



Article

Forecasting Spring Wheat Maturity from UAV-Based Multispectral Imagery Using Machine and Deep Learning Models

Prabahar Ravichandran ¹, Keshav D. Singh ^{1,*}, Harpinder S. Randhawa ¹ and Shubham Subrot Panigrahi ^{1,2}

¹ Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada (AAFC), 5403 1st Avenue South, Lethbridge, AB T1J 4B1, Canada; prabahar.ravichandran@agr.gc.ca (P.R.); harpinder.randhawa@agr.gc.ca (H.S.R.); shubhampanigrahi69@gmail.com (S.S.P.)

² Department of Food Science, University of Arkansas Division of Agriculture, Fayetteville, AR 72704, USA

* Correspondence: keshav.singh@agr.gc.ca

Abstract

Accurate forecasting of crop maturity supports efficient harvest planning and accelerates selection decisions in breeding programs. In spring wheat, maturity is typically assessed through manual scoring late in the season, which limits its usefulness for timely harvest management and early selection decisions in breeding programs. This study evaluated uncrewed aerial vehicle (UAV)-based multispectral imagery for forecasting maturity in spring wheat grown at Lethbridge, Alberta (AB), Canada, during the 2024 and 2025 growing seasons. Thirty cultivars were monitored using seven-band UAV multispectral imagery during grain filling, enabling derivation of core vegetation and senescence-related indices from radiometrically calibrated orthomosaics. Strong correlations ($|r| > 0.85$) were observed between vegetation indices and days remaining to maturity (DRTM), motivating baseline regression models and subsequent evaluation of eleven machine-learning and deep-learning approaches. Among these, support vector regression (SVR) and multi-layer perceptron (MLP) achieved the highest predictive accuracy ($R^2 = 0.95$ – 0.96 ; mean absolute error (MAE) ≈ 1.25 days). Deep learning models achieved performance comparable to machine-learning approaches; however, incorporating spatial information through convolutional neural networks did not improve prediction accuracy. Feature-attribution analysis identified the red, red-edge (RE), and near-infrared (NIR) spectral bands as key predictors, enabling non-destructive, early, and scalable UAV-based maturity forecasting.

Keywords: UAV multispectral imaging; wheat maturity prediction; vegetation indices; machine learning; deep learning; days remaining to maturity



Academic Editor: Ray E. Sheriff

Received: 16 December 2025

Revised: 24 January 2026

Accepted: 3 February 2026

Published: 10 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Wheat is one of the most widely grown cereal crops globally and is a cornerstone of food security and agricultural economies [1]. In Canada, spring wheat is a major commodity crop and a key focus of breeding and agronomic research. High-latitude production regions such as the Canadian Prairies are characterized by short growing seasons and strong sensitivity of phenological development to temperature and moisture variability, particularly during grain filling [2]. These conditions motivate the development of timely, non-destructive approaches for forecasting wheat maturity under Canadian environments.

Accurate forecasting of phenological traits such as maturity is therefore critical for enhancing the efficiency of breeding programs and optimizing agronomic management in

cereal production systems. In spring wheat, maturity timing influences harvest scheduling, grain quality, and cultivar adaptation to local environmental conditions. However, traditional assessment of maturity is based on late-season manual visual scoring, a subjective and labor-intensive approach that lacks scalability and limits its usefulness for early-season decision-making [3]. These limitations underscore the need for rapid, objective, and scalable approaches for forecasting crop maturity. These challenges have motivated increasing interest in leveraging remote sensing technologies to enable earlier, objective, and scalable phenological assessment.

Recent advances in uncrewed aerial vehicle (UAV) technologies and high-resolution multispectral imaging have transformed crop monitoring by enabling non-destructive, high-throughput, and temporally dense observations of canopy dynamics [4–7]. Vegetation indices derived from multispectral bands capture physiological processes associated with photosynthetic activity, chlorophyll degradation, and canopy senescence, key indicators of maturity progression in cereals [8]. These remote sensing capabilities offer a promising alternative to traditional assessment methods by quantifying phenological transitions directly from canopy reflectance patterns. In the Canadian Prairies, where extensive wheat farms operate within narrow harvest windows, early prediction of physiological maturity holds substantial practical value. Accurate forecasts can improve harvest scheduling, grain hauling, and storage logistics while minimizing yield and quality losses from pre-harvest sprouting, frost damage, and reduced falling number [9]. Moreover, early maturity insights support proactive decision-making such as prioritizing vulnerable fields under impending adverse weather, optimizing irrigation or late-season fungicide applications, and planning crop rotations or soil amendments for subsequent seasons.

From a breeding perspective, maturity prediction during the grain-filling stage provides significant advantages. Early estimation enables breeders to classify genotypes into maturity groups well before final phenological scoring, thereby accelerating selection cycles and optimizing resource allocation in multi-environment trials [10]. In addition, UAV-based early assessments facilitate the characterization of genotype $\times$ environment ($G \times E$) interactions by revealing how environmental variability influences phenological development before maturity becomes visually apparent. Because maturity is closely associated with other key agronomic traits, including yield potential, drought tolerance, and lodging resistance, reliable early prediction serves as an integrative framework for balancing trade-offs in breeding and management decisions [11,12].

In contrast to these emerging imagery-based approaches, conventional phenological models primarily estimate maturity using thermal-time accumulation. While traditional thermal-time models rely solely on accumulated temperature (growing degree-days) to simulate phenological development [13–15], imagery-based approaches directly capture canopy-level physiological changes, providing more spatially resolved and integrative estimates of crop maturity and senescence patterns [16–18]. Several studies highlight the potential of UAV sensing for maturity-related traits. Romero and Lopes [19] demonstrated that RGB-derived vegetation indices explained 65% of the variance in heading and maturity predictions for bread wheat. Liu et al. [20] reported that multispectral UAV imaging could predict maize maturity through NDRE-based estimation of grain moisture ($R^2 \approx 0.6$), underscoring the predictive importance of red-edge senescence signals. Hassan et al. [4] further showed that multispectral indices effectively captured temporal senescence dynamics in wheat. However, most existing studies treat maturity as a categorical stage or rely on thermal-time assumptions, leaving a gap for continuous, physiology-based maturity forecasting. Recent integrative LiDAR–multispectral studies have demonstrated that multi-sensor UAV phenotyping can effectively resolve canopy

structural and physiological variation in legumes [21], supporting the broader applicability of imagery-driven maturity forecasting in breeding programs.

Traditional models that estimate days to maturity (DTM), including growing degree-day and beta-function approaches, simulate phenological progress using accumulated temperature and cultivar-specific parameters [13,22]. Although DTM is moderately heritable, its expression is strongly modulated by microclimatic variation and $G \times E$ interactions, limiting the transferability and reliability of thermal-time models across sites, seasons, and cultivars [10]. Temperature-based approaches such as those used by Pullens et al. [15] rely exclusively on heat-unit accumulation and fixed base temperatures, making them suitable for broad scheduling but unable to capture canopy physiological dynamics. Moreover, thermal-time formalisms are sensitive to confounding factors including irrigation effects, evaporative demand, temperature range, and measurement intervals, highlighting inherent limitations in predicting maturity solely from temperature metrics.

To address these challenges, the present study introduces days remaining to maturity (DRTM), a dynamic, image-derived indicator of canopy physiological status. Unlike thermal-time approaches that assume temperature-driven phenological progression, DRTM leverages UAV multispectral reflectance to quantify chlorophyll degradation, red-edge shifts, and canopy senescence. This physiology-based, plot-level measure enables real-time and spatially explicit forecasting of maturity that remains robust across cultivars, environments, and microclimatic conditions. DRTM thus represents a fundamentally different and more phenomics-aligned alternative to temperature-dependent harvest models.

Building on this framework, this study aimed to (i) evaluate the potential of UAV-based multispectral canopy reflectance to predict days remaining to physiological maturity (DRTM) in spring wheat, (ii) compare the performance of vegetation index-based linear models with full-spectrum machine-learning approaches, (iii) assess whether deep-learning models incorporating spatial information improve prediction accuracy relative to conventional machine-learning methods, and (iv) identify the most informative spectral bands contributing to early maturity forecasting through feature-attribution analyses.

2. Materials and Methods

2.1. Experimental Sites, Materials, and Design

Field experiments were conducted at the Fairfield farm experimental site (49°42′32.5″ N, 112°41′31.4″ W) of the Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada. Experiments in the 2024 and 2025 growing seasons were conducted in two adjacent fields located approximately 500 m apart, consistent with the farm's crop rotation system involving fallow and dry bean breeding programs, which is implemented annually to manage soil health and pest pressure. Heritage bread wheat panel consisting of thirty Canadian western spring wheat varieties were selected for evaluation. The experiments followed a randomized complete block design (RCBD) with three replications. Each plot measured 0.92 m × 3 m and was seeded at a rate of 300 seeds per square meter (Figure 1). The list of cultivars and their field arrangement for Range 1 in the 2025 trial is directly provided in the figure for reference.

Daily weather data were obtained from Environment and Climate Change Canada (ECCC) using the nearest long-term meteorological station (Lethbridge; Station ID 2265), located approximately 5 km from the experimental site. Daily records included maximum, minimum, and mean air temperature, precipitation, wind, and derived degree-day variables for the May–September growing period in 2024 and 2025. Missing temperature observations were linearly interpolated to ensure continuous daily time series, while missing precipitation values were assumed to be zero. The experimental plots were equipped with a linear irrigation system to ensure uniform water distribution across treatments;

accordingly, daily water input was calculated as the sum of measured precipitation and applied irrigation. The combined effects of temperature and total water input are shown in Figure 2, illustrating the seasonal weather dynamics observed at the Fairfield experimental site during the 2024 and 2025 growing seasons and highlighting interannual differences that contributed to distinct patterns of phenological development.

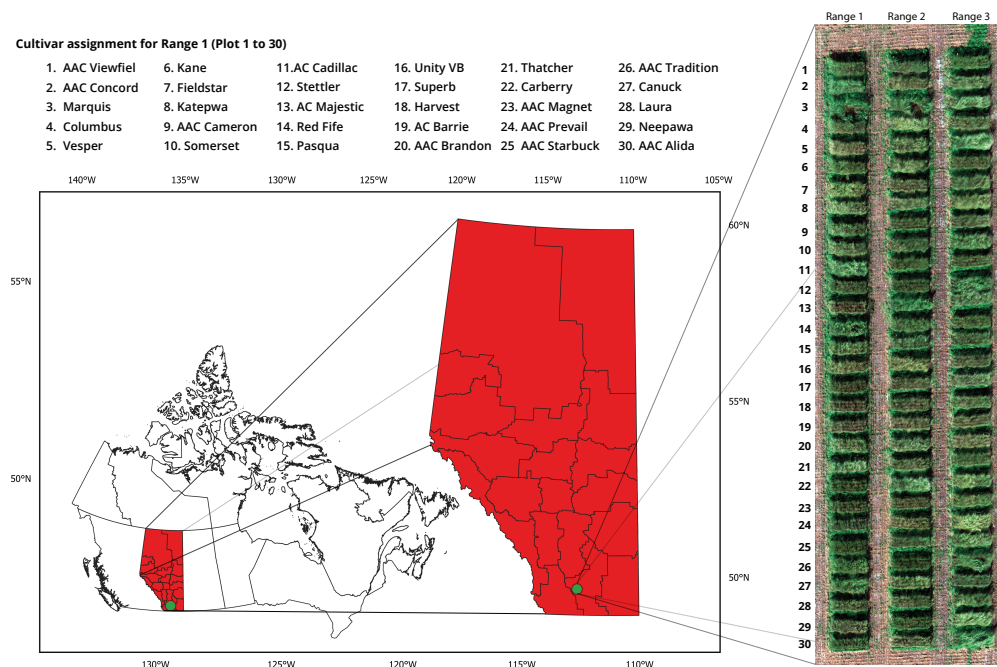


Figure 1. Geographical location of the experimental site at the Fairfield farm, Lethbridge Research and Development Centre, Alberta, Canada, shown at national and provincial scales, along with a representative UAV orthomosaic of the experimental plots. Cultivars planted in Range 1 of the 2025 trial are labeled in the figure for reference.

2.2. Phenological Measurements and Definition of Maturity Metrics

Physiological maturity was determined through field-based visual assessment for each experimental plot following standard spring wheat phenological criteria, corresponding to the completion of grain filling and loss of green tissue. Maturity dates were recorded at the plot level for all cultivars and replications across both growing seasons. Days to maturity (DTM) were defined as the number of days from seeding to the observed physiological maturity date for each plot.

For imagery-based modelling, days remaining to maturity (DRTM) was calculated as the temporal difference between the UAV image acquisition date and the corresponding plot-level maturity date. This formulation enables maturity progression to be represented as a continuous response variable, allowing UAV data acquired at different growth stages to be integrated into a unified modelling framework.

The maturity dates reported in Table 1 represent the median physiological maturity date across all cultivars and replications within each growing season and are provided for descriptive context of the UAV acquisition window only. All DRTM calculations and subsequent modelling analyses were performed using plot-specific maturity dates rather than seasonal median values, thereby preserving cultivar- and replication-level variability in maturity dynamics.

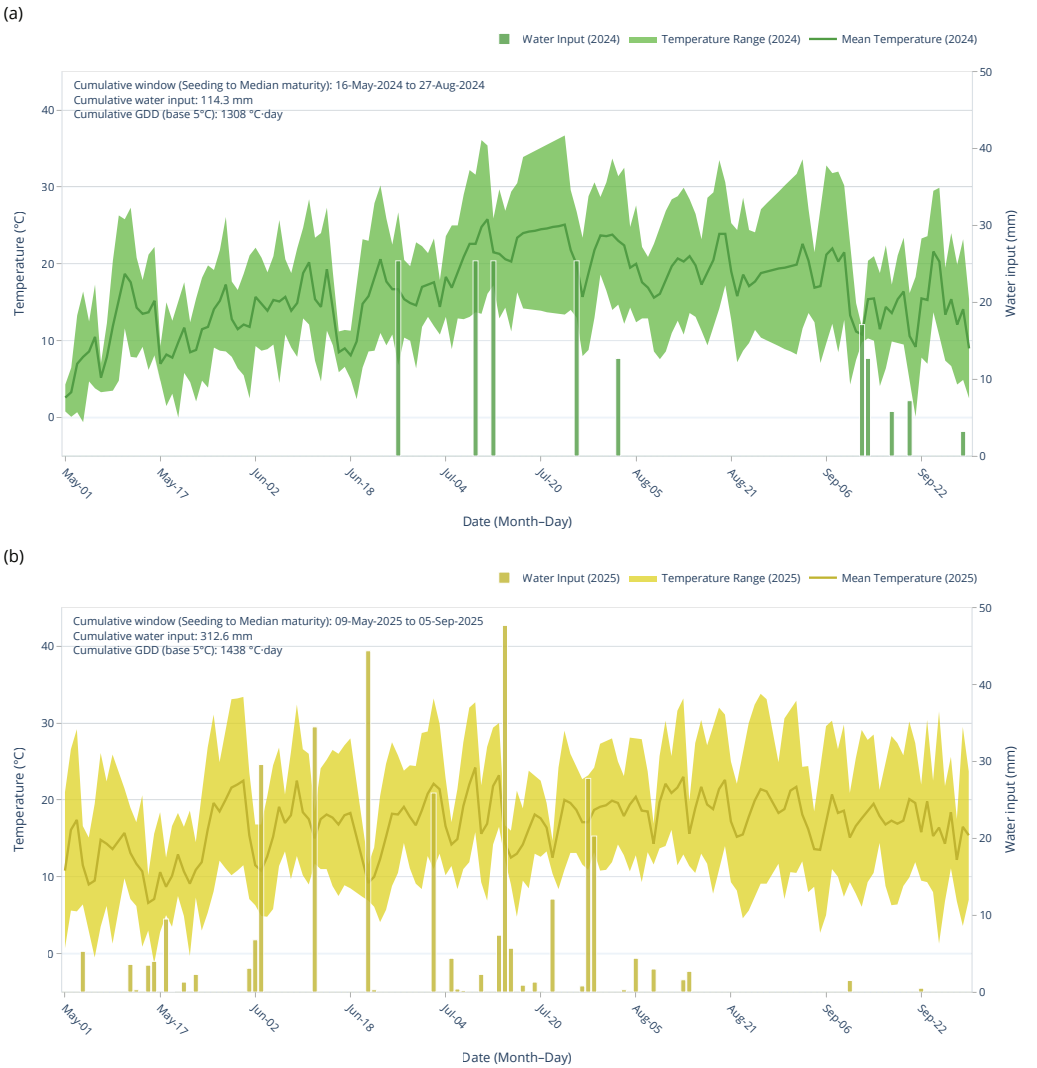


Figure 2. Daily weather trends at the Fairfield experimental site during the 2024 (a) and 2025 (b) growing seasons. Shaded bands show the daily air temperature range (maximum–minimum), solid lines indicate mean daily temperature, and bars represent daily water input (precipitation + irrigation) plotted on the secondary y-axis. Cumulative water input and cumulative growing degree days (base 5 °C) are reported for the seeding-to-median-maturity window shown in each panel.

Table 1. Summary of UAV acquisition dates and corresponding crop phenological stages for the 2024 and 2025 growing seasons.

Dataset	UAV Acquisition Date	Seeding Date	Maturity Date (Median)
1	2 August 2024	16 May 2024	27 August 2024
2	14 August 2024		
3	22 August 2024		
4	13 August 2025	9 May 2025	5 September 2025
5	21 August 2025		

Note: Dataset refers to individual UAV flight acquisitions. Seeding date indicates the planting date for each growing season. Maturity date (median) represents the median physiological maturity date across all cultivars within the experimental range for the corresponding year. Dates are presented in Day Month Year format.

2.3. Data Acquisition and Processing

The unmanned aerial vehicle (UAV) used in this study was a DJI Matrice 300 RTK system equipped with a MicaSense Altum-PT (AgEagle Aerial Systems Inc., Wichita, KS,

USA), high-resolution multispectral sensor, capable of simultaneously capturing RGB, thermal infrared, multispectral, and panchromatic imagery.

Flights were conducted at an altitude of 25 m above ground level with 3.5 m/s flight speed, maintaining 80% frontal and 85% side overlap to ensure sufficient image redundancy for photogrammetric reconstruction. The resulting orthomosaics had a ground sampling distance (GSD) of approximately 1.08 cm/pixel. Radiometric calibration was performed using MicaSense reflectance panels after each flight to ensure spectral consistency. The Downwelling Light Sensor 2 (DLS2) integrated with the Altum-PT was used to record incident light conditions during each mission, allowing for correction of illumination variability across flights. Aerial imagery was collected under sunny, low-wind conditions near solar noon at one- to two-week intervals during the grain-filling stage.

All UAV flight missions were conducted using the DJI RTK-2 system, providing high-precision onboard positioning during image acquisition. In addition, four ground control points (GCPs) were deployed near the corners of the experimental field and maintained consistently throughout the growing season. These GCPs were surveyed using RTK GNSS and incorporated during photogrammetric processing to support accurate co-registration of orthomosaics across acquisition dates in a year. The combined use of DJI RTK-2 positioning and strategically placed GCPs enabled reliable temporal alignment of orthomosaics, allowing the same plot boundary shapefiles (GeoJSON) to be reused for plot-level data extraction across all UAV missions. This approach ensured spatial consistency in reflectance sampling for time-series analysis of crop maturity progression.

Five UAV image datasets were collected overall for the model development and evaluation. Each dataset corresponds to a distinct acquisition date to capture variability in canopy appearance and maturity progression. The acquisition dates for all datasets are summarized in Table 1. The UAV platform used for data acquisition is illustrated in Figure 3, showing the DJI M300 RTK equipped with a Altum-PT multispectral camera and D-RTK 2 base station for high-resolution canopy imaging over the spring wheat field trial.



Figure 3. UAV system with the mounted multispectral camera (**left**) and a D-RTK 2 base station (**right**) for high-resolution canopy imaging over the spring wheat trial.

Image processing and data extraction were performed using Pix4Dmapper, Rasterio, QGIS, and GeoPandas. Orthomosaics were generated in Pix4Dmapper using radiometrically calibrated reflectance data, where white panel calibration values were entered prior to processing to enhance signal quality and reduce noise. The resulting outputs included seven

reflectance bands including, blue, green, red, red edge, near-infrared (NIR), panchromatic, and thermal exported as individual GeoTIFF files.

Using the Rasterio Python package, these individual bands were stacked to create a single multi-band composite GeoTIFF. Plot boundaries were digitized in QGIS to generate plot-level polygons, which were then exported as GeoJSON files. These polygons were re-imported into Python, and the mask function in Rasterio was used to extract plot-level reflectance values. Finally, GeoPandas was used to spatially join the extracted reflectance data with corresponding ground measurements to compute vegetation and thermal indices for subsequent statistical and modeling analyses.

2.4. Statistical Analysis and Modelling

The statistical and predictive analyses were designed to quantify the relationships between UAV-derived spectral features and DRTM. All analyses were conducted using Python (version 3.12) in a high-performance computing (HPC) environment. The overall workflow comprised three main stages: (i) correlation and linear modelling of vegetation indices, (ii) machine-learning model evaluation using reflectance inputs, and (iii) deep neural network (DNN) modelling and feature attribution. The overall workflow used for imagery-based forecasting of maturity is summarized in Figure 4.

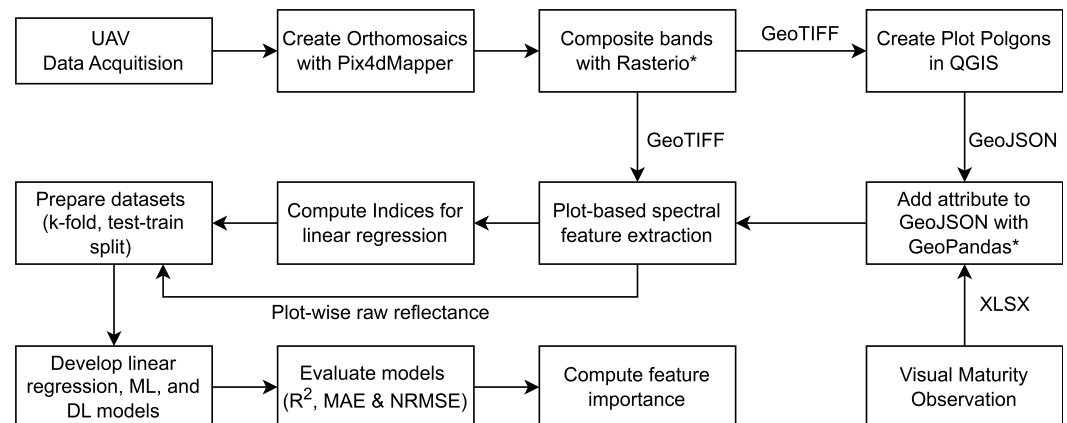


Figure 4. Workflow of the imagery-based forecasting framework for predicting days to maturity in spring wheat. UAV-derived RGB imagery is acquired at multiple time points during the growing season and processed to extract plot-level features, which are integrated with corresponding weather variables and phenological information to train and evaluate machine-learning and deep-learning models for maturity prediction. **Abbreviations:** UAV, unmanned aerial vehicle; ML, machine learning; DL, deep learning; k-fold, k-fold cross-validation; R^2 , coefficient of determination; MAE, mean absolute error; NRMSE, normalized root mean squared error; GeoTIFF, georeferenced raster image format; GeoJSON, geographic vector data format; XLSX, spreadsheet file format. Solid arrows indicate data flow. An asterisk (*) denotes components implemented as Python packages.

2.4.1. Correlation and Linear Regression Analysis

To explore the strength and direction of associations between DRTM and spectral vegetation indices, pairwise Pearson correlation coefficients were computed. Indices listed in Table 2, were grouped into two functional categories: core vegetation indices (e.g., NDVI, GNDVI, NDRE, OSAVI, VARI, SR), representing canopy greenness and vegetation vigor, and chlorophyll/senescence-related indices (e.g., CIRE, NPCI, PSRI, SIPI, MCARI, MCARI1), which are sensitive to pigment concentration and chlorophyll degradation. Correlation matrices were visualized as heatmaps to identify the most responsive indices for maturity prediction.

Table 2. Vegetation, chlorophyll, and senescence-related indices used in this study.

Index Name	Formula	Original Source
Normalized Difference Vegetation Index (NDVI)	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	[23]
Green Normalized Difference Vegetation Index (GNDVI)	$(\text{NIR} - \text{Green}) / (\text{NIR} + \text{Green})$	[24]
Normalized Difference Red Edge Index (NDRE)	$(\text{NIR} - \text{RedEdge}) / (\text{NIR} + \text{RedEdge})$	[25]
Optimized Soil Adjusted Vegetation Index (OSAVI)	$0.16 \times (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red} + 0.16)$	[26]
Visible Atmospherically Resistant Index (VARI)	$(\text{Green} - \text{Red}) / (\text{Green} + \text{Red} - \text{Blue})$	[27]
Simple Ratio (SR)	$\text{RedEdge} / \text{Red}$	[28]
Chlorophyll Index – Red Edge (CI _{RE})	$(\text{NIR} / \text{RedEdge}) - 1$	[29]
Plant Senescence Reflectance Index (PSRI)	$(\text{Red} - \text{Blue}) / \text{RedEdge}$	[30]
Structure Insensitive Pigment Index (SIPI)	$(\text{NIR} - \text{Blue}) / (\text{NIR} + \text{Red})$	[31]
Normalized Pigment Chlorophyll Index (NPCl)	$(\text{Red} - \text{Blue}) / (\text{Red} + \text{Blue})$	[32]
Modified Chlorophyll Absorption Ratio Index (MCARI)	$[(\text{RedEdge} - \text{Red}) - 0.2 \times (\text{Red} - \text{Green})] \times (\text{RedEdge} / \text{Red})$	[33]
Modified Chlorophyll Absorption Ratio Index 1 (MCARI1)	$1.2 \times [2.5(\text{NIR} - \text{Red}) - 1.3(\text{NIR} - \text{Green})]$	[34]

Following correlation analysis, simple linear regression models were fitted using each vegetation index as an individual predictor of days remaining to maturity (DRTM; Equation (1)). Model performance was evaluated using the coefficient of determination (R^2 ; Equation (2)), mean absolute error (MAE; Equation (3)), and normalized root mean squared error (NRMSE; Equation (4)). The combined use of absolute, variance-based, and normalized error metrics enabled a robust assessment of predictive accuracy and comparability across indices. Post-hoc multiple-comparison tests (Tukey's HSD, $p < 0.05$) were performed to statistically discriminate among vegetation indices based on model performance, facilitating identification of top-performing indices and their functional grouping for subsequent feature selection and modelling.

$$\text{DRTM} = \text{Date}_{\text{maturity}} - \text{Date}_{\text{acquisition}} \quad (1)$$

where $\text{Date}_{\text{maturity}}$ is the Julian day corresponding to physiological maturity and $\text{Date}_{\text{acquisition}}$ is the Julian day of UAV image acquisition.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{y_{\max} - y_{\min}} \quad (4)$$

where y_i is the observed value, $\hat{y}_i$ is the predicted value, $\bar{y}$ is the mean of observed values, $y_{\max}$ and $y_{\min}$ are the maximum and minimum observed values, respectively, and n is the sample size.

The R^2 metric quantifies the proportion of variance in the observed data explained by the model, with values approaching 1.0 indicating stronger predictive agreement. MAE provides an interpretable measure of average prediction error in the original units of the response variable (days), while NRMSE expresses prediction error relative to the observed

data range, enabling scale-independent comparison of model performance across indices and datasets. Models exhibiting higher R^2 values and lower MAE and NRMSE values were considered to demonstrate superior predictive performance [35–37].

2.4.2. Machine-Learning Model Development and Evaluation

To capture nonlinear relationships between reflectance data and DRTM, a suite of machine-learning regressors was implemented. The evaluated algorithms included Partial Least Squares Regression (PLSR) [38], Ridge Regression [39], Lasso Regression [40], Elastic Net [41], Support Vector Regression (SVR) [42], K-Nearest Neighbors Regressor (KNN) [43], Gradient Boosting Regressor (GBR) [44], Random Forest Regressor (RF) [45], AdaBoost Regressor [46], Extreme Gradient Boosting (XGBoost) [47], and Multi-Layer Perceptron Regressor (MLP) [48]. Prior to modelling, reflectance features were standardized to ensure consistent scaling across spectral bands. The dataset, comprising 450 (five temporal acquisitions $\times$ 90 plots) observations was randomly partitioned into 80% training and 20% testing subsets, maintaining uniform distribution of maturity values. All models were tuned using an exhaustive grid-search hyperparameter optimization framework with 5-fold cross-validation, in which candidate parameter combinations (Table 3) were evaluated based on cross-validated prediction error, and the configuration minimizing mean absolute error (MAE) was selected as optimal for each model. Model performance was assessed on the independent test set using R^2 and MAE, and Tukey's HSD test was applied to statistically compare mean model accuracies.

Feature importance was derived from each fitted model to evaluate the relative contribution of spectral bands to DRTM prediction. For ensemble and linear regressors, normalized importance scores were extracted from model coefficients or impurity-based measures, then visualized as aggregated heatmaps to highlight consistent spectral sensitivities across algorithms.

2.4.3. Deep Learning Regression and Model Interpretability

A deep neural network (DNN) was developed using PyTorch 2.8.0 to model high-dimensional and non-linear relationships between UAV-derived reflectance and DRTM, following common practice in remote sensing-based regression tasks [49,50]. The network was trained using the Adam optimizer with mean-squared error as the loss function, and performance metrics (R^2 and MAE) were averaged across 100 independent runs to ensure model stability and robustness.

To interpret the spectral contributions learned by the DNN, four attribution techniques—Gradient $\times$ Input, Integrated Gradients, Saliency, and SmoothGrad—were applied to quantify band-level feature importance [51–53]. Normalized attribution scores were averaged over multiple runs to mitigate noise effects, and comparative visualization across methods enabled consistent identification of key spectral regions driving maturity prediction, as recommended for robust model interpretability in deep learning-based remote sensing studies [49].

In addition to the fully connected DNN, a two-dimensional convolutional neural network (2D-CNN) was evaluated to assess whether spatial texture within multispectral canopy patches could further improve DRTM prediction. CNN-based approaches have been widely applied to UAV imagery for crop trait estimation and phenotyping tasks [54,55]. Plot-level image tiles (1 m $\times$ 1 m) were randomly sampled from each orthomosaic using corresponding plot geometries, resized to 45 $\times$ 45 pixels, and used as inputs to a compact CNN trained with Adam optimization and early stopping, following established practices in UAV-based crop modeling [50].

All analyses were implemented in reproducible scripts executed on GPU nodes (128 cores; RAM: 1 TB; GPUs: 2 × NVIDIA Quadro RTX 8000 with 48 GB VRAM each) of the AAFC HPC cluster, ensuring computational efficiency for large-scale UAV datasets.

Table 3. Summary of the hyperparameter search space used for tuning machine-learning models predicting days remaining to maturity (DRTM) from UAV-derived multispectral reflectance.

Model	Hyperparameters Searched
Partial Least Squares Regression (PLSR)	<i>n_components</i> : 1 to $\min(\text{features} , 8)$
Ridge Regression (Ridge)	<i>alpha</i> : 0.01, 0.1, 1, 10, 100
Lasso Regression (Lasso)	<i>alpha</i> : 0.001, 0.01, 0.1, 1.0
Elastic Net	<i>alpha</i> : 0.001, 0.01, 0.1, 1.0; <i>l1_ratio</i> : 0.2, 0.5, 0.8
Support Vector Regression (SVR, RBF)	<i>C</i> : 0.1, 1, 10, 100; <i>gamma</i> : scale, 0.01, 0.001; <i>epsilon</i> : 0.01, 0.1, 0.5
K-Nearest Neighbors Regressor (KNN)	<i>n_neighbors</i> : 3, 5, 7, 9; <i>weights</i> : uniform, distance
Random Forest Regressor (RF)	<i>n_estimators</i> : 200, 400, 800; <i>max_depth</i> : None, 10, 20; <i>min_samples_split</i> : 2, 5, 10
Gradient Boosting Regressor (GBR)	<i>n_estimators</i> : 200, 400; <i>learning_rate</i> : 0.05, 0.1, 0.2; <i>max_depth</i> : 2, 3, 5
AdaBoost Regressor (AdaBoost)	<i>n_estimators</i> : 100, 300, 500; <i>learning_rate</i> : 0.05, 0.1, 0.5, 1.0
XGBoost Regressor (XGBoost)	<i>n_estimators</i> : 300, 600; <i>learning_rate</i> : 0.05, 0.1; <i>max_depth</i> : 3, 5, 7; <i>subsample</i> : 0.8, 1.0; <i>colsample_bytree</i> : 0.8, 1.0
Multi-Layer Perceptron Regressor (MLP)	<i>hidden layer sizes</i> : 50; 100; and 100–50 neurons; <i>activation</i> : ReLU; <i>alpha</i> : 0.0001, 0.001; <i>learning rate</i> : constant, adaptive

3. Results

3.1. Distribution of Maturity Across Years

The distribution of maturity-related traits at the Fairfield site revealed clear year-to-year variation. The distribution of DTM (Figure 5a) showed two distinct clusters corresponding to 2024 and 2025. In 2024, most cultivars reached maturity earlier, with values concentrated between approximately 102 and 106 days after seeding, whereas in 2025, the distribution shifted toward later maturity, ranging from about 118 to 122 days. This difference likely reflects environmental variation between growing seasons, particularly differences in temperature accumulation and precipitation during the grain-filling period as well as potential year-specific genotype responses.

The DRTM histograms (Figure 5b), derived from UAV data acquisition date and DTM, displayed a relatively uniform spread across acquisition dates, with most observations falling between 5 and 30 days before physiological maturity. The substantial overlap between the 2024 and 2025 distributions indicates that flights in both years captured plots at comparable phenological stages, ensuring that canopy reflectance was measured under similar developmental conditions. Collectively, these distributions demonstrate that UAV missions were well timed relative to crop maturity progression, providing temporally consistent datasets for evaluating early-season spectral indicators of maturity.

This alignment across years enhances confidence in the subsequent modeling analyses based on multispectral reflectance features.

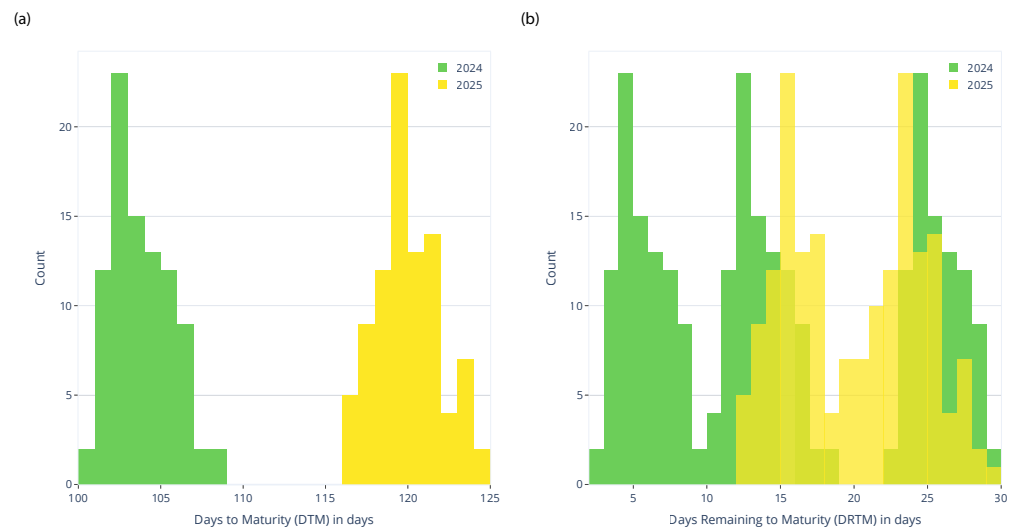


Figure 5. Distribution of maturity traits in spring wheat at the Fairfield experimental site (Lethbridge, Alberta) across two growing seasons (2024 and 2025). (a) Days to maturity (DTM) and (b) Days remaining to maturity (DRTM—Equation (1)) were calculated from DTM and UAV-acquisition dates, respectively.

3.2. Predicting DRTM from Core Vegetation and Chlorophyll/Senescence Related Indices

3.2.1. Correlation Analysis

The heatmap (Figure 6) illustrates the pairwise correlations among DRTM and several vegetation indices derived from UAV imagery. Indices were grouped into two functional categories: Core Vegetation Indices (NDVI, GNDVI, NDRE, OSAVI, VARI, SR), which broadly represent canopy greenness and vegetation vigor, and Chlorophyll/Senescence-related Indices (CIRE, NPCI, PSRI, SIPI, MCARI, MCARI1), which are sensitive to pigment concentration, chlorophyll degradation, and senescence progression. A strong positive correlation ($r > 0.90$) was observed between DRTM and most core vegetation indices, particularly NDVI ($r = 0.96$) and OSAVI ($r = 0.97$), indicating that delayed canopy greenness was associated with longer DRTM.

Conversely, senescence-related indices such as PSRI showed strong negative correlations ($r \approx -0.96$), reflecting their sensitivity to chlorophyll breakdown during crop maturation. SIPI and MCARI1 maintained high positive correlations with DRTM ($r \approx 0.96$ – 0.92), confirming that chlorophyll-sensitive indices also track maturity progression effectively. Overall, the two index groups exhibited complementary relationships with DRTM, demonstrating that both canopy greenness and pigment-based indicators capture physiological processes underlying crop maturity.

3.2.2. Linear Model and Trait Discrimination

To further quantify index performance, a linear model was fitted using DRTM as the response trait and individual vegetation indices as predictors that has Pearson's correlation exceeding ± 0.9 . Post-hoc grouping based on the Tukey test (95% confidence) identified statistically distinct subsets of indices. As shown in Table 4, indices such as OSAVI, PSRI, and NDVI formed the top-performing group with the highest mean R^2 values and lowest MAE, indicating that indices representing overall canopy vigor provided the strongest linear response to maturity variation. In contrast, indices such as VARI and MCARI1 exhibited significantly lower predictive accuracy, suggesting that

pigment- and photosynthetic-efficiency-related indices may capture maturity variation under different canopy or environmental conditions. Overall, both core vegetation and chlorophyll/senescence-related indices contributed valuable but distinct information for predicting DRTM, supporting the use of multiple spectral indicators in maturity modeling.

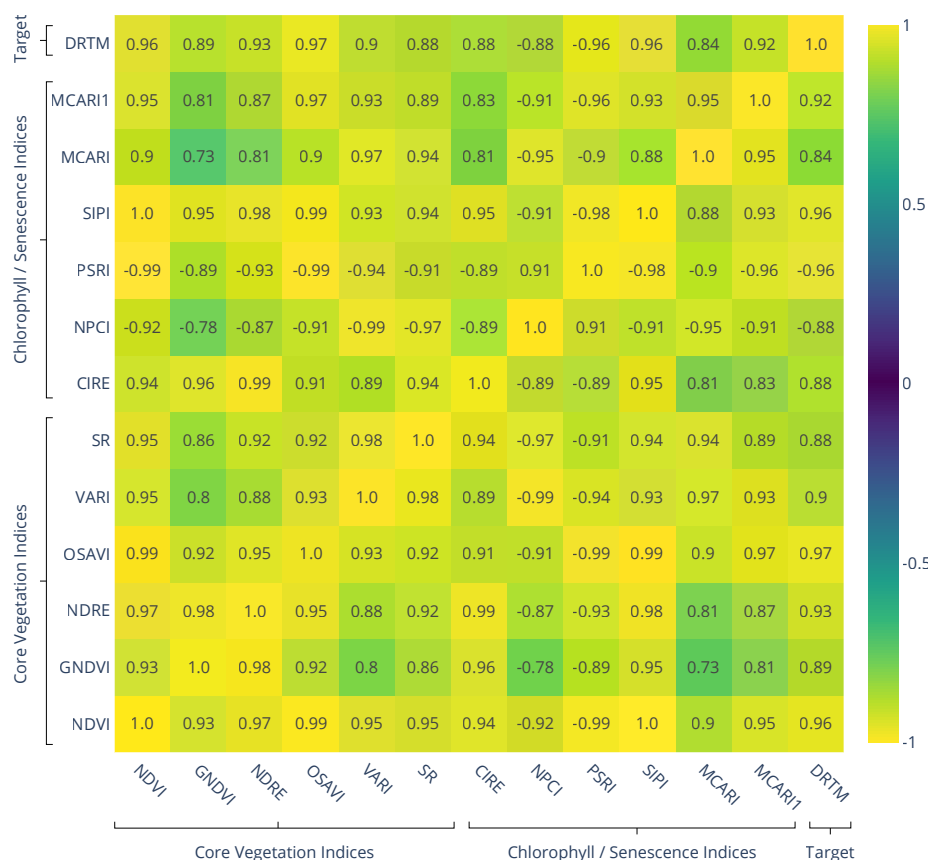


Figure 6. Correlation matrix showing relationships between days remaining to maturity (DRTM) and spectral vegetation indices derived from uncrewed aerial vehicle (UAV) imagery.

Table 4. Predictive performance of individual vegetation indices for estimating days remaining to maturity (DRTM) using simple linear regression. Values are reported as mean $\pm$ standard deviation across cross-validation folds. Different superscript letters within a column indicate statistically significant differences based on Tukey's HSD test ($p < 0.05$).

Index	R^2 (Mean $\pm$ SD)	MAE (Mean $\pm$ SD, Days)	NRMSE (Mean $\pm$ SD)
OSAVI	0.93 $\pm$ 0.01 ^a	1.49 $\pm$ 0.11 ^a	0.074 $\pm$ 0.007 ^a
PSRI	0.92 $\pm$ 0.01 ^a	1.60 $\pm$ 0.08 ^a	0.079 $\pm$ 0.005 ^a
NDVI	0.92 $\pm$ 0.01 ^a	1.63 $\pm$ 0.12 ^a	0.080 $\pm$ 0.007 ^a
SIPI	0.91 $\pm$ 0.01 ^a	1.73 $\pm$ 0.12 ^a	0.085 $\pm$ 0.006 ^a
NDRE	0.86 $\pm$ 0.02 ^b	2.31 $\pm$ 0.20 ^b	0.110 $\pm$ 0.008 ^b
MCARI1	0.83 $\pm$ 0.03 ^b	2.29 $\pm$ 0.24 ^b	0.117 $\pm$ 0.010 ^b
VARI	0.81 $\pm$ 0.01 ^c	2.75 $\pm$ 0.11 ^c	0.127 $\pm$ 0.007 ^c

Note: R^2 is the coefficient of determination; MAE is the mean absolute error expressed in days; NRMSE is the normalized root mean squared error calculated by dividing RMSE by the range of observed DRTM values.

3.3. Comparative Performance of Machine Learning Models for Predicting DRTM from Reflectance Inputs

3.3.1. Model Performance Evaluation

A suite of machine learning algorithms was evaluated for predicting DRTM using UAV-derived reflectance data. Performance metrics, summarized in Table 5, indicate that

the Multi-Layer Perceptron (MLP) and Support Vector Regression (SVR) models achieved the highest predictive accuracy, with mean R^2 values of 0.96 and 0.95 and the lowest MAE values (1.25–1.26), respectively. These models formed the highest statistical grouping (^a) according to Tukey's HSD test ($p < 0.05$), signifying significantly superior performance over other algorithms.

Tree-based ensemble methods (GBR, Random Forest, and XGBoost) showed strong and consistent predictive performance ($R^2 \approx 0.94$) and were statistically comparable to the top-performing models in terms of explained variance (^a). However, their error metrics (MAE and NRMSE) were significantly higher than those of MLP and SVR, placing them in a lower statistical group (^b) for prediction error. In contrast, regularized linear models (Lasso, Ridge, and Elastic Net), along with KNN and PLSR, exhibited slightly reduced predictive accuracy ($R^2 \approx 0.93$) and consistently higher errors, forming a distinct lower-performing group (^b). Overall, these results highlight the advantage of non-linear and multi-layer models for capturing complex relationships between UAV-derived reflectance features and maturity dynamics.

Overall, learning-based (MLP) and kernel-based (SVR) approaches provided the most accurate and robust predictions of DRTM, whereas traditional linear regression models exhibited reduced predictive accuracy and higher error.

3.3.2. Feature Importance Across Machine Learning Models

The relative contribution of each spectral band to model performance was assessed using normalized feature importance scores derived from regression and ensemble models. The heatmap (Figure 7) summarizes the average normalized importance of spectral features across all machine learning algorithms.

Among the tested features, the red band consistently exhibited the highest importance, particularly in tree-based ensemble models such as GBR, and XGBoost, where it accounted for more than 70% of the model's predictive contribution. This highlights the strong sensitivity of these models to canopy reflectance in the red region, which is closely associated with chlorophyll absorption and senescence processes.

Table 5. Predictive performance of machine learning models for estimating days remaining to maturity (DRTM) using UAV-derived multispectral reflectance features. Values are reported as mean $\pm$ standard deviation across cross-validation folds. Different superscript letters within a column indicate statistically significant differences among models based on Tukey's HSD test ($p < 0.05$).

Model	R^2 (Mean $\pm$ SD)	MAE (Mean $\pm$ SD, Days)	NRMSE (Mean $\pm$ SD)
MLP	0.96 $\pm$ 0.01 ^a	1.25 $\pm$ 0.09 ^a	0.059 $\pm$ 0.005 ^a
SVR	0.95 $\pm$ 0.01 ^a	1.26 $\pm$ 0.08 ^a	0.061 $\pm$ 0.004 ^a
GBR	0.94 $\pm$ 0.01 ^a	1.44 $\pm$ 0.09 ^b	0.068 $\pm$ 0.005 ^b
Random Forest	0.94 $\pm$ 0.01 ^a	1.44 $\pm$ 0.07 ^b	0.068 $\pm$ 0.004 ^b
XGBoost	0.94 $\pm$ 0.01 ^a	1.44 $\pm$ 0.09 ^b	0.069 $\pm$ 0.005 ^b
AdaBoost	0.94 $\pm$ 0.01 ^a	1.49 $\pm$ 0.10 ^b	0.070 $\pm$ 0.004 ^b
Lasso	0.93 $\pm$ 0.01 ^b	1.50 $\pm$ 0.08 ^b	0.074 $\pm$ 0.004 ^b
Ridge	0.93 $\pm$ 0.01 ^b	1.50 $\pm$ 0.07 ^b	0.074 $\pm$ 0.003 ^b
Elastic Net	0.93 $\pm$ 0.01 ^b	1.51 $\pm$ 0.07 ^b	0.075 $\pm$ 0.003 ^b
KNN	0.93 $\pm$ 0.02 ^b	1.45 $\pm$ 0.14 ^b	0.074 $\pm$ 0.011 ^b
PLSR	0.93 $\pm$ 0.01 ^b	1.51 $\pm$ 0.08 ^b	0.075 $\pm$ 0.004 ^b

Note: R^2 is the coefficient of determination; MAE is the mean absolute error expressed in days; NRMSE is the normalized root mean squared error calculated by dividing RMSE by the range of observed DRTM values.

In contrast, linear models (Ridge, Lasso, Elastic Net, PLSR) distributed their importance more evenly across bands, with the red, RE, and NIR wavelengths contributing notably to model prediction. Likewise, SVR model also showed high sensitivity to

red, RE, and NIR features, reflecting its ability to capture non-linear responses in the reflectance–maturity relationship.

Overall, the results indicate that red, RE, and NIR spectral regions are the most influential predictors of DRTM across models, aligning with established vegetation reflectance theory and physiological patterns of canopy senescence.

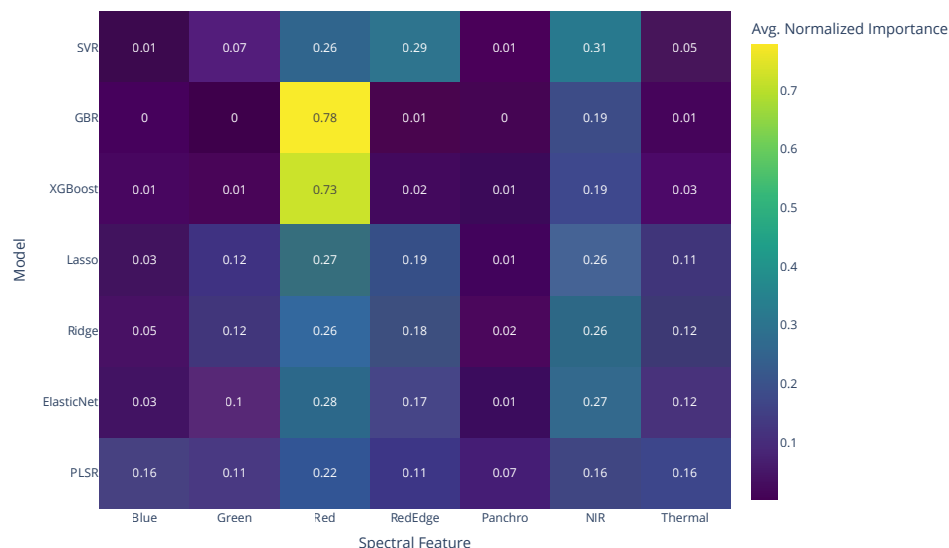


Figure 7. Average normalized feature importance across machine learning models for predicting days remaining to maturity (DRTM).

3.4. Deep Learning Model Performance and Spectral Band Attribution

A deep neural network (DNN) regression model was trained to predict DRTM from UAV-derived reflectance features. The model achieved consistent performance across 100 independent runs, with an average coefficient of determination ($R^2 = 0.96 \pm 0.01$) and a MAE of 1.23 ± 0.08 days. The relationship between predicted and observed DRTM on unseen test dataset was highly linear (Figure 8), indicating excellent predictive agreement across UAV acquisition dates.

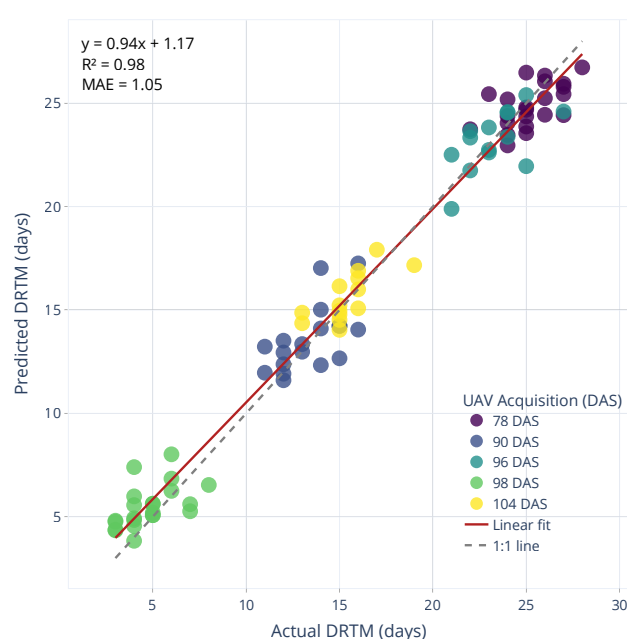


Figure 8. Actual versus predicted DRTM from the best-performing deep learning model, colored by UAV acquisition timing (days after seeding, DAS).

These results demonstrate strong predictive ability, with error values within 1–2 days of field-observed maturity, confirming the model’s capability to capture nonlinear reflectance–maturity relationships. To interpret the DNN, four attribution techniques, Gradient $\times$ Input, Integrated Gradients, Saliency, and SmoothGrad were applied to quantify feature importance across spectral bands (Figure 9).

All attribution methods consistently identified the red, RE, and NIR bands as the most influential predictors of maturity, together contributing approximately 40–45% of the total normalized importance. Moderate importance was attributed to blue, green, and panchromatic bands, while thermal band contributed less, suggesting lower direct sensitivity to maturity-related canopy reflectance changes. The consistency among multiple interpretability approaches confirms the robustness of the learned relationships and indicates that maturity prediction is primarily driven by vegetation vigor and chlorophyll absorption dynamics represented by the red, red-edge, NIR spectral region.

An additional hypothesis tested in this study was that incorporating spatial information from multispectral image patches through a two-dimensional convolutional neural network (2D-CNN) would enhance predictive accuracy by capturing within-plot textural cues related to senescence progression. However, the 2D-CNN did not outperform the feature-based DNN; instead, predictive accuracy declined ($R^2 = 0.92 \pm 0.61$; MAE = 1.62 ± 0.95 days). This suggests that, under the conditions of this dataset, characterized by relatively homogeneous canopies within plots and high-quality reflectance calibration, the spatial patterns present in small canopy patches contributed less predictive signal than aggregated reflectance features. Thus, contrary to the initial expectation, adding spatial information did not improve model performance, highlighting that maturity prediction in this setting is driven primarily by spectral rather than spatial variability.



Figure 9. Feature importance of the deep learning model derived from four attribution methods (Gradient $\times$ Input, Integrated Gradients, Saliency, and SmoothGrad) applied over 100 runs.

4. Discussion

This study demonstrates that UAV-derived multispectral reflectance can reliably forecast spring wheat maturity without reliance on temperature-based or cultivar-specific phenology models. Substantial interannual variability in days to maturity (DTM) at the Fairfield site, with earlier maturity observed in 2024 and markedly delayed maturity in 2025, highlights the strong influence of environmental conditions on phenological development. Such variability is well documented in cereal crops and underscores the known limitations of thermal-time and growing-degree-day approaches under heterogeneous mi-

croclimates and genotype $\times$ environment ($G \times E$) interactions [10,14]. In contrast, the Days Remaining to Maturity (DRTM) framework introduced here provides a direct, observation-driven measure of physiological progression that is inherently responsive to canopy status rather than assumed temperature responses. The consistency with which UAV acquisitions captured comparable phenological stages across years—typically 5–30 days prior to maturity—supports DRTM as a robust reflectance-based maturity indicator that is resilient to interannual climatic variability.

Strong spectral–physiological relationships were observed across all analyses. Correlations between DRTM and core vegetation indices exceeded $|r| > 0.95$ for NDVI and OSAVI, while senescence-sensitive indices such as PSRI exhibited similarly strong inverse relationships ($r \approx -0.96$). These findings align with established physiological understanding: greenness indices reflect canopy chlorophyll content and biomass persistence, whereas senescence indices capture chlorophyll degradation and carotenoid dominance during grain filling [30,56]. Linear modeling further reinforced these relationships, with OSAVI, PSRI, and NDVI achieving the highest predictive performance ($R^2 \approx 0.92$ – 0.93 ; MAE ≈ 1.49 – 1.63 days), while indices such as VARI and MCARI1 performed substantially worse ($R^2 \approx 0.81$ – 0.83). This divergence emphasizes that indices explicitly designed to correct for soil background effects (OSAVI) or senescence dynamics (PSRI) provide complementary and more physiologically relevant signals of maturity progression than purely visible-band indices [26,57]. Although NDVI remains widely used as a proxy for stay-green and senescence traits [8,56], the comparable or superior performance of OSAVI and PSRI suggests they represent robust alternatives for maturity assessment in spring wheat.

Machine learning models leveraging full multispectral reflectance further improved predictive accuracy, indicating that maturity-related information is distributed across multiple spectral bands rather than confined to individual indices. Among the evaluated approaches, multilayer perceptrons (MLP) and support vector regression (SVR) achieved the highest performance ($R^2 = 0.96$ and 0.95 ; MAE = 1.25 – 1.26 days), outperforming tree-based ensembles ($R^2 \approx 0.94$) and regularized linear regressors ($R^2 \approx 0.93$). The superiority of nonlinear models is consistent with previous remote-sensing studies showing that reflectance–phenology relationships are inherently nonlinear, particularly during late grain filling when spectral changes accelerate [19,49]. Spectral importance analyses consistently identified the red, red-edge, and near-infrared (NIR) regions as dominant predictors, collectively accounting for more than 70% of variance in ensemble models and approximately 40–45% of attribution in deep neural networks. These regions are known to be sensitive to chlorophyll concentration, red-edge position shifts, and internal leaf structure changes associated with senescence and maturity [50,58].

Deep learning further strengthened these findings. The fully connected deep neural network (DNN) achieved $R^2 = 0.96 \pm 0.01$ and MAE = 1.23 ± 0.08 across 100 independent runs, demonstrating excellent stability and predictive agreement. Such consistency indicates that the DRTM–reflectance relationship is sufficiently strong to be learned robustly without extensive architectural complexity. In contrast, incorporating spatial information through a 2D convolutional neural network (CNN) did not improve performance; instead, the CNN exhibited reduced accuracy and higher variance ($R^2 = 0.92 \pm 0.61$; MAE = 1.62 ± 0.95). Similar outcomes have been reported in other cereal phenotyping studies where uniform canopy structure and minimal within-plot heterogeneity limited the utility of texture-based features [55]. These results suggest that, under well-managed and spatially homogeneous conditions, maturity signals are dominated by spectral rather than spatial variability. Consequently, simpler spectral models may be preferable to more complex spatial architectures for operational maturity forecasting in such systems.

Collectively, these findings position DRTM as a physiologically meaningful, image-derived trait that bridges the gap between traditional phenological scoring and data-driven forecasting. By relying on direct canopy observations rather than accumulated thermal units, the proposed framework offers improved robustness under variable climatic conditions and provides a scalable foundation for high-throughput phenotyping and precision agriculture applications.

5. Limitations and Future Perspectives

Despite the consistently strong predictive performance achieved in this study, several limitations must be acknowledged to properly contextualize the results and guide future research. First, all experiments were conducted under well-managed, irrigated conditions at a single site. While this design minimized confounding stress effects and enabled clear isolation of spectral-maturity relationships, it also constrains generalizability. Under rainfed or drought-stressed environments, premature leaf senescence induced by water or heat stress may spectrally resemble physiological maturity, thereby weakening the association between reflectance-based indicators and true grain maturity [10,14]. For example, Romero and Lopes [19] reported substantial reductions in predictive performance under rainfed conditions, with R^2 values declining from 0.53 to 0.22 in oats and from 0.63 to 0.60 in wheat. Consequently, the accuracy reported here likely represents an upper bound under optimal moisture availability, and broader multi-environment validation across contrasting water regimes, soil backgrounds, and canopy architectures is required to assess robustness and transferability.

A second important limitation relates to UAV flight timing and temporal sampling density. Although this study emphasizes the value of *early* maturity forecasting, most UAV acquisitions occurred approximately 5–30 days prior to median maturity, rather than during earlier vegetative or heading stages. While this temporal window is agronomically relevant for harvest scheduling and late-season decision-making, it does not fully exploit the potential of UAV-based sensing for long-range forecasting earlier in the season. Phenological prediction accuracy is known to be highly sensitive to both observation timing and revisit frequency, with uncertainty increasing as observations move further from the target developmental stage [59,60]. Although the strong performance observed here indicates that multispectral reflectance contains highly informative physiological signals during late grain filling, future work should evaluate denser and earlier flight schedules spanning stem elongation through heading to quantify how forecast horizon length affects accuracy and to identify the earliest phenological stage at which reliable DRTM prediction becomes feasible [61].

Spatial scale and experimental scope also impose constraints. Data were collected at a single location over two growing seasons, which raises the possibility that models partially learned site- or year-specific patterns, particularly given the distinct maturity distributions observed between 2024 and 2025. While repeated cross-validation demonstrated strong internal consistency, true generalization would be more rigorously assessed using between-year or leave-one-environment-out validation frameworks [62]. Expanding datasets to include multiple locations and seasons would reduce the risk of overfitting and strengthen confidence in model portability across environments.

From a methodological perspective, the limited benefit of incorporating spatial texture through 2D convolutional neural networks highlights another boundary of the current approach. The uniformity of well-managed spring wheat canopies and the use of plot-level reflectance summaries likely reduced the availability of informative spatial patterns, rendering spectral signals dominant. Similar findings have been reported in other cereal phenotyping studies, where mean spectral responses outperformed texture-based features

under homogeneous canopy conditions [55]. Alternative strategies—such as incorporating temporal sequences of images, multi-angle observations, or coupling spectral data with structural metrics from LiDAR—may unlock additional predictive value in more heterogeneous or stress-affected production systems [63].

Despite these limitations, this work establishes Days Remaining to Maturity (DRTM) as a robust, image-derived phenotypic trait with clear utility for breeding programs and precision agriculture. Future research should extend this framework to diverse genotypes and agroecological zones, integrate complementary stress indicators such as canopy temperature and soil moisture, and explore domain adaptation or transfer learning techniques to improve cross-site portability [49]. Importantly, scaling beyond UAV platforms is feasible: satellite sensors including Sentinel-2, Landsat 8/9, and PlanetScope provide red, red-edge, and near-infrared bands that have been widely shown to be sensitive to crop phenology, senescence, and maturity-related processes [64–66]. Multi-scale modeling approaches that fuse UAV and satellite observations could therefore enable continuous, landscape-level forecasting of crop maturity, supporting yield estimation, harvest logistics, and climate-resilient decision-making at regional scales.

6. Conclusions

This study demonstrates that UAV-based multispectral imagery, combined with machine learning and deep learning models, provides a precise, scalable, and operationally robust framework for forecasting days remaining to maturity (DRTM) in spring wheat. By relying exclusively on canopy reflectance information, the proposed approach bypasses the assumptions and site-specific calibration requirements inherent to temperature-based or cultivar-dependent phenology models, while still achieving high predictive accuracy. Across all evaluated methods, top-performing models—including deep neural networks (DNN), multilayer perceptrons (MLP), and support vector regression (SVR)—consistently achieved R^2 values of 0.95–0.96 with mean absolute errors of approximately 1–2 days, underscoring the robustness and repeatability of reflectance-driven maturity forecasting.

Among vegetation indices, OSAVI, PSRI, and NDVI exhibited the strongest linear associations with DRTM ($R^2 \approx 0.92$ – 0.93), reflecting the complementary roles of canopy greenness persistence and senescence dynamics during grain filling. However, models trained on full multispectral reflectance consistently outperformed index-based approaches, confirming that maturity-related information is distributed across multiple spectral bands rather than encapsulated by any single index. Feature importance analyses and deep learning attribution methods converged in identifying the red, red-edge, and near-infrared (NIR) regions as the dominant drivers of predictive performance, consistent with known physiological processes such as chlorophyll degradation, red-edge position shifts, and changes in internal leaf structure during maturation.

An important methodological insight from this work is that incorporating spatial texture through two-dimensional convolutional neural networks did not improve prediction accuracy and instead increased model variance. This finding indicates that, under well-managed and spatially homogeneous canopy conditions, maturity signals are primarily governed by spectral responses rather than within-plot spatial variability. As a result, simpler spectral models may offer superior performance and greater operational reliability compared with more complex spatial architectures in similar production systems.

Overall, the results establish DRTM as a reliable, image-derived phenological trait that bridges traditional maturity scoring and data-driven forecasting. The proposed framework has clear implications for high-throughput phenotyping in breeding programs, where objective and repeatable maturity estimates are critical, as well as for precision agriculture applications such as harvest scheduling and yield risk management. By demonstrating that

accurate maturity prediction is achievable using UAV multispectral imagery alone, this study provides a strong foundation for future efforts aimed at extending reflectance-based maturity forecasting across environments, seasons, and sensing scales.

Author Contributions: P.R.: Data curation, Methodology, Formal analysis, Conceptualization, Investigation, Validation, Visualization, Writing—original draft; K.D.S.: Data curation, Investigation, Validation, Visualization, Writing—review and editing, Funding acquisition, Project administration, Resources, Supervision; H.S.R.: Validation, Visualization, Writing—review and editing, Funding acquisition, Resources; S.S.P.: Investigation, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by AAFC (Ref: AGR-19913) through Western Grains Research Foundation (WGRF-VarD2334), and Saskatchewan Wheat Development Commission (SWDC-276-221123).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors sincerely thank Mark Virginillo, field technician with the Spring Wheat Breeding Program, for his valuable support in field trial setup, visual rating and management. © His Majesty the King in Right of Canada, 2026.

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

References

1. Shiferaw, B.; Smale, M.; Braun, H.J.; Duveiller, E.; Reynolds, M.; Muricho, G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Secur.* **2013**, *5*, 291–317. [\[CrossRef\]](#)
2. Asseng, S.; Ewert, F.; Martre, P.; Rötter, R.P.; Lobell, D.B.; Cammarano, D.; Kimball, B.A.; Ottman, M.J.; Wall, G.W.; White, J.W.; et al. Rising temperatures reduce global wheat production. *Nat. Clim. Change* **2015**, *5*, 143–147. [\[CrossRef\]](#)
3. Dhondt, S.; Wuyts, N.; Inzé, D. Cell to whole-plant phenotyping: The best is yet to come. *Trends Plant Sci.* **2013**, *18*, 428–439. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Hassan, M.A.; Yang, M.; Rasheed, A.; Tian, X.; Reynolds, M.; Xia, X.; Xiao, Y.; He, Z. Quantifying senescence in bread wheat using multispectral imaging from an unmanned aerial vehicle and QTL mapping. *Plant Physiol.* **2021**, *187*, 2623–2636. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Yoosefzadeh-Najafabadi, M.; Singh, K.D.; Pourreza, A.; Sandhu, K.S.; Adak, A.; Murray, S.C.; Eskandari, M.; Rajcan, I. Remote and proximal sensing: How far has it come to help plant breeders? In *Advances in Agronomy*; Sparks, D.L., Ed.; Academic Press: Cambridge, MA, USA, 2023; Volume 181, pp. 279–315. [\[CrossRef\]](#)
6. Grbović, Ž.; Ivošević, B.; Budjen, M.; Waqar, R.; Pajević, N.; Ljubičić, N.; Kandić, V.; Pajić, M.; Panić, M. Integrating UAV multispectral imaging and proximal sensing for high-precision cereal crop monitoring. *PLoS ONE* **2025**, *20*, e0322712.
7. Natarajan, M.; Singh, K.D.; Geddes, C.M.; Shirliff, S.J.; Ravichandran, P.; Wang, H. UAV-based hyperspectral imaging to evaluate plant moisture and desiccant response in lentil (*Lens culinaris*). *Can. J. Plant Sci.* **2025**, *105*, 1–13. [\[CrossRef\]](#)
8. Tschurr, F.; Walter, A.; Liebisch, F. UAV-based monitoring of cereal senescence dynamics. *Remote Sens.* **2024**, *16*, 341. [\[CrossRef\]](#)
9. Hill, B.D.; McGinn, S.M.; Korchinski, A.; Burnett, B. Neural network models to predict the maturity of spring wheat in western Canada. *Can. J. Plant Sci.* **2002**, *82*, 7–13. [\[CrossRef\]](#)
10. Eyshi Rezaei, E.; Siebert, S.; Ewert, F. Climate and management interaction cause diverse crop phenology trends. *Agric. For. Meteorol.* **2017**, *233*, 55–70. [\[CrossRef\]](#)
11. Ravichandran, P.; Singh, K.D.; Noble, S.D.; Soolanayakanahally, R.; Sangha, J.S.; Brauer, E.K.; Molina, O.; Nilsen, K.T.; Randhawa, H.S.; Halcro, K.; et al. Precision phenotyping in wheat: LiDAR-based plant height estimation and lodging classification using uncrewed ground vehicles. *Can. J. Remote Sens.* **2025**, *51*, 2516742. [\[CrossRef\]](#)
12. Ravichandran, P.; Singh, K.D.; Randhawa, H.S.; Dhariwal, R.; Sangha, J.S.; Ellert, B.; Wang, H.; Chegoonian, A.; Natarajan, M. High-Throughput Screening of Wheat Genotypes for Drought Tolerance Using Aerial Thermal Imagery. In Proceedings of the 2025 13th International Conference on Agro-Geoinformatics, Boulder, CO, USA, 7–10 July 2025.
13. Mkhabela, M.; Ash, G.; Grenier, M.; Bullock, P. Testing the suitability of thermal time models for forecasting spring wheat phenological development in western Canada. *Can. J. Plant Sci.* **2016**, *96*, 765–775. [\[CrossRef\]](#)
14. Parent, B.; Tardieu, F. Temperature responses of developmental processes have not been affected by breeding in different ecological areas for 17 crop species. *New Phytol.* **2019**, *214*, 760–774. [\[CrossRef\]](#)

15. Pullens, J.W.M.; Sørensen, C.A.G.; Olesen, J.E. Temperature-based prediction of harvest date in winter and spring cereals as a basis for assessing viability for growing cover crops. *Field Crops Res.* **2021**, *264*, 108085. [[CrossRef](#)]
16. Zhou, J.; Yungbluth, D.; Vong, C.N.; Scaboo, A.; Zhou, J. Estimation of the Maturity Date of Soybean Breeding Lines Using UAV-Based Multispectral Imagery. *Remote Sens.* **2019**, *11*, 2075. [[CrossRef](#)]
17. Anku, K.; Percival, D.; Vankoughnett, M.; Lada, R.; Heung, B. Monitoring and Prediction of Wild Blueberry Phenology Using a Multispectral Sensor. *Remote Sens.* **2025**, *17*, 334. [[CrossRef](#)]
18. Navasca, H.; Bazrafkan, A.; Dariva, F.D.; Kim, J.H.; Worral, H.; Johnson, J.P.; Acharya, S.R.; Piche, L.; Ross, A.; Raymon, G.; et al. Improving estimation of days to maturity in field pea using RGB aerial imagery and machine learning. *Plant Phenome J.* **2025**, *8*, e70038. [[CrossRef](#)]
19. Romero, M.; Lopes, M.S. Heading and maturity date prediction using vegetation indices under contrasting water regimes. *Eur. J. Agron.* **2024**, *160*, 127330. [[CrossRef](#)]
20. Liu, T.; Zhu, S.; Yang, T.; Zhang, W.; Xu, Y.; Zhou, K.; Wu, W.; Zhao, Y.; Yao, Z.; Yang, G.; et al. Maize height estimation using combined unmanned aerial vehicle oblique photography and LIDAR canopy dynamic characteristics. *Comput. Electron. Agric.* **2024**, *218*, 108685. [[CrossRef](#)]
21. Panigrahi, S.S.; Singh, K.D.; Balasubramanian, P.; Wang, H.; Natarajan, M.; Ravichandran, P. UAV-Based LiDAR and Multispectral Imaging for Estimating Dry Bean Plant Height, Lodging and Seed Yield. *Sensors* **2025**, *25*, 3535. [[CrossRef](#)]
22. Saiyed, I.M.; Bullock, P.R.; Sapirstein, H.D.; Finlay, G.J.; Jarvis, C.K. Thermal time models for estimating wheat phenological development and weather-based relationships to wheat quality. *Can. J. Plant Sci.* **2009**, *89*, 429–439. [[CrossRef](#)]
23. Rouse, J.W.; Haas, R.H.; Deering, D.; Schell, J.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Technical Report; Goddard Space Flight Center: Greenbelt, MD, USA, 1973.
24. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [[CrossRef](#)]
25. Barnes, E.M.; Clarke, T.R.; Richards, S.E.; Colaizzi, P.D.; Haberland, J.; Kostrzewski, M.; Waller, P.; Choi, C.; Riley, E.; Thompson, T.; et al. *Coincident Detection of Crop Water Stress, Nitrogen Status and Canopy Density Using Ground-Based Multispectral Data*; USDA: Madison, WI, USA, 2000; pp. 1–15.
26. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [[CrossRef](#)]
27. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [[CrossRef](#)]
28. Birth, G.S.; McVey, G.R. Measuring the Color of Growing Turf with a Reflectance Spectrophotometer. *Agron. J.* **1968**, *60*, 640–643. [[CrossRef](#)]
29. Gitelson, A.A.; Viña, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*. [[CrossRef](#)]
30. Merzlyak, M.N.; Gitelson, A.A.; Chivkunova, O.B.; Rakitin, V.Y. Non-destructive optical detection of pigment changes during leaf senescence. *Physiol. Plant.* **1999**, *106*, 135–141. [[CrossRef](#)]
31. Penuelas, J.; Baret, F.; Filella, I. Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica* **1995**, *31*, 221–230.
32. Peñuelas, J.; Gamon, J.A.; Fredeen, A.L.; Merino, J.; Field, C.B. Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sens. Environ.* **1994**, *48*, 135–146. [[CrossRef](#)]
33. Daughtry, C.S.T.; Walthall, C.L.; Kim, M.S.; de Colstoun, E.B.; McMurtrey, J.E. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [[CrossRef](#)]
34. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies. *Remote Sens. Environ.* **2004**, *90*, 337–352. [[CrossRef](#)]
35. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [[CrossRef](#)]
36. Willmott, C.J. Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* **1982**, *63*, 1309–1313. [[CrossRef](#)]
37. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [[CrossRef](#)]
38. Wold, H. Estimation of principal components and related models by iterative least squares. In *Multivariate Analysis*; Academic Press: New York, NY, USA, 1966; pp. 391–420.
39. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [[CrossRef](#)]
40. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **2018**, *58*, 267–288. [[CrossRef](#)]
41. Zou, H.; Hastie, T. Regularization and Variable Selection Via the Elastic Net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [[CrossRef](#)]

42. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
43. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [[CrossRef](#)]
44. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
45. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
46. Freund, Y.; Schapire, R.E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [[CrossRef](#)]
47. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), San Francisco, CA, USA, 13–17 August 2016*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794. [[CrossRef](#)]
48. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
49. Zhu, X.X.; Tuia, D.; Mou, L.; Xia, G.S.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [[CrossRef](#)]
50. Tao, H.; Feng, H.; Xu, L.; Miao, M.; Long, H.; Yue, J.; Li, Z.; Yang, G.; Yang, X.; Fan, L. Estimation of crop growth parameters using UAV-based hyperspectral remote sensing data. *Sensors* **2020**, *20*, 1296. [[CrossRef](#)] [[PubMed](#)]
51. Simonyan, K.; Vedaldi, A.; Zisserman, A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. *arXiv* **2014**. [[CrossRef](#)]
52. Sundararajan, M.; Taly, A.; Yan, Q. Axiomatic Attribution for Deep Networks. *arXiv* **2017**. [[CrossRef](#)]
53. Smilkov, D.; Thorat, N.; Kim, B.; Viégas, F.; Wattenberg, M. SmoothGrad: Removing noise by adding noise. *arXiv* **2017**. [[CrossRef](#)]
54. Nevavuori, P.; Narra, N.; Lipping, T. Crop yield prediction with deep convolutional neural networks. *Comput. Electron. Agric.* **2019**, *163*, 104859. [[CrossRef](#)]
55. Haghighattalab, A.; González Pérez, L.; Mondal, S.; Singh, D.; Schinstock, D.; Rutkoski, J.; Ortiz-Monasterio, I.; Singh, R.P.; Goodin, D.; Poland, J. Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. *Plant Methods* **2016**, *12*, 35. [[CrossRef](#)]
56. Lopes, M.S.; Reynolds, M.P. Stay-green in spring wheat can be determined by spectral reflectance measurements. *J. Exp. Bot.* **2012**, *63*, 3789–3798. [[CrossRef](#)]
57. Behn, H.; Ballvora, A.; Bendig, J.; Ispizua Yamati, F.R.; Koua, A.P.; Mahlein, A.K.; Mason, A.S.; Rascher, U.; Sadeqi, M.B.; Léon, J. UAV-based multispectral image analysis revealed stay-green haplotypes in wheat specific for different soil nitrogen levels. *BMC Plant Biol.* **2025**, *25*, 1405. [[CrossRef](#)]
58. Clevers, J.G.P.W.; Kooistra, L. Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 574–583. [[CrossRef](#)]
59. Sakamoto, T.; Wardlaw, B.D.; Gitelson, A.A.; Verma, S.B. Temporal dynamics of crop phenology from time-series vegetation index data. *Remote Sens. Environ.* **2018**, *208*, 337–350. [[CrossRef](#)]
60. Zhang, X.; Friedl, M.A.; Schaaf, C.B.; Strahler, A.H.; Hodges, J.C.F.; Gao, F.; Reed, B.C.; Huete, A. Monitoring vegetation phenology using MODIS. *Remote Sens. Environ.* **2003**, *84*, 471–475. [[CrossRef](#)]
61. Wu, C.; Wang, L.; Niu, Z. UAV-based crop phenology detection and its sensitivity to revisit frequency. *ISPRS J. Photogramm. Remote Sens.* **2021**, *178*, 173–186. [[CrossRef](#)]
62. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013. [[CrossRef](#)]
63. Malambo, L.; Popescu, S.C.; Murray, S.C.; Putman, E.; Pugh, N.A.; Horne, D.W.; Richardson, G.; Sheridan, R.; Rooney, W.L.; Avant, R.; et al. Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *64*, 31–42. [[CrossRef](#)]
64. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
65. Mulla, D.J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst. Eng.* **2013**, *114*, 358–371. [[CrossRef](#)]
66. Yue, J.; Li, T.; Shen, J.; Wei, Y.; Xu, X.; Liu, Y.; Feng, H.; Ma, X.; Li, C.; Yang, G.; et al. Winter wheat maturity prediction via Sentinel-2 MSI images. *Agriculture* **2024**, *14*, 1368. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.