establishing an ensemble of decision trees and producing predictions by averaging the results of many trees. For categorical variables, random forests choose the most popular prediction among all the trees (Breiman, 2001). Usually, a single decision tree is programmed to split the feature space at points that create the maximum separation in terms of output classes. Random forests differ from other tree-based methods by enforcing uncorrelated trees by randomly selecting candidate features for splitting and only considering those features in the splitting process (James et al., 2013). This enforces trees to consider different sets of features for splits, resulting in uncorrelated trees with reduced variance (Hastie et al., 2009). The randomness in the algorithm also makes the classifier robust to overfitting (Breiman, 2001). Random forest classifiers are also robust to outliers and noise (Rodriguez-Galiano et al., 2012). Lastly, these algorithms are user-friendly as they have fewer parameters than other machine-learning algorithms (Zhou et al., 2016). The disadvantage of the random forest algorithm is that it is a black-box algorithm that will prevent us from understanding the classification process. However, this disadvantage is also encountered in other machine learning methods that are frequently used.

Hyperparameters are user-selected parameters that determine the structure of a classifier. Two primary hyperparameters of an RF classifier are the number of classification trees (NTREE) and the number of classification variables used at each split (MTRY) (Ming et al., 2016; Rodriguez-Galiano et al., 2012). These hyperparameters determine the structure of each classification tree. The number of classification trees reduces the chance of overfitting as it is increased and limits the value of generalization error (Breiman, 2001). The hyperparameter MTRY is the number of features randomly selected at each node, and it must be set with an integer in the interval [1, M], where M is the number of classification features (Bernard et al., 2009). MTRY determines the level of randomization in the feature selection process in such a way that a smaller MTRY means stronger randomization (Bernard et al., 2009).

2.8.1. Feature selection

When classification feature space is too large for efficient computation, feature selection using random forest importance can be performed (Box 18 in Fig. 2). To obtain the importance of each feature, the Out Of Bag (OOB) sample, which is a set of data points that are not included in the tree-building process, is selected by the algorithm (Genuer et al., 2010). For each tree, the prediction error for the OOB sample is calculated, and then it is repeated after each feature is permuted (Han et al., 2016). The difference between the errors is calculated and averaged for all trees (Han et al., 2016). The mean decrease in accuracy for all features can be used to select features that contribute more to the classifier's accuracy.

The available input feature number has increased to 212 (168 bands + 3 indices for each 14 date intervals + 2 NDVI rates) with the addition of vegetation indices. Considering that this study aims to create a high-resolution crop map covering Türkiye, the computational cost of the classification is decreased through feature selection. After the dataset is processed to its final form, five different random forest algorithms are run using the randomForest package in R (Liaw and Wiener, 2002). Since different classifiers add randomness to the results, it is possible to have a more generalizable opinion with different classifier results. The average of these five random forest classifiers' Mean Decrease Accuracy output is taken for each feature and sorted from highest to lowest to evaluate the contribution of each feature to the classification accuracy.

2.8.2. Hyperparameter tuning

Before the crop mapping step, a random forest classifier that gives the best results on the validation set should be found. To find the best-performing classifier, hyperparameter tuning using grid search is performed (Box 20 in Fig. 2). Grid search is the process where each combination of hyperparameters is used for classification one by one, and the accuracy for these combinations is calculated to find the combination that gives the best accuracy. NTREE options are selected to increase to a maximum of 500 trees because computational cost increases as NTREE increases. In the case of MTRY, apart from some generic numbers, $\sqrt{M}$ suggested by Bernard et al. (2009), which is also the default in the randomForest package, and $(\log_2 M+1)$ suggested by Breiman (2001) are selected for hyperparameter tuning where M is the number of classification features. Even though a decrease in MTRY weakens individual trees, an increase in MTRY reduces the randomness and creates classifiers that are more likely to overfit the training data, which results in lower

Table 1
The numbers of bands and indices in each month in the best performing 120 inputs for each temperature zone classifiers and total of the zones.

Month	Low	Medium	High	Total
March[a]	11	5	5	21
April	13	15	19	47
May	19	20	24	63
June	18	17	12	47
July	16	27	18	61
August	25	19	24	68
September[a]	11	10	10	31
October[a]	6	5	6	17

[a] Features removed from the dataset.

accuracies (Bernard et al., 2009; Rodriguez-Galiano et al., 2012). So, its maximum value is limited to 20 to avoid overfitting. Table 3 shows the hyperparameters used in the grid search. To decide on the best classifier, the overall accuracy of each combination of MTRY and NTREE is calculated for the validation set. After tuning, NTREE = 500 and MTRY = $\sqrt{M}$ is selected as final hyperparameters (Box 21 in Fig. 2).

2.9. Crop mapping

After determining the optimal hyperparameters, a classifier for each temperature zone is created. Each raster file is divided into tiles to make the classification possible with limited computer memory. Later, all bands for each tile are read on R programming and preprocessed to have the same bands and the same form with the training, test, and validation sets. Resultant data frames are classified using one of three classifiers according to each pixel's temperature zone. After separate classifications of temperature zones, all classification results are combined in one data frame for each tile and later this data frame is converted to a raster format to visualize the resultant map.

2.10. Accuracy assessment

Accuracy assessment is necessary for land cover classifications to evaluate the success of the resultant map (Stehman, 1996). The Kappa coefficient, overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA) are calculated for this study (Box 23 in Fig. 2). First of all, an error matrix is formed that compares the ground truth data used for validation and test data (shown in the rows of the confusion matrices in this study) and classified data (shown in the columns of the confusion matrices in this study) (Story and Congalton, 1986). After using the information provided by the error matrix, OA, UA, PA, and Kappa coefficient (Cohen, 1960) are calculated.

A simple error matrix and formulas for the user's and producer's accuracy are given in Table 4, where each class' correctly classified pixel number is given in bold letters (A, D). Columns show the number of classified pixels for each class, while rows show the number of ground truth pixels for each class. To calculate user's accuracy, which refers to how many of the classified pixels of a class actually belong to that class in the ground truth data, the number of correctly classified pixels for a class is divided by the column sum. To calculate the producer's accuracy, which refers to how many ground truth pixels are classified correctly for each class, the number of correctly classified pixels for a class is divided by the row sum. Overall accuracy is the ratio of the total number of correctly classified pixels for all classes to the total number of ground truth pixels. Lastly, the Kappa coefficient is a proportion of accuracy that takes into account the possibility of correctness taking place by chance (Sun, 2011), and its formulation is given in Equations (5a), (5b) and (5c), 5d, and 5e where P_e is agreement between classification and ground truth happening by chance, P_0 is the overall accuracy, $P_{Class\ 1}$ is the probability that all pixels classified as Class 1 randomly, and $P_{Class\ 2}$ is the probability that all pixels classified as Class 2 randomly.

$$Overall\ Accuracy = \frac{A+D}{A+B+C+D} \tag{4}$$

$$P_{Class\ 1} = \left(\frac{A+B}{A+B+C+D}\right) \times \left(\frac{A+C}{A+B+C+D}\right) \tag{5a}$$

$$P_{Class\ 2} = \left(\frac{C+D}{A+B+C+D}\right) \times \left(\frac{B+D}{A+B+C+D}\right) \tag{5b}$$

$$P_0 = Overall\ Accuracy = \frac{A+D}{A+B+C+D} \tag{5c}$$

Table 2
Top 15 highest performing features of the classifiers for each temperature zone.

Rank of Importance	LOW			MEDIUM			HIGH		
	Month	Half of month	Band/Index	Month	Half of month	Band/Index	Month	Half of month	Band/Index
1	May	2nd	B4			Spring rate			Spring rate
2	June	1st	NDVI	April	2nd	NDVI	July	2nd	NDVI
3	August	2nd	B1	May	1st	NDVI	July	1st	NDVI
4	May	2nd	NDVI	July	2nd	B5	August	1st	B6
5	March	2nd	$NDWI_G$	August	1st	B5	May	2nd	B12
6	June	1st	B4	August	1st	B4	May	1st	$NDWI_G$
7	June	2nd	B4	May	1st	$NDWI_G$	May	2nd	NDVI
8	September	2nd	B1	May	!st	$NDWI_M$	July	2nd	$NDWI_G$
9	May	1st	B5	August	1st	B12	May	1st	NDVI
10	May	2nd	B12	July	2nd	B4	July-	1st	$NDWI_G$
11	May	1st	B1	July	2nd	B12	April	2nd	$NDWI_G$
12	May	2nd	B1	April	2nd	$NDWI_G$	May	2nd	$NDWI_G$
13	May	2nd	B5	July	1st	NDVI	July	2nd	B4
14	June	1st	B12	August	1st	NDVI	May	1st	B5
15	May	1st	B4	April	2nd	$NDWI_M$	May	2nd	B11

Table 3
Hyperparameters for random forest.

NTREE	1, 5, 10, 50, 100, 200, 300, 500
MTRY	1, 2, 3, 4, 5,6, 7, 8, 9, 10, 11, 12, 15, 20, $\sqrt{M}$ (Bernard et al., 2009), ($\log_2 M+1$) (Breiman, 2001)

Table 4
A simple error matrix and formulas for User's accuracy and Producer's Accuracy.

		CLASSIFIED			
		Class 1	Class 2	Row Sum	Producer's accuracy
GROUND TRUTH	Class 1	**A**	B	A+B	A/(A+B)
	Class 2	C	**D**	C+D	D/(C+D)
Column Sum		A+C	B+D		
User's accuracy		A/(A+C)	D/(B+D)		

$$P_e = P_{Class\ 1} + P_{Class\ 2} \tag{5d}$$

$$Kappa = \frac{Po - Pe}{1 - Pe} \tag{5e}$$

In addition to the mentioned accuracy measures, the classification map is inspected visually for obvious misclassifications. As the last exercise of accuracy assessment, the area percentages on the classification map were compared in the provinces with the highest production for products for which production data statistics are available. The Republic of Türkiye Ministry of Agriculture and Forestry (TMA) publishes yearly production reports for certain crops that contain information about which provinces have the most production for the year. This information is usually in terms of percent area, while in some reports, it is in tons of production. Two different government-based data sets are mentioned in this study; crop type information at the field scale obtained from farmers' declarations and crop yield information at the provincial scale obtained from the total purchase amounts counted by the state. Here, individual farmer declarations are considered less reliable as farmers receive incentives for certain types of crops and there is no subsequent validation. On the other hand, provincial harvest data collected by the government is considered accurate and reliable as these data reflect the actual crop traded. In the last part of the Results and Discussion chapter, available production data for 2019 is compared to the classification map. First, the pixel number of each crop is obtained using Semi-Automatic Classification Plugin in QGIS for each province. Then, crop percent areas are calculated and sorted by the highest percent area. Provinces with the highest production of a certain crop are compared with the data provided by the Ministry of Agriculture and Forestry, which contains data for only ten provinces with the highest production.

3. Results and Discussion

3.1. Feature selection with random forest importance

After the average of five random forest classifiers' Mean Decrease Accuracy output is taken for each feature, and sorted from highest to lowest, the contribution of each feature to the classification accuracy is evaluated. Even though classifiers give different importance values for each feature, the order of the best-performing first 120 features do not change significantly. Because of this, only the first 120 features with the highest decrease in accuracy are inspected, as it is also observed that after that point, feature importance decreases notably to have much less contribution to accuracy. Table 1 shows the number of bands and indices for each month in the top 120 best-performing features of the classifiers for each temperature zone. As highlighted in the table, features obtained from the median composites of March, September, and October is found less than other months in the top 120 best-performing features, indicating that these months have less contribution to accuracy when compared to other months. These months' data are removed from the dataset before hyperparameter tuning to decrease the computational cost. After eliminating these features, classification is performed using the remaining 152 features (B1, B2, B3, B4, B5, B6, B7, B8, B8A, B9, B11, B12, NDVI, $NDWI_M$; $NDWI_G$ for 10 median composites over the months April, May, June, July and August, and 2 NDVI rates; Box 19 in Fig. 2)

Even though no further elimination is performed over individual bands, to evaluate the importance of indices and the calculated growth rates, best-performing features of temperature zones are also inspected. Table 2 displays the top 15 highest performing features of the classifiers for each temperature zone. While spectral indices are not significant in LTZ, B1 and B4 are more prominent than the other Sentinel-2 bands. In contrast to the LTZ, the best-performing features in MTZ contain nine indices, with the spring rate

contributing the most to classification accuracy. The HTZ follows a similar trend, with spring rate being the best-performing characteristic; in total, ten indices are in the top fifteen best-performing properties. As a result of index contribution to classification accuracy, it is reasonable to conclude that using indices, especially $NDWI_G$ and NDVI, over medium and high-temperature regions with similar climatic circumstances to the study area can improve crop classification accuracy. The integration of the NDVI change rate during crop growth phases is also highly recommended in such regions due to the impact of the spring rate over the classification accuracy of MTZ and HTZ.

3.2. Random forest classifier accuracy on the test set

In this chapter, classifier accuracies of three temperature zones, evaluated using measures mentioned in the accuracy assessment part of the study, are given. As mentioned in the methodology, ground truth data is divided into three subsets, training, testing, and validation. The validation set is used for hyperparameter tuning, while the test set is left out until all tuning is done and final classifiers are obtained.

3.2.1. Low-Temperature Zone

The confusion matrix obtained from the LTZ classifier using the test set is given in Table 5. In the confusion matrices given, each row has been normalized within itself to facilitate color coding for easier visualization of misclassifications. The number of classes in the LTZ is fewer than in the other zones since agricultural activities in this zone are sparser than in the others. Due to this sparse crop production, the training sample size is smaller than in other zones, and this can potentially be the reason for the lower classification accuracy obtained from the LTZ with OA of 89% and Kappa coefficient of 0.88. When land cover classes are inspected individually, barley, bare soil, rice paddy, forest, cherry, chickpea, water, and alfalfa classes reach PAs higher than 90%, showing the classifier's success on these classes. On the other hand, wheat, maize, potato, and urban classes reach unsatisfying PAs. When the confusion matrix is inspected, it can be seen that 94 of 255 wheat pixels are classified as barley. This misclassification can be explained by the similar phenological features of barley and wheat (Zheng et al., 2015). Another confusion occurred between maize and rice paddy, as 76 of 102 maize pixels are classified as rice paddy. Since a significant similarity between the phenological features of these two crops is not observed in this study, the reason behind this confusion should be further examined. Urban pixels are misclassified as bare soil due to a lack of training data for the classifier to identify urban pixels and similar reflectance curves of bare soil and urban pixels (Piyoosh and Ghosh, 2018; Zhang et al., 2015). Despite the classes with low performance, the overall performance of the Low-Temperature Zone is decided to be used for crop mapping.

3.2.2. Medium-Temperature Zone

Medium-Temperature Zone contains more data for classification than the other temperature zones, and it also contains a greater

Table 5
Confusion Matrix, Producer's Accuracy, User's Accuracy, Overall Accuracy and Cohen's Kappa of Low-Temperature Zone.

	Alfalfa	Bare Soil	Barley	Beans (dry)	Cherry	Chickpea	Cow Vetches	Forest	Maize	Oat	Potato	Rice Paddy	Steppe	Suger Beet	Urban	Water	Wheat	PA (%)	UA (%)
Alfalfa	231		4							4	2							96	61
Bare Soil		1942	3								1		34					98	100
Barley	16		1529				26			25			3				37	93	87
Beans (dry)				57							11			3				80	100
Cherry					48			4										92	98
Chickpea						98							1					99	100
Cow Vetches	5		28				502			58	1		3					84	78
Forest					1			1460		1								100	100
Maize									23			76		3				23	100
Oat	26	1	65				72			472			13					73	81
Potato	22	1	1				32	1		21	128		6					60	75
Rice Paddy	7											164						96	68
Steppe	57		40										625					87	90
Suger Beet											17			126				88	86
Urban		3																0	0
Water																40		100	100
Wheat	12	4	94				10			4	11		13	14			93	36	72
																		OA = 0.89	
																		Kappa = 0.88	

Table 6
Confusion Matrix, Producer's Accuracy, User's Accuracy, Overall Accuracy and Cohen's Kappa of Medium-Temperature Zone.

	Alfalfa	Apple	Apricot	Bare Soil	Barley	Beans	Cherry	Chickpea	Cotton	Cow Vetches	Forest	Grape	Green Tea Leaves	Hazelnut	Maize	Melon	Oat	Olive	Onion (Dry)	Pepper	Poppy	Potato	Rapeseed	Red Lentil	Rice Paddy	Steppe	SugerBeet	Sunflower	Tobacco	Tomato	Triticale	Urban	Water	Watermelon	Wheat	PA (%)	UA (%)
Alfalfa	412	5	9		21		1	11	8												4															87	95
Apple	2	956	90				13	20											1		1					3		3								88	82
Apricot	4	12	1122		8		36	20														1				10		8								86	82
Bare Soil				5455																																100	99
Barley	1		5		3868																25										13				163	95	81
Beans	16					600		7	89			1			1				3			1					7	55						2		77	82
Cherry		18	36		2		222	1			18	2		5	1											1		1								72	80
Chickpea			2		107			4848									2																			98	92
Cotton			1	1		5		2	1557						4												1	77								94	91
Cow Vetches					61			2		109							3				36					5									1	50	87
Forest		4									2520			1																						100	99
Grape		6	16									126										8						6								78	88
Green Tea Leaves													57	3																						95	100
Hazelnut											1			656																						100	99
Maize															5559				112								268	10		2						93	99
Melon		1			2	14										184			5	22								6		183		33		4		41	99
Oat			1		30			##									326																		2	69	96
Olive		3	2				1				2							72								2										88	100
Onion (Dry)						1		2							21				521	51						2	7			3						86	80
Pepper				59		44														34										27						21	29
Poppy					28																1489		1												5	98	90
Potato		69	8												1							84														52	89
Rapeseed		1			21			7		16											36		745												6	90	100
Red Lentil			64		42			24				6												42												24	100
Rice Paddy							1																		673											100	100
Steppe			4	2	6			25									8				8					973									1	95	96
SugerBeet						17			36						1	1											423	77		1						75	59
Sunflower		5				1	2	191	8						18											2		2555		5						92	91
Tobacco			1	3	11			2				4							6							1		19	17			2				26	100
Tomato									2										1	2								2		17						96	43
Triticale					78			9	4								2				4					11					81				99	28	86
Urban																															1108					100	97
Water																									2							362				99	100
Watermelon						52			1			4								8							14	1					105			57	95
Wheat					497																56													2504		82	90

OA = 0.91
Kappa = 0.92

Table 7
Confusion Matrix, Producer's Accuracy, User's Accuracy, Overall Accuracy and Cohen's Kappa of High-Temperature Zone.

	Alfalfa	Apple	Apricot	Bare Soil	Barley	Cherry	Chickpea	Cotton	Forest	Grape	Green House	Hazelnut	Lemon	Maize	Mandarin	Melon	Olive	Orange	Peach	Pepper	Potato	Rapeseed	Red Lentil	Rice Paddy	Steppe	Sunflower	Tobacco	Tomato	Urban	Water	Watermelon	Wheat	PA (%)	UA (%)
Alfalfa	271							18		2			2	1			6									4							84	97
Apple		6								7			12		8			15															13	100
Apricot	3		279					1		25			65	1	19		1	8															69	91
Bare Soil				252													3								8								96	100
Barley					266																											116	70	89
Cherry						29				5		1	1		3																		74	97
Chickpea							60						1				69									16						19	36	100
Cotton	6				2			6644						2						1						2		37				1	99	97
Forest									1032								1																100	98
Grape			22					22		1986			4		9		26									6							96	94
Green House											545																		1				100	100
Hazelnut												140	4					1														97	99
Lemon			4		25	1				26			744		56		153	37								18						2	70	72
Maize								54						3141										2		1		2					98	98
Mandarin									18	11			46		1050		14	24								34			1				88	77
Melon								3								29								4		15		3					54	85
Olive									4	8					3		1074	9							1								98	77
Orange										9			140		219		16	527															58	84
Peach										7			13	28	2		1	3	5													1	8	100
Pepper								83												116						6		3					56	98
Potato																					26												100	100
Rapeseed																	1					616											100	100
Red Lentil																							458									248	65	90
Rice Paddy														1										312		1							100	100
Steppe					5												13						49		1797							85	92	100
Sunflower										5					3											2665		6					99	94
Tobacco																										17	11		16				25	100
Tomato								21		3			3				1			1						42		602					89	88
Urban																	13												1438				99	99
Water																														438			100	100
Watermelon																5										8		29			68		62	100
Wheat													1	1			1						2									8435	100	95
																																	OA = 0.94	
																																	Kappa = 0.94	

Table 8
Confusion Matrix, Producer's Accuracy, User's Accuracy, Overall Accuracy and Cohen's Kappa of Combined Temperature Zones.

	Alfalfa	Apple	Apricot	Bare Soil	Barley	Bean (Dry)	Cherry	Chickpea	Cotton	Forest	Grape	Green House	Green Tea Leaves	Hazelnut	Lemon	Maize	Mandarin	Melon	Oat	Olive	Onion (Dry)	Orange	Peach	Pepper	Poppy	Potato	Rapeseed	Red Lentil	Rice Paddy	Steppe	SugerBeet	Sunflower	Tobacco	Tomato	Triticale	Urban	Vetch	Water	Watermelon	Wheat	PA (%)	UA (%)
Alfalfa	914	5	9		25		1	11	26		20				2	1			4	6					4	2					4									88	84
Apple	2	962	9				13	20			7				12		8				1	15			1					3	3									91	82
Apricot	7	102	1401		8		36	20	1		25				65	1	19			1		8				1				1	8									82	84
Bare Soil				7649	3															3						1			42										99	99
Barley	17		5		5663														25					25					3					13		26			316	93	83
Bean (Dry)	16					657		7	89		1					1					3					12					1	55					2		78	83
Cherry		18	36		2		299	1		22	7			6	1	1	3													1	1								75	84
Chickpea			2		107			5006							1				2	69										1	16							19	96	92
Cotton	6		1	1	2	5		2	8201							24							1								1	79		37				1	98	96
Forest		4					1			45631				1					1	1																			100	99
Grape		6	38						22		2112				4		9			26						8					12								94	94
Green House												545																							1				100	100
Green Tea Leaves													57	3																									95	100
Hazelnut										1				796	4							1																	99	99
Lemon			4		25		1				26				744		56			153		37									18							2	70	72
Maize									54							8723					112								78		271	11		4					94	99
Mandarin										18	11				46		1050			14		24									34				1				88	77
Melon		1			2	14			3									213			5			22					4		21		186		33		4		42	97
Oat	26		1	1	95			115											798											13						72			2	71	86
Olive		3	2				1			6	8						3			1146		9								3										97	78
Onion (Dry)						1		2								21					521			51						2	7		3						86	80
Orange											9				140		219			16		527																		58	84
Peach											7				13	28	2			1		3	5																1	8	100
Pepper				59		44			83															150								6		30						40	64
Poppy					28																				1489		1												5	98	90
Potato	22	69	8	1	1					1						1			21							238				6						32				60	82
Rapeseed		1			21			7												1					36		1361									16			6	94	100
Red Lentil			64		42			24			6																	500											248	57	91
Rice Paddy	7						1									1													3939		1								100	98
Steppe	57		4	2	51			25											8	13					8		49		3395										86	92	97
SugerBeet						17			36						1		1									17					549	77		1						79	63
Sunflower		5				1	2	191	8		5					18	3													2		5220		11						95	92
Tobacco			1	3	11			2			4										6									1	36	28			18				25	100
Tomato									23		3				3					1	1			3							44		772						91	71
Triticale					78			9	4										2						4					11					81			99	28	86
Urban				3																13															2546				99	98
Vetch	5				89			2											61						36	1				8						611		1	75	80
Water																													2								840			100	100
Watermelon						52			1		4							5						8							14	9		29				173		59	97
Wheat	12			4	591										1	1			4	1					56	11		2		13	14					10			1032	94	93
																																								OA = 0.92 Kappa = 0.93	

Table 9
Confusion Matrix, Producer's Accuracy, User's Accuracy, Overall Accuracy and Cohen's Kappa of Classification on Test set without Temperature Zones.

	Alfalfa	Apple	Apricot	Bare Soil	Barley	Bean (Dry)	Cherry	Chickpea	Cotton	Forest	Grape	Green House	Green Tea Leaves	Hazelnut	Lemon	Maize	Mandarin	Melon	Oat	Olive	Onion (Dry)	Orange	Peach	Pepper	Poppy	Potato	Rapeseed	Red Lentil	Rice Paddy	Steppe	SugerBeet	Sunflower	Tobacco	Tomato	Triticale	Urban	Vetch	Water	Watermelon	Wheat	PA (%)	UA (%)
Alfalfa	892	2	3		72		1	7	11		18		1						4					1	2	11				6	1	1								1	86	88
Apple	5	939	16				15	16			27				9	3	1			1		2								5		7		1							83	89
Apricot	25	49	1428		9		16	8			81				26	5	10		1	15		6				1				11		22								83	77
Bare Soil				7474				6																						40		1				177				97	99
Barley	21	1	5		5537			3								1			53	3		1			13							14			6		33			402	91	81
Bean (Dry)	16					547		9	78		5					13										70			3		7	103		2						64	78
Cherry		8	38				252			27	45			4	7	3				7		1								2		2							2	63	83
Chickpea			2	1	118			5069											1											2		14							16	97	91
Cotton	9		1		1	6			6630							16		1		2						16					1	1		27					2	99	96
Forest		6					5			5005	1			1						2																				100	98
Grape		3	51					1	22		2107				1		3			30										6		13								94	88
Green House				1								532																								13					97	100
Green Tea Leaves										1			52	7																											87	98
Hazelnut										7				793	2																										99	98
Lemon	4	16	94		27		5	13		38	9			2	651		50			64		38								33		21							1	61	75
Maize					15				55	1						8725					105					1			51		280	17		1					2	94	98
Mandarin		4	34					4		27	4				43		1025			10		25										21			1					86	80
Melon					3	8					6							213						13					4		7	27		187		40				42	94
Oat	16		10		75		1	151			4								723						2					18		13					103		7	64	86
Olive			37		6		3	4		7	8						3			1042		12								59										88	86
Onion (Dry)				1		1		16	1		1					21					475			31							43	10		8						78	82
Orange	1	3	23							6	9				121		191			13		541	1																	2	59	84
Peach		22	11		1		3				58				12	31						5	79																		36	99
Pepper				33		42			54															167		1					1	6		44		24				45	75
Poppy					53																				1463		1													6	96	93
Potato	7				1	4		2	12							10			6	2						1712				4	3	121					2			91	90
Rapeseed	3				14																				3		1421					1							7	98	100
Red Lentil			2		3			87												3								509												280	58	97
Rice Paddy	7									1						1													3939			1								100	98
Steppe	1			4	145			30											7	3					10			16		3445					2				35	93	94
SugerBeet						29			3		1					2		1								44						550	68		10						78	61
Sunflower						14	1	112			9					19	2									8					2	1	5290		8						97	89
Tobacco				2	9			5										1													3		52	19		19					17	100
Tomato	4								1		4									1				3		37					2	103		685						1	81	70
Triticale					90			6											1						1	4				8		2			69				107	24	90
Urban			3	37				2				1								10																2509					98	90
Vetch	5				115														43						22					4		7				616			2	76	82
Water																													2							4		836			99	100
Watermelon			1			50					5							10						7		3					11	27		11				170		58	100
Wheat	3	3	2		560		2	5	25		1				1	9	2		6			10			56		1		10		10			2					1044	94	93
																																									OA = 0.92 Kappa = 0.91	

number of crop types due to high agricultural activity in middle Anatolia. Table 6 shows that the classifier gives 91% OA with a Kappa of 0.92. When land cover classes are inspected individually, it is seen that 16 of 35 land cover classes (bare soil, barley, chickpea, forest, green tea leaves, hazelnut, maize, poppy, potato, rapeseed, rice paddy, steppe, sunflower, tomato, urban, water) give PAs higher than 90% and 6 of 35 land cover classes (wheat, apple, apricot, dry onion, alfalfa, olive) give accuracies between 80% and 90%. There are also classes that are classified with almost complete correctness (bare soil, rice paddy, hazelnut, forest, urban, and water with PA > 99%). On the other hand, some classes give lower user's and producer's accuracies, like tobacco, pepper, and red lentils, which can be due to inadequate training data. Confusion between wheat and barley can also be observed in MTZ as it is in LTZ. According to tons of crop production data provided by the Turkish Statistics Institute (2019), among all products, wheat has the highest (16.9%), and barley has the fourth highest production percentage (6.44%). Although the misclassification of these two crops affects the overall accuracy of the classification map as large areas are covered by wheat and barley, the fact that these two crops have very similar water footprints (Batan, 2021; Bulut and Canbaz, 2022) shows that this misclassification would not affect water management practices as much as misclassification among crops with different water demands.

Another confusion occurs between melon and tomato, as almost half of the melon pixels are classified as tomato pixels. Triticale pixels are also highly misclassified as barley and wheat, potentially due to the fact that triticale is a hybrid of wheat and barley (Ruβwurm and Körner, 2017). Unlike other classes, tomato and sugar beet have much lower UAs than their PAs since most melon pixels are classified as tomato, and more than one-third of the pixels classified as sugar beet belongs to the maize class. The reasons behind these misclassifications should be further investigated.

3.2.3. High-Temperature Zone

High-Temperature Zone also includes regions with high agricultural activities like MTZ. Especially cotton, sunflower, and fruits like olive, lemon, mandarin, grape, and orange crops are cultivated in regions covered by HTZ. Table 7 shows that the HTZ classifier gives the highest accuracy among the classifiers used in the study, with 94% overall accuracy and a Kappa of 0.94. Among 34 landcover classes, 16 classes (sunflower, wheat, rice paddy, hazelnut, bare soil, forest, rapeseed, maize, cotton, potato, greenhouse, grape, steppe, urban, water, olive) reach PAs higher than 90%, and three crop classes (tomato, mandarin, alfalfa) reach accuracies between 80% and 90%. In addition, cotton, sunflower, wheat, rice paddy, forest, rapeseed, potato, greenhouse, urban, and water pixels are classified with more than 99% accuracy. Peach, apple, tobacco, and chickpea crop pixels are mostly misclassified by the classifier, potentially due to the small number of training pixels available for these crop classes. One of the most noticeable misclassifications is between red lentil and wheat, as 248 of 458 red lentil pixels are classified as wheat. This confusion can result from similar growing periods over the same region (Southeastern Anatolia). Another misclassification occurred between citrus fruits orange, lemon, and mandarin, potentially due to the tree structures of the crops. The High-Temperature Zone classifier is decided to be used for the mapping step as it gives very high accuracies and Kappa coefficients.

3.2.4. Combined temperature zones

Confusion matrices of temperature zones are combined to evaluate the accuracy of the whole test set. With this combined confusion matrix, it is possible to see how accurately the whole test data is classified with study area division. The resultant confusion matrix is given in Table 8, which shows 92% OA with a Kappa of 0.93. More than 85% of the test pixels are classified correctly for 23 of 40 land cover classes (barley, sunflower, bare soil, wheat, green tea leaves, rice paddy, tomato, hazelnut, forest, poppy, rapeseed, mandarin, maize, chickpea, cotton, greenhouse, dry onion, steppe, urban, grape, water, alfalfa, olive). In addition to that, it is shown in the confusion matrix that more than 95% of the test pixels are classified correctly for 13 classes (forest, greenhouse, water, urban, bare soil, rice paddy, hazelnut, cotton, poppy, olive, chickpea, sunflower, green tea leaves). On the other hand, PAs lower than or equal to 60% are obtained for classes peach, tobacco, triticale, pepper, melon, red lentil, orange, watermelon, and potato. These classes also reached lower accuracies in separate temperature zones except for potato in HTZ, where all potato test pixels are classified correctly. As discussed for each temperature zone previously, these misclassifications potentially occur due to the inadequate number of training pixels (e.g., for peach and tobacco) or similar spectral and phenological features of different crops. Misclassification of triticale as wheat and barley is an example of confusion due to similar features of different crops as it is a hybrid of wheat and barley. Another misclassification observed in the separate temperature zones is between wheat and barley. Even though the confusion between these crops decreases the accuracies for one of the classes in separate temperature zones, the combined PAs of the crops exceeds 90%. These accuracy measures represent the correct classification of only the predetermined test pixels and not the whole map. Nevertheless, the accuracy obtained from the combination of test sets of temperature zones is satisfactory when it is compared with the results of other large-scale studies (e.g., Yılmaz et al., 2020; Yang et al., 2019; d'Andrimont et al., 2021) and can give an opinion about the accuracy of the classification map.

3.2.5. Comparison with the classifier without the temperature zones

The dataset before the division of temperature zones is also used for classification. The training subset of this dataset contains all

Table 10
Kappa coefficient and overall accuracy for combined temperature zones and classification without temperature zones with improvement in accuracy with temperature zoning.

	CombinedTemperature Zones (%)	Without Temperature Zones (%)	Improvement with Temperature Zoning (%)
Overall Accuracy	92.35	91.54	0.81
Cohen's Kappa	91.86	90.99	0.87

training data used in each temperature zone. So, the total number of training pixels used in each temperature classifier is used for one classifier, which is also the case for test and training pixels. The confusion matrix for this classifier which gives 92% OA with Kappa of 0.91, is given in Table 9. In addition, Table 10 is given for the comparison between the OA and Kappa coefficient obtained using the classifier without the temperature zoning and the confusion matrix of combined temperature zones, and Table 11 shows the PAs for each class for this comparison. OA and Kappa coefficients increase less than %1 when the study area is divided into temperature zones (OA increases by 0.81% and Kappa increases by 0.87%). Even though, on average, the use of temperature zones yielded marginally

Table 11
Producer's Accuracy for Each Class for Combined Temperature Zones and Classification Without Temperature Zones with Improvement in PA with Temperature Zoning.

Class Name	PA for Combined Temperature Zones (%)	PA for Without Temperature Zones (%)	Improvement in PA with Temperature Zoning (%)
Bean (Dry)	77.02	64.13	12.9
Cherry	75.13	63.32	11.81
Tomato	90.82	80.59	10.24
Olive	97.04	88.23	8.81
Lemon	69.79	61.07	8.72
Green Tea Leaves	95.00	86.67	8.33
Tobacco	25.45	17.27	8.18
Onion (Dry)	85.69	78.13	7.57
Oat	71.06	64.38	6.68
Peach	40.09	35.59	4.5
Triticale	28.13	23.96	4.17
Green House	99.82	97.44	2.38
Bare Soil	99.36	97.09	2.27
Alfalfa	88.39	86.27	2.13
Mandarin	87.65	85.56	2.09
Barley	92.94	90.87	2.07
Apple	84.61	82.59	2.02
Poppy	97.77	96.06	1.71
Urban	99.38	97.93	1.44
Watermelon	58.64	57.63	1.02
Water	99.76	99.29	0.48
Hazelnut	99.25	98.88	0.37
Grape	94.41	94.19	0.22
Cotton	98.97	98.76	0.21
Forest	99.84	99.70	0.14
Melon	41.93	41.93	0
Rice Paddy	99.75	99.75	0
Maize	94.27	94.29	-0.02
Potato	90.72	90.77	-0.05
Wheat	93.87	93.98	-0.1
Sugar Beet	77.54	77.68	-0.14
Vetch	75.06	75.68	-0.61
Red Lentil	56.56	57.58	-1.02
Chickpea	95.85	97.05	-1.21
Sunflower	95.50	96.78	-1.28
Steppe	91.81	93.16	-1.35
Orange	57.85	59.39	-1.54
Apricot	81.79	83.36	-1.58
Rapeseed	93.93	98.07	-4.14
Pepper	40.32	44.89	-4.57

approved results, when individual classes are inspected, PAs of 25 classes increase with temperature zones while PAs of 13 classes decrease with temperature zoning.

One unexpected result from the classification without the temperature zones is that even though wheat pixels have different spectral curves in different temperature zones, the classification accuracy does not improve with temperature zones (Table 11). It is also observed that classes with the highest increase in accuracy with temperature zones (dry beans, cherry, tomato, olive, lemon, and green tea leaves) have more than 80% of their training pixels in one temperature zone. In contrast, classes with the highest decrease in accuracy with temperature zones (rapeseed, pepper) have training pixels that are more equally distributed to temperature zones. So, it is possible that classifying a crop in the temperature zone that the crop is mostly cultivated in, rather than classifying it with a classifier that includes areas that the crop is not commonly cultivated, can increase the accuracy of that crop. In addition, dividing the ground truth data into more temperature zones decreases the training pixel number, which can potentially be a reason for decreased accuracy with temperature zoning for classes distributed to temperature zones.

3.3. Visual Inspection of the classification map

As mentioned in the Methodology chapter, a crop classification map is created using random forest classifiers. The resultant map is given in Fig. 5. In this part of the study, the classification map is inspected to identify visible problems. The first problem that draws attention is the misclassification of Salt Lake as urban and greenhouse, as shown in Fig. 6. This confusion potentially occurs due to the similar spectral characteristics of bare soil and urban surfaces (Piyoosh and Ghosh, 2018; Zhang et al., 2015). To overcome this problem, NDBI proposed by Zha et al. (2003) to map urban and built-up areas, dry built-up index (DBI), and dry bare-soil index (DBSI) to map built-up and bare areas in a dry climate proposed by Rasul et al. (2018) can be utilized. In addition to that, the training pixel number of the urban class can be increased for better results.

Another significant misclassification is the abundance of pixels classified as olive in the West. A representative part of the map is given in Fig. 7. The reason for this misclassification is potentially the absence of shrub class or specifically maquis, which is the characteristic plant formation of the Mediterranean climate that has a widespread distribution along the coast and east-west oriented valleys in the Aegean Region (Günal, 2013). In addition to the number of olive pixels, a rectangular region visibly classified differently from its surrounding can be seen in the center of the image. That rectangular region does not contain as many olive pixels as the rest of the image. This patch corresponds to different temperature zones from their surroundings, which may mean the temperature zones should have been smoothed to have a more homogeneous classification.

A large region classified as oat in the northeastern corner of Türkiye can be seen in Fig. 5. The region is also shown in Fig. 8 in more detail. No information in the literature or other sources implies an abundant oat production in the region. So, it can be concluded that the classification is incorrect in the specified region. The reason behind this error should be further investigated.

3.4. Percent Area Comparison with province production data of Ministry of Agriculture and Forestry

Fig. 9 shows a scatter plot with data for the crops for which TMA data is available. This figure indicates a high coefficient of determination ($R^2 = 0.85$), demonstrating that crop areas derived from the classified map are in good agreement with TMA data. F igures A1 through A9 in the Appendix contain more extensive information and discussion on the area comparison. Cotton ($R^2 = 0.98$) and rice paddy ($R^2 = 0.97$) provide the best agreement between the two sets of data. When confusion matrices are evaluated, the

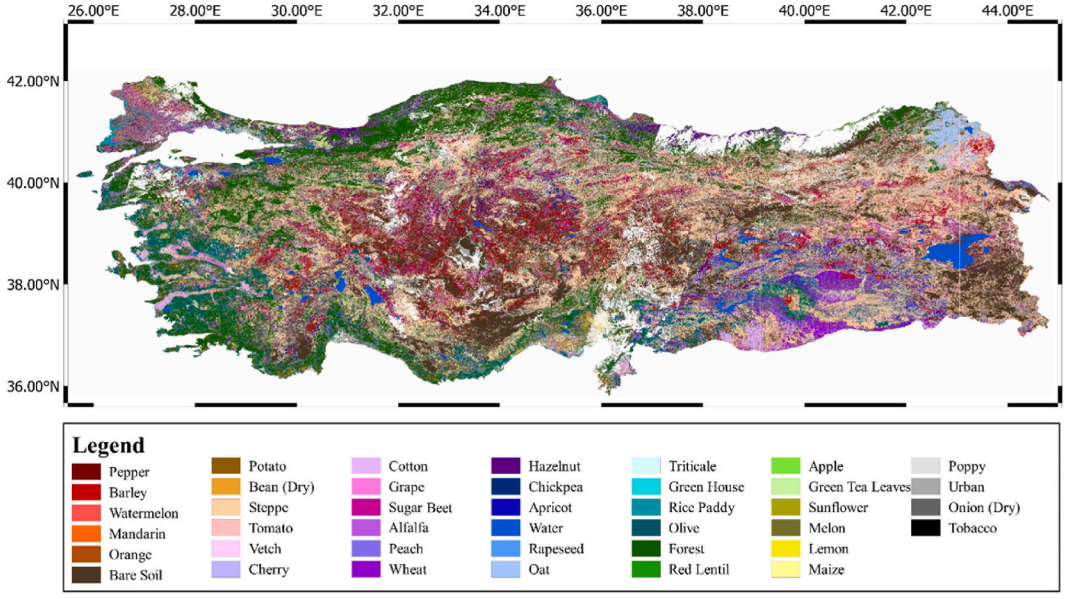

Fig. 5. Resultant Crop Classification Map of Türkiye for the year 2019.

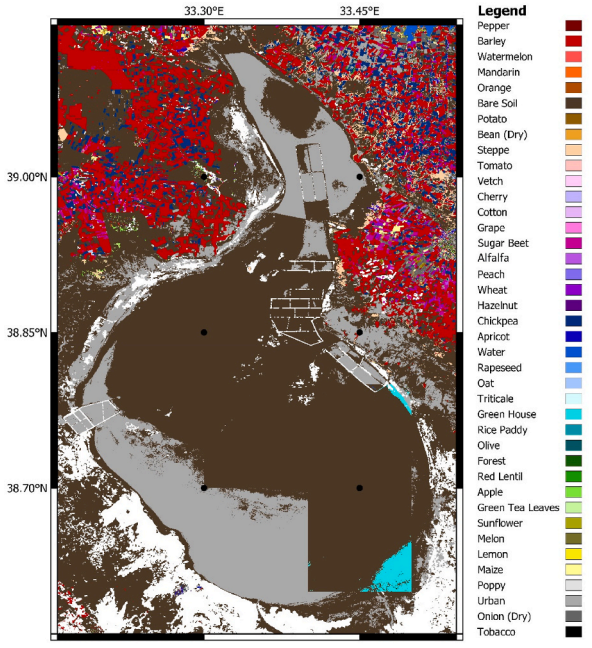

Fig. 6. Misclassification of Salt Lake as urban and greenhouse.

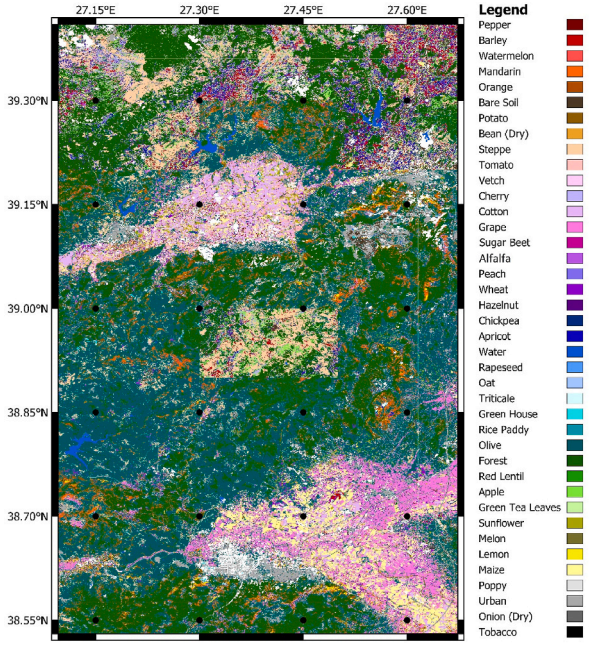

Fig. 7. Misclassification of maquis as olive in west Türkiye.

classifier's performance in distinguishing these crops is also observed, with these crops receiving 94%–100% PAs with test sets throughout all temperature zones that include them. Red lentil ($R^2 = 0.84$), barley ($R^2 = 0.71$), and sugar beet ($R^2 = 0.89$) classes exhibit good agreement between the two datasets. However, potato ($R^2 = 0.55$) and chickpeas ($R^2 = 0.44$) do not exhibit good agreement. When the confusion matrices of all classifiers are examined, it is clear that both crops are misclassified as other crops, with poor producer accuracies. Wheat is observed to have poor agreement between the datasets ($R^2 = 0.002$). Upon reviewing the confusion matrices for all temperature zones, it is observed that barley in LTZ, barley and triticale in MTZ, and barley, red lentil, and steppe classes in HTZ are frequently confused with wheat. This confusion may explain why the classification map has larger and lower percent areas across provinces than the TMA data. Despite some discrepancies between the two data sets in certain classes, the crop

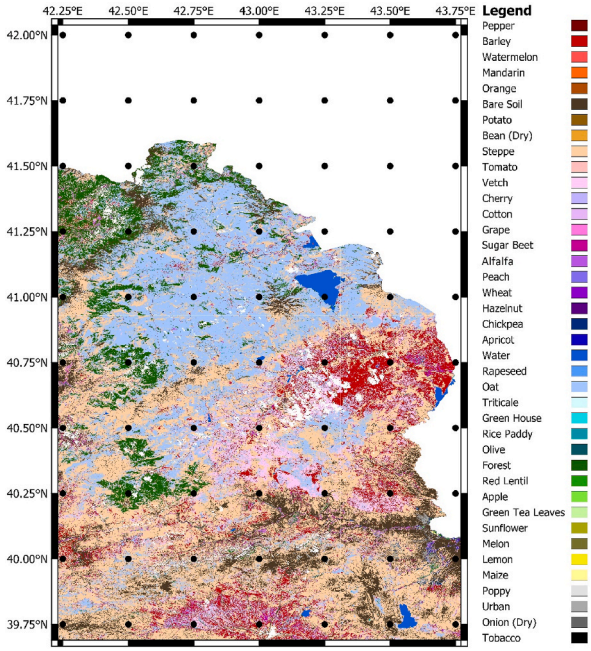

Fig. 8. Misclassification as oat in Northeast Türkiye.

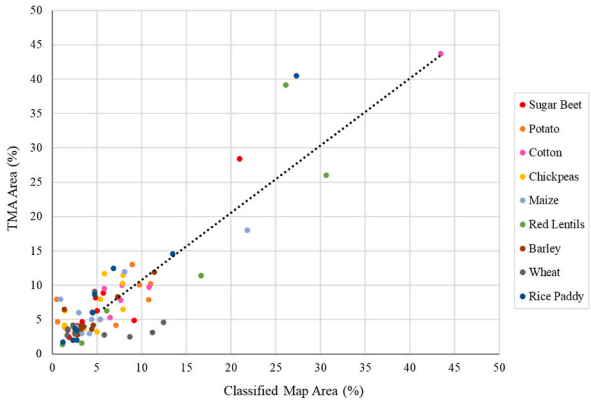

Fig. 9. Percent area comparison between the classified map and Ministry of agriculture and Forestry data for selected crops. The Trendline between the two is also illustrated, with the Corresponding coefficient of determination.

distribution over the study area is well represented in the classified map, with an overall R^2 of 0.85.

4. Conclusion

The purpose of this study is to examine, for the first time, the added value of dividing the study area based on temperature data to produce a more accurate, high-resolution, country-level crop map for 2019 that includes 34 crop and 6 land-cover classes with a spatial resolution of 10 m. For this purpose, the study area is divided into three temperature zones using 20-year mean ERA5-Land 2m air temperature product, and classification is performed over these zones separately with three different random forest classifiers, and for accuracy comparison purposes without the division step with one classifier for the whole study area. All bands except Band 10 of Sentinel-2 Level-2A are reduced to 15-day median images from March 15, 2019, to October 15, 2019. Apart from the spectral bands of Sentinel-2, the Normalized Difference Vegetation Index, Normalized Difference Water Index proposed by Gao (1996), and Normalized Difference Water Index proposed by McFeeters (1996) are calculated to be used as classification features. When contributions of classification features are inspected through random forest importance, it is observed that indices used in the study contributed to the classification accuracy, and they are recommended for similar studies.

When classifiers are applied to the test set, the Low-Temperature Zone reached 89% overall accuracy with a 0.88 Kappa coefficient for 17 land cover classes. The Medium-Temperature Zone reached 91% overall accuracy with a 0.92 Kappa coefficient for 35 land

cover classes, and the High-Temperature Zone reached %94 overall accuracy with a 0.94 Kappa coefficient for 34 land cover classes, giving the best accuracy among the classifiers. Finally, test sets of the temperature zones are combined, and an overall accuracy of 92% with a Kappa coefficient of 0.93 is achieved with this combined test set. When confusion matrices of the whole test set with and without temperature zones are examined, even though overall accuracy differs marginally, it is observed that more classes reach better accuracies with temperature zoning. Considering that the product map is a large-scale map, minor increases in accuracy mean better performance over a large region. In addition, although the division of ground truth data is disadvantageous due to the decrease in the number of training samples for each class, the increase in accuracy with temperature zones shows the importance and success of temperature zoning. Another sign of the success of the crop map is the similarity of crop distribution with the crop area data provided by the Ministry of Agriculture and Forestry. The implemented methodology is also temporally and spatially scalable as long as ground truth and remote sensing data are available since no year-specific measure is considered. In general, wetter/drier years might result in a reduced/increased number of available images, which might affect the accuracy of the resultant map. On the other hand, drier years would yield in reduced vegetation activity for the rain-fed crops, hence reduced temporal difference between the bands. This reduced signal difference in vegetation growth might impact the classification accuracy, but it needs to be separately investigated in a dedicated study.

The process of temperature zoning is not computationally costly, so for the classification of large areas that have multiple climate types, temperature zoning is recommended based on the results obtained from this study. In addition, because the crop map patches occur due to different temperature zones, smoothing of the temperature zones can be recommended for future work. The temperature thresholds selected arbitrarily for this study can be adjusted for better accuracy for future studies also considering the topographic variation. The study area can be divided into zones according to other factors affecting crop growth or crop distribution, like humidity (Heuvelink and Dorais, 2005), elevation (Machado et al., 2002) and precipitation, or a combination of those to improve the performance of study area division. Here, in addition to temperature, coupled used of elevation information is not expectedly improve the classification accuracy as the temperature goes down with the increased elevation (and vice versa) as a result on lapse rate, hence this elevation change related information is already included in the classification algorithm. On the other hand, relative humidity information is not included, hence, it might impact the classification accuracy. Another improvement that would benefit classification accuracy could be the method of handling cloud gaps. As with most optical satellite image processing studies, cloud cover was a major limitation in this study. More complex cloud gap filling algorithms like self-organizing Kohonen maps (SOMs) (Latif et al., 2008) can be considered or radar data can be used for improved accuracy in regions with persistent cloud cover problem. For future work, temperature information can be used as classification input, as it is in the study of Zhang et al. (2022), and results can be compared to a classification in which temperature is used for the division of the study area. The decrease in the number of training samples due to zone division could also potentially affect the accuracy for specific classes. Future research could address these limitations and continue to advance the accuracy and efficiency of crop mapping methodologies.

Overall, this study contributes to the field of agricultural research by highlighting the benefits of temperature zoning and providing a more accurate and informative crop classification map, which can be invaluable for water resource management, agricultural planning, and policymaking.

Ethical statement

All ethical practices have been followed in relation to the development, writing, and publication of the article.

CRediT authorship contribution statement

E. Donmez: Writing – original draft, Visualization, Software, Methodology, Formal analysis. **M.T. Yilmaz:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **I. Yucel:** Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Elif Donmez reports financial support was provided by Deutsche Forschungsgemeinschaft and the Scientific and Technological Research Council of Turkey (TÜBİTAK).

Data availability

The authors do not have permission to share data.

Acknowledgements

We gratefully acknowledge the European Space Agency (ESA) for providing the Sentinel satellite data. We are also grateful to Tarla.io for providing the crop type data used as ground truth. The first author is funded by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under Grant Number 117Y118 between November 2019 and July 2020, and under Grant Number 120R065 between January 2022 and August 2022, and finally by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). SFB 1502/1–2022 - Project number: 450058266, from November 2022.

References

Abdikan, S., Sekertekin, A., Ustuner, M., Sanli, F.B., Nasirzadehdizaji, R., 2018. Backscatter analysis using multi-temporal Sentinel-1 SAR data for crop growth of maize in Konya Basin, Turkey. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci 42, 9–13.

Alganci, U., Sertel, E., Ozdogan, M., Ormeci, C., 2013. Parcel-level identification of crop types using different classification algorithms and multi-resolution imagery in Southeastern Turkey. Photogramm. Eng. Rem. Sens. 79 (11), 1053–1065.

Arias, M., Campo-Bescós, M.Á., Álvarez-Mozos, J., 2020. Crop classification based on temporal signatures of Sentinel-1 observations over Navarre province, Spain. Rem. Sens. 12 (2), 278.

Asam, S., Gessner, U., Almengor González, R., Wenzl, M., Kriese, J., Kuenzer, C., 2022. Mapping crop types of Germany by combining temporal statistical metrics of sentinel-1 and sentinel-2 time series with LPIS data. Rem. Sens. 14 (13), 2981.

Batan, M., 2021. Planning the use of water in Şanlıurfa province, which struggles with drought: water footprint analysis. Journal of the Faculty of Engineering and Architecture of Gazi University 36 (4), 2135–2149.

Bauer, M.E., Hixson, M.M., Davis, B.J., Etheridge, J.B., 1978. Area estimation of crops by digital analysis of Landsat data. Photogramm. Eng. Rem. Sens. 44 (8), 1033–1043.

Bernard, S., Heutte, L., Adam, S., 2009. Influence of hyperparameters on random forest accuracy. International Workshop on Multiple Classifier Systems 171–180.

Braaten, J. (n.d.). Sentinel-2 cloud masking with s2cloudless. Google Earth Engine. Retrieved June 25, 2022, from https://developers.google.com/earth-engine/tutorials/community/sentinel-2-s2cloudless.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Bremer, D.J., Lee, H., Su, K., Keeley, S.J., 2011. Relationships between normalized difference vegetation index and visual quality in cool-season turfgrass: II. Factors affecting NDVI and its component reflectances. Crop Sci. 51 (5), 2219–2227. https://doi.org/10.2135/cropsci2010.12.0729.

Bulut, A.P., Canbaz, G.T., 2022. Sivas ilinde Bugday, Arpa, Seker Pancari ve Aycicegi Uretimi icin Su Ayak izinin Hesaplanmasi. Bilecik Seyh Edebali Universitesi Fen Bilimleri Dergisi 9 (1), 249–255.

Cakirli Akyüz, N., Theuvsen, L., 2021. Organic agriculture in Turkey: status, achievements, and shortcomings. Organic Agriculture 11 (4), 501–517. https://doi.org/10.1007/s13165-021-00362-2.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20 (1), 37–46.

d'Andrimont, R., Verhegghen, A., Lemoine, G., Kempeneers, P., Meroni, M., van der Velde, M., 2021. From parcel to continental scale – a first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. Remote Sensing of Environment 266 (October). https://doi.org/10.1016/j.rse.2021.112708.

European Space Agency. (2018). *Sentinel-2 MSI Level-2A BOA Reflectance*. European Space Agency. https://doi.org/10.5270/S2_-6eb6imz.

Gao, B., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment 58 (3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3.

Genuer, R., Poggi, J.-M., Tuleau-Malot, C., 2010. Variable selection using random forests. Pattern Recogn. Lett. 31 (14), 2225–2236.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth engine: planetary-scale geospatial analysis for everyone. Remote sensing of Environment 202, 18–27.

Griffiths, P., Nendel, C., Hostert, P., 2019. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. Remote Sensing of Environment 220, 135–151. https://doi.org/10.1016/j.rse.2018.10.031.

Günal, N., 2013. Türkiye'de iklimin doğal bitki örtüsü üzerindeki etkileri. Acta Turcica 1, 1–22.

Han, H., Guo, X., Yu, H., 2016. Variable selection using mean decrease accuracy and mean decrease gini based on random forest. In: *2016* 7th Ieee International Conference on Software Engineering and Service Science (Icsess), pp. 219–224.

Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, vol. 2. Springer.

Heuvelink, E., Dorais, M., 2005. Crop growth and yield. Crop Production Science in Horticulture 13, 85.

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational high resolution land cover map production at the country scale using satellite image time series. Rem. Sens. 9 (1), 95.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning, vol. 112. springer, New York, p. 18.

Latif, B.A., Lecerf, R., Mercier, G., Hubert-Moy, L., 2008. Preprocessing of low-resolution time series contaminated by clouds and shadows. IEEE Trans. Geosci. Rem. Sens. 46 (7), 2083–2096.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R. News 2 (3), 18–22.

Machado, S., Bynum, E.D., Archer, T.L., Lascano, R.J., Wilson, L.T., Bordovsky, J., et al., 2002. Spatial and temporal variability of corn growth and grain yield: implications for site-specific farming. Crop Sci. 42 (5), 1564–1576.

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Rem. Sens. 17 (7), 1425–1432. https://doi.org/10.1080/01431169608948714.

Ming, D., Zhou, T., Wang, M., Tan, T., 2016. Land cover classification using random forest with genetic algorithm-based parameter optimization. J. Appl. Remote Sens. 10 (3), 035021 https://doi.org/10.1117/1.jrs.10.035021.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13 (9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.

Myneni, R.B., Hall, F.G., Sellers, P.J., Marshak, A.L., 1995. The interpretation of spectral vegetation indexes. IEEE Trans. Geosci. Rem. Sens. 33 (2), 481–486.

Özür, N., Ataol, M., 2018. Türkiye'de CORINE verilerinin kullanılmasına dair değerlendirme. Çankırı Karatekin Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 9 (2), 110–130.

Palacios-Orueta, A., Khanna, S., Litago, J., Whiting, M.L., Ustin, S.L., 2006. Assessment of NDVI and NDWI spectral indices using MODIS time series analysis and development of a new spectral index based on MODIS shortwave infrared bands. Proceedings of the 1st International Conference of Remote Sensing and Geoinformation Processing 1, 207–209.

Piyoosh, A.K., Ghosh, S.K., 2018. Development of a modified bare soil and urban index for Landsat 8 satellite data. Geocarto Int. 33 (4), 423–442.

R Core Team, 2020. R: a language and environment for statistical computing. https://www.r-project.org/.

Rasul, A., Balzter, H., Ibrahim, G.R.F., Hameed, H.M., Wheeler, J., Adamu, B., et al., 2018. Applying built-up and bare-soil indices from Landsat 8 to cities in dry climates. Land 7 (3), 81.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogrammetry Remote Sens. 67 (1), 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002.

Rußwurm, M., Körner, M., 2017. Multi-temporal land cover classification with long short-term memory neural networks. Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci. 42.

Sensoy, S., Demircan, M., Ulupinar, Y., Balta, I., 2008. Climate of Turkey. Turkish State Meteorological Service, p. 401.

Siebert, S., Ewert, F., 2012. Spatio-temporal patterns of phenological development in Germany in relation to temperature and day length. Agric. For. Meteorol. 152 (1), 44–57. https://doi.org/10.1016/j.agrformet.2011.08.007.

Steduto, P., Faurès, J.-M., Hoogeveen, J., Winpenny, J., Burke, J., 2012. Coping with water scarcity: an action framework for agriculture and food security. FAO Water Reports 16, 78.

Stehman, S., 1996. Estimating the kappa coefficient and its variance under stratified random sampling. Photogramm. Eng. Rem. Sens. 62 (4), 401–407.

Story, M., Congalton, R.G., 1986. Accuracy assessment: a user's perspective. Photogramm. Eng. Rem. Sens. 52 (3), 397–399.

Sun, S., 2011. Meta-analysis of Cohen's kappa. Health Serv. Outcome Res. Methodol. 11 (3–4), 145–163. https://doi.org/10.1007/s10742-011-0077-3.

Turker, M., Arikan, M., 2005. Sequential masking classification of multi-temporal Landsat7 ETM+ images for field-based crop mapping in Karacabey, Turkey. Int. J. Rem. Sens. 26 (17), 3813–3830. https://doi.org/10.1080/01431160500166391.

Turkish Statistics Institute, 2018. Bitkisel Üretim Istatistikleri, 2018. Retrieved June 16, 2022, from. https://data.tuik.gov.tr/Bulten/Index?p=Bitkisel-Uretim-Istatistikleri-2018-27635.
Turkish Statistics Institute, 2020. Bitkisel üretim verileri. from. https://www.tarimorman.gov.tr/sgb/Belgeler/SagMenuVeriler/BUGEM.pdf. (Accessed 16 June 2022).
Undesa, P., 2017. World Population Prospects: the 2017 Revision.
Wallace, J.S., 2000. Increasing agricultural water use efficiency to meet future food production. Agric. Ecosyst. Environ. https://doi.org/10.1016/S0167-8809(00)00220-6.
Yang, N., Liu, D., Feng, Q., Xiong, Q., Zhang, L., Ren, T., Zhao, Y., Zhu, D., Huang, J., 2019. Large-scale crop mapping based on machine learning and parallel computation with grids. Rem. Sens. 11 (12), 1–22. https://doi.org/10.3390/rs11121500.
Yılmaz, İ., İmamoğlu, M., Özbulak, G., Kahraman, F., Aptoula, E., 2020. Large scale crop classification from multi-temporal and multi-spectral satellite images. 2020 28th Signal Processing and Communications Applications Conference (SIU), pp. 1–4.
Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. Int. J. Rem. Sens. 24 (3), 583–594.
Zhang, C., Chen, Y., Lu, D., 2015. Mapping the land-cover distribution in arid and semiarid urban landscapes with Landsat Thematic Mapper imagery. Int. J. Rem. Sens. 36 (17), 4483–4500.
Zhang, L., Gao, L., Huang, C., Wang, N., Wang, S., Peng, M., et al., 2022. Crop classification based on the spectrotemporal signature derived from vegetation indices and accumulated temperature. International Journal of Digital Earth 1–27.
Zhang, S., Tao, F., 2013. Modeling the response of rice phenology to climate change and variability in different climatic zones: comparisons of five models. Eur. J. Agron. 45, 165–176. https://doi.org/10.1016/j.eja.2012.10.005.
Zhang, T., Su, J., Liu, C., Chen, W.-H., Liu, H., Liu, G., 2017. Band selection in Sentinel-2 satellite for agriculture applications. 2017 23rd International Conference on Automation and Computing (ICAC), pp. 1–6.
Zheng, B., Myint, S.W., Thenkabail, P.S., Aggarwal, R.M., 2015. A support vector machine to identify irrigated crop types using time-series Landsat NDVI data. Int. J. Appl. Earth Obs. Geoinf. https://doi.org/10.1016/j.jag.2014.07.002.
Zhou, X., Zhu, X., Dong, Z., Guo, W., 2016. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. The Crop Journal 4 (3), 212–219.
Turkish Statistics Institute (2019). Bitkisel üretim istatistikleri, 2019. Retrieved June 16, 2022, from https://data.tuik.gov.tr/Bulten/Index?p=Bitkisel-Uretim-Istatistikleri-2019-30685.
QGIS.org, 2024. QGIS Geographic Information System. In: QGIS Association. https://qgis.org/en/site/getinvolved/faq/index.html.