Smart Agricultural Technology 12 (2025) 101650

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Field crop mapping using machine learning and multi-sensor satellite fusion: toward dynamic agricultural monitoring

Nechama Z. Brickner [a], Lior Fine [a,b], Offer Rozenstein [b], Tarin Paz-Kagan [a,*]

[a] French Associates Institute for Agriculture and Biotechnology of Dryland, The Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boqer Campus 8499000, Israel

[b] Institute of Soil, Water and Environmental Sciences, Agricultural Research Organization– Volcani Institute, HaMaccabim Road 68, Rishon LeZion 75359, Israel

ARTICLE INFO

Keywords:
Hierarchical classification
Crop type mapping
Sentinel-1 and sentinel-2
Random forest
Phenology
Crop rotation

ABSTRACT

Monitoring crop dynamics with precision is vital for food security and sustainable agricultural management. Yet, conventional monitoring approaches often lack sufficient spatial resolution, revisit frequency, and scalability, particularly in heterogeneous and fragmented agricultural landscapes. This study introduces a novel hierarchical remote sensing framework that integrates Sentinel-1 SAR and Sentinel-2 multispectral data with machine learning to generate high-resolution, multi-season crop type maps across extensive agricultural regions. Focusing on the western Negev in Israel (2018–2024), we developed a three-tier Random Forest classification workflow specifically designed for field crop mapping. The workflow integrates multi-temporal spectral, phenological, and SAR-derived features in a stepwise approach: (1) agricultural land cover types (94% overall accuracy), (2) wheat identification (95% accuracy), and (3) classification of 13 field crop types (81% accuracy), including key crops such as wheat, corn, and cotton. The hierarchical structure improved classification precision and enabled robust generalization across years, facilitating tracking of crop rotation, land-use intensity, and field-level management practices. Importantly, the model captured climate-driven phenological gradients in wheat, revealing spatial variability in crop development patterns that conventional methods often overlook. The novelty of this study lies in its scalable, multi-sensor classification framework that fuses radar, optical, and phenological information to support operational, dynamic crop monitoring. By coupling remote sensing outputs with national GIS infrastructure, this approach offers a cost-effective, transferable solution to advance precision agriculture, promote climate adaptation, and guide sustainable land-use planning at regional and national scales.

1. Introduction

Agricultural land is the primary source of food, providing approximately 90 % of global food calories and 80 % of protein and fats [1]. It covers approximately 33 % of the Earth's land surface, encompassing about 1.24 billion hectares. It supports a wide range of ecosystem services, including the provision of food, fuel, and fiber, which require accurate monitoring of agricultural cropland [2]. Yet accurately and automatically mapping crops at national to global scales remains a major challenge, particularly in the context of field crops that are highly dynamic in both space and time [3]. Among the most essential crops are wheat, rice, and maize, which serve as dietary staples for billions worldwide. Accurate field crop mapping at national and global scales is therefore vital for monitoring food security, anticipating supply chain disruptions, and guiding policy interventions. It also enables timely responses to climate-induced yield variability, pest outbreaks, and water resource challenges, enhancing resilience in both local farming systems and international markets. Despite agriculture's global importance, our understanding of its spatial and temporal dynamics remains limited, especially in areas experiencing rapid land-use transitions or climatic stress [4]. Traditional monitoring systems often lack the spatial resolution, revisit frequency, and scalability required to capture agricultural dynamics effectively, particularly in field crop systems. This monitoring gap limits timely decision-making and hinders the assessment of sustainability outcomes, especially in complex, fragmented, or data-scarce regions.

To address these limitations, advanced technologies such as satellite remote sensing and machine learning (ML) have become indispensable for agricultural monitoring and management [5]. These technologies enable automated mapping of cropping patterns, land-use changes, and

* Corresponding author.
E-mail address: tarin@bgu.ac.il (T. Paz-Kagan).

https://doi.org/10.1016/j.atech.2025.101650
Received 28 August 2025; Received in revised form 17 November 2025; Accepted 17 November 2025
Available online 18 November 2025

related environmental impacts [6], providing a scalable, cost-effective, and data-driven foundation to support sustainable intensification, adaptive land management, and resilience to climate variability [7]. However, agricultural land use frequently changes over short timeframes, as multiple crops may be grown and harvested within a single year, each with distinct phenological cycles [3,8]. This high temporal and spatial variability underscores the need for frequent, high-resolution crop classification maps at the field level to facilitate informed decision-making in agricultural policy and resource allocation [6].

Despite this need, most agricultural maps today rely on governmental GIS databases or records from agrarian organizations [8,9]. These data are often based on infrequent or limited-scale field surveys [10]. Such surveys are typically costly, labor-intensive, and have a restricted geographical scope [11]. Therefore, integrating remote sensing and ML techniques is crucial for developing automated, accurate, and timely crop mapping models. When properly calculated and validated with field data, these methods can enable the production of automated crop maps at regional to national scales, thereby enhancing the capacity for effective agricultural monitoring and decision-making. Numerous models and operational frameworks have been developed for crop mapping using satellite data, particularly in countries with large-scale agricultural monitoring programs. For instance, the United States and Canada produce annual 30-meter-resolution crop maps that serve as benchmarks for remote-sensing-based agricultural monitoring [3,10]. In the U.S., the Cropland Data Layer (CDL), produced by the USDA's National Agricultural Statistics Service, has provided annual crop type maps since 2008. The CDL integrates imagery from multiple sensors, including Landsat, MODIS, and Sentinel-2 (S2), using state-level classification models [12]. In Canada, the Annual Crop Inventory (ACI) has been generated by Agriculture and Agri-Food Canada (AAFC) since 2009, using both optical (Landsat, S2) and synthetic aperture radar (Synthetic Aperture Radar - SAR, RADARSAT-2) data [13]. These national-scale products demonstrate the feasibility and utility of remote sensing for consistent and standardized agricultural monitoring. Recent studies increasingly leverage S2 imagery due to its high spatial (10–20 m), temporal (5-day), and spectral resolution, making it highly suitable for monitoring dynamic cropping systems [8,10,14,15]. Nonetheless, Landsat imagery remains valuable, especially for long-term time-series analysis, given its continuous archive dating back to the 1980s [16].

A growing number of studies utilize multi-sensor fusion, combining data from various platforms to improve classification accuracy and robustness. Common combinations include S2 and Sentinel-1 (S1, SAR) [17], S2 and Landsat 8 [3], or more complex fusions like S2, S1, and PlanetScope [18] and S2, S1, and Landsat 8 [9]. Research consistently shows that multi-source integration enhances classification performance, particularly in cloud-prone regions, by increasing the frequency of usable observations [9,18]. For example, Rao et al. [18] reported that over half of the S2 images during their study period were cloud-obscured, resulting in only six usable scenes, while S1 provided 17 cloud-independent acquisitions. Similarly, Orynbaikyzy et al. [17] demonstrated that incorporating SAR data improved the spatial transferability of crop classification models to new geographic regions. Another promising development is the use of hierarchical classification models, which first categorize broad land cover types and subsequently refine classification into specific crop types. This strategy is particularly effective in handling imbalanced datasets and large numbers of crop classes, thereby improving model accuracy and generalization [15,16].

Despite progress, several limitations persist in current crop-mapping approaches. Many national-scale systems, such as the CDL and ACI, are tailored to specific geographic and agronomic contexts, which can limit their transferability to other regions or cropping systems, particularly in arid, heterogeneous, or data-scarce landscapes [19]. Existing classification approaches often focus on dominant crops or single time points, limiting their ability to capture crop diversity, rotations, and climate-driven phenological shifts at the field scale [10,18]. These models often rely on extensive, high-quality training data, which may not be readily available, consistent, or up to date in many parts of the world [19]. Furthermore, their reliance on predefined crop calendars and static land cover classifications [20] can reduce flexibility in capturing non-traditional, short-season, or multi-cycle cropping patterns, which are common in dryland and intensively managed systems [21]. Maintaining high classification accuracy is also challenging in environments where crop types exhibit spectral similarity or where fields are small, fragmented, and frequently changing, as is the case in many agricultural dryland regions. These limitations underscore the urgent need for adaptive, transferable, and data-efficient crop-mapping frameworks that deliver reliable outputs under variable environmental conditions and limited field data availability. In this study, we addressed these challenges by implementing a multi-sensor satellite fusion approach, combined with ML, tailored to the complex characteristics of dryland agriculture in Israel's western Negev. Our approach was applied in regions characterized by small, fragmented fields, limited ground-truth data, and highly dynamic crop management practices, demonstrating the potential of flexible, scalable models to improve crop classification in challenging agroecosystems.

Building on the need for adaptive, data-efficient crop-mapping approaches, satellite time-series (TS) data have emerged as essential tools for capturing crop development dynamics across the growing season [22]. Crop phenology can be detected from multi-temporal satellite imagery, ranging from single-scene classifications to dense TS [8,16,17]. Using TS data enhances classification accuracy by capturing key phenological changes, enabling models to distinguish between crops based on their temporal profiles. However, TS analysis presents several challenges, including cloud contamination, inconsistent revisit frequencies, and data gaps. To address these issues, pre-processing techniques such as outlier removal, gap filling (e.g., linear interpolation), and smoothing algorithms (e.g., Savitzky-Golay, Whittaker filter, and Fourier transforms) are applied [3,23]. Moreover, TS data supports the identification of multiple crop cycles within a year, enabling the detection of crop rotations and field transitions, which are crucial for mapping highly dynamic agricultural systems, such as those in dryland regions.

To further enhance classification, vegetation indices (VIs), mathematical combinations of spectral bands, are used to highlight specific crop traits, such as chlorophyll content or canopy density [6,24]. From these VI-TS, phenological features are extracted to characterize growth cycles more precisely. Features such as Start of Season (SOS), End of Season (EOS), Maximum of Season (MOS), Length of Season (LOS), and Seasonal Integral (area under the VI curve) provide crop-specific signatures that can be used to distinguish between crop types [8,16,25]. Various methods exist to derive these features, including threshold-based approaches, moving averages, and percent amplitude techniques [8,22,25,26]. Incorporating phenological features enhances model accuracy, particularly when distinguishing between crops with overlapping spatial or spectral characteristics but distinct seasonal behaviors. In this study, these phenological indicators, derived from VI-TS, are central to classifying crops in small, fragmented, and highly dynamic dryland fields. A knowledge gap remains in applying VI-TS and phenological metrics to assess wheat performance across aridity gradients. VI-TS improves crop classification by capturing temporal growth patterns, especially when using phenological features (e.g., SOS, EOS, MOS) [8,16,27]. However, to the best of our knowledge, no studies have leveraged them to explore spatial variations in wheat development across varying local and regional climatic gradients and vulnerable drylands. This study addresses this gap using robust VI-TS analysis along a dryland aridity gradient.

The study aimed to develop an automated crop mapping framework capable of detecting short-term fluctuations in the agricultural dynamics of field crops driven by climatic variability in the western Negev, Israel. The specific objectives were: (1) To develop a multi-sensor, automated crop hierarchical classification approach for mapping agricultural land

cover, field crops, and wheat, with a particular focus on the western Negev. (2) To identify and characterize changes in agricultural land use, including shifts in crop types, phenological deviations, or temporary field abandonment related to environmental conditions. (3) To examine intra- and inter-annual variability in wheat phenology along an aridity gradient using VI-TS analysis, with the aim of disentangling the effects of climatic factors. Our research hypothesizes that the high-resolution, multi-sensor VI-TS and ML models will enable accurate crop classification and interannual monitoring of field crops, even in spatially heterogeneous and dynamic dryland environments. Moreover, we hypothesize that wheat fields will exhibit delayed phenological development, with later SOS, MOS, and EOS dates, due to shifting climatic conditions. Ultimately, a clear north-to-south productivity gradient will be evident in wheat production across the aridity gradient, with lower productivity observed in drier regions. This gradient will be reflected in VI-derived metrics, such as MOS values and LOS, and will be further shaped by spatial heterogeneity in climate impacts and agricultural practices.

The novelty of this study lies in several methodological and operational advances. First, we developed a MOS-based cross-index and MOS-aligned SAR feature selection approach that synchronizes optical and radar inputs at phenological peaks, enhancing classification across crop types. Second, we employed an expert-driven feature selection strategy informed by agronomic knowledge, enabling us to identify biologically meaningful predictors beyond purely statistical rankings. Third, we tested a hierarchical Random Forest (RF) workflow across seven years (2018–2024), demonstrating robust multi-year generalization in a fragmented dryland environment. Collectively, these innovations provide a transferable framework for operational crop monitoring at regional to national scales.

2. Materials and methods

2.1. Study area

The study area spans approximately 1700 km^2 in the western Negev, bounded by Ashkelon to the north, Route 6 to the east, and the border with Egypt to the south (Fig. 1). Of this area, more than 1000 km^2 is used for agriculture, with an elevation range from sea level to 300 m [28]. The region experiences a strong north–south rainfall gradient, with average annual precipitation ranging from 500 mm in the north to just 100 mm in the south [29]. The climate spans three zones: hot-summer Mediterranean (Csa), hot semi-arid (BSh), and hot desert (BWh), based on the Köppen–Geiger classification [30]. The Aridity Index (AI) ranges from 0.077 to 0.235, categorizing the region as arid to semi-arid [31]. Approximately 25,000 agricultural plots in the area include orchards, row crops, greenhouses, and fallow land. Wheat, sown between early November and late December, occupies about 432 km^2 (~40 %) of agricultural land. Optimal sowing occurs in mid-November when soil moisture is available. However, due to the unpredictability of rainfall, many farmers sow in dry soil, especially in years with delayed rainfall [32].

The Western Negev is a key agricultural region in Israel, particularly known for its diverse field crop production, making it an ideal case study for examining the effects of climatic aridity and external stressors on crop classification and productivity. Notably, the region is also recognized as the wheat granary of the State of Israel, due to its extensive wheat cultivation and its critical role in national grain production [33]. Globally, wheat is a staple crop for over 35 % of the population, accounting for approximately 20 % of daily caloric and protein intake [34]. Although significant research has addressed wheat genetics, drought resilience, and yield modeling, a knowledge gap remains in

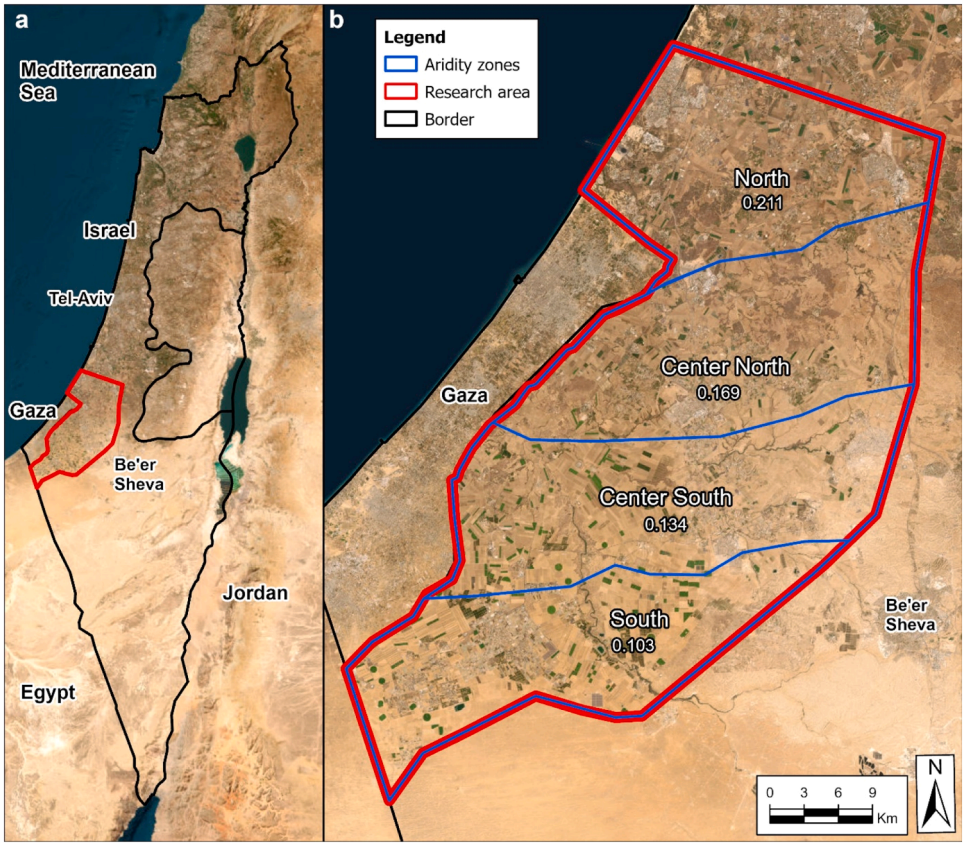

Fig. 1. The study area is in the western Negev region of Israel. (a) A polygon representing the study area in Israel, and (b) a zoomed-in image of the study area, where the red polygon is the study area and the blue is the aridity zones. The names of each zone are listed below the corresponding mean aridity index.

applying VI-TS and phenological metrics to assess wheat performance across aridity gradients. In Israel, wheat yields are relatively low, about 2–3 t/ha, compared to countries like Egypt, which exceeds 6 t/ha due to extensive irrigation [35]. Within Israel, productivity is higher in the northern regions, while yields tend to be lower in the southern, more arid areas. Israel's wheat production is primarily rainfed and constrained by limited arable land, making it particularly vulnerable to climate variability, including droughts and heatwaves.

2.2. Experimental design

The experimental design of this study comprises five key components, which together enable the development and validation of an automated crop-mapping framework tailored for field crops in the Western Negev region (Fig. 2). **Step 1: Data Inputs** - The study uses two primary types of input data: vector and raster. Vector data were obtained from official Geographic Information System (GIS) layers provided by the Israeli Ministry of Agriculture (MOAG), datasets purchased from three professional sources that map agricultural fields, and a dedicated field survey conducted in 2024. Raster data includes satellite imagery from S1 and S2, as well as high-resolution aerial imagery. Satellite data served as the core input for classification modeling, while aerial imagery was used to label agricultural plots for training of land cover types. **Step 2: Satellite Preprocessing** - The S2 optical imagery underwent a multi-stage preprocessing workflow that included: cloud masking, VIs calculation, parcel-based time-series extraction, temporal gap filling and smoothing, feature extraction, including phenological metrics (e.g., SOS, LOS), and feature selection. In parallel, S1 SAR imagery was processed to generate monthly and quarterly median composites, and mean backscatter values were extracted for each parcel to capture structural changes in crop fields. **Step 3: Hierarchical Classification** - A classification approach that follows a predefined class taxonomy or tree structure, where broader classes are subdivided into more specific ones at successive levels. A three-level hierarchical ML framework was developed to enhance classification accuracy in this complex agricultural landscape, encompassing agricultural land cover classification, wheat classification, and specific field crop identification. This hierarchical approach enables progressive refinement in classification, thereby enhancing model robustness and interpretability. **Step 4: Post-Processing and Spatiotemporal Analysis** - This phase involved several analyses to derive actionable insights from the classification outputs, including the generation of annual crop classification maps, change detection to assess shifts in land use or crop types, Mann-Kendall (MK) trend analysis to identify statistically significant changes in agricultural dynamics over time, and analysis of wheat cultivation patterns. **Step 5: Wheat Analysis Across the Aridity Gradient** - This final step evaluates wheat performance under varying climatic conditions. It integrates classified wheat plots, climate data (rainfall and temperature), AI-determined aridity zones, and median NDVI time series per zone. This step aims to assess phenological development, productivity, and resilience of wheat in response to climate gradients and external disturbances, offering a deeper understanding of crop-environment interactions in arid and semi-arid systems. Each work stage described above is detailed in the following sections, where we outline the methodology, data processing workflows, classification models, and analytical procedures used in this study (Fig. 2).

2.3. Step 1: data inputs

2.3.1. Vector data – GIS plots, and field survey

The vector data were primarily derived from the GIS database maintained by the MOAG, which includes geospatial information for approximately 200,000 agricultural plots nationwide. This database provides annual historical datasets dating back to 2017. It is regularly updated using data submitted by farmers, the Plants Production and Marketing Board, and official surveys, the most recent of which was conducted in 2019. Plot updates are based on two main criteria: (1) the crop type, which reflects changes in cultivation, and (2) plot boundary delineation. However, despite these updates, plot-level information for field crops is often outdated or inconsistent, limiting its reliability for direct use in classification. As a result, for this study, only plot boundaries from the GIS database were used from 2018 to 2023, with the 2023 plot boundaries extended for use in 2024. Additionally, the MOAG GIS department identified and extracted approximately 2000 highly reliable

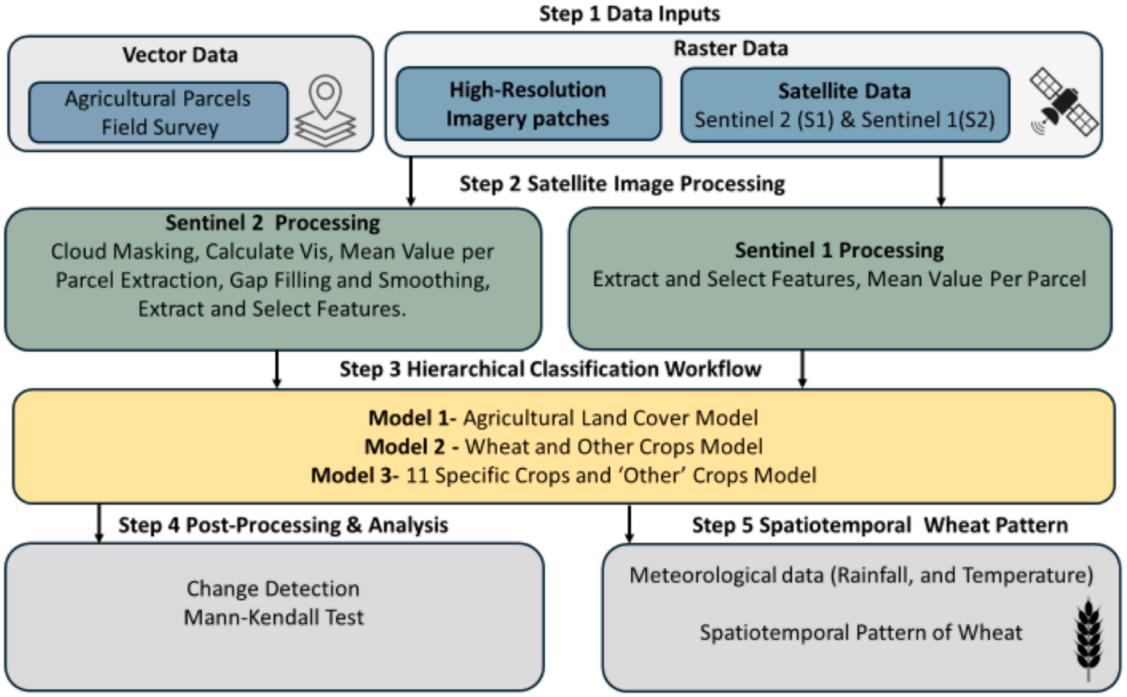

Fig. 2. The experimental design is divided into five main stages: (1) data inputs, (2) satellite image processing, (3) hierarchical classification workflow, (4) post-processing analysis, (5) spatiotemporal wheat pattern.

field crop plots based on confirmed farmer reports from 2018 to 2023. These records were used to generate training data, including 249–489 plots per year, of which 141–227 were wheat plots. To supplement this dataset and ensure up-to-date ground truth for the 2024 season, a dedicated field survey was conducted in collaboration with a professional mapping company.

For this study, a dedicated field survey was conducted in 2024 in collaboration with a professional mapping company to collect up-to-date reference data on crop locations. The survey was performed in two phases, winter and summer, to capture seasonal variations and ensure comprehensive coverage of plots and crop types throughout the year. Survey plots were distributed across multiple climate regions of Israel to achieve spatially representative coverage. It was also stratified by crop type (14 categories). Special efforts targeted underrepresented crops such as melon, onion, and pumpkin. Nevertheless, the limited sample sizes for some rare crops remain a potential source of bias. These field observations were recorded as point data (XY coordinates), and 1694 plots were added. Plot boundaries were then delineated from this point data using a combination of manual digitization and spatial joining with the MOAG crop layer. Satellite imagery from multiple timepoints (before, during, and after the survey period) was used to validate crop growth stages and refine boundary accuracy.

To further enhance the robustness of model training, we integrated additional datasets purchased from three professional sources that map agricultural fields in Israel between 2018 and 2024. These included two Israeli companies providing farm management software and Kanat, an insurance fund for agricultural natural disasters in Israel. All external datasets were validated against field observations and cross-checked with the Ministry of Agriculture GIS layers before use. Table S1 summarizes the combined datasets by year, showing the growing volume and diversity of validated training data with a total of 12,205 validated plots incorporated into the analysis (Table S1).

2.3.2. Raster data -satellite imagery

The primary data sources used in this research were satellite imagery from S2 and S1, both of which are part of the Copernicus program operated by the European Space Agency (ESA). These satellites are designed as two-satellite constellations positioned 180° apart to ensure high temporal coverage. S2 provides multi-spectral optical imagery, capturing detailed spectral information about vegetation, while S1 supplies SAR data, which is particularly useful for structural analysis and monitoring under all weather conditions. The S2 satellite is equipped with the Multi-Spectral Instrument (MSI), which includes 13 spectral bands with spatial resolutions of 10 m, 20 m, or 60 m, and a revisit frequency of 5 days (or 10 days per satellite). The 10 m bands include blue, green, red, and near-infrared (NIR), while the 20 m bands include red-edge bands, additional NIR, and two shortwave infrared (SWIR) bands. The 60 m bands provide information on aerosol, water vapor, and cirrus clouds. S1, on the other hand, is an active radar sensor that operates in the C-band (5.54 cm wavelength), enabling it to capture imagery day and night, regardless of cloud cover or lighting conditions. S1 supports several imaging modes, including single (HH or VV) and dual (HH+HV or VV+VH) polarizations. The mode used in this study is the Interferometric Wide Swath (IW) mode, which provides dual-polarization data (VV and VH) and is ideal for monitoring terrestrial surfaces. Together, the fusion of optical (S2) and radar (S1) datasets provides complementary information on crop spectral properties and structural characteristics, enabling robust classification and monitoring of agricultural dynamics.

2.4. Step 2: satellite preprocessing and processing

2.4.1. S2 optical imagery preprocessing

S2 imagery is organized in a tile format, with nine tiles covering the entire territory of Israel and two specifically covering the study area. This research utilized the S2_SR_Harmonized collection in Google Earth Engine (GEE), which includes Level-2A surface reflectance products that have been pre-processed through atmospheric correction. The analysis period spans from October 2017 to September 2024, corresponding to Israel's hydrological year, and comprises a total of 3760 images. To ensure high-quality data, tiles with more than 75 % cloud cover were excluded using a cloud filter [9]. Cloud and shadow pixels were further removed using the s2cloudless algorithm [36]. The images were then scaled by dividing pixel values by 10,000 to normalize reflectance values between 0 and 1. Outlier values below 0 or above 1 were clipped accordingly. Subsequently, a set of spectral indices was calculated (Table S2). These indices are mathematical combinations of reflectance values at specific wavelengths, designed to enhance the visibility of vegetation and soil characteristics [6]. Among these, VIs are particularly effective for assessing crop growth, health, biomass, and chlorophyll content [24]. In this study, eight spectral indices were computed for each image and employed in both time-series and classification analyses. The definitions and formulations of these VIs are presented in Table S2.

2.4.2. S1 SAR imagery preprocessing

S1 provides SAR data through the S1_GRD product available on GEE. Due to variations in orbit direction and ground geography, images of the same location may appear different. S1 offers two polarization modes: VV (vertical-vertical) and VH (vertical-horizontal). In addition to these original polarizations, three additional bands were derived. The first two are the ratios between the polarizations, VV/VH and VH/VV. The third band derived is the Radar Vegetation Index (RVI) [37], calculated as follows, Eq (1):

$$RVI = \frac{4 * VH}{VH + VV} \quad \text{(Eq 1)}$$

One approach to utilizing SAR data is to use image composites [9,14,38]. Typically, a speckle filter removes the salt-and-pepper noise commonly observed in SAR images before creating the composite image [9,14]. However, Lindsay et al. [38] found that averaging the pixel values over time eliminated the need for speckle filtering. For this research, quarterly and monthly median image composites were used. Median composites were selected instead of averages to reduce the salt-and-pepper effect. The quarterly composites were divided into the following periods: October–December, January–March, April–June, and July–September. Quarterly composites utilized only the ascending orbit, whereas monthly composites were generated for both ascending and descending orbits. These composites provided consistent, temporally representative SAR datasets for input into classification models and time-series analysis.

2.4.3. Data processing: filling gaps and smoothing

Following the initial storage of S2 data, post-processing was applied to refine the TS by removing outliers, filling data gaps, and smoothing temporal fluctuations. The original dataset was acquired at nominal 5-day intervals, but frequent cloud contamination resulted in missing dates. NDVI was the first index processed due to its well-established value range and predictable response to vegetation dynamics [39]. Typically, NDVI TS shows a clear growth trend, rising as vegetation develops, peaking near full canopy, and eventually plateauing or declining depending on crop phenology. Mid-season anomalies (e.g., sudden drops) are often attributed to unfiltered clouds or shadow artifacts. To standardize temporal resolution, the TS data were converted to a 365-day daily format, with missing values marked as NA. Two filtering functions were applied to identify and remove outliers. The first function flags values falling more than 0.05 below a 30-day moving average. The second removed values, which deviated by more than 0.1 from their immediate neighbors, were excluded after NAs were removed. These outliers were replaced with NA, followed by linear interpolation to bridge the gaps [3].

For gaps at the TS extremities, backfill and front-fill techniques were applied using the first and last valid image values. For smoothing,

various algorithms were considered, including asymmetric Gaussian, double logistic, Savitzky-Golay, Fourier-based, and Whittaker filters [23]. The Whittaker smoothing algorithm was ultimately selected for its superior performance and efficiency [40]. This method, a penalized least squares approach, balances fidelity and smoothness by minimizing the roughness of the fitted curve. It uses a single parameter, λ (lambda), to control the smoothing level: higher values produce smoother curves, while lower values retain more variability [40]. Compared to other tested algorithms, the Whittaker smoother preserved NDVI peak values, which are essential for accurate phenological analysis. This smoothing workflow was then extended to all spectral bands and VIs. However, instead of recalculating outliers for each index, the outlier positions identified in the NDVI series were reused. This decision was based on the tailored NDVI thresholds, which are not directly transferable due to differing index ranges. Subsequent steps, including gap-filling and Whittaker smoothing, remained consistent across all indices.

2.4.4. *Data processing -Feature extraction and selection from S2 and S1*

Key phenological features, such as the SOS, MOS, and EOS, offer valuable insights into vegetation growth cycles, crop productivity, and environmental responses, and are frequently used in crop classification models [8]. Since the TS dataset is structured by day, using all daily values as model inputs would yield an excessively high number of features. Therefore, specific informative features were extracted to reduce dimensionality while preserving relevant temporal information [41]. From the NDVI and Plastic Greenhouse Index (PGHI) time series, two fundamental features were extracted for each year: the annual maximum value and the corresponding day of year (DOY). The maximum NDVI was used to distinguish bare fields from vegetated classes, such as orchards and row crops, while the PGHI maximum helped differentiate covered crops (e.g., greenhouses) from vegetation and bare soil. Additionally, the NDVI-TS was aggregated to half-monthly means, yielding 24 temporal segments per year. This aggregation enables temporal pattern recognition, which is particularly useful for differentiating between crop types. For instance, orchards typically maintain high NDVI values year-round, whereas row crops exhibit sharp seasonal peaks and extended periods of low NDVI before sowing and after harvest. Phenological features were extracted from the NDVI-TS to describe crop growth phases. In fields with multiple cropping cycles per year, the TS was first segmented into two separate series, each representing a crop cycle. A dual crop was defined by the presence of two NDVI peaks separated by at least 75 days and a local minimum [8]. In such cases, values associated with the secondary crop cycle were reset to 0.1 before analysis. From each time series (or cycle), 15 phenological features were derived: SOS, MOS, and EOS: their dates, DOY, and NDVI values, LOS: difference between EOS and SOS in DOY, Enhancement Period (EP) and Reduction Period (RP) in DOY, Seasonal Integral (Area Under the Curve; AUC), along with EP and RP integrals, These metrics reflect the intensity and timing of crop development and senescence phases [22,26]. To calculate SOS and EOS, a moving average and a reverse moving average were applied, each using a 15-day window. SOS is defined as where the moving average curve intersects the raw NDVI-TS, and EOS is where the reverse moving average intersects the TS. However, if either intersection occurs after the NDVI exceeds a threshold of 0.2, that threshold is used instead to define the respective phenological marker [25] (Fig. 3).

In this study, phenological features were extracted primarily from the NDVI-TS due to NDVI's well-documented relationship with vegetation dynamics [39] and its standardized, interpretable value range. While other VIs offer sensitivity to specific vegetation properties, such as red-edge reflectance, chlorophyll absorption, or water content, they often have variable value ranges and less consistent temporal profiles, making phenological feature extraction more complex [24]. Therefore, instead of extracting full phenological metrics from all VIs, we focused on using the MOS date, as determined from the NDVI-TS, to extract corresponding reflectance or index values from other bands and VIs. These included S2 bands B2 (Blue), B3 (Green), B4 (Red), and B8 (NIR), as well as the 6 vegetation indices (Table S2). This approach allowed us to capture complementary vegetation information at peak growth while maintaining methodological consistency.

S1-SAR features were derived from plot-level mean values [9], aggregated into both quarterly and monthly composites. From the quarterly composites, three key bands, VV, VH, and their ratio (VV/VH), were retained. These quarterly datasets were used primarily for

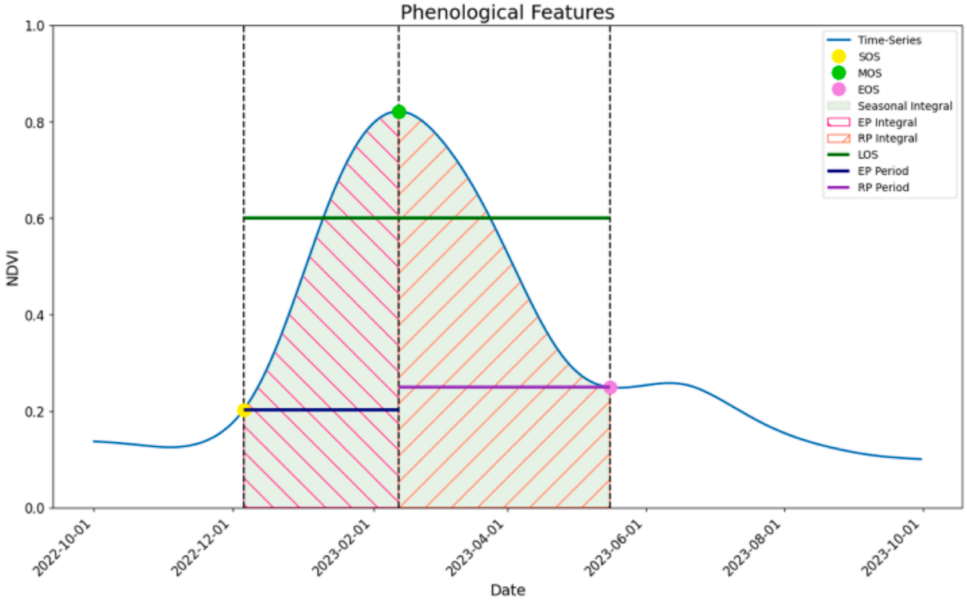

Fig. 3. Phenological features extracted from the NDVI time series (NDVI-TS): Start of Season (SOS) – yellow dot, End of Season (EOS) - pink dot, Maximum of Season (MOS) - green dot. The Length of Season (LOS) is represented by the green horizontal line, the Enhancement Period (EP) by the dark blue line, and the Reduction Period (RP) by the purple line. The Seasonal Integral (also referred to as the Area Under the Curve, AUC) is represented by the green shaded area. The EP Integral is marked with pink diagonal lines, and the RP Integral with orange diagonal lines. These features reflect the timing and magnitude of crop development and senescence within a growing season.

agricultural land cover classification. In contrast, the monthly composites were utilized for modeling specific crop types and included a broader range of features: VV, VH, VV/VH, VH/VV, and the RVI. To reduce the dimensionality of the monthly feature set while preserving relevance to crop phenology, only the SAR values from the month of MOS, as identified from the NDVI time series (NDVI-TS), were selected for final modeling.

In this study, we adopted an expert-guided approach and statistical feature selection, rather than relying solely on conventional statistical methods. This strategy enabled the selection of features specifically aligned with the objectives of each classification task: (1) agricultural land cover classification and (2) crop-type discrimination. For agricultural land cover classification, we prioritized features that maximize separability among land cover classes. In contrast, for crop-specific classification, we emphasized features associated with peak vegetation growth, which are most informative for distinguishing between crop types. Table S3 summarizes the selected features for each classification task. To optimize model performance and reduce redundancy, feature selection was conducted using two complementary techniques: the Boruta algorithm to identify the most informative predictors, and the Shapley Additive exPlanations (SHAP) approach to assess feature importance [42–44]. The Boruta algorithm, a wrapper around RF, identifies all relevant features by comparing the importance of actual predictors against that of randomly permuted "shadow" variables (Fig. S1). This approach is particularly effective for high-dimensional datasets, enhancing both model performance and interpretability [42]. Boruta analysis was implemented using the Boruta package in R [43]. To verify the robustness of our expert-driven feature selection, we also performed independent validation using SHAP analyses (Fig. S2). SHAP values were computed to interpret the contribution of individual features to model predictions. SHAP provides a unified, model-agnostic framework for quantifying feature influence using cooperative game theory, enabling the assessment of each variable's impact on classification outcomes, including its magnitude and direction. This feature selection approach offers two key advantages: (1) It incorporates domain knowledge, ensuring that the features reflect meaningful biophysical or phenological processes together with standard statistical patterns. (2) It reduces the risk of overfitting and enhances model interpretability, as the selected features are directly linked to agricultural and seasonal dynamics relevant to the study.

2.5. Step 3: hierarchical crop classification

2.5.1. Selecting classification models

This study evaluated three ML models, including RF, Support Vector Machine (SVM), and AdaBoost, using the Scikit-Learn Python library. Preliminary testing on the training data revealed that all three models performed comparably; however, the RF model consistently achieved the highest accuracy and robustness. Given its widespread application and proven effectiveness in crop classification [17], RF was selected for the final classification workflow. RF is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes for classification tasks [45]. It is particularly effective for remote sensing applications, such as crop type classification, due to its ability to handle high-dimensional, non-linear, and heterogeneous data [46]. RF also provides insights into feature importance, enabling the identification of variables that contribute most to the model's accuracy. Key hyperparameters in RF include the number of trees (estimators), maximum tree depth, and minimum samples per split. While increasing the number of trees can improve performance, it may also increase computational time. To optimize the RF model, a grid search was performed to identify the optimal hyperparameter combination. Despite its high performance, one limitation of RF is its interpretability; the decision-making process is not easily visualized, making it a "black box" model [47]. Preliminary grid-search optimization tested variations in tree depth (6–12) and the number of estimators (100–500). Accuracy gains were marginal (<1 %), so standardized parameters (100 trees, maximum depth = 8) were adopted across all models to balance performance and computational efficiency.

This study implemented a hierarchical classification framework comprising three sequential RF models (Fig. 4). The framework was designed to reflect the inherent hierarchical structure of agricultural land cover classification, beginning with broad agricultural land cover types and narrowing to specific crop types. This also allowed the use of different features for each classification task. Additionally, the hierarchical approach was also effective in handling data imbalance by partitioning the dataset into targeted subsets for each classification stage. In this system, the output of each model serves as the input for the subsequent model. The workflow progresses from general agricultural land cover classification (Model 1) to binary classification of wheat versus other crops (Model 2), and finally to multiclass classification of 11 specific crop types (Model 3).

***Hierarchical -Model 1*:** The first model was designed to classify four major agricultural land cover types: orchards, covered areas, bare fields, and row crops. We employed an RF classifier with 100 trees and a maximum depth of 8, utilizing 40 features derived from TS-VIs and SAR quarterly mean composites (Table S3). The labeled dataset was divided into about 66 % for training (with cross-validation), 17 % for validation, and 17 % for test data to assess model performance (Table S4).

Hierarchical -Model 2: The second model was applied exclusively to the "row crops" class identified by Model 1, with the goal of distinguishing wheat from all other row crops. Due to class imbalance, with wheat being overrepresented, all non-wheat crops were aggregated into a single "other" class. This resulted in a dataset containing 4775 wheat-labeled plots and 7419 plots of other crops. The RF classifier was configured with 100 trees, a maximum depth of 8, and 32 input features (Table S3). The data were split into approximately 69 % for training via cross-validation, 17 % for validation, and 14 % for testing on independent data to evaluate model performance (Table S5).

Hierarchical -Model 3: The third model further classified the "other" row crops identified in Model 2 into 11 specific crop types: chickpea, corn, cotton, melon, onion, peanut, potato, pumpkin, sunflower, tomato, and watermelon, with any remaining crops grouped as "other." The RF model used the same configuration (100 trees, depth = 8, 32 features). To mitigate data imbalance, a maximum of 150 training points per crop class was used. Crops with fewer than 150 available samples were used in full. The final dataset includes 7419 plots, split into about 70 % (calibration), 17 % (validation), and 13 % (test) to assess model performance (Table S6). To evaluate model robustness across years, we conducted explicit cross-year validation using 2023 as an independent test year. For Model 1, the 2023 dataset (≈17.5 % of all data) was reserved for testing, with the remaining 2018–2022 data split 80/20 for training and validation. For Models 2 and 3, the 2023 dataset (≈14 % and ≈13 % of the total data, respectively) was used for testing, while the 2018–2022 and 2024 datasets were used for training and validation, with an approximate 70/15/15 split. This design allowed independent testing of model generalization under varying climatic and management conditions (Tables S4-S6).

2.5.2. Statistical analysis- accuracy metrics

To evaluate model performance, several metrics were used, including Producer's Accuracy (PA), User's Accuracy (UA), Overall Accuracy (OA), F1 Score, and the Kappa Statistic [48]. PA (recall) measures the proportion of correctly classified samples within each class relative to the total number of samples in that class. This metric assesses the model's ability to accurately identify instances of a given class. The UA (precision) quantifies the proportion of correctly classified samples within a class relative to the total number of samples assigned to that

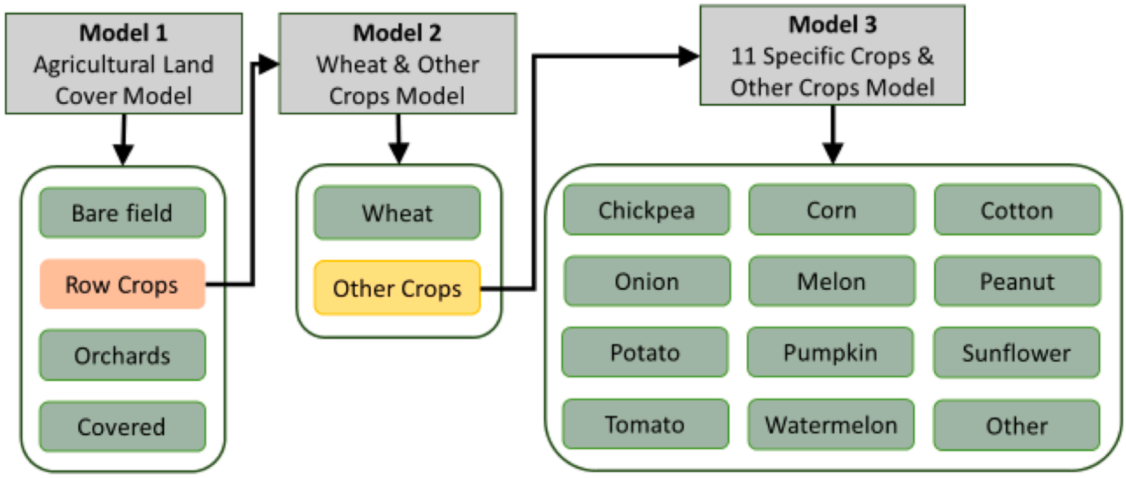

Fig. 4. Hierarchical classification workflow for crop mapping, composed of three sequential Random Forest models. Model 1 distinguishes between agricultural land cover types. Model 2 classifies wheat (including barley and oats) versus other crops within agricultural pixels. Model 3 further classifies the remaining non-wheat crops into specific crop types (e.g., cotton, corn, sunflower). The output of each model serves as input for the next, forming a top-down decision-making process.

class. It indicates the likelihood that a classified point truly belongs to the assigned class. OA is the proportion of correctly classified samples relative to the total number of samples. The F1 score, a harmonic mean of precision and recall, balances the correctness of positive predictions with the ability to identify actual positive cases. It ranges from 0 to 1, where 1 indicates perfect classification, and 0 signifies poor performance. The Kappa statistic evaluates classification performance by measuring agreement between predicted and actual labels while accounting for chance agreement. Kappa values range from −1 to 1, with 1 indicating perfect agreement, 0 signifying agreement equivalent to random chance, and negative values suggesting worse than random performance.

2.6. Step 4: post-processing and spatiotemporal analysis

We applied post-classification change detection to examine transitions between agricultural land cover classes over time. Post-classification comparison is a widely used method for detecting changes in land cover. It is typically presented in a "from–to" format, which identifies the specific nature of class transitions between two time points [49]. Using the output from the first classification model (agricultural land cover classes), each pair of consecutive years was compared. The analysis included only plots with identical IDs and geometries across the years to ensure reliable spatial and temporal alignment. This approach enabled the identification of stable classes, land-use conversions, and possible disturbances or shifts in agricultural practices within the study area.

To evaluate temporal patterns and detect trends in land cover dynamics, the MK trend test was applied. The MK test is a nonparametric statistical method designed to assess monotonic trends in time-series data without assuming any specific distributional form [50,51]. However, the standard MK test does not account for seasonality or autocorrelation, which can lead to biased results. Therefore, modified versions such as the Seasonal MK test were also employed [52]. The trend analysis was performed using the pyMannKendall Python package [53] and included three distinct tests: (1) Annual Category Trends: The original MK test was used to evaluate trends in the total number of plots classified into each agricultural land cover category across the study period. (2) Monthly Crop Classification Trends**:** A Seasonal MK test was used to assess temporal trends in the number of classified plots per crop type for each month. The monthly classification results were based on phenological segmentation using SOS and EOS dates. (3) NDVI Trend in Persistent Wheat Fields: A focused seasonal MK test was applied to wheat plots that remained consistently classified as wheat across all seven study years (2018–2024). Only plots with stable plot IDs and geometries were included to ensure consistency. The objective was to detect trends in NDVI values over time while avoiding interference from inter-crop variability. Since each crop has unique phenological characteristics and NDVI ranges, isolating wheat ensures that the trend analysis accurately reflects long-term changes in wheat productivity and vegetation vigor.

2.7. Step 5: wheat analysis across aridity gradient

The Western Negev presents a unique opportunity to study wheat phenology and productivity across an aridity gradient. To assess this, we utilized an AI raster dataset for the region [54]. Based on the AI values, the study area was divided into four distinct aridity zones (Fig. 1) using natural breaks classification in ArcGIS Pro. Climate data, including temperature and rainfall, were collected for the agricultural years 2018–2024. Temperature data were obtained from 10 monitoring stations and rainfall data from 45 stations, provided by the Israel Meteorological Service (IMS). Temperature values were processed as daily mean temperatures (°C), and rainfall was recorded as daily precipitation (mm). For each aridity zone, climate data was averaged across the corresponding meteorological stations to derive zone-level climate profiles. To examine the effects of aridity on wheat growth, we analyzed the median NDVI time series of plots classified as wheat in each aridity zone. A one-way analysis of variance (ANOVA) was conducted to test statistically significant differences in phenological metrics between zones and years. A post hoc Tukey's Honest Significant Difference (HSD) test [55] was applied to assess pairwise differences across zone–year combinations. The following six phenological parameters were evaluated: SOS, DOY, MOS DOY, EOS DOY, LOS, and NDVI value at MOS, as well as AUC. All statistical analyses were performed in Python using the StatsModels package.

3. Results

3.1. Crop classification models

Model 1 – Agricultural Land Cover Classification: The first model classified the data into 4 agricultural land cover types using 40 features (Table S3). These features included yearly maximum value of NDVI and PGHI, the Day of Year (DOY) corresponding to the maximum NDVI, the annual mean PGHI value, half-monthly mean NDVI values, and quarterly mean values of the VV, VH, and VV/VH Sentinel-1 SAR bands. The feature importance from Model 1 shows the relative contribution of each

variable to classification accuracy across the two feature selection approaches tested (Fig. S2). Among all features, the maximum NDVI value and the yearly average PGHI emerged as the most influential. The model achieved strong performance metrics, including an OA of 94.87 %, an F1 Score of 94.87 %, and a Kappa Score of 91.21 on the test dataset. These results demonstrate the model's robustness in distinguishing among agricultural land cover classes using a combination of optical and radar-derived features.

The confusion matrix reveals that misclassifications predominantly occur between orchards and row crops, as well as between covered and bare fields (Fig. S3). Additionally, a subset of covered field samples was misclassified as either orchards or row crops. The confusion between orchards and row crops is likely due to the phenological similarity between deciduous trees and spring crops during certain periods of the growing season. Furthermore, orchards with dense weed growth between tree rows exhibit spectral signatures like those of row crops, which contribute to classification errors. Misclassifications between covered and bare fields can be explained by net coverings that show color and reflectance properties like those of bare soil. These covers often display low PGHI values, making them difficult to distinguish from uncovered soil. Additionally, NDVI signals can sometimes penetrate certain cover types, further complicating the differentiation process. Finally, the misclassification of covered plots as orchards or row crops likely results from partial visibility of vegetation signals through the coverings, leading to spectral overlap with actual vegetation classes. Table 1 presents the PA and UA for the test dataset. The results indicate high accuracy, with the lowest PA at 82.24 % and the lowest UA at 88.0 % in the bare fields.

Fig. 5 presents annual agricultural land cover maps for the study period, illustrating the spatial distribution of different land cover types. Most of the area is dominated by field crops (~57 %), while bare fields (~7 %) and orchards (~17 %) are more scattered. Orchards tend to be clustered, whereas most covered plots (~19 %) are in the southern part of the study area, often near settlements such as Moshavim, Kibbutzim, and other residential zones. Figure S4 displays the number of plots and total area for each agricultural land cover type across the study period. The data shows that the number and area of orchards and covered plots remain relatively stable over time. In contrast, bare fields and row crops exhibit greater interannual variability. A notable pattern is the inverse relationship between row crops and bare fields, as the number or area of one increases, the other tends to decrease. This trend is particularly evident between 2018 and 2022. However, in 2024, although the number of field crop plots increased relative to 2023, their total area decreased, while the total area of bare fields increased. This discrepancy indicates that bare field plots in 2024 were, on average, larger than in previous years. When analyzing the numbers of plots, orchards, and covered plots, consistent levels are observed, averaging around 4,200 and 4,600 plots annually. In contrast, bare field plots vary more significantly, ranging from 1,500 to 3,200 plots. Despite this, the total area of orchards, covered plots, and bare fields is relatively similar. However, orchards consistently occupy a larger area, suggesting that the average orchard plot is larger than the average covered plot.

Model 2 – Wheat and Other Crops Classification - The second model was designed to distinguish wheat plots from other crop types. The dataset was categorized into two classes: wheat and other. The other class serves as an intermediate category, which will be further classified into specific crop types in the next modeling step (Model 3). This model was built using 32 features, including phenological metrics derived from the NDVI TS (SOS, MOS, EOS), both DOY and NDVI values, LOS, EP, RP, Area Under the Curve (AUC), and its derivatives AUC_EP and AUC_RP. In addition, spectral features included MOS value-based indices on the NDVI date, MOS for B2, B3, B4, and six vegetation indices (Fig. S3). Mean monthly SAR band values: VV, VH, VV/VH, VH/VV, and RVI, for both ascending and descending modes, corresponding to the MOS timing (Table S3). Among all features, SOS-DOY, VH-ASC, and MOS-DOY were identified as the most influential in classification performance based on the two feature selection approaches tested (Fig. S2). The model achieved high classification accuracy, with an OA of 95.17 %, an F1 score of 95.18 %, and a Kappa score of 90.25 on the test dataset. Figure S5 presents the confusion matrix for the test set, demonstrating excellent classification performance: Wheat plots were correctly classified in 96.78 % of cases, and other plots were correctly classified in 93.93 % of cases. Table 2 presents the PA and UA for each class for the test set. The results indicate high classification performance, with both PA and UA exceeding 92.55 %. These results confirm the model's strong performance in distinguishing wheat from other row crops, ensuring reliable classification for further use.

Model 3 – Classification of 11 Specific Crops and an "Other" Category - The third model classified agricultural plots into 12 field crop categories, comprising 11 specific crops and one "other" category. This model used the same 32 features as Model 2, with the most influential being MOS-DOY, EOS-DOY, and SOS-DOY, as determined by the two feature selection approaches tested (Fig. S2). The model achieved the following performance metrics on the test set: OA of 81.15 %, F1 Score of 79.43 %, and Kappa Score of 78.25 %. The confusion matrix for the test dataset (Fig. S6) shows high classification accuracy for several crops, including cotton (98.17 %), peanut (93.33 %), and watermelon (97.5 %). These classes showed minimal misclassification. However, lower accuracies were observed for crops such as melon (10.64 %), pumpkin (29.41 %), and the tomato class (56.0 %). A significant portion of misclassifications involved the 'other' category. For example, Onion plots were misclassified as 'other' in 11.0 % of cases, and Chickpea: 13.51 %. PA and UA for each class are summarized in Table 3. Key insights from the classification performance include high PA and UA (>80 %) for the following crops: Corn, Cotton, Onion, and Peanut. High PA but lower UA are observed for the following crops: Sunflower (PA = 94.44 %, UA = 73.91 %) and Watermelon (PA = 97.5 %, UA = 60.94 %). This indicates that while actual plots were accurately identified, many predicted instances were misclassified as these crops. Lower PA but relatively high UA for: Melon: PA = 10.64 %, UA = 83.33 %; Pumpkin: PA = 29.41 %, UA = 90.91 %; Melon showed a notably low PA (10.64 %) but a relatively high UA (83.33 %), suggesting that while actual melon plots were often misclassified, the predictions labelled as melon were generally correct. Notably, 30.04 % of melon plots were misclassified as watermelon, highlighting similarities in spectral or phenological characteristics (Fig. S6). Frequent misclassifications between classes included: 'Other' → Watermelon (4.59 %), Cotton (1.02 %), Potato (1.53 %); Specific crops → 'Other': Chickpea (13.51), Corn (8.94), Melon (23.4 %), Onion (11.0 %), Potato (20.88 %), Tomato (20.0 %), Pumpkin (5.88 %) (Fig. S6). These findings underscore the challenge of accurately classifying diverse crops with overlapping phenology, highlighting the need for refined class definitions or additional data layers (Fig. 6).

Using the classification results and key phenological features, specifically the SOS and EOS dates, seven classification maps were generated, one for each agricultural year (defined as October to September). Each map consists of 12 panels, representing monthly snapshots, thereby capturing the spatiotemporal dynamics of crop phenology throughout the year. Fig. 7 illustrates the 2024 classification map, showing how agricultural plots transition across crop classes over time. The map highlights the seasonal dominance of different crop types:

Table 1
Producers' and Users' Accuracy (%) for the four agricultural land cover classes test dataset, assessing classification performance for each class in terms of omission and commission errors.

Class	Producer's Accuracy %	User's Accuracy %
Bare Field	82.24	88.0
Covered	92.86	94.03
Orchards	94.09	89.74
Row Crops	96.84	97.42

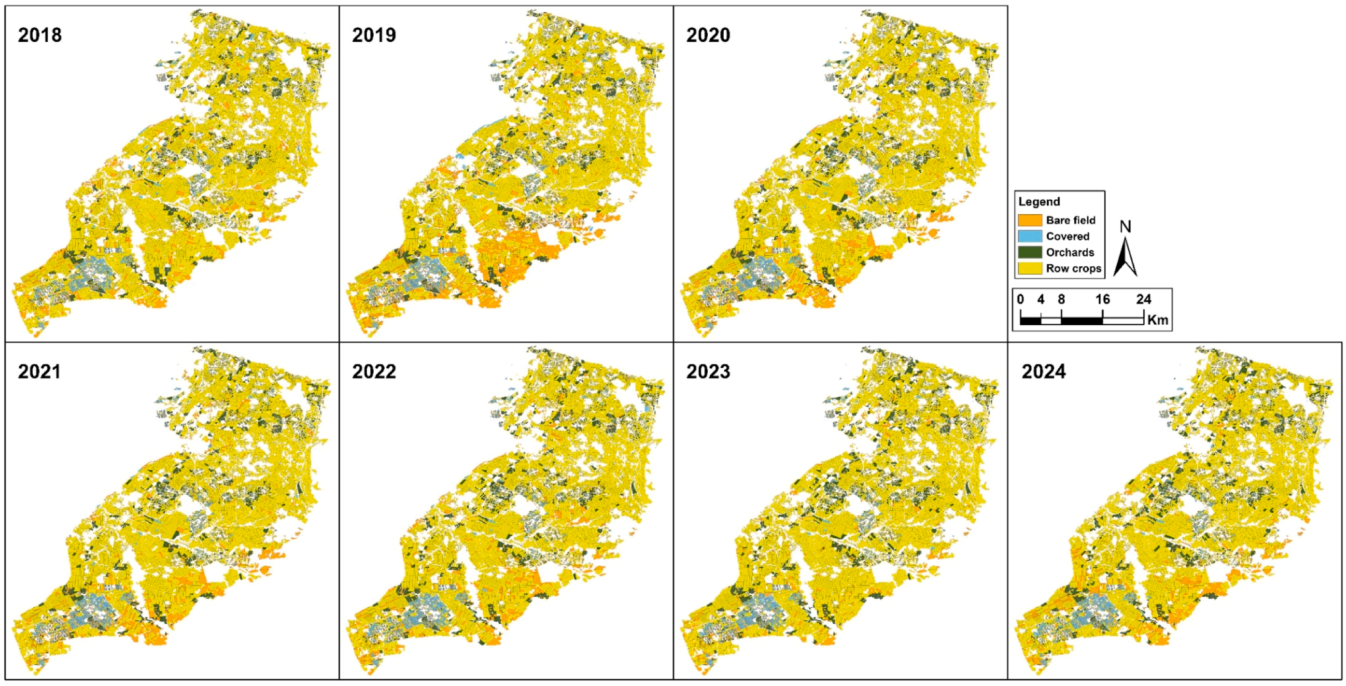

Fig. 5. Annual agricultural land cover maps for the 2018–2024 agrarian years, showing the spatial distribution of major agricultural land cover types across the study area.

Table 2
Producers' and Users' Accuracy (%) for the **wheat** and intermediate **'other'** crop classes test dataset, evaluating the classification performance of Model 2 in terms of omission and commission errors.

Class	Producer's Accuracy %	User's Accuracy %
Other	93.93	97.39
Wheat	96.78	92.55

Table 3
Producers and Users' Accuracy (%) for the 11 specific crops and one "other" category test dataset, evaluating the classification performance of Model 3 in terms of omission and commission errors.

Class	Producer's Accuracy %	User's Accuracy %
Chickpea	62.16	76.67
Corn	89.43	94.83
Cotton	98.17	96.41
Melon	10.64	83.33
Onion	82.0	87.23
Other	89.23	70.85
Peanut	93.33	87.5
Potato	79.12	83.72
Pumpkin	29.41	90.91
Sunflower	94.44	73.91
Tomato	35.9	90.32
Watermelon	97.5	60.94

wheat appears prominently as a winter crop. The 'other' category comprises crops that are cultivated throughout the year, reflecting their phenological and crop diversity. Most specific crop types are predominantly grown in the spring and summer. When a plot is not actively cultivated, i.e., falls outside its SOS–EOS window, it is classified as a bare field, emphasizing periods of fallow or land preparation. Figure S7 presents a monthly summary of the number of plots by agricultural cover type and specific crop class across the study period. The top panel shows that wheat and 'other' consistently constitute the largest share of cultivated plots annually. These phenological transitions correlate with temporal patterns in wheat (yellow) and 'other' (green) classes. In addition, the transition between active cultivation and fallow periods is reflected in the bare field category, which peaks during the summer months and declines during winter, when winter crops, such as wheat, dominate. While the total number of plots per class remains relatively stable across years, the temporal distribution within each year reveals subtle trends, which are further analysed in the following section.

3.2. Change detection of the Western Negev area

Change-detection analysis was conducted to evaluate interannual transitions among agricultural land cover classes. The assessment was restricted to plots that retained the same ID and geometry across consecutive years, ensuring consistency in spatial units for temporal comparison. Table S7 summarizes the number of plots that either changed class or remained unchanged for each land cover type across year pairs. Fig. 8 spatially visualizes the distribution of these class transitions. The results indicate that the most frequent changes occurred between bare fields and row crops, with bidirectional transitions observed across nearly all years. These two classes exhibit the highest interannual variability. Spatially, Fig. 8 shows that most changes are concentrated in the southern region of the study area, with an additional cluster of high change activity along the southeastern boundary. The total number of plots exhibiting land cover change ranged from 2492 to 3477 plots per year, representing approximately 11 % to 15 % of all plots. In contrast, orchards and covered plots showed high temporal stability, with minimal annual transitions. These findings underscore the dynamic nature of field crop land use in the Western Negev, particularly in relation to management practices and seasonal agricultural decisions, while also highlighting the stability of perennial and infrastructure-associated land uses. When training in 2018–2022 and 2024 for Models 2–3 and testing in 2023, overall accuracy and F1-scores declined by less than 2–3 % relative to within-year validation, these results confirm that the hierarchical workflow maintains strong predictive capability and generalization across different years and environmental conditions.

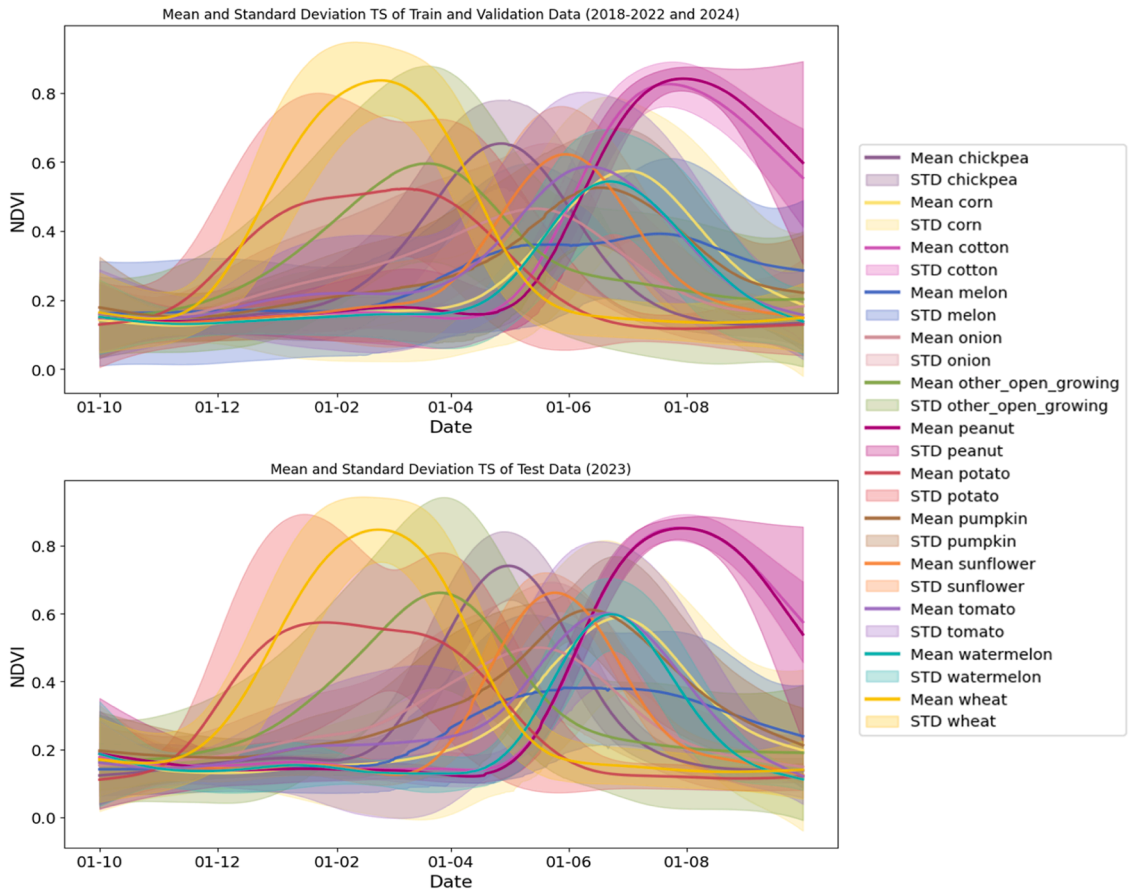

Fig. 6. Mean and standard deviation of the time series (TS) NDVI values for the training and validation data (on top) and the test data (on the bottom), grouped by crop type across all years. Note: The "wheat" class includes barley and oats.

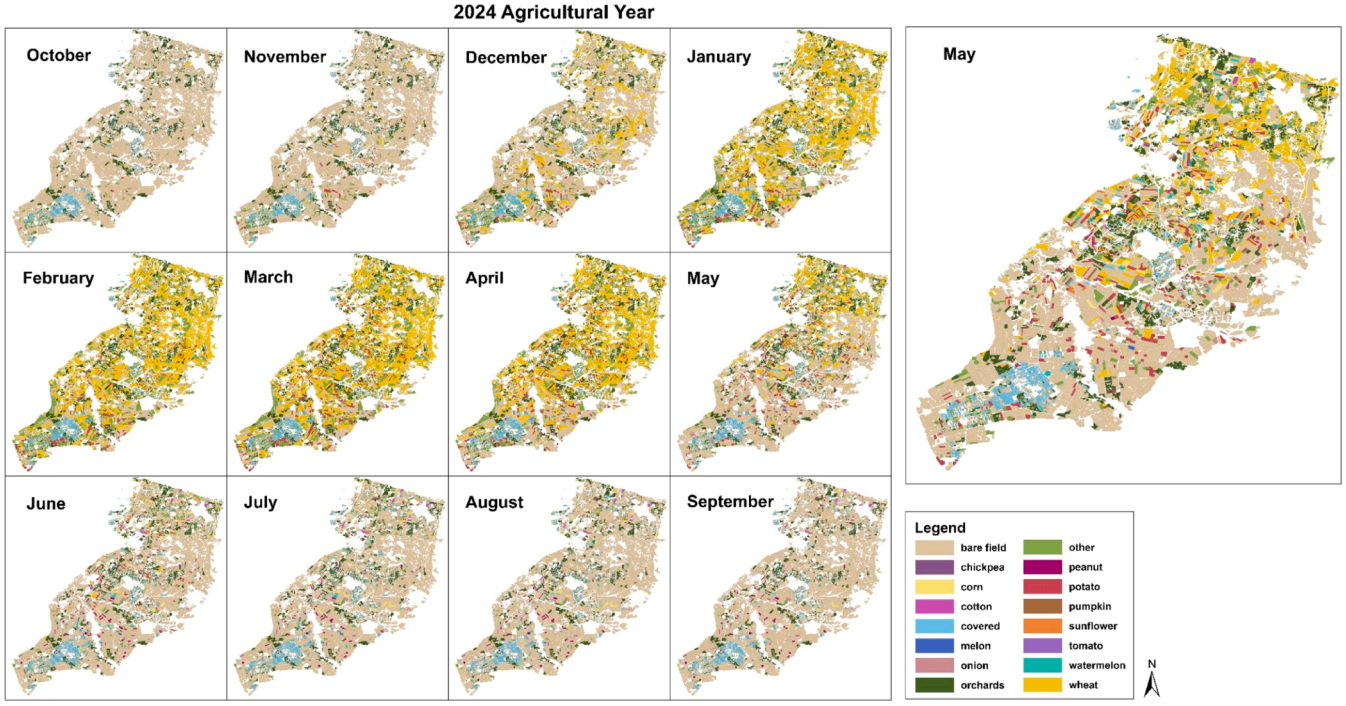

Fig. 7. Monthly classification map for the 2024 agricultural year, generated using Start of Season (SOS) and End of Season (EOS) dates for each plot. The map illustrates the spatiotemporal distribution of crop types across months, highlighting seasonal transitions in cultivation status. Note: The wheat class includes barley and oats.

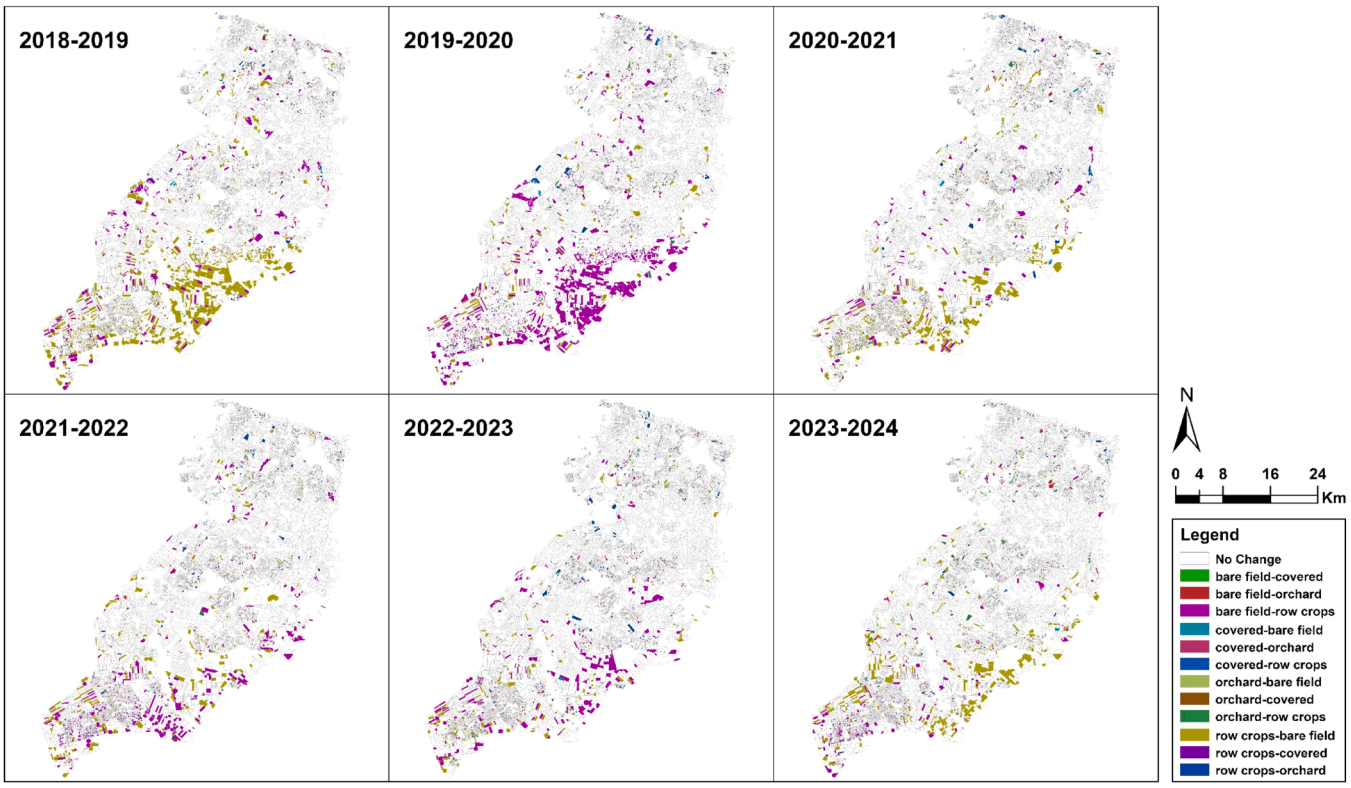

Fig. 8. Spatial distribution of changes between consecutive years. Only plots with the same ID and geometry for pairs of successive years were included in the calculation. The legend format follows: Change from X to Y.

3.3. Mann-Kendall trend analysis of the western negev area

Multiple MK trend tests were conducted on different subsets of the classification results, using both the original MK and the seasonal MK approaches. In the first test, the original MK method was applied to seven annual observations to assess temporal trends in the number of plots classified into each land cover category over the study period. The results revealed statistically significant ($p < 0.05$) increasing trends for orchards and corn, while no significant trends were detected for the remaining classes (Table 4). The second analysis employed the seasonal MK test to detect trends at a monthly scale, based on classification results aligned with SOS and EOS dates. This test identified significant upward trends in covered plots and orchards ($p < 0.05$). In contrast, significant decreasing trends were observed for chickpeas and watermelons. The remaining classes did not exhibit statistically significant trends (Table 4). The next trend analysis focused specifically on wheat, the most prevalent field crop in the study area, to evaluate long-term vegetation dynamics using the seasonal MK test on NDVI values. The analysis included plots classified as wheat (including barley and oats) across all seven agricultural years, with consistent plot ID and geometry.

Table 4
Results of the original and seasonal Mann-Kendall tests assessing trends in the number of plots by crop type over the study period. The table includes the test statistics, significance level (p-value), and trend direction (increasing, decreasing, or no trend) for each crop.

	Original MK				Seasonal MK			
Crop	Trend	Z-score	Tau	p-value	Trend	Z-score	Tau	p-value
Bare field	no trend	−1.202	−0.429	0.230	no trend	−0.824	−0.079	0.410
Chickpea	no trend	−1.802	−0.619	0.072	decreasing	−2.106	−0.183	0.035*
Corn	increasing	2.886	0.952	0.004**	no trend	1.410	0.111	0.159
Cotton	no trend	0.901	0.333	0.368	no trend	−0.238	−0.024	0.812
Covered	no trend	1.502	0.524	0.133	increasing	5.680	0.524	0.000***
Melon	no trend	0.615	0.238	0.539	no trend	−0.541	−0.044	0.589
Onion	no trend	−0.601	−0.238	0.548	no trend	−1.739	−0.163	0.082
Orchards	increasing	2.103	0.714	0.035*	increasing	7.761	0.714	0.000***
Other	no trend	0.601	0.238	0.548	no trend	0.260	0.028	0.795
Peanut	no trend	0.000	0.000	1.000	no trend	0.393	0.032	0.694
Potato	no trend	0.000	0.048	1.000	no trend	−1.141	−0.107	0.254
Pumpkin	no trend	0.000	0.048	1.000	no trend	−1.119	−0.099	0.263
Sunflower	no trend	−0.901	−0.333	0.368	no trend	−1.286	−0.107	0.199
Tomato	no trend	−0.354	−0.200	0.724	no trend	−0.309	−0.024	0.757
Watermelon	no trend	−1.502	−0.524	0.133	decreasing	−2.385	−0.214	0.017*
Wheat	no trend	1.202	0.429	0.230	no trend	1.649	0.155	0.099

* $p < 0.05$,.
** $p < 0.01$,.
*** $p < 0.001$.

A total of 492 plots met these inclusion criteria. The results revealed that 251 plots (51.01 %) exhibited a significant increasing trend in NDVI ($p < 0.05$), 172 plots (34.95 %) showed a significant decreasing trend, and 69 plots (14.02 %) demonstrated no significant trend. These findings highlight the phenological variability and dynamic changes that occur even within the same crop class, reflecting underlying influences such as climate variability, soil conditions, and agronomic practices.

3.4. Spatiotemporal patterns of wheat in the western negev area

To examine the influence of climatic conditions on wheat dynamics, the study area was divided into four aridity zones: North, Center North, Center South, and South (Fig. 1). For each zone, the NDVI time series of plots classified as wheat (including barley and oats) were analyzed alongside rainfall and temperature data from 2018 to 2024 (Fig. 9). Fig. 9 consists of five panels: the first four display the median NDVI, rainfall, and temperature trends across the study years for each zone. The fifth panel compares median NDVI time series across all zones,

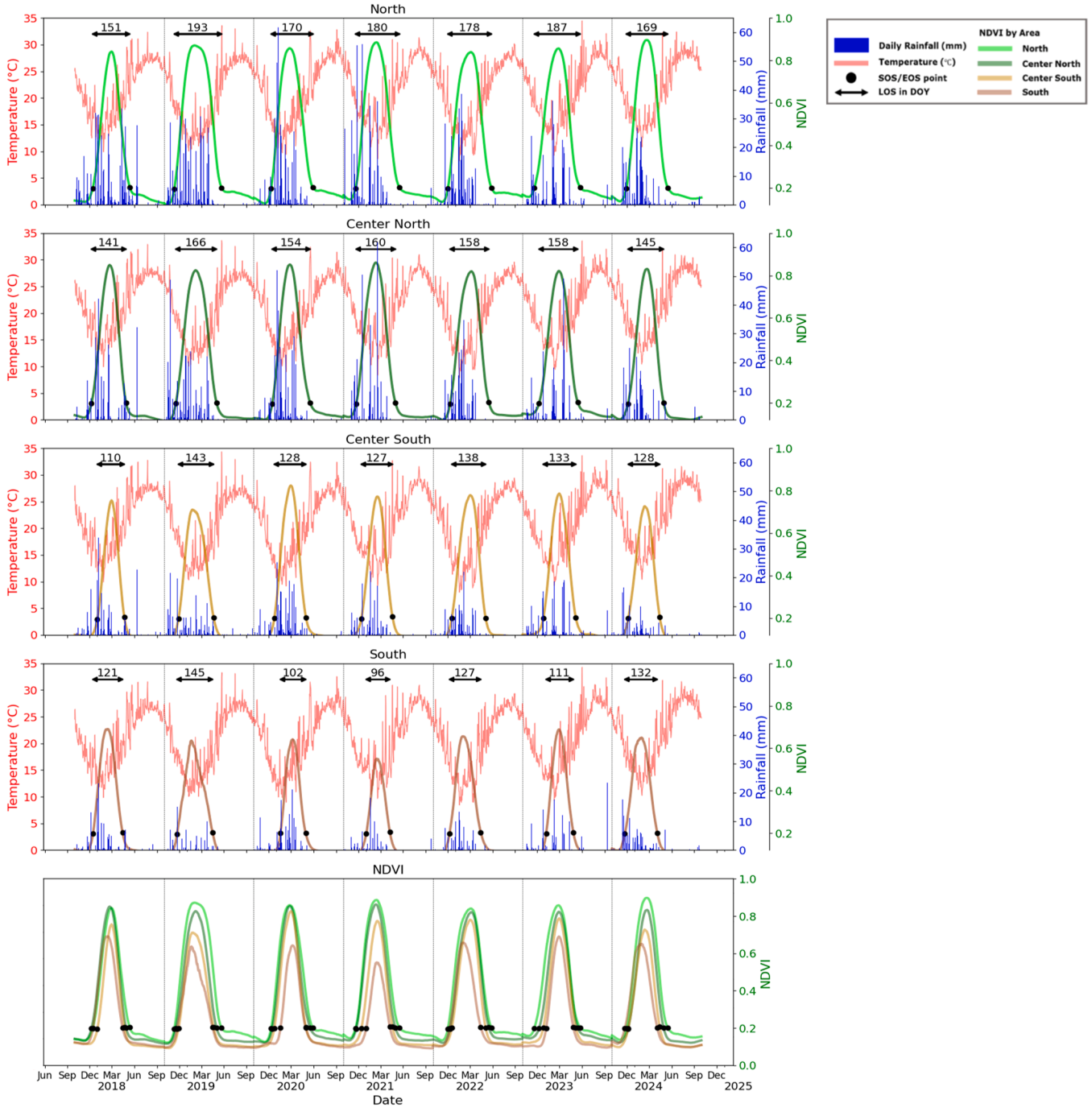

Fig. 9. Time-series data from the 2018–2024 agricultural years, presenting daily mean temperature (°C), daily rainfall (mm), and median NDVI for plots classified as wheat (including barley and oats). The data are shown for four aridity zones: North, Center North, Center South, and South (Panels 1-4). Each panel illustrates the seasonal dynamics and climatic variability specific to its zone. The fifth panel compares the median NDVI time series across all zones using color-coded lines. Black circles indicate the median Start of Season (SOS) and End of Season (EOS) dates, while horizontal arrows represent the Length of Season (LOS) for each aridity zone. It highlights phenological differences and productivity patterns along the aridity gradient.

illustrating phenological differences. The results clearly demonstrate an aridity gradient from north to south. The North zone receives significantly more rainfall, while temperatures remain relatively consistent across zones. In terms of growing season length (LOS), the North zone exhibits the longest LOS, followed by the Center North. The Center South and South zones have shorter growing seasons. Phenological patterns also shift along the gradient. The South zone consistently shows earlier SOS and EOS, and lower median NDVI values, indicating lower productivity. In contrast, the North zone has the highest NDVI values, except in 2018 when the Center North surpassed it.

To further assess the influence of space and time, six phenological features (SOS DOY, MOS DOY, EOS DOY, LOS, MOS value, and AUC) were compared across aridity zones and years (Fig. 10). Results show strong spatial differences aligned with the aridity gradient in all features except the MOS DOY.

4. Discussion

This study demonstrates the feasibility and effectiveness of integrating multi-sensor satellite data with ML in a hierarchical classification framework to produce accurate, automated crop maps in a data-scarce, arid agricultural region. By leveraging phenological features derived from dense VIs-TS, structural information from SAR, and high-resolution spatial data, the model achieved high classification

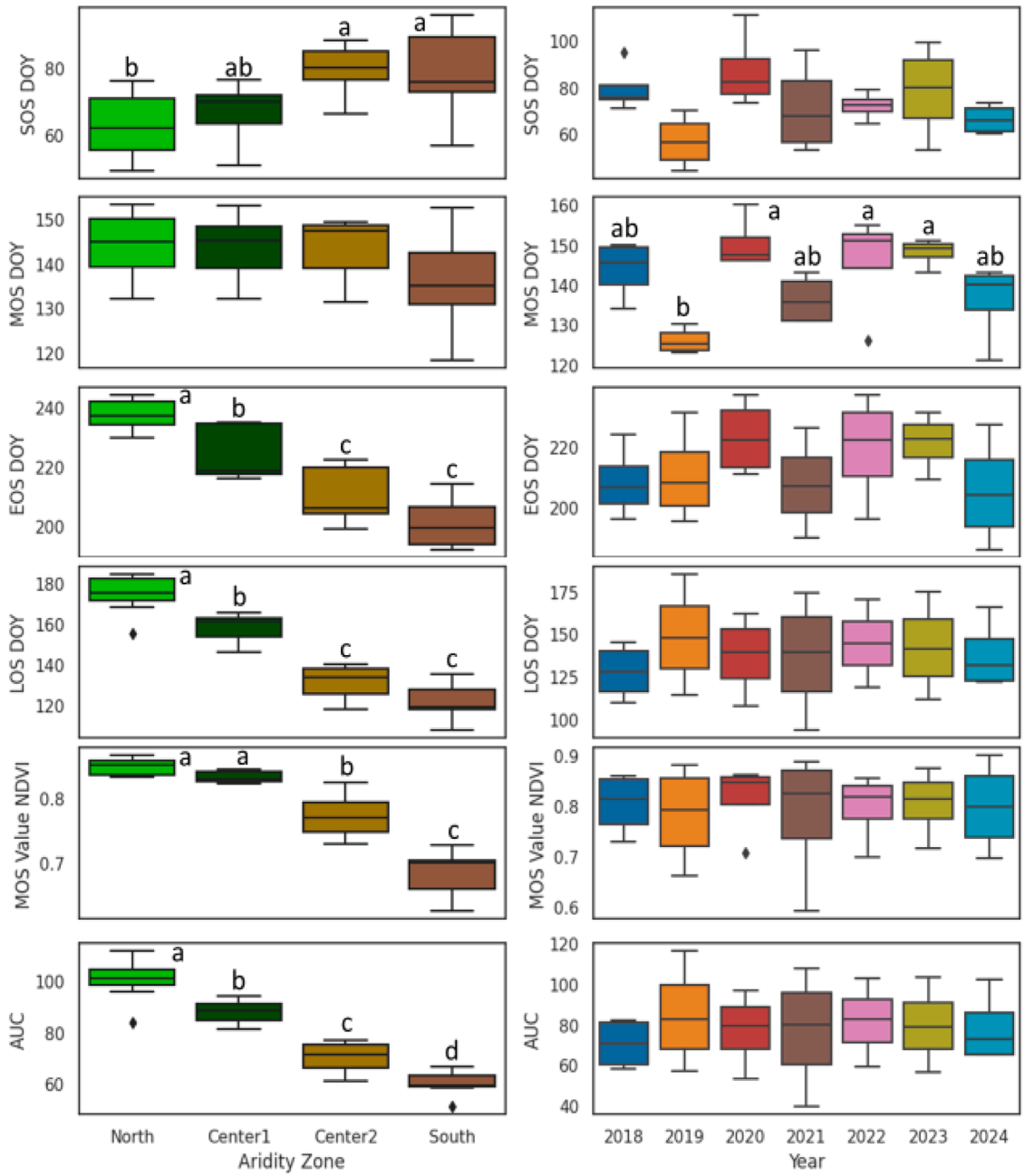

Fig. 10. Boxplots illustrate the distribution of six key phenological parameters for wheat plots (including barley and oats): Start of Season (SOS), Max of Season (MOS), End of Season (EOS), Length of Season (LOS), MOS NDVI value, and Area Under the Curve (AUC). Left panel: Parameters grouped by aridity zone, averaged across all years. Right panel: Parameters grouped by year, averaged across all zones. Note that aridity-based grouping does not consider interannual variation, and year-based grouping does not consider aridity zone variation. These comparisons highlight spatial and temporal influences on wheat phenology across the Western Negev.

performance across multiple crop types, including wheat, one of Israel's most important staple crops. The results underscore the value of satellite-based crop monitoring in capturing dynamic land-use patterns and crop phenology, especially in heterogeneous dryland systems where traditional agricultural mapping approaches often fall short. These findings not only validate the proposed methodology but also highlight its scalability and transferability to other regions facing similar climatic and data limitations. Furthermore, the identification of statistically significant land-use trends and the spatial gradient in wheat productivity provides important insights for agricultural planning and climate resilience strategies at both regional and national levels. Israel's agricultural sector, particularly the Western Negev, is dominated by field crops, many of which follow dual- or multi-season cropping cycles, especially in rapidly changing environments. As a result, the development of annually updated classification maps is critical to enabling timely, evidence-based decision-making. Such maps can serve as a valuable tool for decision-making when accurate and agricultural information is crucial for ensuring food security, effective resource allocation, and informed strategic planning.

4.1. Hierarchical classification framework

We opted for traditional ML approaches due to the relatively small size of the available dataset, which was insufficient to support the data-intensive requirements of deep learning models. Among conventional algorithms, the RF classifier has consistently demonstrated strong performance in crop classification tasks, particularly when working with multi-source remote sensing data [3,9,16,17,27]. We developed a hierarchical workflow composed of three RF-based models. This hierarchical approach enabled progressive classification, starting with broad agricultural land cover types, then distinguishing wheat crops, and finally identifying specific crop types. Such a structure was particularly well-suited for this classification task as each model can use different features. The first model, which has non-vegetated classes, lacks distinct seasonal spectral patterns, limiting the utility of time-series phenology for their classification. The second and third models classify crops using phenological features, which are most informative for vegetated classes. The hierarchical structure also facilitated better class balance, especially for wheat classification in the second model. This was important because wheat is the dominant crop in the Western Negev and a key focus of the study, given its relevance to national food security. The second model achieved a high OA of 95 %, underscoring its effectiveness in reliably separating wheat from other crop types. A notable limitation of the hierarchical modeling approach is the potential for error propagation between classification levels. For instance, if a crop is misclassified at an early stage (e.g., as non-agricultural), it cannot be corrected later in subsequent classification errors or cumulative classification errors. Despite this drawback, several studies have shown that hierarchical classification frameworks often outperform flat, non-hierarchical models by improving both interpretability and overall accuracy [15].

4.2. Limitations of training data and classification challenges

Sufficient and robust training data are essential for developing accurate classification models. Accordingly, a significant amount of time was invested in creating and validating training datasets for all three models in this study: Model 1 for classifying agricultural land use and Models 2 and 3 for distinguishing specific crop types. Despite this substantial effort and the high overall accuracy achieved, several limitations remain. For Model 1, training data were generated by manually labeling high-resolution imagery. However, a key challenge lies in the fundamental difference between what is visible in high-resolution images and what satellite sensors detect. High-resolution imagery provides a static snapshot of a single moment in time and cannot reliably represent dynamic classes such as bare fields or field crops without additional temporal context [56]. To overcome this, the study incorporated the annual maximum NDVI value from satellite imagery to better capture vegetation presence and phenological cycles.

For Models 2 and 3, which focused on crop-specific classification, training data were collected from across Israel, not exclusively from the Western Negev study area. This introduced potential variability in crop phenology due to regional differences in rainfall, temperature, and planting schedules. While this could affect model performance in the study region, previous research suggested that including data from broader geographic areas can enhance model robustness and generalizability [3,9,17]. In fact, most large-scale crop classification models were developed over areas much larger than Israel without significant loss of accuracy. Lastly, some crop classes had relatively small training sample sizes, which likely affected the model's ability to distinguish them accurately. Limited representation may have hindered the algorithm's capacity to learn class-specific features, leading to increased misclassification in minority classes. Model 3 exhibited relatively low accuracy for several minor crop types, such as melon and pumpkin (~10–30 % PA), primarily due to spectral and phenological similarity with watermelon and 'other', and the limited number of available training samples. To mitigate these limitations, future work should expand the ground survey dataset, integrate additional vegetation indices sensitive to chlorophyll and canopy water content, and apply finer temporal segmentation (e.g., 10-day composites) to capture intra-seasonal growth dynamics better and improve class separability.

4.3. Feature selection strategy and importance

We adopted a targeted feature selection strategy, prioritizing variables expected to provide meaningful and biologically relevant information rather than using the full suite of available features. For example, after time-series extraction and preprocessing (including gap-filling and smoothing), each spectral band or VIs initially yielded 365 daily values per year. However, incorporating all 365 features into the model would have been excessive and potentially detrimental due to redundancy and overfitting. Even limiting the selection to the dates of S2 acquisitions would still produce 73 features per index, an impractically large number. To streamline the feature set while maintaining relevance to the classification tasks, we focused on extracting key statistical and phenological features. These included the maximum and mean annual values for each VIs, half-monthly mean NDVI values, and phenological features such as SOS, MOS, and EOS. Importantly, full phenological metrics were derived solely from NDVI due to its consistent, interpretable temporal profile. For all other VIs, only a single representative value was extracted, specifically, the value on the MOS date. To the best of our knowledge, this approach, differentiating feature extraction between NDVI and other VIs, has not been previously applied in crop classification studies. Existing studies often focus on a single VI (e.g., NDVI or EVI) for phenological extraction [16,26], or combine all raw VI values with phenological features [26]. Others extract phenological features for multiple VIs but primarily compare classification performance across indices [8]. The two feature selection model results confirmed that SOS-DOY, MOS-DOY, and MOS-aligned SAR VH consistently ranked among the most important predictors across models, further supporting the validity and interpretability of our expert-based feature selection approach.

In the first classification model, which distinguished between four agricultural land cover types (bare fields, covered fields, field crops, and orchards), 40 features were used. Among these, the annual maximum NDVI and the mean annual PGHI were the most significant, contributing approximately 17 % and 12 %, respectively, to model performance. These results align with expectations: NDVI is highly sensitive to vegetation and is essential for distinguishing vegetation from non-vegetated areas. PGHI, in turn, captures the spectral responses of greenhouse materials, such as plastics and netting, helping to distinguish covered structures from bare fields. Similarly, the importance of annual NDVI summary features has been reported by Le et al. [57] in land use/land

cover classification studies. The second model, which distinguished wheat from other field crops, and the third model, which classified 12 specific crops, both used 32 features. In the second model, SOS-DOY, VH-ASC, and MOS-DOY were the most influential, each contributing approximately 17 %, 13 %, and 10 % respectively. This finding is consistent with wheat's unique phenology as a winter crop, in contrast to the spring and summer phenology of other crops. In the third model, feature importance was more evenly distributed, with EOS-DOY contributing about 9 %, followed by MOS-DOY, the MOS value of MCARI, and SOS-DOY at around 7 % each. This distribution reflects the overlapping phenologies among different crops, necessitating the use of multiple features for accurate classification.

Our results align with those of Htitiou et al. [26], who noted that SOS, MOS, and EOS dates and NDVI and EVI2 values were the most influential features for most crops. They also found that red-edge-based features were particularly informative for fallow land. Similarly, Goldberg et al. (2021) [8] showed that MOS DOY was the most critical phenological feature, with date-based features generally outperforming their corresponding VI values. Blickensdörfer et al. [9] highlighted the importance of SAVI, NDVI, and NIR bands (each contributing 8–10 %) in crop classification, and observed that SAR bands, especially VH and VV, contributed more significantly than RVI or VH/VV. These findings are consistent with our observations across all three models.

4.4. Accuracy assessment

The OA of the three classification models developed in this study was 94.87 %, 95.17 %, and 81.15 %, respectively. These results indicate strong model performance across different classification levels. The first model, which categorized agricultural land cover types, achieved an OA comparable to those reported in previous land use/land classification studies [58,59], demonstrating the effectiveness of our approach for large-scale land cover mapping. The second model, focused on classifying wheat, achieved a notably high OA of 95.17 %, consistent with other studies that evaluated wheat classification across different phenological stages. For example, Liu et al. [60] reported OA values ranging from 92 % to 96.9 % using an RF model, with the jointing to heading stage identified as the most suitable period for wheat classification. Our model's performance supports these findings and highlights the benefits of phenology-informed feature extraction for crop-type differentiation, especially for winter crops such as wheat. The third model classified 12 specific crop types and attained an OA of 81.15 %, aligning well with results from similar multi-class crop mapping studies [3,9,14,17,27]. However, direct comparisons across studies are challenging due to differences in crop classes, geographic regions, and validation strategies. For example, Blickensdörfer et al. [3], Griffiths et al. [9], and Orynbaikyzy et al. [3,9,17] included a grassland class and more granular cereal classes (e.g., separating winter and spring cereals), along with other crops that are not commonly grown in our study area. Goldberg et al. [8], who also conducted crop classification research in Israel, classified six crop types (barley, carrot, chickpea, cotton, wheat, and 'other' category) and reported an OA of 68 % using their best-performing model (XGBoost with phenological features extracted from NDII). As in our study, wheat consistently achieved the highest PA and UA, which is a pattern commonly observed in the literature [8,18,27]. This is likely due to wheat's unique phenological profile as a winter crop, which contrasts with the spring and summer growth cycles of most other crops. Similarly, cotton and peanut also exhibited high classification accuracy in our model, likely due to their distinct and late-season phenology, which reduces spectral overlap with other summer crops.

4.5. Change detection of the Western Negev area

The most significant land cover transitions identified in this study occurred between bare fields and row crops, reflecting frequent shifts in both directions. In contrast, orchards and covered plots demonstrated relative temporal stability (Figs. 5 and 8). These patterns are consistent with expected agricultural management practices. Row crops are often subject to seasonal rotation with fallow (bare) periods, while orchards, by nature, are long-lived perennial systems. Covered structures, such as greenhouses or shade nets, are typically semi-permanent, with only the coverings (e.g., plastic or netting) periodically replaced. The annual rate of land cover ranged between 11 % and 16 % of plots, underscoring the dynamic nature of field-based agriculture in the study area. However, two key limitations should be considered when interpreting these transition dynamics. First, classification errors may introduce bias, resulting in either overestimation or underestimation of land cover change rates. Second, the analysis was limited to plots that maintained consistent identifiers and geometries across consecutive years. Plots that changed ID, geometry, or inclusion status were excluded from the transition analysis, potentially leading to underrepresentation of certain transitions. To address this limitation, future work could incorporate advanced field segmentation techniques that dynamically detect changes in plot boundaries and configurations over time. This would reduce uncertainty and enable the inclusion of evolving field geometries caused by land reallocation, plot merging or splitting, and the installation or removal of agricultural infrastructure (e.g., irrigation systems, greenhouses).

Field segmentation also plays a critical role in aligning classification outputs with real-world agricultural management units. Accurate segmentation enhances the interpretability of results by ensuring that classification labels align with actual field boundaries, thereby facilitating direct application by farmers, land managers, and policymakers. Moreover, segment-based classification reduces within-field noise and edge effects often present in pixel-based methods, leading to more stable and realistic monitoring outcomes [8]. When combined with time-series analysis and high-resolution imagery, field segmentation enhances the spatial and temporal resolution of land-use assessments, enabling more precise monitoring of crop rotations, management intensity, and responses to environmental or economic pressures. By integrating segmentation with geometry tracking and change detection, future studies can offer more accurate, robust, and actionable insights into agricultural dynamics at scale.

Similar studies have explored agricultural land transitions in other regions. For example, in the Egyptian coastal zone, El-Hattab [61] reported a substantial reduction in palm and fruit tree cultivation and a simultaneous expansion of field crops and urban areas between 2000 and 2013, driven by economic considerations such as crop profitability. In the Nile Delta, agricultural land loss to urban development has been extensively documented, prompting policy-driven efforts to reclaim bare land for cropland [62]. Xu et al. [63] examined long-term cropland dynamics across Egypt, Ethiopia, and South Africa, noting annual increases in cropland area and shifts from natural vegetation or pasture to cultivated fields. In Alanya, Turkey, a region known for greenhouse cultivation, İnalpulat and Genç [59] observed modest greenhouse expansion alongside reductions in traditional open-field agriculture due to urban growth.

The added value of the present study lies in its fine-scale, plot-level monitoring of land cover transitions over multiple years, using high-resolution satellite imagery and object-based classification. Unlike many previous studies that rely on aggregated land-use categories or coarse regional trends, our approach enables the detection of subtle yet agriculturally meaningful transitions, particularly those related to crop rotation and fallow periods. By leveraging consistent plot geometry and unique identifiers, the methodology minimizes noise from classification artifacts and geometric mismatches, enabling robust and repeatable change detection. This framework is especially well-suited for intensively managed agricultural regions, where short-term decisions, such as planting delays, crop switching, or greenhouse removal, can significantly alter land cover. The methodology's precision and scalability make it valuable not only for academic research but also for operational monitoring by agricultural agencies. It provides a transferable tool for

assessing the impacts of economic pressures, policy shifts, and climatic variability on land use, thereby supporting informed decision-making for sustainable agriculture, land planning, and resource allocation. While our current framework is implemented annually, future work could advance toward near-real-time crop monitoring. With the increasing availability of cloud-based platforms and automated workflows, it is feasible to integrate continuous updates of GIS plot boundaries with automated ingestion of Sentinel-1/2 imagery. Once established, such a system could produce updated classification maps within 2–3 days of new images becoming available. This capability would support operational applications such as early detection of crop failures, rapid response to extreme weather events, and adaptive management of irrigation and fertilization schedules. Transitioning from annual to near-real-time monitoring, therefore, represents an important next step for expanding the operational value of the framework.

4.6. Trend analysis of the Western Negev area focusing on wheat

The TS trend analysis revealed significant increases in the number of plots cultivated with corn, wheat, and orchards, both on annual and monthly scales. Monthly assessments showed a substantial increase in covered plots, accompanied by a notable decline in bare fields, as well as a decrease in the cultivation of chickpeas and watermelons (Table 4). These dynamics are particularly relevant for food security, as reductions in the cultivation of specific crops may signal vulnerabilities in agricultural sustainability and market supply chains.

Among 492 plots that were consistently classified as wheat for seven consecutive years, 251 exhibited a significant increasing trend in NDVI values, suggesting enhanced productivity. Conversely, 172 plots showed a decreasing trend, and 69 plots displayed no significant change. Interestingly, almost all the consistently classified wheat plots were in the northern part of the study area, which is typically more productive due to favourable climatic and edaphic conditions. However, the variability observed at the individual plot level highlights the heterogeneity of productivity trends, even within high-performing sub-regions. In contrast to such broader regional assessments, the present study offers a fine-scale, plot-level perspective on productivity trends, capable of detecting spatial heterogeneity and localized changes even within high-yielding zones. By combining multi-year classification consistency with phenology-informed remote sensing, this approach enables a more accurate assessment of vegetation dynamics and the early detection of yield trends. This level of detail is vital for informing targeted interventions, guiding adaptive farm management, and supporting precision agriculture strategies. Moreover, integrating multi-temporal satellite data with crop-specific phenological metrics enriches our understanding of how environmental variability and agricultural practices interact to shape productivity. These insights are particularly valuable in arid and semi-arid regions, such as the Western Negev, where agriculture must be resilient to increasing climatic and market uncertainties. The framework developed here can be scaled or transferred to other areas facing similar challenges, offering a practical tool for evidence-based decision-making in sustainable land-use planning.

4.7. Spatiotemporal pattern of wheat in the Western Negev area

The primary findings of this study highlight a clear north–south aridity gradient that strongly influences wheat productivity in the Western Negev. Specifically, the northern and central-northern zones demonstrated significantly higher productivity than the central-southern and southern zones. This is evidenced by longer LOS, higher MOS, and greater AUC values in the north. Additionally, the SOS in the northern zone occurs later than in the southern zone but earlier than in the central-southern zone, closely aligning with the SOS timing in the central-northern region. These findings are consistent with those of Mechiche-Alami & Abdi [64], who reported that pixels with longer LOS and earlier SOS typically corresponded to increased AUC, indicating greater vegetation productivity. This pattern is especially meaningful in the context of climate change, as numerous studies predict increasing aridity across North Africa and southern Europe under a range of future climate scenarios [65,66]. Within this framework, the current findings suggest that wheat productivity is likely to decline, particularly in the southern parts of the study area, where the climatic stress is expected to intensify.

Understanding the spatial relationship between aridity and wheat phenology is essential for predicting regional vulnerabilities and aiding climate-resilient agricultural planning. Although the current analysis is based on a seven-year series characterized by significant variability within and between years, particularly in rainfall, it still shows clear and consistent spatial patterns. However, the limited timeframe highlights the importance of long-term monitoring to detect multi-decade changes better and produce more reliable projections. Importantly, this study introduces a new spatial link between phenological productivity metrics and aridity gradients, providing a scalable and transferable framework for identifying future vulnerable areas. This approach is especially useful for water-sensitive crops, such as wheat, offering early warning insights that can guide adaptive management, inform land-use decisions, and enhance food security planning amid increasingly variable climate conditions. The classification output generated by our framework has direct applications for both agronomic management and policy planning. Crop-rotation maps can guide soil fertility management and pest-control strategies, while phenology monitoring provides early indicators for optimizing irrigation scheduling and fertilizer application. At larger scales, analyzing multi-year crop dynamics, such as wheat distribution across aridity zones, can inform food security planning and help governmental agencies assess agricultural resilience under changing climatic conditions.

5. Conclusion

This study introduces a comprehensive and innovative framework for monitoring crop dynamics and productivity in dryland agricultural systems, leveraging multi-sensor satellite data, object-based classification, and phenology-based indicators. In contrast to conventional mapping approaches, our methodology uniquely integrates fine-resolution spatial analysis with time-series vegetation metrics and environmental gradients to capture both land-use transitions and intra-seasonal crop performance. The use of phenological metrics derived from vegetation indices provides a biologically grounded lens for assessing agricultural responses to climate variability and management practices. The novelty of this study lies in its ability to detect subtle yet ecologically and economically significant shifts in crop development, providing insights that are often overlooked by traditional monitoring systems. By revealing spatial patterns across aridity zones and identifying emerging trends in field crop behavior, the research provides critical information to anticipate risks, support food security, and guide adaptive land-use planning. The framework is not only robust and data-efficient but also scalable and transferable to other regions facing similar challenges, particularly those experiencing intensifying water scarcity and climate uncertainty. As such, it represents a significant step forward in operationalizing automated, evidence-based agricultural monitoring for policy and planning under conditions of global change.

Funding

We thank the Ministry of Agriculture and Food Security for its support of this study. This work was funded under Grant No. [304081400].

Data availability

Sentinel-1 and Sentinel-2 imagery are freely available from Copernicus Data Space Ecosystem (h ttps://dataspace.copernicus.eu/ accessed on 13-Jul-2025). Map products are intended to become open to

the public via MOAG, and in the meantime, are available upon request.

Declaration of AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT to enhance readability and language. All content generated with the assistance of this tool was subsequently reviewed and edited by the author(s), who take full responsibility for the final content of the publication.

Ethics statement

Not applicable: This manuscript does not include human or animal research.

If this manuscript involves research on animals or humans, it is imperative to disclose all approval details.

CRediT authorship contribution statement

Nechama Z. Brickner: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lior Fine:** Writing – review & editing, Methodology, Data curation, Conceptualization. **Offer Rozenstein:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Tarin Paz-Kagan:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We gratefully acknowledge the Israel Ministry of Agriculture and Food Security for funding and supporting this research. We thank Omer Ben-Asher and Erez Osman for their valuable contributions. Special appreciation is extended to our colleague, Dr. Vladislav Moshe Dubinin, for his indispensable assistance throughout the study. We also thank Avi Atanelov and Adi Edri for their dedication and support during the research process. Finally, we are grateful to KANAT (the Israeli Insurance Fund for Natural Risks in Agriculture) and the collaborating farmers for their support and for providing the training data used in the study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.atech.2025.101650.

Data availability

Data will be made available on request.

References

[1] E.S. Cassidy, P.C. West, J.S. Gerber, J.A. Foley, Redefining agricultural yields: from tonnes to people nourished per hectare, Environ. Res. Lett. 8 (3) (2013), https://doi.org/10.1088/1748-9326/8/3/034015.

[2] P. Potapov, et al., Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century, Nat. Food 3 (1) (2022) 19–28, https://doi.org/10.1038/s43016-021-00429-z. Jan.

[3] P. Griffiths, C. Nendel, P. Hostert, Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping, Remote Sens. Environ. 220 (2019) 135–151, https://doi.org/10.1016/j.rse.2018.10.031. Jan.

[4] P. Meiyappan, M. Dalton, B.C. O'Neill, A.K. Jain, Spatial modeling of agricultural land use change at global scale, Ecol. Modell. 291 (2014) 152–174, https://doi.org/10.1016/j.ecolmodel.2014.07.027. Nov.

[5] F.J. López-Andreu, J.A. López-Morales, M. Erena, A.F. Skarmeta, J.A. Martínez, Monitoring system for the management of the common agricultural policy using machine learning and remote sensing, Electronics (Switzerland) 11 (3) (2022), https://doi.org/10.3390/electronics11030325. Feb.

[6] R.P. Sishodia, R.L. Ray, S.K. Singh, Applications of remote sensing in precision agriculture: a review, Remote Sens. (Basel) 12 (19) (2020) 1–31, https://doi.org/10.3390/rs12193136. Oct.

[7] L. Lipper, et al., Climate-Smart Agriculture for Food Security, Nature Publishing Group, 2014, https://doi.org/10.1038/nclimate2437. Jan. 01.

[8] K. Goldberg, I. Herrmann, U. Hochberg, O. Rozenstein, Generating up-to-date crop maps optimized for sentinel-2 imagery in israel, Remote Sens. (Basel) 13 (17) (2021), https://doi.org/10.3390/rs13173488. Sep.

[9] L. Blickensdörfer, M. Schwieder, D. Pflugmacher, C. Nendel, S. Erasmi, P. Hostert, Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany, Remote Sens. Environ. 269 (2022), https://doi.org/10.1016/j.rse.2021.112831. Feb.

[10] N. You, et al., The 10-m crop type maps in Northeast China during 2017–2019,, Sci. Data 8 (1) (2021), https://doi.org/10.1038/s41597-021-00827-9. Dec.

[11] Y. Xie, Z. Sha, M. Yu, Remote sensing imagery in vegetation mapping: a review, J. Plant Ecol. 1 (1) (2008) 9–23, https://doi.org/10.1093/jpe/rtm005. Mar.

[12] H. Li, et al., Automated In-season crop-type data layer mapping without ground truth for the conterminous United States based on multisource satellite imagery, IEEE Trans. Geosci. Remote Sens. 62 (2024) 1–14, https://doi.org/10.1109/TGRS.2024.3361895.

[13] T. Fisette, et al., AAFC Annual Crop Inventory Status and challenges, in: Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics), IEEE, 2013, pp. 270–274.

[14] C. Luo, et al., Using time series sentinel-1 images for object-oriented crop classification in google earth engine, Remote Sens. (Basel) 13 (4) (2021) 1–19, https://doi.org/10.3390/rs13040561. Feb.

[15] M.O. Turkoglu, et al., Crop mapping from image time series: deep learning with multi-scale label hierarchies, Remote Sens. Environ. 264 (2021), https://doi.org/10.1016/j.rse.2021.112603. Oct.

[16] H. do Nascimento Bendini, et al., Detailed agricultural land classification in the Brazilian cerrado based on phenological information from dense satellite image time series, Int. J. Appl. Earth Obs. Geoinf. 82 (2019), https://doi.org/10.1016/j.jag.2019.05.005. Oct.

[17] A. Orynbaikyzy, U. Gessner, C. Conrad, Spatial transferability of random forest models for crop type classification using Sentinel-1 and Sentinel-2, Remote Sens. (Basel) 14 (6) (2022), https://doi.org/10.3390/rs14061493. Mar.

[18] P. Rao, et al., Using sentinel-1, sentinel-2, and planet imagery to map crop type of smallholder farms, Remote Sens. (Basel) 13 (10) (2021), https://doi.org/10.3390/rs13101870. May.

[19] S. Ge, J. Zhang, Y. Pan, Z. Yang, S. Zhu, Transferable deep learning model based on the phenological matching principle for mapping crop extent, Int. J. Appl. Earth Obs. Geoinf. 102 (2021), https://doi.org/10.1016/j.jag.2021.102451. Oct.

[20] C. Boryan, Z. Yang, R. Mueller, M. Craig, Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program, Geocarto Int. 26 (5) (2011) 341–358, https://doi.org/10.1080/10106049.2011.562309. Aug.

[21] B. Wu, et al., Challenges and Opportunities in Remote Sensing-Based Crop monitoring: a Review, Oxford University Press, 2023, https://doi.org/10.1093/nsr/nwac290. Apr. 01.

[22] P. Jönsson, L. Eklundh, TIMESAT - a program for analyzing time-series of satellite sensor data, Comput. Geosci. 30 (8) (2004) 833–845, https://doi.org/10.1016/j.cageo.2004.05.006. Oct.

[23] Y. Shao, R.S. Lunetta, B. Wheeler, J.S. Iiames, J.B. Campbell, An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data, Remote Sens. Environ. 174 (2016) 258–265, https://doi.org/10.1016/j.rse.2015.12.023. Mar.

[24] J. Xue, B. Su, Significant remote sensing vegetation indices: a review of developments and applications, Hindawi Ltd. (2017), https://doi.org/10.1155/2017/1353691.

[25] S. Belda, et al., DATimeS: a machine learning time series GUI toolbox for gap-filling and vegetation phenology trends detection, Environ. Model. Softw. 127 (2020), https://doi.org/10.1016/j.envsoft.2020.104666. May.

[26] A. Htitiou, A. Boudhar, Y. Lebrini, R. Hadria, H. Lionboui, T. Benabdelouahab, A comparative analysis of different phenological information retrieved from Sentinel-2 time series images to improve crop classification: a machine learning approach, Geocarto Int. 37 (5) (2022) 1426–1449, https://doi.org/10.1080/10106049.2020.1768593.

[27] K. Luo, L. Lu, Y. Xie, F. Chen, F. Yin, Q. Li, Crop type mapping in the central part of the North China Plain using Sentinel-2 time series and machine learning, Comput. Electron. Agric. 205 (2023), https://doi.org/10.1016/j.compag.2022.107577. Feb.

[28] O. Rozenstein, A. Karnieli, Comparison of methods for land-use classification incorporating remote sensing and GIS inputs, Appl. Geogr. 31 (2) (2011) 533–544, https://doi.org/10.1016/j.apgeog.2010.11.006. Apr.
[29] Israel Meteorological Service, Average rain 1991-2020," annual average rain 1991-2020, Accessed: Jan. 30, https://ims.gov.il/he/node/1256, 2025.
[30] H.E. Beck, N.E. Zimmermann, T.R. McVicar, N. Vergopolan, A. Berg, E.F. Wood, Present and future köppen-geiger climate classification maps at 1-km resolution, Sci. Data 5 (2018), https://doi.org/10.1038/sdata.2018.214.
[31] J.-F. Bastin et al., "The extent of forest in dryland biomes," 2017. [Online]. Available: https://www.science.org.
[32] J.D. Klein, et al., Establishment of wheat seedlings after early sowing and germination in an arid mediterranean environment, Agron. J. 94 (3) (2002) 585–593, https://doi.org/10.2134/agronj2002.5850.
[33] D. Haim, M. Shechter, P. Berliner, Assessing the impact of climate change on representative field crops in Israeli agriculture: a case study of wheat and cotton, Clim. Change 86 (3–4) (2008) 425–440, https://doi.org/10.1007/s10584-007-9304-x. Feb.
[34] T. Ben Hassen, H. El Bilali, Impacts of the Russia-Ukraine war on global food security: towards more sustainable and resilient food systems? Foods 11 (15) (2022) https://doi.org/10.3390/foods11152301. Aug.
[35] S. Asseng, et al., Can Egypt become self-sufficient in wheat? Environ. Res. Lett. 13 (9) (2018) https://doi.org/10.1088/1748-9326/aada50.
[36] Zupance, "Improving cloud detection with machine learning," medium. Accessed: Feb. 01, 2025. [Online]. Available: https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13.
[37] R. Nasirzadehdizaji, F.B. Sanli, S. Abdikan, Z. Cakir, A. Sekertekin, M. Ustuner, Sensitivity analysis of multi-temporal Sentinel-1 SAR parameters to crop height and canopy coverage, Appl. Sci. (Switz.) 9 (4) (2019), https://doi.org/10.3390/app9040655. Feb.
[38] E. Lindsay, et al., Multi-temporal satellite image composites in Google Earth Engine for improved landslide visibility: a case study of a glacial landscape, Remote Sens. (Basel) 14 (10) (2022), https://doi.org/10.3390/rs14102301. May.
[39] C.J. Tucker, "Red and photographic Infrared l,lnear combinations for monitoring vegetation," 1979.
[40] P.H.C. Eilers, "A perfect smoother," Jul. 15, 2003. doi: 10.1021/ac034173t.
[41] H.B. Sandya, H.P. Kumar, H. Bhudiraja, S.K. Rao, Fuzzy rule based feature extraction and classification of time series signal, Int. J. Soft Comput. Eng. (IJSCE) 3 (2) (2013) 2231–2307.
[42] M.B. Kursa, W.R. Rudnicki, Feature selection with the Boruta package, J. Stat. Softw. 36 (2010) 1–13.
[43] M.B. Kursa, W.R. Rudnicki, M.M.B. Kursa, Package 'Boruta,', R. J. 15 (2018).
[44] S.M. Lundberg, P.G. Allen, S.-I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017) [Online]. Available, https://github.com/slundberg/shap.
[45] L. Breiman, "Random forests," 2001.
[46] A.D. Kulkarni, B. Lowe, Random forest algorithm for land cover classification, Int. J. Recent Innov. Trends Comput. Commun. (2016) [Online]. Available, http://hdl.handle.net/10950/341http://www.ijritcc.org.
[47] C. Gómez, J.C. White, M.A. Wulder, Optical remotely sensed time series data for land cover classification: a review, Elsevier B.V. (2016), https://doi.org/10.1016/j.isprsjprs.2016.03.008.
[48] A.E. Maxwell, T.A. Warner, Thematic classification accuracy assessment with inherently uncertain boundaries: an argument for center-weighted accuracy assessment metrics, Remote Sens. (Basel) 12 (12) (2020), https://doi.org/10.3390/rs12121905. Jun.
[49] D. Lu, P. Mausel, E. Brondízio, E. Moran, Change detection techniques, Int. J. Remote Sens. 25 (12) (2004) 2365–2401, https://doi.org/10.1080/0143116031000139863. Jun.
[50] H.B. Mann, "Nonparametric tests against trend," 1945. [Online]. Available: https://www.jstor.org/stable/1907187.
[51] M. Kendall, Rank correlation measures, Charles Griffin, London 202 (1975) 15.
[52] R.M. Hirsch, J.R. Slack, A nonparametric trend test for seasonal data with serial dependence, Water. Resour. Res. 20 (6) (1984) 727–732, https://doi.org/10.1029/WR020i006p00727.
[53] Md. Hussain, I. Mahmud, pyMannKendall: a python package for non parametric Mann Kendall family of trend tests, J. Open. Source Softw. 4 (39) (Jul. 2019) 1556, https://doi.org/10.21105/joss.01556.
[54] A. Trabucco and R. Zomer, "Global aridity index and potential evapotranspiration (ET0) climate database v2," 2018.
[55] J.W. Tukey, "The problem of multiple comparisons. Multiple comparisons.," 1953.
[56] A.E. Maxwell, T.A. Warner, B.C. Vanderbilt, C.A. Ramezan, Land cover classification and feature extraction from National Agriculture Imagery Program (NAIP) Orthoimagery: a review, Am. Soc. Photogramm. Remote Sens. (2017), https://doi.org/10.14358/PERS.83.10.737.
[57] T.D.H. Le, L.H. Pham, Q.T. Dinh, N.T.T. Hang, T.A.T. Tran, Rapid method for yearly LULC classification using Random Forest and incorporating time-series NDVI and topography: a case study of Thanh Hoa province, Vietnam, Geocarto Int. 37 (27) (2022) 17200–17215, https://doi.org/10.1080/10106049.2022.2123959.
[58] M.A. Friedl, et al., MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets, Remote Sens. Environ. 114 (1) (2010) 168–182, https://doi.org/10.1016/j.rse.2009.08.016. Jan.
[59] M. İnalpulat, L. Genç, Short-term change detection and Markov chain prediction of greenhouse areas in Alanya, Turkey using sentinel-2 imageries, Eur. J. Sci. Technol. (2021), https://doi.org/10.31590/ejosat.1019033. Dec.
[60] S. Liu, et al., The accuracy of winter wheat identification at different growth stages using remote sensing, Remote Sens. (Basel) 14 (4) (2022), https://doi.org/10.3390/rs14040893. Feb.
[61] M.M. El-Hattab, Applying post classification change detection technique to monitor an Egyptian coastal zone (Abu Qir Bay), Egypt. J. Remote Sens. Space Sci. 19 (1) (2016) 23–36, https://doi.org/10.1016/j.ejrs.2016.02.002. Jun.
[62] T.M. Radwan, G.A. Blackburn, J.D. Whyatt, P.M. Atkinson, Dramatic loss of agricultural land due to urban expansion threatens food security in the Nile Delta, Egypt, Remote Sens. (Basel) 11 (3) (2019), https://doi.org/10.3390/rs11030332. Feb.
[63] Y. Xu, et al., Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: experiments from three sites in Africa, Remote Sens. Environ. 218 (2018) 13–31, https://doi.org/10.1016/j.rse.2018.09.008. Dec.
[64] A. Mechiche-Alami, A.M. Abdi, Agricultural productivity in relation to climate and cropland management in West Africa, Sci. Rep. 10 (1) (Feb. 2020) 3393, https://doi.org/10.1038/s41598-020-59943-y.
[65] D. Carvalho, S.C. Pereira, R. Silva, A. Rocha, Aridity and desertification in the Mediterranean under EURO-CORDEX future climate change scenarios, Clim. Change 174 (3–4) (2022), https://doi.org/10.1007/s10584-022-03454-4. Oct.
[66] C.E. Park, et al., Keeping global warming within 1.5 °c constrains emergence of aridification, Nat. Clim. Chang. 8 (1) (2018) 70–74, https://doi.org/10.1038/s41558-017-0034-4. Jan.