remote sensing

MDPI

Article

Mapping Anti-Hail Net Systems in Apple Orchards Using Multisensor Time Series and Machine Learning

Danielle Elis Garcia Furuya [1,*], Édson Luis Bolfe [1,2], Taya Cristo Parreiras [2], Victória Beatriz Soares [2], Franco da Silveira [1], Jayme Garcia Arnal Barbedo [1], Thiago Teixeira Santos [1] and Luciano Gebler [3]

[1] Brazilian Agricultural Research Corporation, Embrapa Agricultura Digital, Campinas 13083-886, SP, Brazil; edson.bolfe@embrapa.br (É.L.B.); franco.da.silveira@hotmail.com (F.d.S.); jayme.barbedo@embrapa.br (J.G.A.B.); thiago.santos@embrapa.br (T.T.S.)
[2] Programa de Pós-Graduação em Geografia, State University of Campinas (Unicamp), Campinas 13083-970, SP, Brazil; t234520@dac.unicamp.br (T.C.P.); v206942@dac.unicamp.br (V.B.S.)
[3] Brazilian Agricultural Research Corporation, Embrapa Uva e Vinho, Vacaria 95200-970, RS, Brazil; luciano.gebler@embrapa.br
* Correspondence: danielle.furuya@colaborador.embrapa.br or daniellegarciafuruya@gmail.com

Highlights

What are the main findings?

- This study demonstrates that anti-hail net systems in apple orchards can be mapped using multisensor satellite time series and artificial intelligence.
- Net color differences (white, black, and no net) were identified with high classification accuracy using Random Forest and 1DCNN models.

What are the implications of the main findings?

- The integration of HLS and Sentinel-2 data provides a robust framework for monitoring orchard management practices in heterogeneous agricultural landscapes.
- Soil- and moisture-related spectral indices (especially BSI and NDWI) play a key role in improving classification performance across hierarchical mapping levels.

Abstract

Apple orchards are increasingly adopting anti-hail nets to mitigate climate risks; however, these structures alter canopy reflectance and pose challenges for remote sensing. This study presents an operational framework to map apple orchards under different netting conditions in Vacaria, Brazil. Multisensor surface reflectance data from Sentinel-2 and Harmonized Landsat and Sentinel-2 were used to generate dense spectral index time series combined with field observations. Five spectral indices, Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil-Adjusted Vegetation Index (SAVI), Normalized Difference Water Index (NDWI), and Bare Soil Index (BSI), were evaluated individually and in combination within a hierarchical classification framework. Random Forest (RF) and one-dimensional convolutional neural networks (1DCNN) were applied as complementary machine learning approaches. RF showed more stable performance across hierarchical levels, while indices contributed differently depending on scale: BSI and NDVI were more effective at broader levels, whereas EVI and SAVI were critical for discriminating net colors. To our knowledge, this is the first study applying multisensor time series and machine learning to map anti-hail net systems in apple orchards.

Keywords: Random Forest; 1DCNN; Harmonized Landsat and Sentinel-2 (HLS); agricultural production; land use land cover

Academic Editors: Enrico Corrado Borgogno Mondino, Filippo Sarvia and Samuele De Petris

Received: 13 March 2026
Revised: 28 April 2026
Accepted: 5 May 2026
Published: 8 May 2026

1. Introduction

Apple is one of the most economically important fruit crops worldwide, with a strong presence in both temperate and subtropical regions. However, the sector faces growing challenges associated with climate variability and change, including temperature shifts, altered precipitation regimes, and an increased frequency of extreme weather events [1,2]. Among these threats, hailstorms represent one of the most damaging factors, capable of severely reducing yield and fruit quality [3].

To minimize these risks, protective structures such as anti-hail nets have been adopted in several apple-producing regions over the last decades [4]. These nets reduce direct hail damage but also influence the orchard microclimate by modifying solar radiation, water, temperature, humidity, and wind dynamics [3,5,6]. While effective in reducing production losses, anti-hail nets also alter canopy reflectance, complicating the monitoring of apple orchards through remote sensing techniques.

Remote sensing (RS) has been increasingly applied to perennial crops to support crop mapping, phenological monitoring, and yield estimation [7,8]. Several methods, sensors, and analytical frameworks have been tested, ranging from vegetation indices and spectral–temporal analysis to advanced machine learning and deep learning approaches [9–12]. Nevertheless, perennial crop mapping remains challenging due to spectral heterogeneity, complex canopy structures, and management practices. When protective nets are present, the challenge is further intensified as the spectral response of the canopy becomes masked or modified [13]. Despite the expansion of hail net use, few studies have addressed their explicit detectability in remote sensing applications, with most research focusing instead on agronomic and physiological impacts.

The use of multisensor time series offers new opportunities to overcome these limitations. Harmonized Landsat and Sentinel-2 (HLS) products combine the high temporal resolution of Landsat and Sentinel-2, enabling denser spectral index time series (SITS) and improving the capacity to capture subtle canopy differences and management effects [14,15]. Sentinel-2 data remain valuable due to their finer spatial resolution and well-established use in vegetation studies [16,17]. Together, these datasets can provide complementary insights into the detectability of anti-hail nets and their variants.

Spectral indices such as Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil-Adjusted Vegetation Index (SAVI), Normalized Difference Water Index (NDWI), and Bare Soil Index (BSI), have long been used to assess vegetation vigor, soil exposure, and water content [18,19]. When applied as a time series, they enhance the ability to monitor phenological dynamics and land-use transitions. However, the choice and number of indices can affect classification outcomes. Open questions remain regarding the minimum number of indices required, their relative importance across sensors, and their stability under different machine learning frameworks.

A recent review by Furuya et al. (2025) [20] highlighted the absence of studies applying artificial intelligence algorithms to map anti-hail net systems in apple orchards. The analysis revealed that only one article had addressed this subject using remote sensing, specifically Sentinel-2 imagery, without integrating machine or deep learning approaches. This gap represents a critical research opportunity, as the application of robust classification algorithms combined with spectral time series can enhance the automatic detection and characterization of netted orchards. Filling this gap is essential to advance methodological frameworks for perennial crop monitoring and to support agricultural management under protective systems. Although peer-reviewed studies on anti-hail net mapping in apple orchards using ML/DL are absent, a recent conference contribution [21] demonstrated the feasibility of applying Random Forest and Sentinel-2 imagery for net detection in Vacaria (RS, Brazil), achieving an overall accuracy of 0.76. This supports the potential of

ML approaches for this task, while also highlighting the need for more robust algorithms and multi-sensor frameworks.

Brazil stands out as one of the most important apple-producing countries in the Southern Hemisphere. The municipality of Vacaria, located in the state of Rio Grande do Sul, is a strategic production hub, consistently ranking among the country's top apple producers [22]. In this region, anti-hail nets are widespread, making it an ideal case for evaluating their detectability through remote sensing.

To address the identified gaps, this study evaluates the feasibility of mapping apple orchards cultivated under anti-hail net systems of different colors (black and white), together with other land-use and land-cover classes within a heterogeneous landscape. The analysis relies on SITS, derived from HLS and Sentinel-2 data combined with machine learning. To achieve this goal, the study focuses on four main aspects: (i) the potential of vegetation indices and machine learning to detect apple production under anti-hail nets, including their color variants; (ii) the influence of spatial/temporal resolution and dataset choice (HLS vs. Sentinel-2) on classification performance; (iii) the role, stability, and interpretability of vegetation indices across algorithms (Random Forest vs. 1D Convolutional Neural Networks); and (iv) the trade-offs between accuracy and computational cost.

2. Materials and Methods

To achieve this study's objectives, a stepwise approach was adopted (Figure 1) ranging from field data acquisition to accuracy assessment. Each step will be detailed in the following subsections.

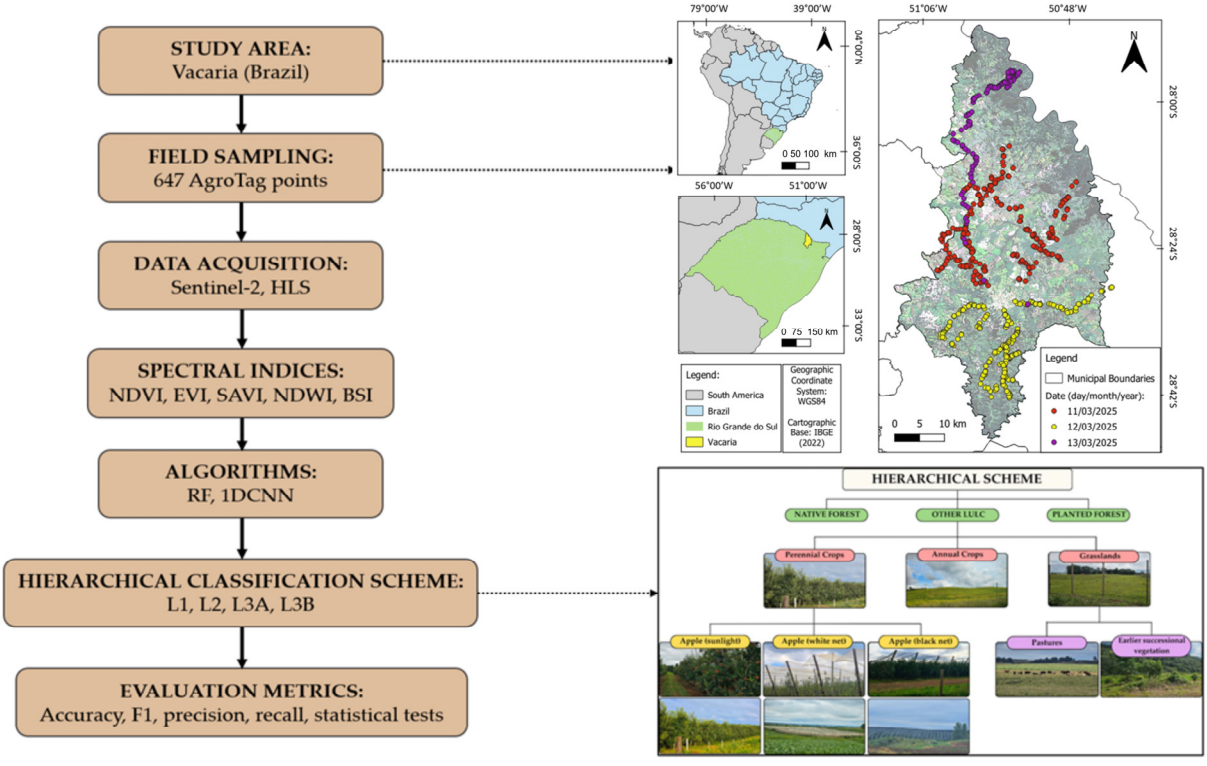

Figure 1. The flowchart illustrates the methodological framework applied in this study, from study area definition and field data collection to spectral index processing, hierarchical classification, and performance evaluation.

2.1. Study Area Characterization

Vacaria is a municipality in the northeastern part of Rio Grande do Sul (RS) State, in southern Brazil (Figure 2), situated within the Atlantic Forest biome with an area of

approximately 2125 km^2. Vacaria is Brazil's second-largest apple-producing municipality, with approximately 254,000 metric tons harvested in the 2021/2022 season [23,24]. Since 2010, the municipality has produced roughly 37% of Rio Grande do Sul's apple output, and about 17% of Brazil's total. The northeastern region of Rio Grande do Sul and the southern region of Santa Catarina stand out in terms of Brazil's apple production (Figure 2), owing to their favorable temperature and relief conditions.

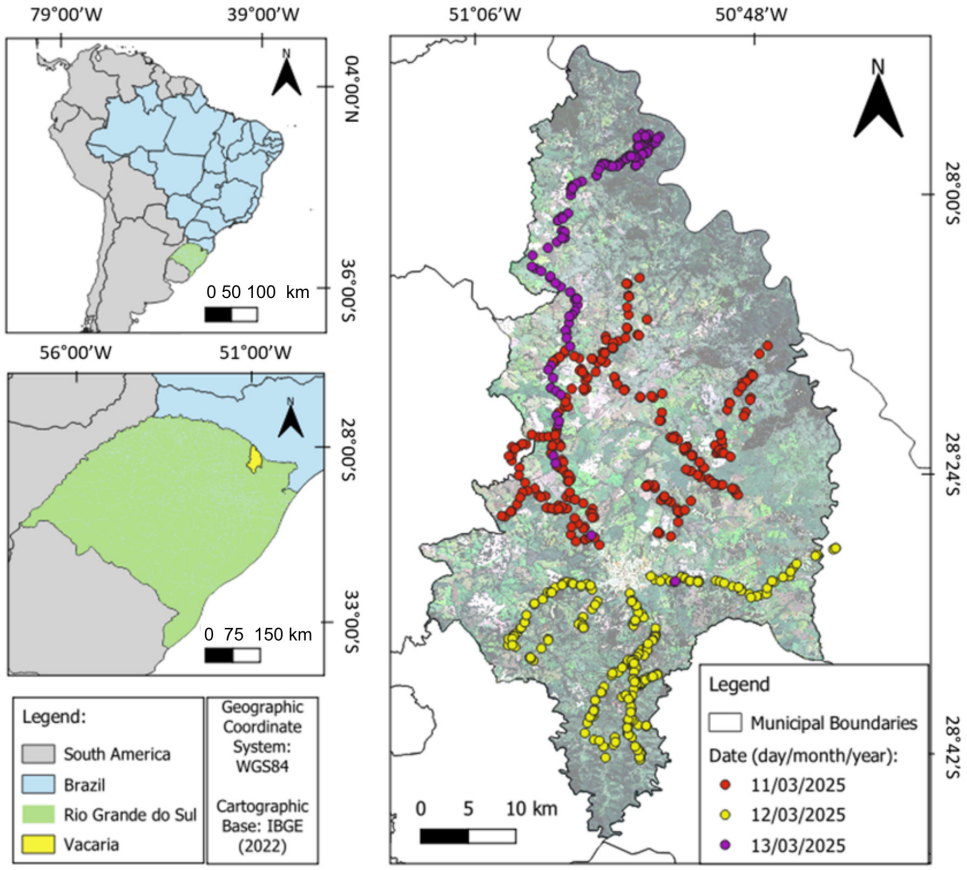

Figure 2. Study area in Vacaria, Brazil, with 647 field points collected using the AgroTag mobile application (Version 7.9.2) in March 2025. Colors represent sampling days: March 11 (red), March 12 (yellow), and March 13 (purple).

Vacaria's climate is classified as temperate with mild summers and no distinct dry season (Cfb) according to the Köppen classification. The climatological normal for annual accumulated precipitation is 1826.5 ± 303.6 mm, while the mean annual air temperature is 16.6 ± 0.46 °C, averaging 11 °C in winter (June–September) and 19.9 °C in summer (December–March). Between 1991 and 2020, precipitation showed an increasing trend of +9.37 mm per year (R^2 = 0.27), and temperature displayed a warming trend of +0.02 °C per year (R^2 = 0.38). This characterization was done using data from the Climate Hazards Center InfraRed Precipitation with Station (CHIRPS) [25], and from the European Centre for Medium-Range Weather Forecasts (ERA5) [26].

Vacaria presents altitudes ranging from 550 to 997 m (Figure 3A). The relief is predominantly gently undulating and undulating, representing most of the municipal territory, while flat areas also occupy a considerable portion. More rugged landscapes, such as strongly undulating and mountainous areas, occur in smaller proportions and are mainly associated with native forests and natural grasslands. The predominant soils in the municipality include Brown Latosols, Humic Cambisols, and Red Latosols (Figure 4), which are

distributed according to relief conditions and influence agricultural suitability across the region [27].

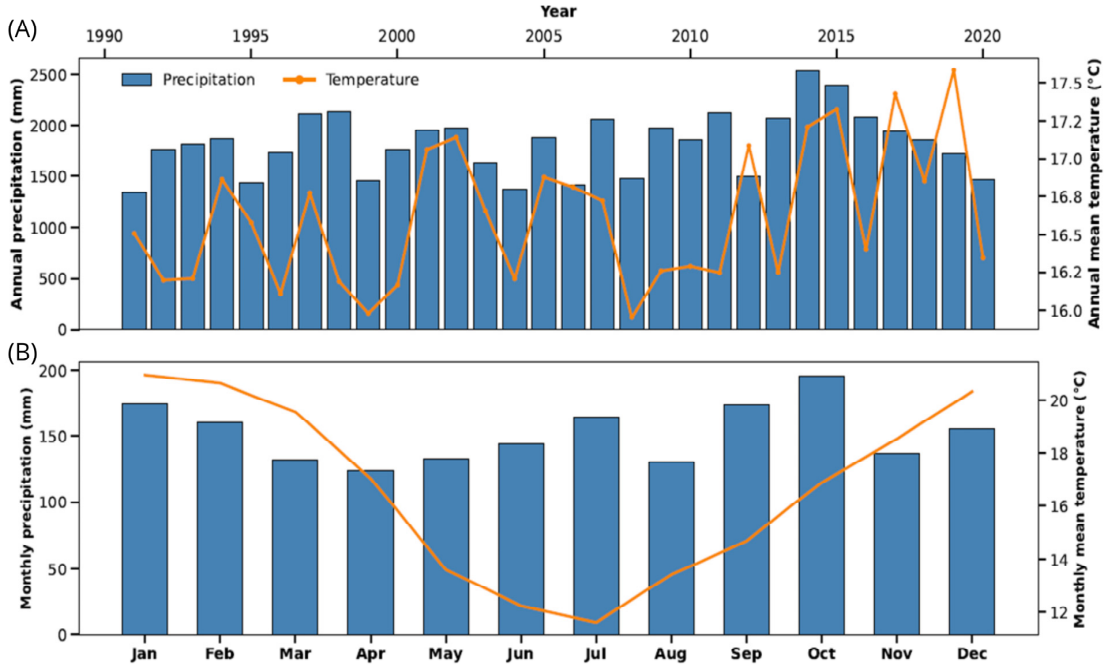

Figure 3. Annual and monthly precipitation and mean temperature in Vacaria/RS (1990–2020). The graph was prepared by the authors using data from [25,26], obtained via Google Earth Engine (GEE).

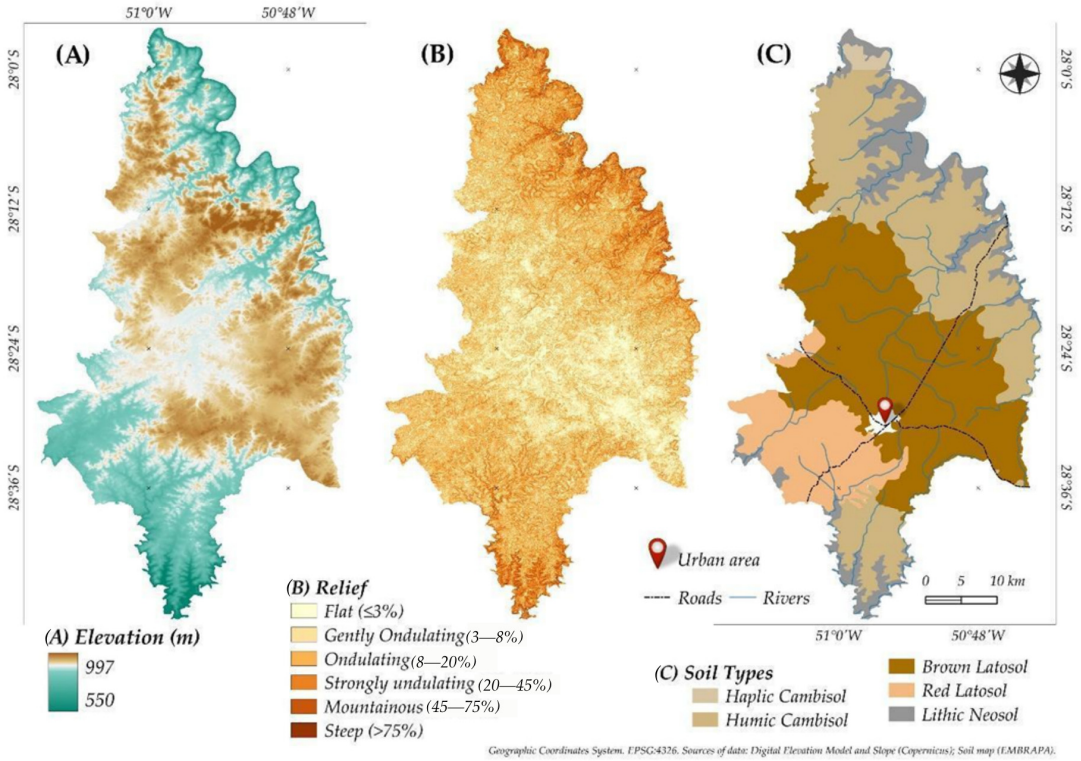

Figure 4. (**A**) Elevation, (**B**) relief, and (**C**) soil types in the Vacaria, Rio Grande do Sul, Brazil. The relief classes were defined according to the terrain inclination in percentage, as established by Embrapa [28].

Vacaria has diversified agricultural production. In 2024, the municipality harvested more than 107,000 hectares, including 67,400 ha of soybeans, 11,600 ha of corn, 7750 ha of

wheat, 9100 ha of oats, 4000 ha of beans, and 6672 ha of apples. Other crops and fruits are also cultivated, such as grapes (117 ha), peaches, and pears, the latter two mainly grown in greenhouses [23]. In addition, the region hosts significant livestock activities, both beef and dairy cattle [27].

2.2. Covariates and Remote Sensing Datasets

Mapping apple orchards covered by protective nets represents a novel approach in automatic crop mapping, as the synthetic material of these coverings introduces additional spectral complexity. These structures alter canopy reflectance by modifying illumination intensity, scattering patterns, and background contribution. The partial shading and anisotropic reflection they produce can distort vegetation signals by influencing photosynthetic activity and promoting differential vegetative growth [3]. To account for these effects, this study employed a set of complementary spectral indices (Table 1) designed to capture distinct surface and illumination characteristics. Vegetation activity was assessed through the NDVI (Normalized Difference Vegetation Index) [29], the EVI (Enhanced Vegetation Index) [30], and the SAVI (Soil-Adjusted Vegetation Index) [31]. EVI was applied to mitigate NDVI's tendency to saturate under high-biomass conditions, which may arise from shading-induced vegetative vigor, and to reduce sensitivity to illumination and atmospheric variability. The SAVI was used to account for variations in soil exposure resulting from differences in canopy density and shading patterns between net-covered and uncovered apple orchards. The NDWI (Normalized Difference Water Index) was employed to capture potential increases in canopy water content and the reduced moisture loss associated with the moderated microclimate under protective nets [3]. Finally, the BSI (Bare Soil Index) was added to detect soil exposure and structural differences related to apple orchard management. Integrating these indices is expected to yield a more comprehensive characterization of the spectral behavior of apple orchards under protective structures, supporting the development of scalable and robust models for production mapping in synthetically modified environments.

Table 1. Spectral indices used in the study.

Spectral Index	Characteristics	Equation	Reference
NDVI	Measures the quantity and vigor of green vegetation, indicating biomass and photosynthetic activity	$\frac{NIR-R}{NIR+R}$	Rouse (1974) [29]
EVI	Improves vegetation detection in dense areas, reducing saturation and atmospheric influence.	$G \times \frac{(NIR-R)}{(NIR+C_1 \times R-C_2 \times B+1)}$	Gao et al. (2000) [30]
SAVI	Adjusts for the influence of exposed soil in areas of low vegetation cover.	$\frac{NIR-R}{(NIR+R+L)} \times (1+L)$	Huete (1988) [31]
NDWI	Highlights the presence and water content in vegetation and soil.	$\frac{NIR-SWIR1}{NIR+SWIR1}$	Gao et al. (1996) [32]
BSI	Highlights areas of exposed soil in contrast to vegetation and water	$\frac{(SWIR+R)-(NIR+B)}{(SWIR+R)+(NIR+B)}$	Rikimaru et al. (2002) [33]

In EVI: G = 2.5; C1 = 6; C2 = 7.5. In SAVI, L = 0.5. NIR = Near-Infrared band; SWIR = Shortwave Infrared band; SWIR1 = Shortwave Infrared band 1; Red = Red band; Blue = Blue band; and Green = Green band.

This study used Sentinel-2 Multispectral Instrument—MSI (European Space Agency (ESA) and HLS imagery to compare the influence of spatial and temporal resolution on apple orchards mapping. The goal was not only to detect and characterize the net structures but also to understand how finer spatial detail (10 m in Sentinel-2) versus higher temporal frequency (up to two days in HLS) affects the ability to isolate net-covered pixels from neighboring land-use and land-cover (LULC) classes. Sentinel-2 surface reflectance Level 2 data were used to derive the five spectral indices using Google Earth Engine (GEE).

For HLS, two collections were considered: HLS.S30 (S2 surface reflectance) and HLS.L30 (Landsat surface reflectance), also acquired using GEE. The study period extended from 1 January 2024, to 31 December 2024. All data with cloud cover < 90% for both datasets in the period were incorporated into the analysis. This threshold was adopted to exclude only nearly completely contaminated scenes, preserving as many valid pixels as possible. It ensures greater temporal density of observations and reduces gaps in the series. The criterion also minimizes bias toward wetter periods, when cloud occurrence is higher, by avoiding the systematic removal of entire dates. Thus, partially cloud-free areas are still exploited, maintaining the spatial and temporal consistency needed for interpolation.

We chose to use remote sensing data from 2024 to define the LULC classes identified in the field in March 2025, based on three premises: (i) major land-use changes occurring up to the survey date would already be captured in 2024 imagery; (ii) more stable land-use classes in Vacaria, such as apple orchards, pastures, and natural vegetation, have a low probability of recent conversion; and (iii) although there are more dynamic classes (e.g., annual crops) that can vary from year to year, using 2024 data reduces the risk of missing these changes. Finally, the Sentinel constellation underwent a significant change in late 2024, with the replacement of Sentinel-2A by the newly launched Sentinel-2C [34]. Therefore, using data before these changes avoids potential fluctuations in revisit frequency and calibration associated with this transitional period.

Initial preprocessing included cloud, shadow, snow, and high-aerosol masking based on the SCL (for S2) and the Fmask (for HLS) bands. For each image, reflectance bands were mapped to a standard set of spectral channels, after which the indices were calculated. Each image was tagged with acquisition date, sensor (S30 or L30 in the HLS case), and cloud coverage metadata. Because the HLS dataset presents irregular acquisition intervals, the median revisit frequency was used to define a fixed temporal grid for resampling. Over the study area, this median revisit frequency was approximately three days.

In the Google Collaboratory environment, daily observations were first aggregated by date; for HLS, Landsat-30 (L30) was prioritized when both sensors were available. This was necessary because the HLS framework uses the Landsat Operational Land Imager (OLI)-United States Geological Survey (USGS) as the reference standard for bandpass adjustment and Bidirectional Reflectance Distribution Function (BRDF) normalization. Landsat products offer higher geolocation accuracy and calibration stability compared to pre-GRI Sentinel-2 acquisitions [34]. Each dataset was then mapped to its native regular grid: Sentinel-2 to a 5-day grid (1 January 2024 to 31 December 2024) and HLS to a 3-day grid over the same period (median temporal resolution). Each series was reindexed to this grid and filled using linear time interpolation (method "time" in pandas library version 2.2.2), followed by forward and backward filling to eliminate remaining gaps at the edges.

All interpolated series were smoothed via discrete wavelet transform using Daubechies-4 (db4) with soft-thresholding of detail coefficients at $\sigma/2$ (series standard deviation), followed by reconstruction (waverec). This configuration represents a balanced smoothing setup, as the $\sigma/2$ soft-threshold effectively suppresses high-frequency noise while preserving phenological shape. More aggressive thresholds (e.g., $\sigma/3$ or $\sigma/4$) tend to over-smooth vegetation signals, flattening seasonal dynamics and reducing the interpretability of short-term variations. The Daubechies-4 (db4) wavelet was selected for its intermediate support length, offering an optimal balance between denoising performance and phenological fidelity. Shorter wavelets (e.g., db2) may leave residual noise, while higher-order functions (e.g., db6 or db8) typically oversmooth and distort short-term vegetation dynamics. Overall, the adopted db4–$\sigma/2$ configuration provided a stable compromise that minimized noise without altering genuine temporal patterns.

Therefore, considering the fixed temporal resolutions of the datasets (5-day for Sentinel-2 and 3-day for HLS), each Sentinel-2 index yielded 73 valid observations annually, while each HLS index yielded 122 observations per year. The comparative average signatures of S2 and HLS for all target classes in this study for each spectral index during the analysis period are displayed in Figure 5.

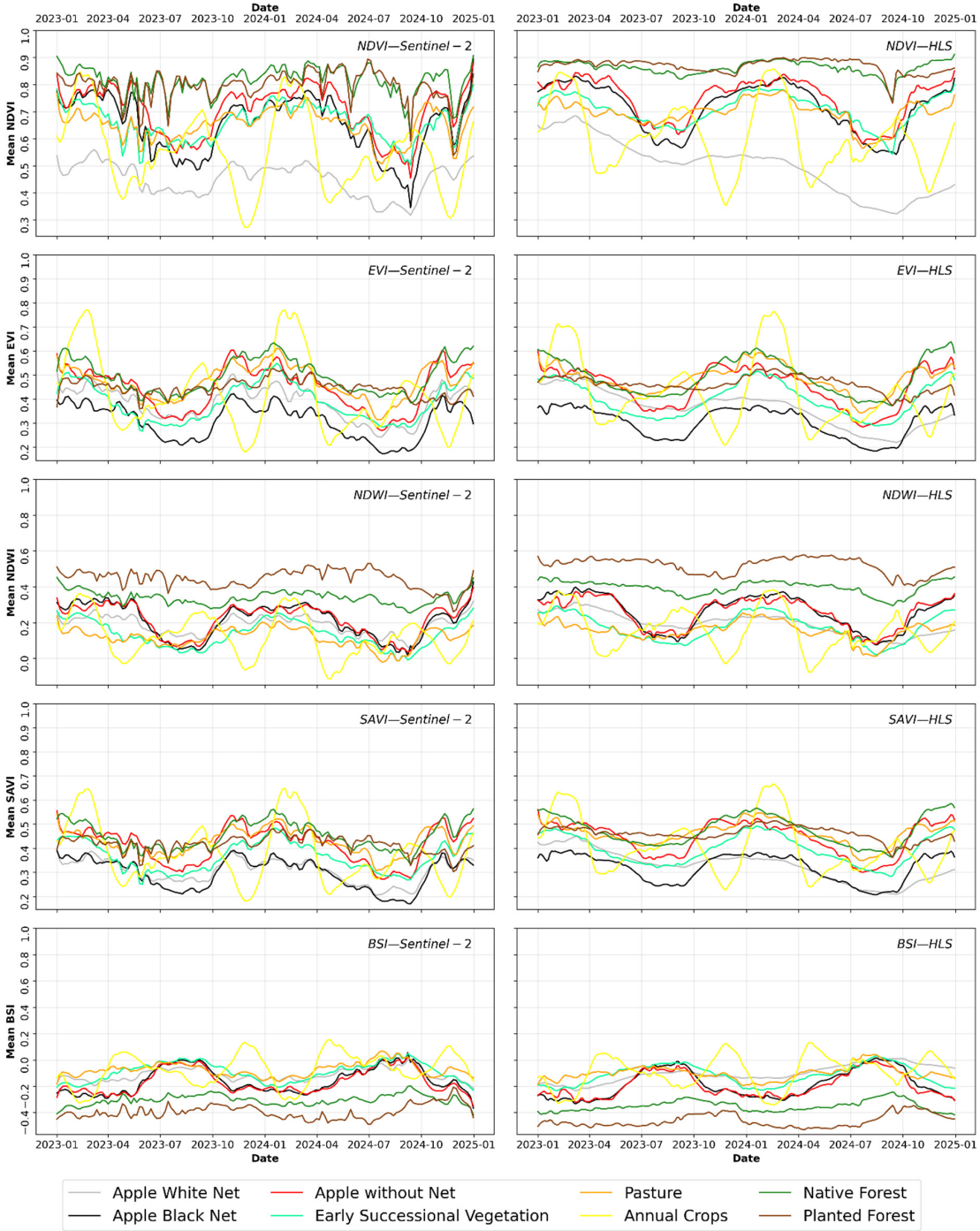

Figure 5. Mean temporal signatures by class and index from 2023 to 2024: Comparing S2 and HLS temporal–spectral behavior. NDVI—Normalized Difference Vegetation Index; NDWI—Normalized Difference Water Index; BSI—Bare Soil Index; EVI—Enhanced Vegetation Index; and SAVI—Soil-Adjusted Vegetation Index.

2.3. Sampling Strategies and Classification Scheme

To build the training and validation dataset, two complementary approaches were combined: (i) a three-day field campaign to collect ground samples along major roads, and (ii) remote sampling based on a temporal visual inspection of satellite data. The procedures are described below.

The field campaign took place 11–13 March 2025 to collect in situ data. The survey focused on apple orchards, both with and without protective nets, and also covered other agricultural land uses, including grain and cereal crops and pastures. Although fruit production sites for grapes, pears, strawberries, and blueberries were visited, these were excluded from analysis due to their limited area and use of greenhouses.

A total of 647 sampling points were collected within the municipality using the AgroTag application [35] (Figure 2). This app enables the collection of georeferenced data through standardized LULC forms, supports photographic documentation, and provides access to high-resolution basemaps. The plots visited were vectorized as polygons, from which up to three random points were sampled per feature. This approach was adopted for two main reasons: (i) to increase the volume of training and validation data, a crucial step for ensuring the proper performance of machine learning (ML) and deep learning (DL) models; and (ii) to better capture the spectral variability of LULC classes, as considerable variations can occur even within a single production plot. To mitigate spatial autocorrelation, an anti-spatial leak strategy was adopted, as described in the following section.

In addition to the agricultural areas sampled in the field, remote sampling was carried out for classes not visited in situ or with insufficient field data, namely native forest, planted forest (silviculture), pasturelands, and early-successional vegetation, the latter representing herbaceous–shrubby cover in transitional landscapes between forest remnants and agricultural lands.

To achieve this, sample points were identified using 10 m true-color Sentinel-2 mosaics for March 2025 (to align with the field campaign) and historical high-resolution Google Earth imagery using Quantum GIS (version 3.30.2, QGIS Development Team. Open Source Geospatial Foundation Project). To improve the representativeness of the training dataset and enhance the model's generalization capacity, the remote sampling area was limited by a 20 km buffer beyond the municipal boundaries. This broader search radius increased the diversity of LULC conditions available for sampling, reducing spatial bias and supporting the extrapolation of model predictions to adjacent areas with similar environmental and spectral characteristics. The final total number of sampling points was 3685, as shown and described in Figure 6.

Based on this dataset, a hierarchical classification scheme was developed to gradually isolate pixels related to apple production. This allows for the identification and characterization of apple cultivation, while also supporting the assessment of model performance in classifying other LULC types within the study area. At the first level (L1), the landscape was separated into native forest (NF), planted forest (PF), and other LULC (OT), creating a broad distinction between forests and human-modified areas. The second level (L2) refined the "other LULC" category into perennial crops (PC), annual crops (AC), and grasslands (GR), highlighting differences in vegetation structure and phenology. Finally, at the third level (L3A and L3B), the classification reached the target classes: perennial crops were subdivided into apple production under sunlight (AS) and two classes of apple production under protective anti-hail nets: white net (AWN), and black net (ABN). Moreover, grasslands were separated into pasturelands (PT) and earlier successional vegetation (ESV). This stepwise approach, illustrated in Figure 7, allowed us to progressively isolate pixels related to apple cultivation. This ensured not only the identification of orchards but also

the characterization of management practices, which are essential for monitoring apple production in the region.

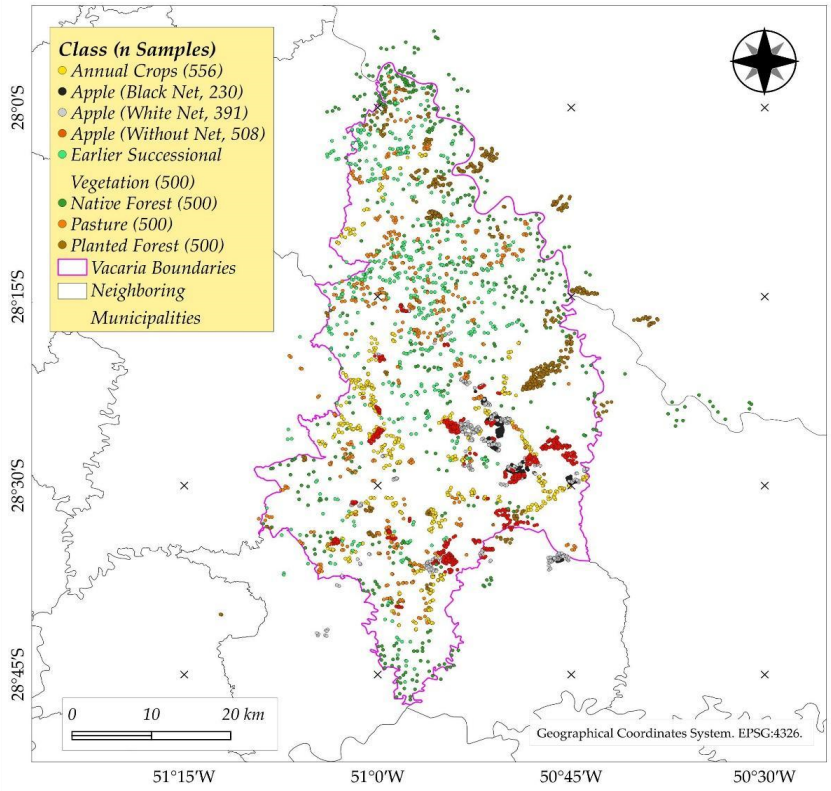

Figure 6. Spatial distribution of land-use and land-cover sample points in Vacaria region, Rio Grande do Sul, Brazil. The dataset includes apple orchards with and without anti-hail nets, combining field data collected in March 2025 with remote inspection of satellite imagery from the same period. The samples were collected using a 20 km buffer around the boundaries of Vacaria as a coverage limit.

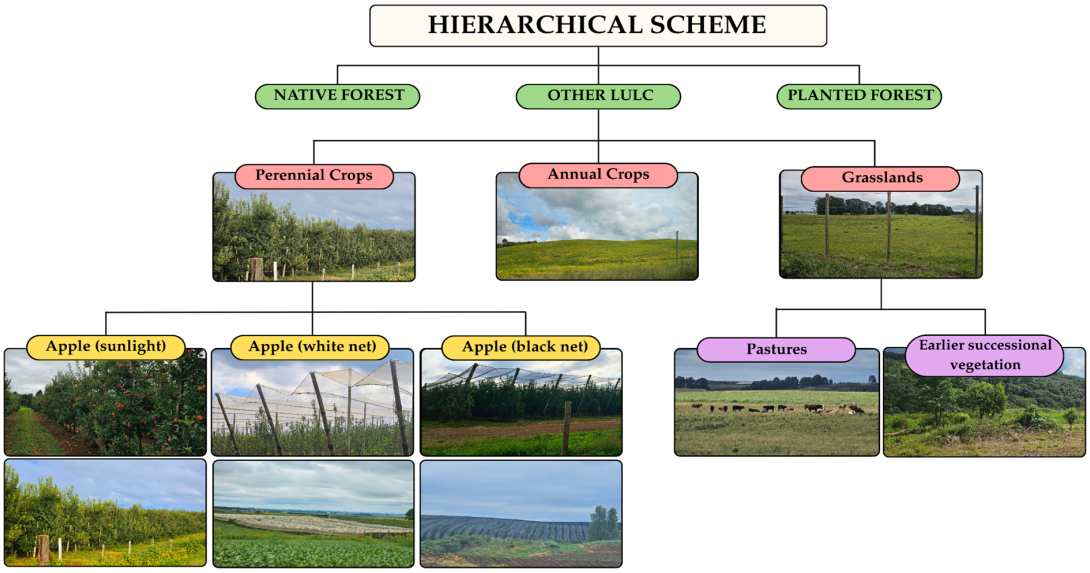

Figure 7. Hierarchical land use and land cover (LULC) classification scheme. The structure is organized into three main levels: (i) native forest, planted forest, and other LULC; (ii) perennial crops, annual crops, and grasslands/pastures; and (iii) subclasses of perennial crops: apple (sunlight), apple (white net), apple (black net); and subclasses of grasslands, pastures and earlier successional vegetation.

The adoption of a hierarchical classification approach was essential for ensuring thematic consistency and improving model interpretability across multiple spatial and spectral scales. This structure allows the progressive isolation of apple-related classes while preserving contextual information about land uses. Including non-apple categories at broader levels (e.g., native and planted forests, annual and perennial crops, and grasslands) was crucial to represent the full land-use heterogeneity of the study area and to avoid confusion between spectrally similar classes. Such an approach captures the complexity of LULC systems, reduces ambiguities in classification boundaries, and provides a scalable framework that enhances both thematic coherence and cross-regional comparability. This methodological choice aligns with widely adopted hierarchical frameworks in LULC mapping, which emphasize multi-level organization as a means of improving accuracy and interpretability across classification tasks [36–38].

2.4. Variable Combinations, Data Partitioning, and Learning Algorithms

To assess the contribution of the S2 and HLS datasets to the detection and characterization of all LULC classes, particularly those associated with AS, AWN, and ABN, a systematic evaluation using the five selected spectral indices (NDVI, NDWI, EVI, SAVI, and BSI) was implemented. Each index was tested individually and in all possible non-redundant combinations (pairs, triplets, quartets, and a quintet), resulting in 31 distinct input configurations. The analysis aimed not only to evaluate the relative contribution of each index at all classification levels, but also to identify the role of the underlying surface properties they represent (e.g., water, exposed soil, vegetation vigor) in the outcomes. In addition, we examined whether the number of indices affected model performance to determine the minimum index set required for optimal results at each level, thereby improving overall model interpretability.

Because up to three samples were collected within the same plot, an additional step was required to prevent spatial leakage and inflated accuracy due to spatial autocorrelation between nearby points. To control for this effect across all classification levels, a hybrid spatial filtering strategy was implemented, combining grid-based blocking, within-cell minimum-distance pruning, and grouped cross-validation. Each sample was first assigned to a square grid cell defined in projected coordinates (EPSG:31983). Within each cell and class, a greedy cKDTree filter removed points located closer than the minimum spacing, ensuring no two proximate observations remained. This procedure ensured spatial independence both within and across folds: if a sample was included in the training set, no other from the same grid cell could appear in the test set.

After spatial filtering, a strict balance was enforced by limiting each class to the minority class count and applying hierarchical stratification. This preserved proportional representation at two levels: (i) across all target classes within each classification level, maintaining inter-class balance, and (ii) within composite classes containing multiple subclasses, retaining their internal proportions in both training and testing sets. The resulting datasets were thus balanced and spatially independent. They were then partitioned into five fixed folds for cross-validation using StratifiedGroupKFold, ensuring both class balance and group integrity under an 80/20 train/test ratio per fold. Although this procedure reduced the overall number of samples, it still provided sufficient data for robust classification (Table 2).

Table 2. Distribution of training and testing samples (≈80/20 split *) across the five fixed cross-validation folds for each class and classification level.

Level	Classes	Number of Samples: Train/Test				
		Fold 1	**Fold 2**	**Fold 3**	**Fold 4**	**Fold 5**
1	Native Forest—NF	188/51	194/45	193/46	191/48	190/49
	Planted Forest—PF	193/46	185/54	194/45	192/47	192/47
	Other—OT	192/47	195/44	186/53	191/48	192/47
2	Perennial Crops—PC	225/50	220/55	219/56	221/54	215/60
	Annual Crops—AC	220/55	213/62	222/53	217/58	228/47
	Grasslands—GR	215/60	226/49	219/56	222/53	218/57
3A	Apple Sunlight—AS	89/21	87/23	86/24	87/23	91/19
	Apple White Net—AWN	85/25	86/24	91/19	86/24	92/18
	Apple Black Net—ABN	88/22	91/19	86/24	90/20	85/25
3B	Pasture—PT	226/65	245/46	231/60	231/60	231/60
	Earlier Successional Veg.—ESV	239/52	221/70	236/55	234/57	234/57

* Although folds were defined using a nominal 80/20 train–test ratio, the actual proportions may vary slightly due to the group-based constraint of the StratifiedGroupKFold algorithm. Each grid cell was treated as an indivisible group, so the exact ratio depended on the number and size distribution of grid cells assigned to each fold.

Two machine learning algorithms were selected to represent distinct methodological paradigms: a traditional ensemble-based model (Random Forest, RF) and a deep-learning model (1D Convolutional Neural Network, 1DCNN). RF [39] was chosen as a baseline because it remains one of the most reliable and computationally efficient algorithms for crop mapping [8,14], especially under limited sample size and heterogeneous conditions. It provides robust performance, interpretability through variable importance, and resilience to noise and overfitting [40]. 1DCNN, which has yielded good results in agricultural mapping [41–43], was adopted to exploit the temporal and spectral dependencies across the multivariate index sequences, while maintaining a lightweight architecture suitable for the available training sample size.

Regarding hyperparameterization, the RF classifier was implemented using the square root of the total number of predictors as the number of features for each split, following the original recommendation for classification tasks [39]. A minimum of two samples per leaf node was defined to prevent overfitting on small subsets. At the same time, the maximum tree depth was limited to 20 to constrain model complexity and improve generalization. Class weights were adjusted through balanced subsampling to account for residual class imbalance. The number of trees was fixed at 200, as preliminary tests indicated no meaningful accuracy gains beyond this threshold, while ensuring stable feature importance estimates and manageable computational cost. A full grid search on this parameter was avoided, as this increases computational efficiency without improving model performance. Parallel processing was enabled, and reproducibility was ensured by setting a fixed seed.

1DCNN was implemented using TensorFlow, with the final architecture detailed in Table 3. The 1DCNN architecture adopted in this study was designed to be lightweight and computationally efficient, as deeper models tend to provide only limited accuracy gains at substantially higher computational costs [41–43]. In contrast to multi-block or residual architectures commonly used for large-scale crop mapping, this model employs only two convolutional layers followed by batch normalization and global average pooling, resulting in fewer parameters and faster convergence. This compact configuration was also

motivated by the relatively small training sample size and the smoothed temporal profiles of the vegetation indices to ensure stable performance and reduce overfitting risks.

Table 3. 1DCNN architecture and training parameters, reproduced in Levels 1, 2, 3A, and 3B.

Steps	Layers and Parameters	Output
Input	Multivariate sequence, shape = (T, F)	(T, F)
1	Conv1D(64, k = 5), padding = "same", ReLU, L2 = 1×10^{-4} → BatchNorm	(T, 64)
2	Conv1D(64, k = 5), padding = "same", ReLU, L2 = 1×10^{-4} → BatchNorm	(T, 64)
3	GlobalAveragePooling1D	64
4	Dense(64) + ReLU → Dropout (0.3)	64
5	Dense(3) + Softmax	3
	Hyperparameters and Protocols (Training Stage)	
Optimizer	Adam (lr = 0.001), clipnorm = 1.0	
Loss/Metric	Sparse Categorical Cross-Entropy/Accuracy	
Regularization	L2 = 1e-4 in conv/dense areas; Dropout = 0.3	
Callbacks	EarlyStopping (patience = 30, restore_best_weights = True); ReduceLROnPlateau (factor = 0.5, patience = 10)	
Batch/Epochs	batch_size = 32; maximum. 200 epochs	
Balancing	class_weight to mitigate imbalance	

All experiments were executed in the Google Colab environment running Python 3.12.12 on Linux 6.6.105+, with eight logical (four physical) CPU cores and 54.8 GB of RAM. No GPU acceleration was used during model training. The main libraries used were NumPy 2.0.2, pandas 2.2.2, scikit-learn 1.6.1, and TensorFlow 2.19.0. Reported training times correspond exclusively to model fitting stages, measured per fold and averaged across repetitions under identical computational conditions.

2.5. Performance Metrics and Statistical Analysis

A performance assessment was conducted consistently across all classification levels and algorithms (RF and 1DCNN). For each test, per-fold artifacts were stored, including confusion matrices and summary metrics derived from the independent test sets. Model performance was primarily evaluated using overall accuracy and macro-averaged F1-score, computed at the ID level to ensure correspondence with spatially independent validation units. The macro-averaged F1-score provided a balanced assessment across classes regardless of their relative frequency. Confusion matrices from all folds were then aggregated to derive class-specific metrics for each class (Specificity and Sensitivity), providing a complementary view of omission and commission errors. All equations related to these metrics are shown below. For each model and variable combination, the fitting time per fold was recorded, allowing for comparisons of computational efficiency between algorithms. This approach ensured a consistent and comparable evaluation framework across models and hierarchical levels, balancing interpretability and the robustness of the accuracy estimates.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

where TP = True Positive; TN = True Negative; FP = False Positive; and FN = False Negative [44].

$$\mathrm{F1}_{\mathrm{macro}} = \frac{1}{K}\sum_{k=1}^{K} F1_k \quad (2)$$

where K is the total number of classes and

$$\mathrm{F1} = \frac{2\,(Precision \times Recall)}{Precision + Recall} \quad (3)$$

where

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall/Sensitivity = \frac{TP}{TP + FN} \quad (5)$$

$$Specificity = \frac{TN}{TN + FP} \quad (6)$$

Models were ranked through a hierarchical, statistically grounded procedure designed to identify not only the best results, but also the most robust, efficient, and parsimonious configurations. The ranking relied on four complementary decision levels:

1. Primary Performance Metric: The Macro-F1-score was computed from spatially independent validation folds. Macro-F1 balances precision and recall across classes, thus preventing bias toward majority classes and providing a reliable estimate of overall classification quality.
2. Statistical Significance Testing: Since all models were evaluated over identical folds, pairwise comparisons between the top-performing model and all other combinations were conducted using the Wilcoxon signed-rank test. The test evaluated whether the observed differences in Macro-F1 were statistically significant ($p < 0.05$).
3. Parsimony Criterion (Number of Spectral Indices): If models remained equivalent in both performance and training time, a third tie-breaking rule favored the combination using the fewest spectral indices. This approach emphasizes model simplicity and interpretability.
4. Computational Efficiency Criterion (Training Time): In cases where statistical equivalence was observed, the mean training time per fold (fit time in seconds) was used as a secondary ranking criterion. The model requiring less computational time was prioritized.

The primary goal of this study was not to compare algorithms directly, but to assess the feasibility of mapping apple production under anti-hail nets using time-series imagery from different datasets. To facilitate comparison, model performances across classifiers and datasets were analyzed to assess how data scale and heterogeneity influenced each model. The Wilcoxon signed-rank test was used to assess if the median difference between paired performance samples differed from zero. Macro-averaged F1-scores from the 1DCNN and RF models trained on identical variable combinations were compared.

To assess whether specific indices consistently improved performance across combinations, multi-index models were decomposed into index occurrences. For each occurrence of index i within any combination C (with $i \in C$), the corresponding macro-F1 (averaged across folds) was recorded and pooled across datasets and algorithms. Descriptive statistics (mean, standard deviation, frequency) were computed per index to provide an occurrence-based summary of utility. In addition, marginal gains were estimated using a leave-one-in/out approach. The additive gain measured how much the macro-F1 improved when a given index was added to a combination, while the removal impact captured the loss in macro-F1 when that index was removed. These values were averaged across datasets and algorithms to assess each index's overall contribution and consistency.

3. Results

3.1. Overall Classification Performance

Across the four classification levels, RF and 1DCNN both performed well, with macro-F1 > 0.85 in most scenarios, indicating strong discrimination among LULC classes, including apple orchards with and without anti-hail nets. Differences between models and sensors varied with the thematic complexity of each level.

At L1 (Native Forest, Planted Forest, Other), performance was highest and most consistent. RF averaged 0.956 on HLS and 0.952 on S2, with very low standard deviation (<0.02), reflecting the ease of spectral separation, even when including the Other class, which aggregates six distinct subclasses of LULC. 1DCNN also performed well, with averages of 0.891 (HLS) and 0.932 (S2), but with greater variability. When limited to one or two indices, its performance dropped significantly. At this level, RF retained an advantage in stability and accuracy, especially on HLS data (Figure 8).

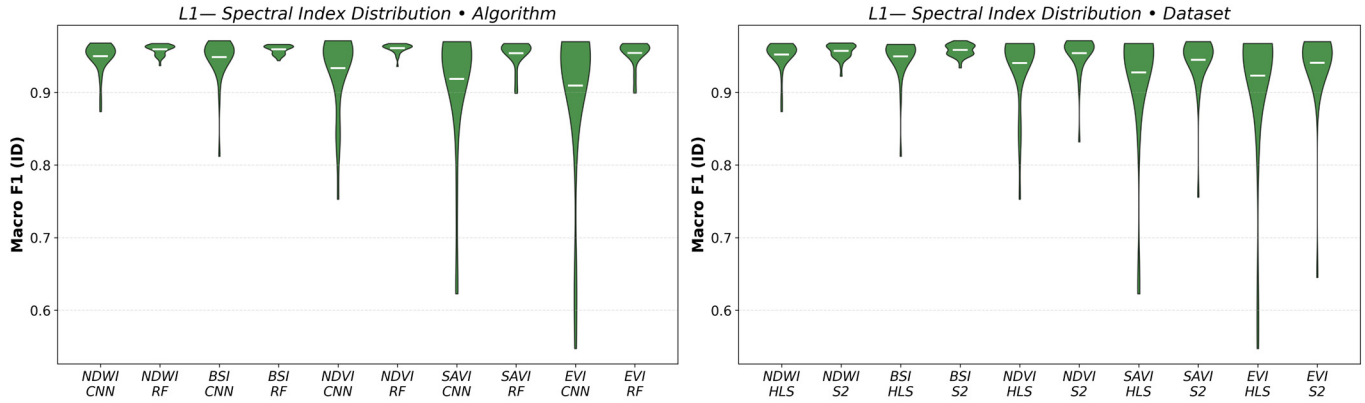

Figure 8. Comparative distribution of index performance by algorithm (**left**) and sensor (**right**) (Level 1—native forest vs. planted forest vs. other).

At L2 (Perennial Crops, Annual Crops, and Grasslands), both models exhibited similar averages (0.90), representing a slight decrease compared to L1. 1DCNN averaged 0.904 (HLS) and 0.901 (S2), while RF reached 0.893 (HLS) and 0.902 (S2). Performance variation across folds remained low for all configurations, with standard deviations ranging between 0.012 and 0.019. Greater dispersion was observed for the S2-derived 1DCNN results (std = 0.076), indicating higher variability across spectral index combinations at this classification level (Figure 9).

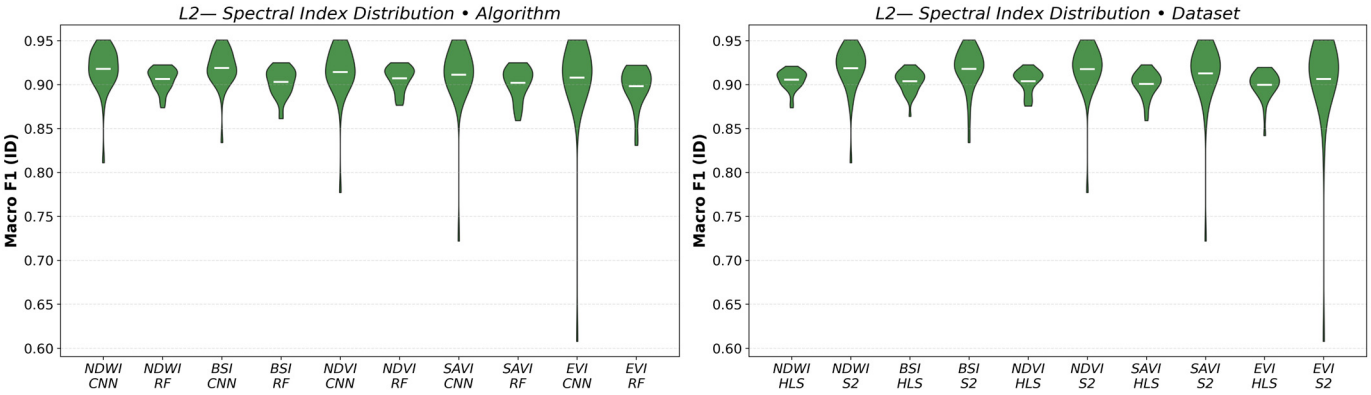

Figure 9. Comparative distribution of index performance by algorithm (**left**) and dataset (**right**) (Level 2—perennial crops vs. annual crops vs. grasslands).

At L3A (apple under sunlight, white net, and black net), which represents the core innovation and primary focus of this study, the results demonstrated encouraging potential despite a broader dispersion in performance (Figure 10). The RF maintained a high average Macro-F1 of 0.874 (HLS) and 0.903 (S2), with moderate variability (standard deviations ≈ 0.05–0.06), while 1DCNN achieved averages of 0.760 (HLS) and 0.842 (S2). Although some models, particularly those using fewer than three indices, showed lower

performance, several combinations reached maximum Macro-F1 values above 0.95, underscoring the strong discriminative capacity of the proposed approach. These results confirm that, even in a highly challenging scenario involving subtle management differences and partial canopy cover, both models were able to capture class-specific spectral-temporal signatures effectively.

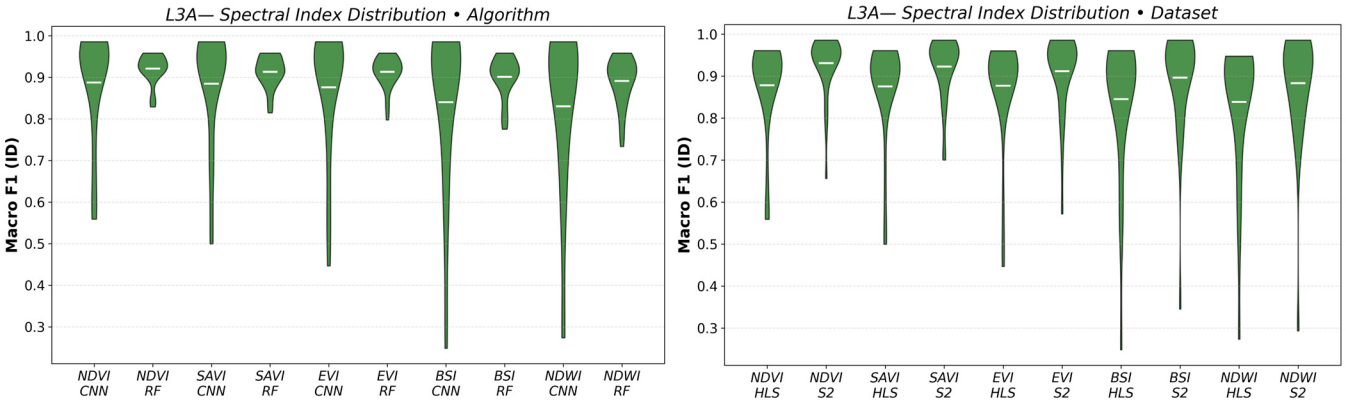

Figure 10. Comparative distribution of index performance by algorithm (**left**) and sensor (**right**) (Level 3A—apple at sunlight vs. apple with white net vs. apple with black net).

Finally, at L3B (Pastures and Earlier Successional Vegetation), both models performed strongly, maintaining high Macro-F1 despite the close spectral–temporal signatures of the two classes (Figure 11). 1DCNN achieved averages of 0.904 (HLS) and 0.854 (S2), while RF reached 0.938 (HLS) and 0.908 (S2). Standard deviations remained low for RF (≈0.03) and moderate for 1DCNN (≈0.12–0.16). These results are particularly relevant given the inherent difficulty of distinguishing cultivated and natural grasslands, which often share similar canopy structure, phenological timing, and reflectance responses across both dry and wet seasons. Peak scores of up to 0.97 Macro-F1 for 1DCNN (HLS) and 0.96 for RF (HLS) indicate that both models captured subtle structural and temporal cues sufficient for reliable discrimination.

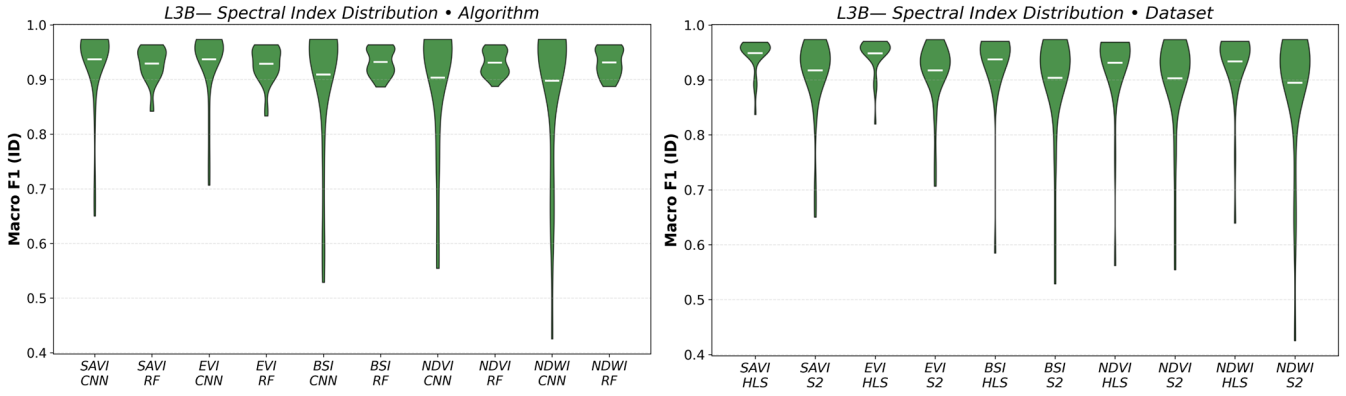

Figure 11. Comparative distribution of index performance by algorithm (**left**) and sensor (**right**) (Level 3B—pasture vs. earlier successional vegetation).

The ten best-performing models for each classification level are summarized in Table 4, ranked according to a multi-criteria framework that integrates predictive accuracy, statistical significance, and model efficiency (see Materials and Methods). Ranking was primarily based on the Macro-F1-score, computed from spatially independent validation folds to ensure robustness across classes.

Table 4. Top 10 best-ranked models for classification levels 1, 2, 3, and 4, considering performance (Macro-F1), significance (after Wilcoxon paired test), efficiency (Fit time), and parsimony (number of indices).

Level	Rank	Dataset	Algorithm	Combination	Macro-F1	Accuracy	Fit Time (s)
L1	1	S2	1DCNN	BSI+NDVI	0.9714	0.9724	17.093
	2	S2	1DCNN	BSI+EVI+NDVI+SAVI	0.97	0.9704	19.035
	3	S2	1DCNN	BSI+EVI+NDVI+NDWI	0.9682	0.9686	15.781
	4	S2	1DCNN	EVI+NDVI+NDWI	0.9679	0.9686	14.68
	5	HLS	RF	EVI+NDVI+NDWI+SAVI	0.9675	0.9679	1.706
	6	HLS	RF	BSI+SAVI	0.9663	0.9665	1.159
	7	HLS	RF	BSI+EVI+NDVI+NDWI+SAVI	0.9663	0.9665	1.809
	8	HLS	RF	BSI+NDVI+SAVI	0.9662	0.9664	1.351
	9	HLS	RF	BSI+EVI+NDVI+SAVI	0.9662	0.9664	1.415
	10	S2	1DCNN	BSI+EVI+NDVI	0.9662	0.9662	17.037
L2	1	S2	1DCNN	BSI+EVI+NDVI+NDWI+SAVI	0.9508	0.9565	21.723
	2	S2	1DCNN	BSI+EVI+NDVI	0.947	0.9526	19.754
	3	S2	1DCNN	BSI+EVI+NDVI+SAVI	0.943	0.9488	21.883
	4	S2	1DCNN	EVI+NDVI+NDWI	0.942	0.9485	19.218
	5	S2	1DCNN	BSI+NDVI+SAVI	0.9414	0.9485	19.077
	6	S2	1DCNN	NDVI+NDWI+SAVI	0.9399	0.9465	20.742
	7	S2	1DCNN	BSI+EVI+NDVI+NDWI	0.9396	0.9466	18.641
	8	S2	1DCNN	BSI+EVI+NDWI+SAVI	0.9391	0.9447	20.73
	9	S2	1DCNN	BSI+NDVI+NDWI+SAVI	0.939	0.9466	22.201
	10	S2	1DCNN	EVI+NDWI+SAVI	0.9379	0.9445	21.357
L3A	1	S2	1DCNN	EVI+NDVI+NDWI+SAVI	0.9857	0.9848	25.823
	2	S2	1DCNN	BSI+EVI+NDVI+NDWI+SAVI	0.9857	0.9848	26.275
	3	S2	1DCNN	BSI+NDVI+NDWI+SAVI	0.9815	0.9809	27.56
	4	S2	1DCNN	EVI+NDVI+SAVI	0.9785	0.9808	33.978
	5	S2	1DCNN	NDVI+SAVI	0.9764	0.9771	27.693
	6	S2	1DCNN	BSI+EVI+NDVI+SAVI	0.9763	0.9771	23.817
	7	S2	1DCNN	BSI+EVI+NDWI+SAVI	0.9725	0.9731	29.91
	8	S2	1DCNN	EVI+NDVI	0.9701	0.9693	28.227
	9	S2	1DCNN	BSI+EVI+NDVI+NDWI	0.9682	0.9696	27.743
	10	S2	1DCNN	BSI+NDWI+SAVI	0.9661	0.969	29.826
L3B	1	S2	1DCNN	BSI+NDWI+SAVI	0.9736	0.9747	25.43
	2	S2	1DCNN	BSI+EVI+NDVI+NDWI+SAVI	0.9735	0.9747	25.491
	3	HLS	1DCNN	BSI+EVI+NDWI	0.9704	0.9708	29.214
	4	S2	1DCNN	BSI+NDVI+SAVI	0.9694	0.9706	26.715
	5	S2	1DCNN	BSI+NDVI+NDWI+SAVI	0.9691	0.9707	25.964
	6	S2	1DCNN	EVI+NDVI+NDWI+SAVI	0.969	0.97	26.394
	7	HLS	1DCNN	BSI+EVI+NDVI+SAVI	0.9687	0.9691	26.041
	8	S2	1DCNN	BSI+EVI+NDWI+SAVI	0.9685	0.9694	27.68
	9	S2	1DCNN	BSI+EVI+NDVI+SAVI	0.9685	0.9694	28.977
	10	HLS	1DCNN	BSI+NDVI+SAVI	0.967	0.9674	26.387

Across 496 model-index combinations, 110 Level 1 combinations achieved $\geq$0.90 overall accuracy (for both HLS and Sentinel-2, using 1DCNN or RF). At Levels 2, 3A, and 3B, the counts were 87, 72, and 94, respectively. Results were consistent when evaluated with macro-F1, reinforcing that both data sources (HLS, Sentinel-2) and classifiers (1DCNN, RF) are well suited to hierarchical crop mapping.

3.2. Classification Performance by Dataset

Across datasets, both algorithms achieved slightly higher mean performance with Sentinel-2 than with HLS, with mean Macro-F1 of 0.8825 versus 0.8647 for 1DCNN and 0.9164 versus 0.9152 for RF (gains of 2.1% and 0.1%, respectively). However, at Level 3B (PT vs. ESV), HLS yielded better results than Sentinel-2 for both algorithms. RF achieved 0.938 $\pm$ 0.033 with HLS versus 0.909 $\pm$ 0.054 with Sentinel-2, and 1DCNN also performed better under HLS (0.904 $\pm$ 0.140 vs. 0.854 $\pm$ 0.178). These gains likely stem from HLS's higher temporal density, which strengthens phenological signal continuity and reduces noise, enhancing class separability when seasonal dynamics are dominant.

This is particularly relevant when dealing with classes that have similar spectral profiles throughout the year, as in the case of grasslands (PT and ESV).

Differences in spatial and temporal resolution had complementary effects on classification performance. Sentinel-2 data improved separability in tasks sensitive to canopy structure and management features, such as detecting apple production under anti-hail nets (L3A) or distinguishing perennial, annual, and grassland classes (L2). In these cases, 1DCNN models trained on Sentinel-2 achieved a higher overall Macro-F1 than with HLS (+1.04% in L2, and +10.9% in L3A). In contrast, HLS performed better in temporally driven analyses, as its higher revisit frequency enhanced phenological signal continuity and reduced noise in time series.

Finer spatial detail reduces mixed-pixel effects and enhances spectral–temporal signals. Conversely, the denser revisit frequency of HLS (~3-day composites from L30 + S30) favored temporarily driven tasks, particularly at Level 3B. At coarser scales (Level 1), performances were similar. This suggests that Sentinel-2's finer spatial resolution (10 m) and revisit frequency (~5 days) already provide sufficient spatial–temporal detail for coarse-level discrimination. In contrast, the denser composites of HLS (~3 days) are advantageous when temporal continuity is critical.

Figures 8–11 compare the mean Macro-F1 scores of both datasets across classification levels and models. While Sentinel-2 retained a slight overall advantage at coarser levels (L1–L3A), HLS had a clear advantage at L3B, highlighting the complementary roles of spatial and temporal resolution on classification outcomes.

Table 5 illustrates how data origin affects class-specific discrimination. 1DCNN displayed more dispersion (F1 ranging from 0.75 to 0.82) than RF, which once again performed particularly well with HLS (F1 > 0.87 across all subclasses). In more homogeneous classes, like Apple (white net), both algorithms performed better with Sentinel-2 (F1 ≈ 0.88–0.92). However, pairwise comparisons between the algorithms yielded $p > 0.1$, indicating no statistically significant difference between the models under this sensor.

Table 5. Classification performance (Precision, Recall, F1, and Support) of RF and 1DCNN for apple orchards and other LULC classes across hierarchical levels (L1–L3B). ΔF1 vs. RF indicates the difference in F1-score between 1DCNN and RF models. *p*-values were obtained using paired Wilcoxon signed-rank tests on index combinations.

Level	Model	Class	Precision	Recall	F1	ΔF1 vs. RF	*p* (Wilcoxon)
L1	1DCNN	NF	0.9252	0.8919	0.9051	−0.043	6.15×10^{-5}
	1DCNN	OT	0.9265	0.9647	0.9436	−0.022	4.16×10^{-8}
	1DCNN	PF	0.9173	0.8939	0.9006	−0.047	3.42×10^{-6}
	RF	NF	0.9631	0.9340	0.9482		
	RF	OT	0.9548	0.9765	0.9655		
	RF	PF	0.9452	0.9502	0.9476		
L2	1DCNN	AC	0.9398	0.9488	0.9437	−0.005	0.221
	1DCNN	GR	0.8663	0.8480	0.8544	0.017	5.18×10^{-6}
	1DCNN	PC	0.9183	0.9151	0.9153	0.008	0.000131
	RF	AC	0.9274	0.9702	0.9483		
	RF	GR	0.8793	0.8014	0.8372		
	RF	PC	0.9011	0.9138	0.9071		
L3A	1DCNN	ABN	0.8505	0.8164	0.8174	−0.079	0.0275
	1DCNN	AS	0.8087	0.8699	0.8187	−0.06	0.0534
	1DCNN	AWN	0.8936	0.7952	0.8196	−0.075	0.493
	RF	ABN	0.8966	0.8969	0.8965		
	RF	AS	0.8760	0.8826	0.8782		
	RF	AWN	0.8970	0.8923	0.8944		
L3B	1DCNN	PT	0.8983	0.9013	0.8974	−0.035	0.908
	1DCNN	ESV	0.8838	0.8733	0.8762	−0.039	0.825
	RF	PT	0.9181	0.9469	0.9322		
	RF	ESV	0.9345	0.8972	0.9153		

3.3. Classification Performance by Algorithm

Overall, RF models achieved higher average performance (0.9158) than 1DCNN models (0.8736) and exhibited lower variability (std = 0.0436 versus 0.1421 for 1DCNN). On the other hand, 1DCNN models achieve higher peak performance. This difference was driven primarily by the weaker 1DCNN performance with one (Macro-F1 of 0.6721) or two indices (0.8664). In contrast, RF averaged 0.8694 and 0.9117 under the same conditions. With three or more indices, 1DCNN outperformed RF, showing higher peak scores with comparable stability. The mean Macro-F1 for 1DCNN increased to 0.9347, 0.9515, and 0.9540 with three, four, and five indices, respectively. Notably, it reached maximum values of 0.9857 for both the four- and five-index combinations, with an average standard deviation as low as 0.024. In contrast, over the same range (three to five indices), RF averaged Macro-F1 scores between 0.9295 and 0.9406, with maxima up to 0.9675.

Considering both performance metrics and Wilcoxon signed-rank results, RF consistently outperformed 1DCNN across most hierarchical levels (Table 6). The largest gaps occurred at L1 (NF, PF, OT) and L3A (AS, AWN, ABN), where RF achieved a mean Macro-F1 of 0.956 ± 0.024 and 0.874 ± 0.064, respectively, while 1DCNN trailed at 0.891 ± 0.150 and 0.760 ± 0.243. These differences were statistically significant ($p < 0.001$). At Level 2 (L2), RF retained a small but consistent advantage ($\Delta = +1.15$ pp, $p < 0.001$). It is worth noting that training time differs substantially between algorithms. On average, Random Forest fits were approximately 21.5 times faster than 1DCNN models (1.10 s vs. 23.66 s), confirming the substantially higher computational cost of convolutional models.

Table 6. Macro-F1 performance comparison between RF and 1DCNN models across levels and datasets, and statistical significance using the Wilcoxon paired test comparing the algorithms.

Level	Dataset	RF Macro-F1 ± SD	1DCNN Macro-F1 ± SD	Δ 1DCNN − RF (%)	*p*-Value	Sig.
L1	HLS	0.9557 ± 0.0241	0.8907 ± 0.1498	−6.51	3.21×10^{-17}	***
L1	S2	0.9519 ± 0.0225	0.9320 ± 0.0912	−1.99	0.06051	ns
L2	HLS	0.8928 ± 0.0239	0.9043 ± 0.0212	1.15	9.47×10^{-9}	***
L2	S2	0.9021 ± 0.0417	0.9013 ± 0.1036	−0.08	4.27×10^{-7}	***
L3A	HLS	0.8738 ± 0.0637	0.7600 ± 0.2430	−11.38	2.07×10^{-4}	***
L3A	S2	0.9031 ± 0.0716	0.8425 ± 0.2100	−6.06	0.2738	ns
L3B	HLS	0.9384 ± 0.0329	0.9039 ± 0.1395	−3.45	1.97×10^{-1}	ns
L3B	S2	0.9085 ± 0.0543	0.8544 ± 0.1782	−5.4	0.06367	ns

ns = not significant; *** = significant levels.

Considering class-wise metrics, RF maintained higher F1 in almost all classes, notably Apple (black net) (F1 = 0.896) and Apple (white net) (F1 = 0.894), as shown in Table 5. 1DCNN was marginally worse, with its largest drop in Apple (sunlight) (F1 = 0.818), reflecting the greater spectral confusion between exposed surfaces and canopy variations. Although RF is superior on average, the performance differences are narrower in more homogeneous subclasses, with Wilcoxon test showing a significant gap ($p < 0.05$) only for Apple (black net).

3.4. Classification Performance by Spectral Index

The evaluation of marginal gains from spectral indices across classification levels revealed consistent patterns linked to vegetation vigor, soil background reflectance, and water content, with sensor-specific differences between Sentinel-2 and HLS. At Level 1 (native vegetation vs. silviculture vs. other land uses), the marginal gain analysis consistently included NDWI and BSI among the most influential indices. In the HLS–1DCNN configuration, NDWI (+0.11) and BSI (+0.10) showed the highest average gains, followed

by NDVI (+0.07). In the Sentinel-2–1DCNN configuration, the same indices (BSI +0.06; NDWI +0.05) were also associated with high average gains, but with greater variability. By contrast, RF models showed smaller marginal effects, with NDVI, NDWI, and BSI averaging only +0.01 to +0.02. Overall, these patterns indicate that indices related to soil exposure (BSI) and moisture (NDWI) contribute consistently across sensors. At the same time, vegetation vigor metrics (NDVI, SAVI, EVI) showed smaller and more variable effects at this classification level.

Graphs were generated to examine how average Macro-F1 changes with and without each index (bootstrapped ablations), across levels, sensors, and architectures, using 120 permutations per scenario. Figure 12 displays the L1 variation with BSI.

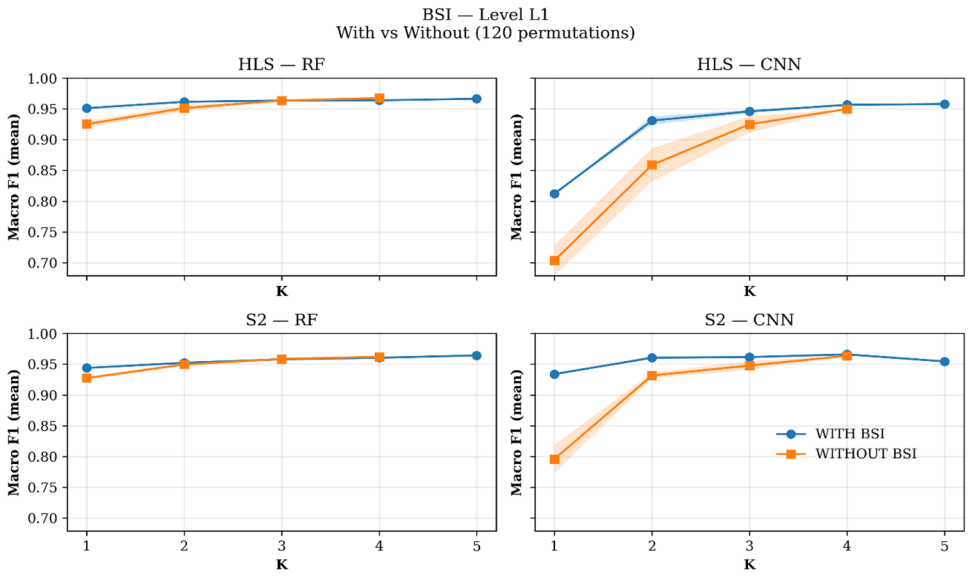

Figure 12. Impact of the Bare Soil Index (BSI) on Level 1 classification (Native Forest, Planted Forest, Others) based on bootstrap analysis.

At L2 (Perennial Crops, Annual Crops, Grasslands), NDWI, BSI, and NDVI under Sentinel-2 + 1DCNN achieved the highest mean marginal gains (≈+0.05–0.06) but also the greatest variability (std ≈ 0.08–0.10), indicating that their contributions were not consistent across all combinations. SAVI and EVI followed the same pattern, with modest mean gains yet marked dispersion. Random Forest also produced modest gains (≈+0.02), but with low variability, indicating a stable, though limited, influence. HLS 1DCNN results were small in magnitude and relatively stable. Overall, NDWI, BSI, and NDVI emerged as the most influential indices. Their impact was more heterogeneous under Sentinel-2 1DCNN, whereas RF offered smaller but steadier gains. Figure 13 shows the performance variation at L2 with and without NDWI.

At Level 3A (apple orchards under sunlight, black anti-hail nets, or white nets), the highest marginal gains were associated with vegetation vigor indices. In HLS 1DCNN models, EVI (+0.22), NDVI (+0.22), and SAVI (+0.21) stood out, followed by BSI (+0.12) and NDWI (+0.11). Sentinel-2 1DCNN presented a similar pattern, with NDVI (+0.18), SAVI (+0.17), and EVI (+0.14) ranking highest, while BSI (+0.11) and NDWI (+0.08) contributed moderately. In contrast, RF models produced smaller but more stable effects, with NDVI, SAVI, and EVI leading on both sensors (≈+0.04–0.08). Overall, vegetation indices (NDVI, SAVI, EVI) emerged as the most informative for distinguishing orchard management types, with 1DCNN capturing stronger but more variable gains than RF. Figure 14 shows the impact of NDVI on L3A model performance.

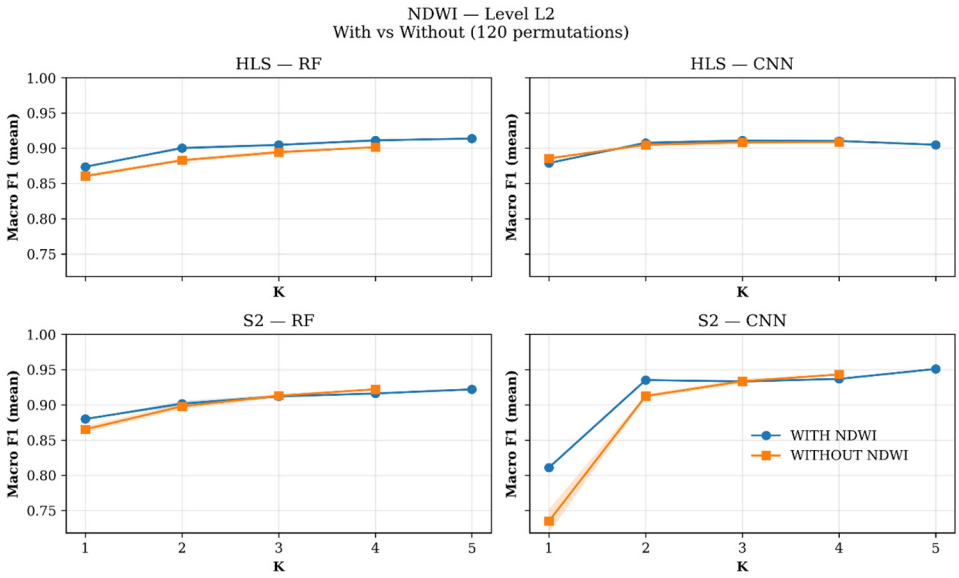

Figure 13. Impact of the Normalized Difference Water Index (NDWI) at Level 2 (Perennial Crops, Annual Crops, Grasslands) based on bootstrap analysis.

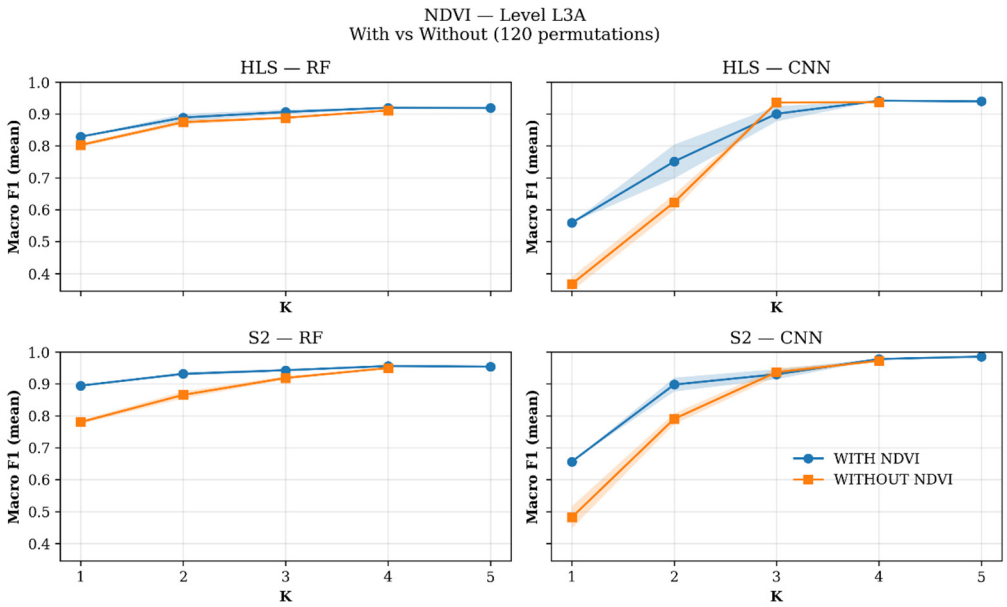

Figure 14. Impact of the Normalized Difference Vegetation Index (NDVI) on Level 3A classification (Apple under sunlight, under white anti-hail nets, and under black nets) based on bootstrap analysis.

At Level 3B (pastures vs. earlier successional vegetation), the marginal gain analysis indicated that SAVI (+0.16) and EVI (+0.16) under Sentinel-2 1DCNN were the strongest contributors, followed by BSI (+0.10). NDVI (+0.09) and NDWI (+0.07) also provided information, but with lower average effects. HLS 1DCNN presented similar trends but with smaller values (SAVI and EVI ≈+0.10), while Random Forest models across both sensors produced modest gains (≈+0.02). Variability was highest under 1DCNN, reflecting the spectral similarity between classes and the dependence on index combinations. Overall, vegetation vigor indices (SAVI and EVI) and soil exposure metrics (BSI) were the most influential in distinguishing pastures from early successional vegetation. However, their effects were less stable compared to those of other levels.

The results show that each classification level benefits from specific spectral indices aligned with surface characteristics. Broader levels (L1 and L2) with greater spectral

heterogeneity respond more strongly to indices sensitive to soil and moisture variation (e.g., NDWI and BSI). In contrast, finer levels (L3A and L3B) with more stable vegetation structures, such as perennial orchards and grasslands, depend more on vegetation indices (NDVI, SAVI, and EVI). This pattern underscores the complementary role of indices and the need for a balanced representation of surface properties. 1DCNN models generally delivered larger marginal gains but with higher variability, reflecting sensitivity to index composition. RF models, on the other hand, produced smaller yet steadier improvements, particularly at L3A and L3B. Sentinel-2 tended to yield higher gains with 1DCNN at the finer levels, whereas HLS provided more consistent results at the broader levels (L1 and L2), highlighting a trade-off between separability and robustness across models and sensors. Figure 15 shows the impact of SAVI on L3B model performance.

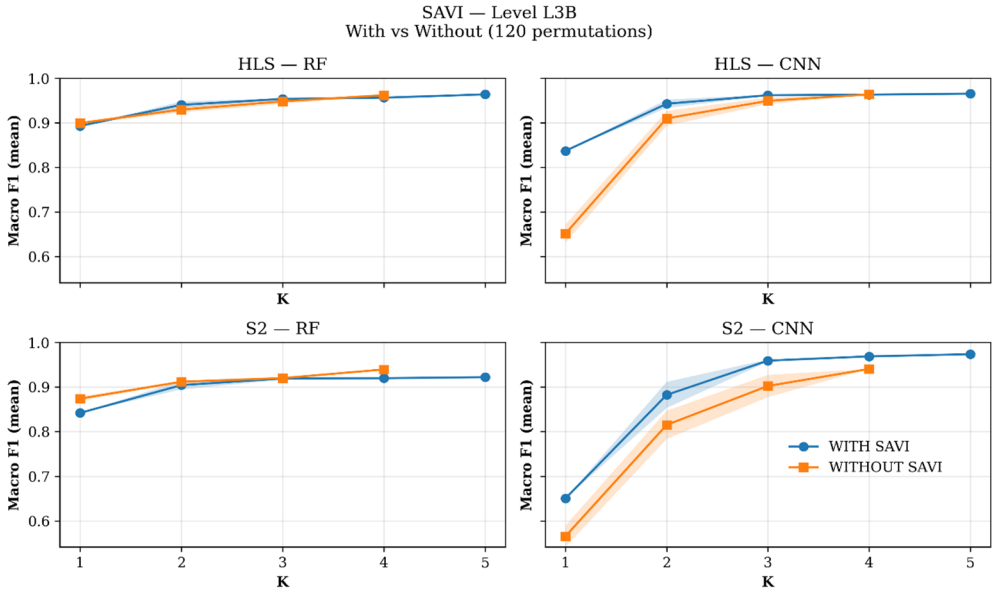

Figure 15. Impact of the Soil-Adjusted Vegetation Index (SAVI) at Level 3B (Pastures and Earlier Successional Vegetation) based on bootstrap analysis.

3.5. Impact of the Number of Indices on Classification Performance

Figure 16 summarizes the relationship between the number of spectral indices and model performance across all classification levels. The results reveal that, as additional indices are included, there is a consistent improvement in performance. However, the magnitude and stability of these gains are highly dependent on algorithm and dataset.

At L1 (Native Forest, Planted Forest, Other), performance significantly improves as more indices are introduced, indicating that broader-scale classifications benefit strongly from the combined representation of multiple surface properties. With a single index, results are highly variable (mean = 0.85, std = 0.12), especially with 1DCNN. Performance stabilizes when two indices are combined (mean = 0.93) and reaches saturation from three indices onward (mean $\approx$ 0.95, std $\leq$ 0.02). This pattern reveals an apparent diminishing-returns effect: the first additional indices, typically combining soil and moisture information, bring substantial gains, while further additions primarily reinforce model stability. The small variation from four to five indices indicates that L1 classifications achieve near-optimal performance when both soil (BSI) and water content (NDWI) conditions are adequately represented.

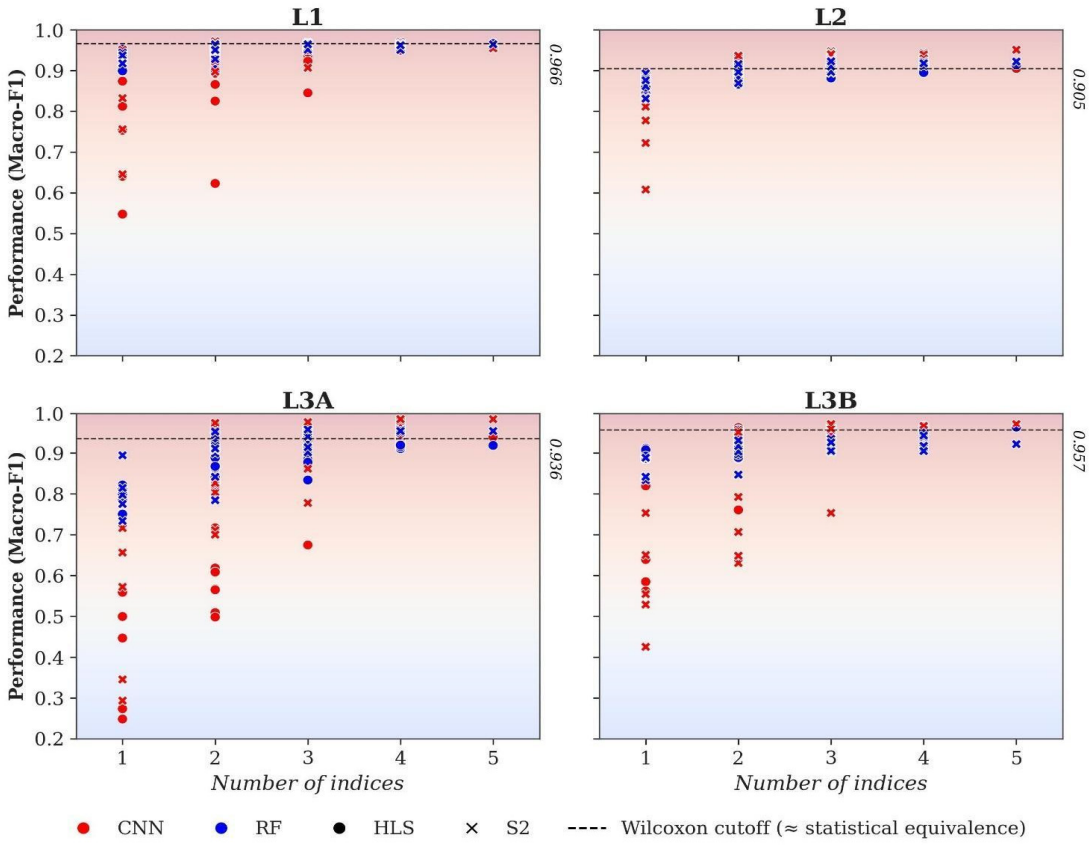

Figure 16. Performance of RF and 1DCNN using HLS and S2 as a function of the number of spectral indices in the feature space. The Wilcoxon cutoff line indicates the threshold above which there is no statistically significant difference between the models.

A similar pattern, albeit less pronounced, is observed for L2 (Perennial Crops, Annual Crops, Grasslands). Models using a single index achieve moderate accuracy (mean = 0.84, std = 0.07), highlighting the instability of isolated spectral features, especially with 1DCNN and S2. The inclusion of a second index substantially improves results (mean = 0.90) and drastically reduces variability, confirming the benefit of combining soil and moisture information at intermediate scales. From this point, gains are incremental: mean Macro-F1 rises from 0.91 (three indices) to 0.92 (five), while standard deviation remains below 0.02. This trend suggests that classifications approach saturation once indices capture both substrate reflectance (BSI) and vegetation moisture balance (NDWI), underscoring the complementary roles of spectral descriptors in transitional land-cover contexts.

At Level 3A (apple orchards under sunlight, black anti-hail nets, or white nets), performance shows a sharp, non-linear improvement as additional indices are added. Specifically, models relying on a single index display low accuracy and high variability (mean = 0.63, std = 0.21), indicating that apple production under anti-hail nets cannot be adequately separated with isolated indices. However, the inclusion of two or three indices improves performance (mean = 0.82–0.92), reduces variance, and captures complementary vegetation and soil features. Performance stabilizes from four indices onward (mean ≈ 0.95, std ≈ 0.02), indicating that once the models balance greenness (EVI, NDVI) with residual soil signal (SAVI), additional indices provide only marginal gains, as evidenced by the learning curves in Figure 17.

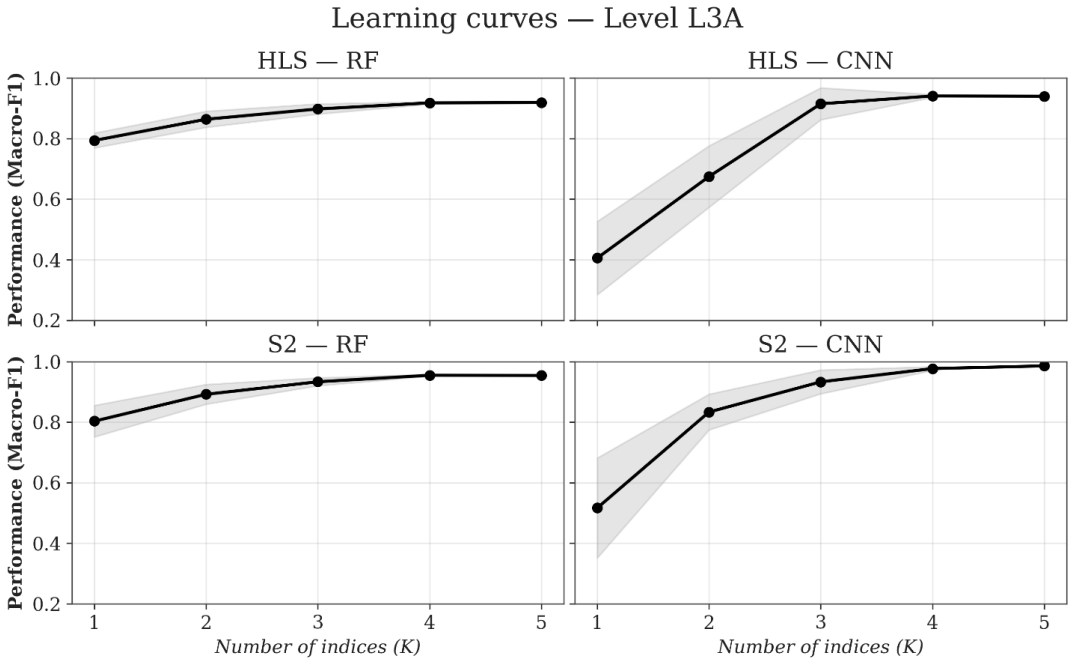

Figure 17. Learning performance across sensors and models for Level L3A, average ± 95% CI.

At Level 3B (pastures vs. earlier successional vegetation), the trend is similar to that in L3A, but progresses more incrementally. Variability is high with one or two indices (mean = 0.76–0.90, std < 0.16), reflecting the difficulty in differentiating vegetation vigor and density with limited spectral descriptors. Accuracy increases when the third index is added (0.94), plateauing afterwards, with a marked drop in standard deviation (≈0.02). This plateau suggests that PT-ESV discrimination becomes robust once multiple vegetation-sensitive indices (EVI, NDVI, SAVI) are incorporated, indicating that stable plant cover is best captured by the combined spectral signatures of chlorophyll activity and canopy density.

Overall, the results across all levels converge on a clear pattern: performance gains peak when indices capturing different surface properties, such as soil brightness (BSI), moisture (NDWI), and vegetation vigor (EVI, SAVI, and NDVI), are jointly included. Broader-scale problems (L1–L2) rely more on soil and moisture variability to define class boundaries, while finer mappings (L3A–L3B) increasingly depend on vegetation indices that capture the structural and physiological differences within stable covers. As shown in Figure 16, this transition from broad to fine spectral dependence highlights the complementary role of the indices and the need for a balanced feature composition to maximize both separability and robustness across classification levels.

3.6. Runtime Analysis and Computational Efficiency

In terms of computational efficiency, training time increased with the number of input indices, especially for 1DCNN models. RF remained highly stable, increasing from 1.00 ± 0.34 s with one index to 1.27 ± 0.64 s with five indices, an almost linear 27% rise. By contrast, 1DCNN exhibited much higher and more variable training costs, ranging from 22.6 ± 11.3 s (one index) to 24.3 ± 6.9 s (two indices), then stabilizing at around 23 s for three or more indices. Although the increase in runtime as more indices were added was modest, the difference between models remained large, with 1DCNNs being almost 22 times slower than RF on average. These findings confirm that while 1DCNNs benefit from additional indices, their computational overhead favors compact, complementary spectral sets that balance representational richness and efficiency across classification levels.

3.7. Visual Classification

The visual classification results were organized by hierarchical level (L1 to L3B) to support the interpretation of the quantitative metrics and assess the spatial consistency of model predictions (Figure 18).

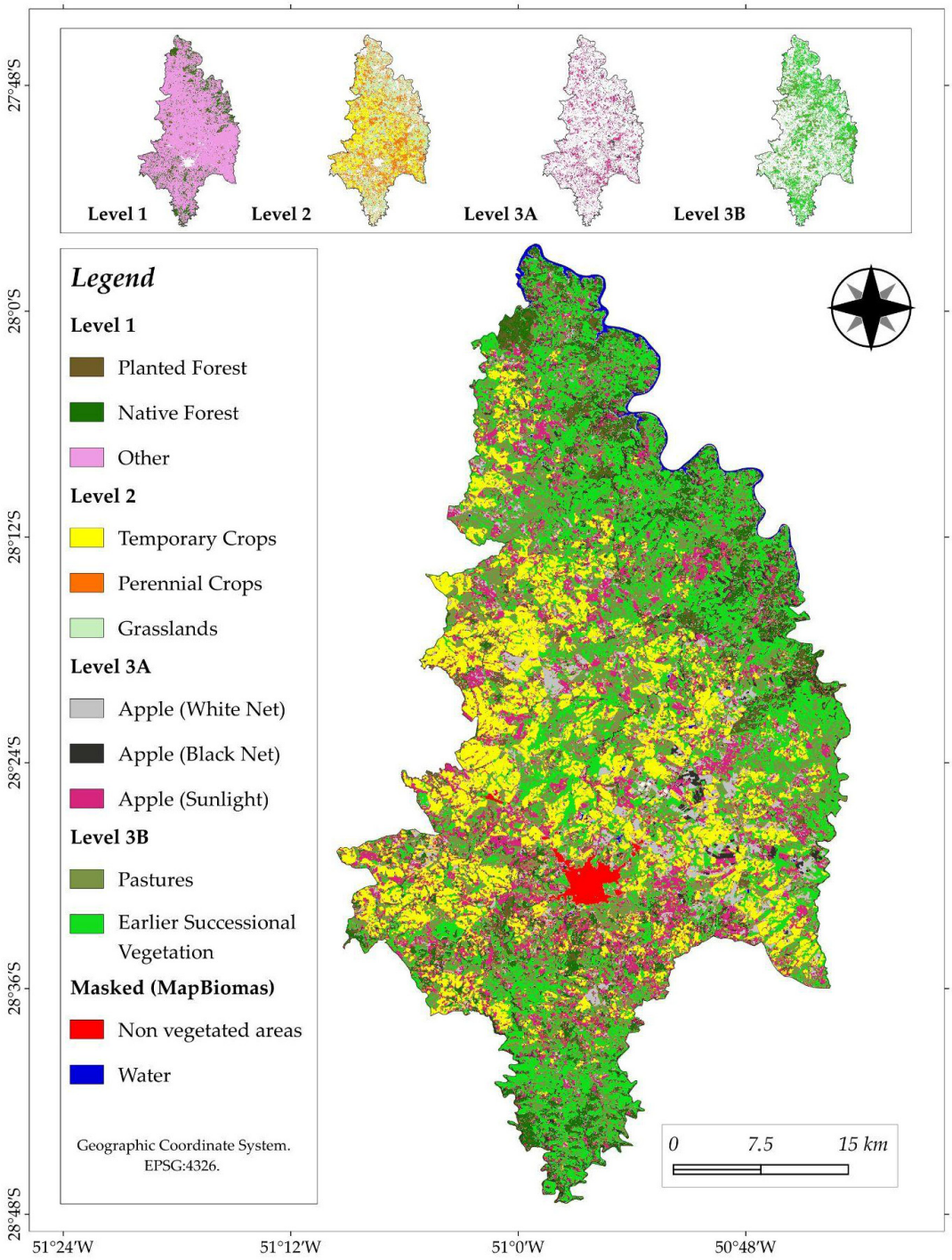

Figure 18. Hierarchical LULC classification maps (Levels L1–L3B) in Vacaria, generated using the best-performing HLS models. Each level refines the previous one, isolating apple orchards under white, black, and no nets, as well as other LULC classes.

The final classification maps for each level were generated using the best-performing HLS (Harmonized Landsat Sentinel-2) models. These models were selected based on macro-F1 scores from cross-validation. HLS was chosen because its performance matched the

Sentinel-2 (S2) models and offered higher computational efficiency for large-scale spatial prediction. At Level 1, the RF (BSI + NDVI + NDWI + SAVI) model distinguished planted forest (PF), native forest (NF), and other land uses (OT). Water and urban areas were masked using the 2024 MapBiomas Project [45]. The other class included categories analyzed in subsequent levels: perennial crops (PC), temporary crops (TC), and grasslands (GR).

Each subsequent level was constrained to the pixels of interest identified in the previous level, ensuring hierarchical consistency. At Level 2, the 1D CNN (BSI + NDVI + SAVI) segmented the OT class into PC, TC, and GR, from which only perennial (PC) pixels advanced to Level 3A. There, the 1D CNN (BSI + NDVI + SAVI) model further classified perennial areas into apple with white net (AWN), apple with black net (ABN), and apple under sunlight (AS) crops. For Level 3B, restricted to grasslands (GR) from Level 2, the 1D CNN (BSI + EVI + NDWI) model separated pastures (PA) from early successional vegetation (ESV). All outputs were refined using a 3 × 3 spatial mode filter, enhancing spectro-temporal coherence and reducing residual noise.

Figure 19 presents examples of the classification of black and white anti-hail nets using Sentinel-2 imagery in Vacaria. To improve the visualization of the examples, a spatial filter was applied to remove all classified patches smaller than 1 hectare. This filtering step helps highlight the main orchard areas under anti-hail net systems and facilitates the visual interpretation of the classification results. From the figure, it is also possible to observe that the model misclassified some humid areas associated with other crops as apple orchards. This limitation may be related to spectral similarities and should be further investigated in future studies.

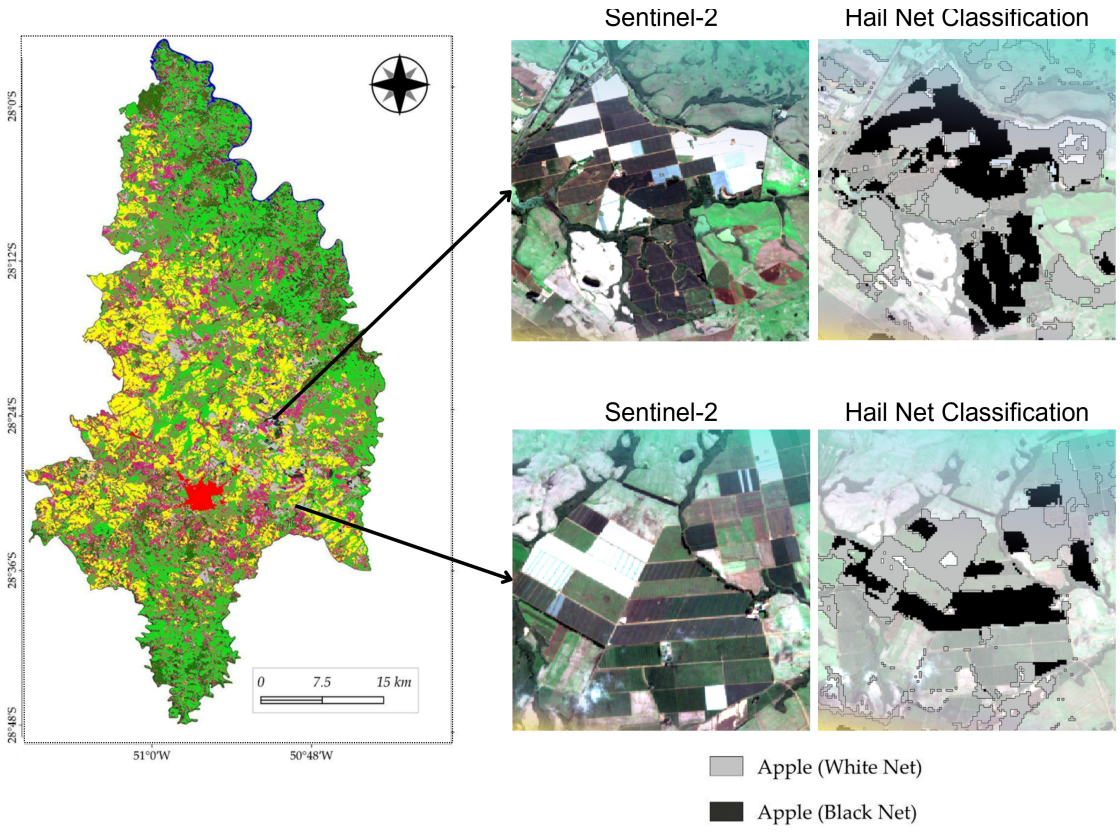

Figure 19. Example of visual classification map showing apple orchards under white and black anti-hail nets in Vacaria. The figure shows two examples with Sentinel-2 images from 2024 (**left**) and the respective classifications obtained for white and black nets (**right**).

4. Discussion

4.1. Hail Impact and Anti-Hail Net Systems in Apple Orchards

Hailstorms represent one of the most severe abiotic hazards to apple production, causing substantial yield and economic losses due to fruit bruising, defoliation, and branch damage, opening the door to diseases [6,46]. In regions frequently affected by convective storms, such as southern Brazil and parts of central Europe, the adoption of anti-hail nets has become the most effective strategy to ensure crop protection and production stability [13,47]. These structures can physically shield canopies from hail impact, drastically reducing fruit losses and improving orchard resilience to increasingly variable climatic conditions [48].

Beyond their protective role, anti-hail nets induce microclimatic and physiological adjustments in the orchard environment. The interception of direct solar radiation modifies the balance between photosynthetically active radiation (PAR) and diffuse light, typically reducing incident PAR by 20–40% depending on the net color and mesh density [46]. Although this reduction might suggest a limitation in photosynthesis, several studies indicate compensatory effects through improved light distribution and higher efficiency in canopy photosynthetic use [13]. Additionally, nets alter air temperature and humidity dynamics, creating more stable conditions that can reduce plant stress and evapotranspiration [6,49].

The color and optical properties of the nets also play a critical role in fruit quality. Photoselective nets, particularly white or pearl-colored ones, transmit a broader spectrum of diffuse light, enhancing fruit size, firmness, and biochemical composition, while dark or black nets tend to reduce sunburn and excessive heating [6,47]. These variations demonstrate that anti-hail nets are not merely protective structures, but agro-technical components that actively modulate orchard microclimate and radiation quality to influence fruit physiology, color development, and post-harvest attributes [46,48]. The most frequently used hail nets around the world remain white and black [20].

From an agronomic perspective, the expansion of net-covered orchards has brought both productivity gains and new challenges. In regions like Vacaria, the increase in areas under protective structures contributes to more consistent yields and improved fruit classification for export markets [47,49]. However, this widespread use also introduces environment and management implications, including changes in orchard energy balance, pollinator activity, and soil microclimate [6,46]. Consequently, monitoring the spatial extent, spectral properties, and long-term effects of anti-hail nets has become essential for evaluating both their agronomic benefits and their environmental footprint.

4.2. Remote Sensing for Classification of Net-Covered Orchards

In this context, advanced monitoring approaches become essential, especially considering that the visual similarity between orchards with and without nets makes traditional field surveys labor-intensive and subjective. The integration of remote sensing, spectral indices, machine learning and deep learning provides a scalable means to map net coverage, analyze its spatial distribution, and assess its impact on canopy reflectance and phenological patterns. Understanding these effects is crucial for optimizing orchard management strategies and for supporting digital agriculture frameworks that balance productivity, fruit quality, and sustainability.

Across all levels, differences in spatial and temporal resolution between HLS and Sentinel-2 shaped classification performance in complementary ways. With 10 m pixels, Sentinel-2 consistently improved separability in tasks that depend on fine-scale canopy structure and management effects, such as differentiating anti-hail net systems and their color variants (Level 3A), and perennial from annual and grassland classes (Level 2). In these settings, 1DCNN with Sentinel-2 achieved a higher macro-F1 than HLS, indicating

that finer spatial detail reduces mixed-pixel effects and enhances the signal captured by spectral-index time series. Conversely, the higher revisit frequency of HLS (~3-day composite of L30 and S30) benefited temporally driven tasks (e.g., Level 3B) and, with algorithms less sensitive to spatial texture (Random Forest), delivered comparable or superior results to Sentinel-2 (e.g., +2.2 pp macro-F1 at Level 3B). At the broadest level (Level 1), performances were similar, with small advantages alternating between datasets depending on the algorithm, suggesting that spatial detail (S2) and temporal density (HLS) can compensate for each other when class contrasts are strong. Overall, these results highlight a trade-off: Sentinel-2's spatial resolution favors net detection and color characterization, whereas HLS's temporal density can enhance separability where phenological trajectories dominate. These findings support a multisensor strategy that matches spatial and temporal resolutions to the target classes and algorithm choice (1DCNN vs. RF).

The observed trade-offs between Sentinel-2 and HLS align with reports on other perennial systems. Sentinel-2, with its 10 m resolution, has been shown to outperform coarser-resolution sensors in tasks requiring fine-scale crop delineation, particularly in orchards and vineyards where canopy management practices strongly influence reflectance [9,17]. In contrast, HLS has been consistently associated with improvements in capturing temporal dynamics, such as crop development stages or stress responses, due to its denser revisit frequency [14,15]. The results confirm this complementarity, as the two datasets emphasize different aspects of class separability: Sentinel-2 highlights structural detail, whereas HLS captures phenological continuity.

The behavior of vegetation reinforces these findings. At broader levels (e.g., L1 and L2), indices sensitive to soil and moisture (BSI, NDWI) benefited more from HLS's temporal density, which stabilizes seasonal trajectories and reduces noise during transitional phases. Conversely, indices capturing canopy structure and density (EVI, SAVI) were more informative with Sentinel-2, whose finer spatial resolution mitigates the mixed-pixel effects caused by heterogeneous orchard layouts and netting. Thus, the relevance of specific indices is not intrinsic but conditioned by the dataset's spatial–temporal resolution, highlighting the importance of sensor–index interactions in perennial crop monitoring.

4.3. One-Dimensional Convolutional Neural Network Performance and Limitations in Remote Sensing

These findings also point to broader implications for the remote sensing of perennial crops under protective structures. The ability to detect apple production under anti-hail net systems, particularly distinguishing between black and white nets, was strongly dependent on indices that adjust for canopy density and soil background (EVI, SAVI). These gains were most pronounced with Sentinel-2. This indicates that protective management can alter the relative importance of indices, shifting the balance from traditional greenness-based measures (NDVI) toward indices that account for structural and radiative modifications. For the monitoring of perennial crops in general, combining Sentinel-2 and HLS in multisensor approaches may yield more robust outcomes, with the former capturing fine-scale management signatures and the latter enhancing phenological stability. This strategy aligns with recent calls for integrating spatial and temporal detail into crop monitoring frameworks, particularly for high-value fruit crops where management practices directly affect both productivity and detectability.

1DCNN delivered consistent results across all tested configurations without extensive tuning. However, its performance remains sensitive to key hyperparameters, as variations in dropout rates, kernel sizes, and hidden layer dimensions can substantially affect accuracy and stability [42]. Another limitation lies in the intrinsic nature of convolutional architectures: while 1DCNNs efficiently capture short- to mid-term temporal dependencies, they are less capable of modeling long-range or season-spanning relationships, which

may reduce transferability across years or cropping cycles. Furthermore, the model relies heavily on input preprocessing, such as gap-filling, smoothing, and temporal alignments, since irregular or noisy time series can degrade feature extraction quality. Although the use of batch normalization and dropout improves robustness, interpretability remains limited compared to tree-based models, as the learned filters represent abstract temporal patterns rather than directly observable variables [50]. Thus, 1DCNN's performance depends on careful parameter calibration, consistent preprocessing, and validation across independent datasets. Caution is required when generalizing results to other regions, sensors, or temporal settings. Despite its strong performance, the 1DCNN model lacks inherent interpretability compared to Random Forest, particularly regarding the contribution of individual spectral indices over time. Future studies should explore explainability techniques, such as SHAP (SHapley Additive exPlanations), to better understand the temporal dynamics learned by deep learning models.

As illustrated in Figures 16 and 17, although the 1DCNN occasionally surpassed RF in specific configurations, its performance strongly depended on the number of indices used. The learning curve revealed that RF maintained stable performance, with macro-F1 scores above 0.70 starting from a single index and gradually improving as more indices were added. In contrast, 1DCNN began with lower macro-F1 values below 0.30 at Level 3A, and only achieved superior results to RF when two or more indices were combined (Figure 16). This pattern suggests that while 1DCNN benefits from richer spectral inputs, its performance is more sensitive to feature dimensionality and training configuration. Moreover, RF achieved competitive results at a lower computational cost and shorter training times, reinforcing its suitability as a practical and efficient algorithm that remains a strong option for solving problems in agriculture [51–53]. Future studies should evaluate the spatial transferability of the proposed models through cross-region validation in adjacent apple-producing areas.

5. Conclusions

This study addresses a critical research gap: to our knowledge, no previous work has applied artificial intelligence and remote sensing to map anti-hail net systems in apple orchards. By combining multisensor spectral index time series with machine and deep learning, we demonstrated the feasibility of automatically detecting orchards under different netting conditions. The results showed that RF outperformed 1DCNN, particularly at finer classification levels, achieving significant improvements in distinguishing canopy management practices. Sentinel-2 provided advantages for fine-scale detection, while the higher temporal density of HLS enhanced temporal stability and class separability at broader classification levels. Together, the complementary characteristics of both datasets provide valuable insights for designing monitoring strategies in perennial cropping systems, especially for planning apple orchard mapping and management programs.

The analysis also revealed that soil and water related indices (BSI, NDWI) were more influential at general classification levels, whereas vegetation vigor indices (NDVI, EVI, SAVI) were decisive for finer-scale orchard differentiation. These findings emphasize that the relevance of each index depends on both spatial and temporal resolution, offering practical guidance for selecting spectral indices according to sensor characteristics and study scale in future studies.

By integrating field and remote samples, a hierarchical classification framework was established, offering a robust and replicable methodological pathway for perennial crop monitoring under protective structures. However, the performance of deep learning models such as 1DCNN remains highly dependent on proper hyperparameter tuning and preprocessing, which must be carefully optimized to ensure model generalization.

The results demonstrate the potential of the proposed framework for consistent mapping of apple orchards under anti-hail nets, supporting long-term monitoring of management practices in heterogeneous agricultural landscapes. This framework can also guide the development of digital tools for orchard management and decision-making, particularly benefiting small and medium-sized producers. Future research should expand on these findings by testing additional machine and deep learning algorithms, integrating complementary satellite, UAV, and ground-based imagery, and extending the analysis to orchards under other hail net colors. Additionally, future studies should investigate the seasonal variability in spectral responses across different phenological stages of apple orchards, as well as incorporate meteorological variables as auxiliary features, which may further improve the understanding and classification of net-covered systems. Such advances will broaden the applicability and scalability of the proposed framework for perennial crop monitoring.

Author Contributions: Conceptualization, D.E.G.F., É.L.B., T.C.P. and V.B.S.; methodology, D.E.G.F., T.C.P. and V.B.S.; software, D.E.G.F., T.C.P. and V.B.S.; validation, D.E.G.F., T.C.P. and V.B.S.; formal analysis, D.E.G.F., T.C.P. and V.B.S.; investigation, D.E.G.F., É.L.B., T.C.P. and V.B.S.; resources, D.E.G.F., É.L.B., T.C.P., V.B.S., J.G.A.B., T.T.S. and L.G.; data curation, D.E.G.F., É.L.B., T.C.P., V.B.S. and L.G.; writing—original draft preparation, D.E.G.F., T.C.P. and V.B.S.; writing—review and editing, D.E.G.F., É.L.B., T.C.P., V.B.S., F.d.S., J.G.A.B., T.T.S. and L.G.; visualization, D.E.G.F., T.C.P. and V.B.S.; supervision, É.L.B.; project administration, É.L.B.; funding acquisition, D.E.G.F., É.L.B., T.C.P., V.B.S., F.d.S., J.G.A.B., T.T.S. and L.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the São Paulo Research Foundation (FAPESP), grant numbers #2022/09319-9, #2024/05205-4 (D.E.G.F.), #2024/13150-5 (T.C.P.), #2025/01750-0 (V.B.S.), and #2023/12215-3 (F.d.S.); and National Council for Scientific and Technological Development (CNPq) -Research Productivity Fellowships, #302963/2025-1 (É.L.B.).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank Nicolas Brandt (Emater, Vacaria, Rio Grande do Sul) for his valuable collaboration and assistance.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

1DCNN	One-Dimensional Convolutional Neural Network
ABN	Apple Black Net
AC	Annual Crops
AS	Apple Production under Sunlight
AWN	Apple White Net
BSI	Bare Soil Index
CCD-AD	Center of Science for Development in Digital Agriculture
Cfb	Temperate oceanic climate (Köppen climate classification)
ESV	Earlier Successional Vegetation
EVI	Enhanced Vegetation Index
GR	Grasslands
HLS	Harmonized Landsat and Sentinel-2
L1	Level 1
L2	Level 2
L3A	Level 3A

L3B	Level 3B
LULC	Land Use Land Cover
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NF	Native Forest
NIR	Near-Infrared band
OT	other LULC
PAM	Municipal Agricultural Production
PAR	Photosynthetically Active Radiation
PC	Perennial Crops
PF	Planted Forest
pp	percentage points
PT	Pasturelands
RF	Random Forest
S2	Sentinel-2
SAVI	Soil-Adjusted Vegetation Index
SITS	Spectral Index Time Series
SWIR	Shortwave Infrared band
SWIR1	Shortwave Infrared band 1

References

1. Bacelar, E.; Pinto, T.; Anjos, R.; Morais, M.C.; Oliveira, I.; Vilela, A.; Cosme, F. Impacts of Climate Change and Mitigation Strategies for Some Abiotic and Biotic Constraints Influencing Fruit Growth and Quality. *Plants* **2024**, *13*, 1942. [CrossRef]
2. Chen, R.; Wang, J.; Li, Y.; Bai, R.; Huang, M.; Zhang, Z.; Zhao, L.; Qu, Z.; Liu, L. Higher Risk of Spring Frost under Future Climate Change across China's Apple Planting Regions. *Eur. J. Agron.* **2024**, *159*, 127288. [CrossRef]
3. Kumar, A.; Negi, M.; Joshi, Y.; Dangi, G.; Sharma, D.P.; Sharma, K.C. Anti-Hail Nets under Hailstorm Incidence: Impact on Apple Orchard Dynamics. *N. Z. J. Crop Hortic. Sci.* **2024**, *53*, 1308–1328. [CrossRef]
4. Guo, J.; Guo, Y.; Tong, P.; Wang, X.; Wang, J. Effects of Different Coverage Years of Hail-Proof Nets on Environment, Leaf Traits and Fruit Quality in Apple Orchards. *Horticulturae* **2025**, *11*, 198. [CrossRef]
5. Szabó, A.; Tamás, J.; Nagy, A. The Influence of Hail Net on the Water Balance and Leaf Pigment Content of Apple Orchards. *Sci. Hortic.* **2021**, *283*, 110112. [CrossRef]
6. Mir, M.A.; Verma, P.; Sharma, N.C.; Sharma, N.; Sarma, U. Apple (*Malus* × *domestica* Borkh.) Production and Quality in Response to Anti-Hail Nets. *Int. J. Biometeorol.* **2024**, *68*, 927–938. [CrossRef]
7. Ali, A.M.; Abouelghar, M.; Belal, A.A.; Saleh, N.; Yones, M.; Selim, A.I.; Amin, M.E.S.; Elwesemy, A.; Kucher, D.E.; Maginan, S.; et al. Crop Yield Prediction Using Multi-Sensors Remote Sensing. *Egypt. J. Remote Sens. Space Sci.* **2022**, *25*, 711–716. [CrossRef]
8. Joshi, A.; Pradhan, B.; Gite, S.; Chakraborty, S. Remote-Sensing Data and Deep-Learning Techniques in Crop Mapping and Yield Prediction: A Systematic Review. *Remote Sens.* **2023**, *15*, 2014. [CrossRef]
9. Trentin, C.; Ampatzidis, Y.; Lacerda, C.; Shiratsuchi, L. Tree Crop Yield Estimation and Prediction Using Remote Sensing and Machine Learning: A Systematic Review. *Smart Agric. Technol.* **2024**, *9*, 100556. [CrossRef]
10. Muruganantham, P.; Wibowo, S.; Grandhi, S.; Samrat, N.H.; Islam, N. A Systematic Literature Review on Crop Yield Prediction with Deep Learning and Remote Sensing. *Remote Sens.* **2022**, *14*, 1990. [CrossRef]
11. Vidican, R.; Mălinaș, A.; Ranta, O.; Moldovan, C.; Marian, O.; Ghețe, A.; Ghișe, C.R.; Popovici, F.; Cătunescu, G.M. Using Remote Sensing Vegetation Indices for the Discrimination and Monitoring of Agricultural Crops: A Critical Review. *Agronomy* **2023**, *13*, 3040. [CrossRef]
12. Gottuk, J.; Stuenzi, S.M.; Runge, A.; Boike, J. Assessing Midsummer Snow-Free Land Surface Albedo Variability across Multiple Arctic Sites Using the Harmonized Landsat and Sentinel-2 Product. *Sci. Remote Sens.* **2025**, *12*, 100283. [CrossRef]
13. Mupambi, G.; Anthony, B.M.; Layne, D.R.; Musacchi, S.; Serra, S.; Schmidt, T.; Kalcsits, L.A. The Influence of Protective Netting on Tree Physiology and Fruit Quality of Apple: A Review. *Sci. Hortic.* **2018**, *236*, 60–72. [CrossRef]
14. Parreiras, T.C.; Santos, C.D.O.; Bolfe, É.L.; Sano, E.E.; Leandro, V.B.S.; Bayma, G.; Silva, L.A.P.; Furuya, D.E.G.; Romani, L.A.S.; Morton, D. Dense Time Series of Harmonized Landsat Sentinel-2 and Ensemble Machine Learning to Map Coffee Production Stages. *Remote Sens.* **2025**, *17*, 3168. [CrossRef]
15. Liu, X.; Xie, S.; Yang, J.; Sun, L.; Liu, L.; Zhang, Q.; Yang, C. Comparisons between Temporal Statistical Metrics, Time Series Stacks and Phenological Features Derived from NASA Harmonized Landsat Sentinel-2 Data for Crop Type Mapping. *Comput. Electron. Agric.* **2023**, *211*, 108015. [CrossRef]

16. Lastovicka, J.; Svec, P.; Paluba, D.; Kobliuk, N.; Svoboda, J.; Hladky, R.; Stych, P. Sentinel-2 Data in an Evaluation of the Impact of the Disturbances on Forest Vegetation. *Remote Sens.* **2020**, *12*, 1914. [CrossRef]
17. Misra, G.; Cawkwell, F.; Wingler, A. Status of Phenological Research Using Sentinel-2 Data: A Review. *Remote Sens.* **2020**, *12*, 2760. [CrossRef]
18. Uyar, N. Index-Driven Soil Loss Mapping across Environmental Scenarios: Insights from a Remote Sensing Approach. *Sustainability* **2025**, *17*, 7913. [CrossRef]
19. Zeng, Y.; Hao, D.; Huete, A.; Dechant, B.; Berry, J.; Chen, J.M.; Joiner, J.; Frankenberg, C.; Bond-Lamberty, B.; Ryu, Y.; et al. Optical Vegetation Indices for Monitoring Terrestrial Ecosystems Globally. *Nat. Rev. Earth Environ.* **2022**, *3*, 477–493. [CrossRef]
20. Furuya, D.E.G.; Bolfe, É.L.; da Silveira, F.; Barbedo, J.G.A.; da Silva, T.L.; Romani, L.A.S.; Castanheiro, L.F.; Gebler, L. Hail Netting in Apple Orchards: Current Knowledge, Research Gaps, and Perspectives for Digital Agriculture. *Climate* **2025**, *13*, 203. [CrossRef]
21. Silva, V.F.; Gebler, L.; Dompieri, M.H.G. Mapeamento de Telas Antigranizo em Pomares de Maçã em Vacaria-RS. In Proceedings of the 19th Congresso Interinstitucional de Iniciação Científica (CIIC 2025), Bento Gonçalves, Brazil, 26–28 August 2024.
22. Furuya, D.E.G.; Bolfe, É.L.; Parreiras, T.C.; Barbedo, J.G.A.; Santos, T.T.; Gebler, L. Combination of Remote Sensing and Artificial Intelligence in Fruit Growing: Progress, Challenges, and Potential Applications. *Remote Sens.* **2024**, *16*, 4805. [CrossRef]
23. IBGE (Instituto Brasileiro de Geografia e Estatística). Produção de Maçã. 2025. Available online: https://www.ibge.gov.br/explica/producao-agropecuaria/maca/rs (accessed on 26 October 2025).
24. Righi, E.; Putti, G.L.; Adams, A.; Sanhueza, R.M.V.; Monteiro, R. I Congresso de Inovação e Tecnologia da Maçã (CITEM). *Agapomi* **2024**, *359*, 8–9. Available online: http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1163933 (accessed on 18 January 2025).
25. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A.; et al. The Climate Hazards Infrared Precipitation with Stations (CHIRPS)—A New Environmental Record for Monitoring Extremes. *Sci. Data* **2015**, *2*, 150066. [CrossRef]
26. Copernicus Climate Change Service (C3S). ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate. 2017. Available online: https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview (accessed on 26 October 2025).
27. Silva, G.B.S.; Bolfe, E.L.; Gava, G.J.C.; Moraes, J.F.L.; Silveira, J.M.C.; De Maria, I.C.; Pires, R.C.M.; Sano, E.E.; Parreiras, T.C.; Santos, C.O.; et al. *Diagnóstico Agroambiental dos Municípios com os Distritos Agrotecnológicos do Projeto Semear Digital*; Embrapa Meio Ambiente: Jaguariúna, Brazil, 2025; Volume 142, 53p.
28. Santos, H.G.; Jacomine, P.K.T.; Anjos, L.H.C.; Oliveira, V.A.; Lumbreras, J.F.; Coelho, M.R.; Almeida, J.A.; Araujo Filho, J.C.; Oliveira, J.B.; Cunha, T.J.F. *Sistema Brasileiro de Classificação de Solos*, 5th ed.; Embrapa: Brasília, Brazil, 2018.
29. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. *Monitoring Vegetation Systems in the Great Plains with ERTS. NASA Tech. Rep.*; NASA: Washington, DC, USA, 1974. Available online: https://ntrs.nasa.gov/citations/19740022614 (accessed on 5 November 2025).
30. Gao, X.; Huete, A.R.; Ni, W.; Miura, T. Optical–Biophysical Relationships of Vegetation Spectra without Background Contamination. *Remote Sens. Environ.* **2000**, *74*, 609–620. [CrossRef]
31. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
32. Gao, B.C. NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [CrossRef]
33. Rikimaru, A.; Roy, P.S.; Miyatake, S. Tropical Forest Cover Density Mapping. *Trop. Ecol.* **2002**, *43*, 39–47.
34. Ju, J.; Zhou, Q.; Freitag, B.; Roy, D.P.; Zhang, H.K.; Sridhar, M.; Neigh, C.S. The Harmonized Landsat and Sentinel-2 Version 2.0 Surface Reflectance Dataset. *Remote Sens. Environ.* **2025**, *324*, 114723. [CrossRef]
35. Spinelli-Araujo, L.; Parreiras, T.C.; Vicente, L.E.; Bolfe, É.L. Aplicabilidade e Eficácia de Tecnologias Digitais Móveis em Levantamentos de Campo para Dados de Agropecuária. In *Proceedings of the XX Simpósio Brasileiro de Sensoriamento Remoto (SBSR)*; Galoá Proceedings, 2023; Instituto Nacional de Pesquisas Espaciais (INPE): Florianópolis, Brazil, 2023. Available online: https://proceedings.science/p/164118?lang=pt-br (accessed on 26 October 2025).
36. Li, J.; Zhang, B.; Huang, X. A Hierarchical Category Structure Based Convolutional Recurrent Neural Network (HCS-ConvRNN) for Land-Cover Classification Using Dense MODIS Time-Series Data. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *108*, 102744. [CrossRef]
37. Parreiras, T.C.; Bolfe, E.L.; Chaves, M.E.D.; Sanches, I.D.A.; Sano, E.E.; Victoria, D.D.C.; Bettiol, G.M.; Vicente, L.E. Hierarchical Classification of Soybean in the Brazilian Savanna Based on Harmonized Landsat Sentinel Data. *Remote Sens.* **2022**, *14*, 3736. [CrossRef]
38. Demirkan, D.Ç.; Koz, A.; Düzgün, H.Ş. Hierarchical Classification of Sentinel-2A Images for Land Use and Land Cover Mapping and Its Use for the CORINE System. *J. Appl. Remote Sens.* **2020**, *14*, 026524. [CrossRef]
39. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
40. Biau, G.; Scornet, E. A Random Forest Guided Tour. *Test* **2016**, *25*, 197–227. [CrossRef]
41. Tufail, R.; Tassinari, P.; Torreggiani, D. Deep Learning Applications for Crop Mapping Using Multi-Temporal Sentinel-2 Data and Red-Edge Vegetation Indices: Integrating Convolutional and Recurrent Neural Networks. *Remote Sens.* **2025**, *17*, 3207. [CrossRef]

42. Pham, V.-D.; Tetteh, G.; Thiel, F.; Erasmi, S.; Schwieder, M.; Frantz, D.; van der Linden, S. Temporally Transferable Crop Mapping with Temporal Encoding and Deep Learning Augmentations. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *129*, 103867. [CrossRef]
43. Ma, Y.; Zhen, Z.; Li, F.; Feng, F.; Zhao, Y. An Innovative Lightweight CNN1D Model for Efficient Monitoring of Large-Scale Forest Composition: A Case Study of Heilongjiang Province, China. *GISci. Remote Sens.* **2023**, *60*, 2271246. [CrossRef]
44. Minaee, S.; Boykov, Y.; Porikli, F.; Plaza, A.; Kehtarnavaz, N.; Terzopoulos, D. Image Segmentation Using Deep Learning: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3523–3542. [CrossRef]
45. MapBiomas Project—Annual Mapping of Land Cover and Use in Brazil. Collection 9. Available online: https://mapbiomas.org (accessed on 20 October 2025).
46. Bosco, L.C.; Bergamaschi, H.; Marodin, G.A.B. Solar Radiation Effects on Growth, Anatomy, and Physiology of Apple Trees in a Temperate Climate of Brazil. *Int. J. Biometeorol.* **2020**, *64*, 1969–1980. [CrossRef]
47. Amarante, C.V.T.; Steffens, C.A.; Argenta, L.C. Yield and Fruit Quality of 'Gala' and 'Fuji' Apple Trees Protected by White Anti-Hail Net. *Sci. Hortic.* **2011**, *129*, 79–85. [CrossRef]
48. Blanke, M.M. Advances in the Sustainable Use of Plastics in Horticulture—Perspectives, Innovations, Opportunities, and Limitations. *Sustainability* **2023**, *15*, 11629. [CrossRef]
49. Manja, K.; Aoun, M. The Use of Nets for Tree Fruit Crops and Their Impact on the Production: A Review. *Sci. Hortic.* **2019**, *246*, 110–122. [CrossRef]
50. Xue, H.; Fan, Y.; Dong, G.; He, S.; Lian, Y.; Luan, W. Crop Classification in the Middle Reaches of the Hei River Based on Model Transfer. *Sci. Rep.* **2024**, *14*, 28964. [CrossRef]
51. Ramos, A.P.M.; Osco, L.P.; Furuya, D.E.G.; Gonçalves, W.N.; Santana, D.C.; Teodoro, L.P.R.; Silva Junior, C.A.; Capristo-Silva, G.F.; Li, J.; Baio, F.H.R.; et al. A Random Forest Ranking Approach to Predict Yield in Maize with UAV-Based Vegetation Spectral Indices. *Comput. Electron. Agric.* **2020**, *178*, 105791. [CrossRef]
52. Zhao, Z.; Li, Y.; Zhang, X.; Li, Z.; Sun, Y.; Gu, X.; Cai, H. A Meta-Learning Method for Recommending Random Forest Model to Simulate Land Use Efficiency in Maize–Soybean Intercropping. *Comput. Electron. Agric.* **2025**, *239*, 110961. [CrossRef]
53. Wei, P.; Ye, H.; Nie, C.; Zhang, Y.; Liu, R. Remote Sensing Estimation of Nitrogen Content in Scenes of Different Crop Types Based on the Random Forest Algorithm. *Comput. Electron. Agric.* **2025**, *231*, 109987. [CrossRef]