



Article

A Machine Learning Framework for Classifying Thermal Stress in Bean Plants Using Hyperspectral Data

Lucas Prado Osco ^{1,*} , Érika Akemi Saito Moriya ² , Bruna Coelho de Lima ³ , Ana Paula Marques Ramos ² , José Marcato Júnior ⁴ , Wesley Nunes Gonçalves ⁵ , Lúcio André de Castro Jorge ⁶ , Veraldo Liesenberg ⁷ , Jonathan Li ⁸ , Ademir Sérgio Ferreira de Araújo ⁹, Nilton Nobuhiro Imai ² and Fábio Fernando de Araújo ³

- ¹ Environment and Regional Development Program, University of Western São Paulo, Rod. Raposo Tavares, km 572, Limoeiro, Presidente Prudente 19067-175, Brazil
 - ² Department of Cartographic Science, São Paulo State University, R. Roberto Simonsen, 305, Presidente Prudente 19060-900, Brazil
 - ³ Agronomy Program, University of Western São Paulo, Rod. Raposo Tavares, km 572, Limoeiro, Presidente Prudente 19067-175, Brazil
 - ⁴ Faculty of Engineering, Architecture, and Urbanism and Geography, Federal University of Mato Grosso do Sul, Av. Costa e Silva, Campo Grande 79070-900, Brazil
 - ⁵ Faculty of Computer Science, Federal University of Mato Grosso do Sul, Av. Costa e Silva, Campo Grande 79070-900, Brazil
 - ⁶ National Research Center of Development of Agricultural Instrumentation, Brazilian Agricultural Research Agency, R. XV de Novembro, 1452, São Carlos 13560-970, Brazil
 - ⁷ Forest Engineering Department, Santa Catarina State University, Av. Luiz de Camões, 2090, Lages 88520-000, Brazil
 - ⁸ Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada
 - ⁹ Department of Agricultural Engineering and Soil Science, Federal University of Piauí, Teresina 64049-550, Brazil
- * Correspondence: lucasosco@unoeste.br; Tel.: +55-18-3229-1000

Abstract

Rising global temperatures pose a significant threat to agricultural productivity, making the early detection of plant stress crucial for minimizing crop losses. While hyperspectral remote sensing is a powerful tool for monitoring plant health, the precise spectral regions and most effective machine learning models for detecting thermal stress remain an open research question. This study presents a robust framework that utilizes eight state-of-the-art machine learning algorithms to classify the reflectance response of thermal-induced stress in two cultivars of bean plants. Our controlled experiment measured hyperspectral data across two growth stages and three stress conditions (pre-stress, during stress, and post-stress) using a spectroradiometer. The results demonstrate the high performance of several algorithms, with the Artificial Neural Network (ANN) achieving an impressive 99.4% overall accuracy. A key contribution of this work is the identification of the most contributory spectral ranges for thermal stress discrimination: the green region (530–570 nm) and the red-edge region (700–710 nm). This framework is a feasible and effective tool for modelling the hyperspectral response of thermal-stressed bean plants and provides critical guidance for future research on stress-specific spectral indices.

Keywords: remote sensing; thermal stress; machine learning



Academic Editor: Eugenio Cavallo

Received: 13 September 2025

Revised: 21 October 2025

Accepted: 28 October 2025

Published: 7 November 2025

Citation: Osco, L.P.; Moriya, É.A.S.; de Lima, B.C.; Ramos, A.P.M.; Júnior, J.M.; Gonçalves, W.N.; Jorge, L.A.d.C.; Liesenberg, V.; Li, J.; de Araújo, A.S.F.; et al. A Machine Learning Framework for Classifying Thermal Stress in Bean Plants Using Hyperspectral Data.

AgriEngineering **2025**, *7*, 376.

<https://doi.org/10.3390/agriengineering7110376>

Copyright: © 2025 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) license

(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The continuous rise in global temperatures, with average temperatures already approximately 1 °C higher than the pre-industrial era [1], poses significant challenges to

agriculture. Projections suggest an increase of an additional 2 °C by 2050 [2]. This scenario necessitates new methods for crop management, particularly in regions with limited irrigation resources. Several countries do not have resources available to irrigate crops frequently, and temperature will be a key factor in their management [3,4]. Therefore, modelling the thermal stress effects on plants or cultivars is an important and current issue for multiple research areas [5,6].

Temperature changes can impact plants in their vegetative growth, photosynthesis performance, water distribution within the xylem, and their reproductive stage [7]. At elevated temperatures, plants increase their evapotranspiration mechanism, requiring higher amounts of water to be absorbed [6,8], increasing their demand for irrigation. As such, thermal stress in plants rapidly increases in response to high temperatures [4,9], and the early detection of these effects may be a key factor to assist the decision-making process in agricultural systems, avoiding potential crop losses.

Remote sensing is an important data source for studying vegetation in general. Datasets can be acquired at different spectral scales, like RGB, multispectral, and hyperspectral, using distinct platforms, such as proximal (or terrestrial), aerial, or orbital. Hyperspectral sensors are being used in a wide range of applications. For example, researchers have used them to detect biotic stresses, like diseases [10] and pests [11], in various crops. Other studies have focused on quantifying plant constituents, such as nitrogen and carbon [12], and chlorophyll content [13], or even predicting the concentration of heavy metals like cadmium [14]. This technology, in combination with others, is also vital for characterizing key vegetation parameters and general plant traits [15] and is particularly effective for modelling abiotic stresses, like water [16,17] and saline stress [18].

Hyperspectral data are complex to analyse, since a huge amount of information is collected into each experiment, often resulting in high-dimensional data to be evaluated. As a result, robust and intelligent methods are required to extract information from these datasets. Principal Components Analysis (PCA) and Partial Least-Squares Regression (PLSR) represent traditional approaches to process hyperspectral data [6,19] and can deal with the problem of high-dimensionality. But, to improve estimations and analysis of such complex data, machine learning algorithms have been proposed as a robust and intelligent approach in recent years [17,20].

As a subfield of artificial intelligence, machine learning algorithms are generally classified into supervised, unsupervised, and reinforced learners. Supervised algorithms require examples of data previously labelled by a human, using it as a reference sample for learning and validating the algorithm's performance [21]. Unsupervised algorithms do not require labelled data and are left to model it with their characteristics [22]. Reinforced learning algorithms seek to analyse and interpret the dataset and correct themselves through a process of retro-feed and a reward definition [23].

In remote sensing, machine learning algorithms have been established as a robust and intelligent approach for analysing complex spectral data. The effectiveness of this approach has been demonstrated in comprehensive reviews covering both general plant stress imaging and specific challenges like crop water stress [17]. Specific applications include using ML models to estimate photosynthetic capacities from leaf reflectance and to model nutrient content in diverse ecosystems, like alpine grasslands [24] and citrus plantations [25,26], for instance. However, the effectiveness of machine learning algorithms for modelling thermal stress on plants with hyperspectral data still is, to the best of our knowledge, an open research issue.

As a related approach, the investigation of water-deficit response using leaf reflectance in the Visible (VIS), Near-Infrared (NIR), and Short-Wave Infrared (SWIR) spectral regions is well-documented. These studies have been conducted across a variety of agricultural

contexts, including vineyards [27], maize fields [28], soil-less tomato crops [29], and even urban turfgrass systems [3], often with the goal of developing more sensitive, image-derived stress indices. For the evaluation of temperature effects, studies were carried out mostly in the thermal infrared spectral region using methods like PCA and PSLR [30,31]. While previous studies have shown that physiological responses to thermal stress are significant in the VIS and NIR regions [6,29], they have not identified the specific wavelengths responsible for these changes.

Some recent studies have succeeded in applying machine learning algorithms for vegetation analysis. The Random Forest (RF) algorithm was able to model the water-stress spectral response of a vineyard with an 83.3% accuracy [27]. In a similar study [16], the water-induced stress in lettuce plants was modelled with an Artificial Neural Network (ANN) over two weeks, and the ANN was able to detect, with 80% accuracy, spectral differences from day one, and returned a 92.4% accuracy at the end of the experiment, when signs of stress were already becoming visible. In this regard, the high accuracy returned by machine learning algorithms in the reported studies is something to be considered. Recent advances in computational techniques have demonstrated the effectiveness of machine learning approaches for thermal stress characterisation in agricultural applications [32], while comprehensive reviews highlight the growing importance of machine vision technologies for stress detection in precision agriculture [33].

The evaluation of thermal stress in plants through hyperspectral remote sensing is little explored in the literature. One example is the Cao et al. (2019) investigation. These authors [4] analysed different spectral Vegetation Indices (VIs) to detect thermal stress using their correlation with photosynthetic parameters of plants. Some spectral indices were able to measure the stress response, but with a low or moderate coefficient of determination (R^2) of 0.54 and 0.71, respectively. This approach's accuracy is outstandingly lower compared to the performances returned by machine learning algorithms in other stress investigations [16,27].

Since machine learning has been the focus of recent publications in the hyperspectral data domain, it is important to assess its performance in detecting thermal stress. Furthermore, machine learning, when combined with a ranking feature, can return a metric used as an indication of individual input variables' contribution, as explained by [34]. This type of analysis offers some potential in the hyperspectral domain. Here, we propose a framework to model the reflectance response of thermal-induced stress in bean plants with multiple machine learning algorithms and combine it with a ranking approach. To address this research gap, we propose a novel framework that models the hyperspectral reflectance response of thermal-induced stress in bean plants using a combination of multiple machine learning algorithms and a ranking feature approach. Our contributions are two-fold: we first identify the most effective machine learning algorithms for this specific task, and second, we define the precise spectral regions most sensitive to thermal-stress response, providing valuable insights for future research and applications.

2. Materials and Methods

The proposed framework is arranged into five main phases (Figure 1). In Phase 1, the experiment was appropriately designed in controlled conditions (laboratory) to consider two bean plant species at three soil characteristics. In Phase 2, the spectral data of leaves at three periods were measured: pre-stress, during stress, and post-stress. This phase was repeated twice, once in the reproductive stage and once in the vegetative stage of both species. In Phase 3, the spectral data were processed, converting the radiance signal to leaf reflectance. The low signal-to-noise ratio was removed from the collected spectral data. In Phase 4, the dataset ($n = 1724$) was separated into three parts: training/validation (80%)

and testing (20%). Finally, in Phase 5, the performance of eight machine learning algorithms was measured for classifying the spectral wavelengths according to the measured periods (pre-, during, and post-stress).

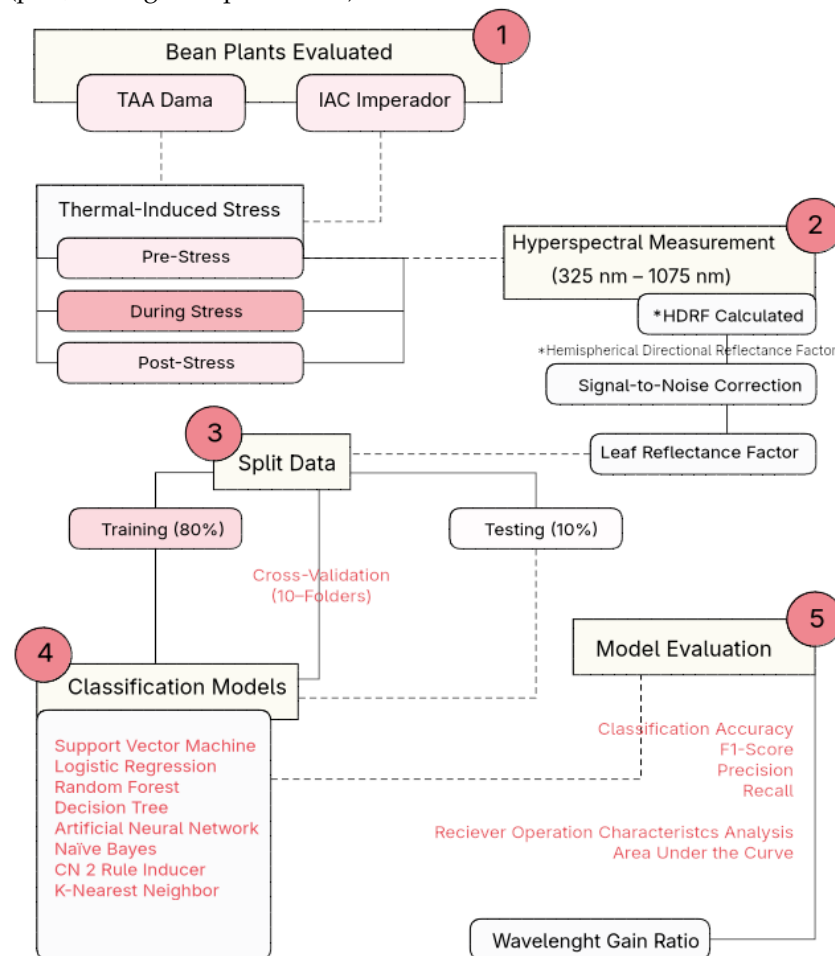


Figure 1. The experiment designed to model the hyperspectral reflectance measurements of thermal-induced stress in bean plants for two growth stages (reproductive and vegetative) through machine learning algorithms.

2.1. Experimental Setup

The experiments were initially carried out in a plant growth chamber (Fitotron®, Weiss Technik Limited, Leicestershire, UK) under controlled conditions to evaluate the effects of the thermal stress on two bean plants' growth stages: reproductive and vegetative. Two Carioca-bean (*Phaseolus vulgaris*) cultivars were used: TAA Dama and IAC Imperador. Both cultivars were placed in pots ($n = 36$) with 2 kg of agricultural soil (pH = 5.9, CaCl_2 0.01 mol L^{-1} , $43.9 \text{ mg} \cdot \text{dm}^{-3}$ of P (Mehlich⁻¹), $2.7 \text{ mmol} \cdot \text{dm}^{-3}$ of K, $25.3 \text{ mmol} \cdot \text{dm}^{-3}$ of Ca, $5.3 \text{ mmol} \cdot \text{dm}^{-3}$ of Mg, and $14.3 \text{ mmol} \cdot \text{dm}^{-3}$ of H + Al). The pots were placed in the growth chamber, where light, humidity, and temperature were monitored.

The experimental design was the same for both species. The plants were grown in pots of equal size, with soil from the same area, in controlled and uniform climatic conditions. These conditions were simulated by the Fitotron® chamber, where temperature, photoperiod, and humidity were equally shared between them. The following growth conditions were established within the chamber: temperature of 20/25 degrees Celsius at night/day and a photoperiod of 8/16 h until the beginning of the reproductive period, with controlled irrigation in luminous conditions ($350 \text{ W} \cdot \text{m}^{-2}$). At the beginning of the reproductive period (R2–R3), thermal stress was established with an increase of 8 degrees at

night/day (28/33 Celsius degrees), lasting 10 days. After this period, the climatic conditions of the chamber were restored.

The second experiment (vegetative stage) was implemented under the same conditions as the first one (reproductive stage), only with the difference that the period of thermal stress was established at the vegetative stage of plants (V3–V4). Both experiments lasted with plant conduction ending at 55 days. Each plant had its spectral readings evaluated before, during, and after stress. This experimental design allowed us to create an heterogeneous dataset, which is important to evaluate the machine learning algorithms' performance in real-life scenarios, building up robust models.

2.2. Spectral Measurement and Data Pre-Processing

The spectral measurements of the leaves were performed by an ASD FieldSpec[®] HandHeld 2 spectroradiometer (Analytical Spectral Devices, Inc., Boulder, CO, USA) with a 1-degree aperture Field of View (FoV), recording wavelengths from 325 to 1075 nm with a spectral resolution of 1.6 nm. The spectral responses of the plants were measured in a dark room covered by dark surfaces, minimizing light interference. Before plant measurement, a Lambertian calibration plate (Spectralon[®] plate) (Labsphere, Inc., North Sutton, NH, USA) was recorded with the ASD equipment. The plants were removed from the chamber and immediately moved to the darkroom. Leaf spectral behaviour was measured at a consecutive number of times for each plant in different leaves. The leaves at an intermediate phase, i.e., neither young nor old, were selected from the middle third of the plant.

Since the experiment was conducted in a dark room with minimal light interference, the reflectance values measured by the equipment were registered at lower than normal levels. The light source used to illuminate the leaves was positioned at a fixed distance between the vases, and the spectroradiometer was also fixed at a variable distance of 12 to 14 cm from the leaves. This variation happened mostly to accommodate differences between plants and leaf distribution, even though most of the plants were similar in structure. The calibration plaque was used multiple times during the experiment, within half an hour between measurements, as specified by its manufacturer. The plaque was also placed upon the same pedestal as the plants, and the distance between the light source (halogen lamp) and the spectroradiometer was adjusted to simulate the distance from the leaves.

A halogen lamp was prepared to illuminate the target, at a 45° angle, while the spectroradiometer was placed on the opposite side, also at a 45° angle. The measurement was repeated during pre-stress, during stress, and post-stress periods, and in two plant stages: reproductive and vegetative. The data were manually divided based on the moment of the measurement, being classified into pre-stress, during stress, and post-stress. This was performed for both species and growth stages (vegetative and reproductive stages). The spectroradiometer recorded the radiance from the leaf that reaches the sensor. To transform this measurement into the leaf reflectance factor, the leaf (i.e., target) radiance was divided by the calibration plate (i.e., reference) radiance. The equipment has a known calibration factor K , which must be multiplied by the result of the first operation [35]. In total, 1724 wavelengths were recorded and used to compile the models. Each of those spectrums was labelled according to its analysis period: pre-stress, during stress, and post-stress.

The experiment was conducted at different stages and periods, which could lead to potential differences between spectra measurements. To remove possible artefacts, the equipment conducted multiple observations and calculated the average of 10 spectra before returning the given values for each leaf. After this, a smoothing technique is applied with a 7-by-7 filter to reduce noise between wavelengths. Lastly, to compensate for changes in the optical path length and reduce scattering issues, the Standard Normal Variate (SNV) was

performed to normalise the data by subtracting each spectrum from its mean and dividing it by its standard deviation. Such procedures were necessary for preparing the data before additional analysis.

2.3. Machine Learning Analysis

From a machine learning perspective, our goal is to solve a supervised multi-class classification problem. Let our dataset be $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$, where N is the total number of spectral measurements ($N = 1724$). Each input sample x_i is a feature vector representing a single hyperspectral measurement. This vector is defined as $x_i = (r_1, r_2, \dots, r_D) \in \mathbb{R}^D$, where r_j is the leaf reflectance value at the j -th spectral band, and D is the total number of spectral bands (wavelengths) recorded by the spectroradiometer from 325 to 1075 nm. Each corresponding label y_i belongs to a discrete set of three thermal stress conditions, such that $Y = \{\text{Pre-Stress, During-Stress, Post-Stress}\}$. The primary objective is to learn a mapping function, or classifier, $f : X \rightarrow Y$. This function is trained on a subset of the data, $\mathcal{D}_{\text{train}}$, and its purpose is to accurately predict the stress label y_{new} for any new, previously unseen hyperspectral sample x_{new} . The effectiveness of the learned function f is then evaluated on a separate test set, $\mathcal{D}_{\text{test}}$, to estimate its generalisation performance.

To ensure a robust evaluation of the models' generalisation capabilities and to prevent data leakage, the dataset partitioning was performed at the pot level. The 36 pots used in the experiment were randomly divided into two independent groups: a training set comprising 80% of the pots (29 pots) and a testing set with the remaining 20% (7 pots). This pot-level split resulted in a training dataset containing 1380 spectral measurements and a testing dataset containing 344 measurements. All spectral samples from the training pots were used exclusively for model training and validation, while all samples from the testing pots were reserved for the final performance evaluation. This pot-level splitting strategy guarantees that the models were tested on entirely new plants that they had not been exposed to during the training phase. This approach provides a more realistic and conservative estimate of the models' performance in a real-world scenario, as it tests their ability to generalize to unseen individuals rather than simply recognising the unique spectral characteristics of plants seen during training.

The cross-validation method, with 10 k-folders, was adopted to train and validate each algorithm. In this method, one of the folders is separated and used as a validation dataset, while the other folders are used to train the model. This is repeated 10 times until all the folders have been used once. The validation dataset is used to calculate the performance of the classifiers during the hyperparametrisation process of the algorithms. For each algorithm, we repeated this procedure 10 times, resulting in a total of 100 repetitions. The open-source software Weka 3.9.5 was used in our analysis, which runs in its Java library.

The hyperparametrisation process considered the individual characteristics of the algorithms, like the number of trees and nodes, the number of neurons, interactions, and function degree, among others. The criteria were defined, but they did not return any practical gains for the Overall Accuracy (OA) since they only increased the processing time needed. The parameters of the methods applied here have been set to their respective library default values, except those described below (Table 1). Details of the parameters and criteria defined by the hyperparametrisation of the eight evaluated algorithms are summarised in Table 1. The machine learning algorithms used were as follows: CN2 Rule Induction (CN2), k-Nearest Neighbors (kNN), Logistic Regression (LR), Naive Bayes (NB), ANN, RF, SVM, and Decision Tree (DT), as they represent the state-of-the-art machine learning algorithms and are available on Weka 3.9.5 software.

Table 1. Algorithms evaluated in this study and their hyperparameter and criteria values.

Algorithm	Hyperparameter	Criteria
CN2	Rule Ordered Measurement: Laplace Accuracy	Beam = 4 Rule Coverage (minimum) = 5 Rule Length (maximum) = 10
kNN	Euclidean Distance	k-Neighbours = 5
LR	Regularisation Strength	Ridge (L2) C = 1
NB	None	None
ANN	Activation Linear Adam Solver Regularisation a = 0.0001	Neurons (Hidden Layer) = 512 Number of Hidden Layers = 2 Epochs = 200
SVM	Radial Basis Function (RBF) Kernel $\exp(-g x-y ^2)$ g = automatic SVM Type Cost = 1 Regression Loss = 50.00	Tolerance = 0.0010 Interaction (limit) = 500
RF	Number of Trees Nodes	Trees = 200 Nodes (maximum) = 5
DT	Number of Leaves Trees depth	Leaves (minimal) = 2 Tree-depth (maximum) = 100

The CN2 classifier is an algorithm designed to classify a set of conditions that has been established. This algorithm is based on the same function as the AQ and the ID3 algorithms, and it uses a rule set similar to AQ, but can handle noisy data like the ID3 [36,37]. The kNN verifies how closely a given data point is in a feature space. In classification tasks, it uses a set of weights and distance metrics to classify an object by a plurality vote of its neighbours [38]. The LR algorithm is based on a sigmoid function, where it handles classification problems with the concept of probability. Although it has been most used for binary classification, LR can be modified to run in an all-against-one approach and evaluate differences among more classes [39]. NB is also a probabilistic approach based on Bayes' theorem. Since it uses a naïve approach, it can completely disregard the correlation between the input variables [39,40].

The ANN machine learning algorithm is inspired by biological neurons and works in multiple connections capable of learning linear and non-linear models. It uses hidden layers to perform the classification task in a feed-forward manner and depends on an activation function for the hidden layer and a solver for weight optimisation [40,41]. In this study, two hidden layers were implemented, while the linear activation function was used for the final layer, since once this function is used, no additional layers are necessary for the network. The SVM is a method that separates an attribute space within a hyperplane. It then calculates a linear function while maximising the margins between instances of different classes [42]. RF and DT both rely on the idea that accuracy can be improved when implementing the results of combined independent predictions. While DT is based on an individual tree with different numbers of nodes and leaves, the RF model is based on multiple DTs [43].

Regarding testing, for which we used 20% of the whole dataset, this set had not been “seen” before by the learners. This ensured the measurement of the actual performance of each machine learning algorithm in the proposed task. The classification result of each algorithm was evaluated with traditional metrics like overall accuracy, F1-score, precision, recall, and specificity. The Receiver Operating Characteristics (ROC) curve and its respective

Area Under the Curve (AUC) values for each class (pre-stress, during stress, and post-stress) were also calculated to compare algorithms' performance.

The evaluation of these metrics allowed the construction of a rank-based approach [34] with the most suitable algorithms for modelling the thermal-stress response of the hyperspectral data. Thus, the contribution of each wavelength to the classification task performed by the algorithms was calculated. This contribution was obtained from the wavelength gain ratio used by the most suitable models (which returned a higher metric than 90% in overall accuracy). For that, values of the gain ratio were plotted and were used to define the spectral regions more responsible for indicating the thermal-stress effect of the induced stress.

A qualitative approach was used to compare the mean spectral wavelengths of each period (pre-, during, and post-stress) in the spectral regions returned by the gain ratio. This comparison was important to ascertain the spectral behaviour of each period and its differences. Aiming to deal with the high dimensionality of the spectral data, a dimension reduction analysis based on PCA was implemented as a pre-processing step. However, during this experimental phase, no satisfactory results were observed, mainly because it has a low impact on most algorithms' performances. This approach could also impact the ranking analysis implemented in the latter stage, and so the full reflectance data was preferred.

3. Results

3.1. Impact of Thermal Stress into Reflectance Measurements

The performance of multiple machine learning algorithms was evaluated for modelling the hyperspectral reflectance measurement of two bean plant cultivars under three periods of thermal-induced (pre-stress, during stress, and post-stress). This experimental variation was important to produce different spectral behaviours, resulting in a heterogeneous dataset. Our results demonstrated the high effectiveness of a selective group (ANN, CN2, DT, and SVM) of the tested algorithms to classify the wavelengths at their different stress moments (pre-stress, during stress, and post-stress), reaching an overall accuracy of 99.4% in some cases.

The algorithms ANN, CN2, DT, and SVM achieved higher overall accuracies in separating spectral measurements from the pre-, during, and post-thermal-induced stress periods, which reached 99.4%, 94.7%, 93.6%, and 90.1%, respectively (Table 2). The NB, LR, and RF performed worse, returning almost a random guess type of analysis (i.e., chances next to 50%). These algorithms operated in different ways, with the advantage of being non-parametric [44]. The combinations and logical operations executed by the algorithms of ANN and CN2 are appropriate to evaluate imperfect and noisy data [37].

Table 2. Results of the classification task performed by the algorithms.

Algorithm	OA (%)	F1-Score (%)	Precision (%)	Recall (%)	Specificity (%)
CN2	94.77	94.78	94.90	94.70	97.20
kNN	87.20	87.01	88.20	87.20	93.30
LR	59.30	55.16	61.40	59.30	80.90
NB	63.37	61.46	66.70	63.40	81.60
ANN	99.42	99.42	99.40	99.40	99.70
RF	65.70	63.21	73.30	65.70	83.10
SVM	90.11	90.01	91.70	90.10	95.60
DT	93.60	93.64	93.80	93.60	97.00

Although the DT algorithm may be unstable in some classification tasks [43], it showed high performance in classifying the three periods evaluated in this study. The SVM algorithm, although presenting high overall accuracy (90.3%) (Table 2), performed worse in discriminating the “post-stress” phase (Figure 2) and presented errors in classifying the post-stressed bean plant spectral wavelengths. The ANN algorithm presented the best performance, being able to classify both “pre-stress” and “during-stress” classes with almost complete accuracy (Figure 2).

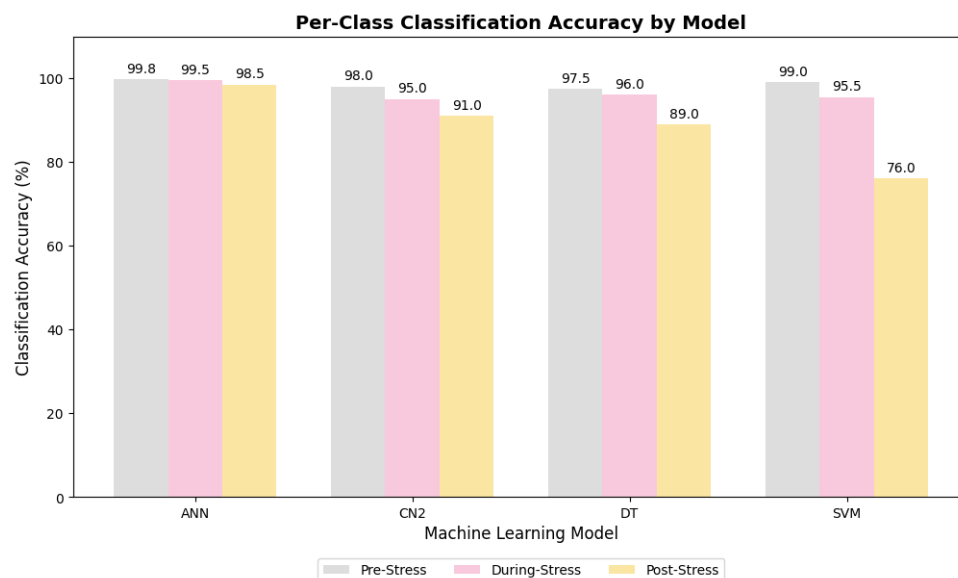


Figure 2. Per-class classification accuracy for the four best-performing models. Each bar represents the accuracy for correctly classifying samples belonging to the pre-stress, during stress, or post-stress conditions.

The performance analysis of the ANN, CN2, DT, and SVM algorithms through the ROC graphic and AUC values demonstrated that, in general, they presented better results for the “pre-stress” condition (Figure 3). ANN’s ROC graphic shows how this algorithm was able to differentiate all classes with an AUC of 0.999. The algorithms of CN2 and DT presented lower true-positive rates with “during-stress” and “post-stress” classes, respectively. The SVM, however, was worse, returning higher false-positive rates for these two classes.

The ANN had a great performance even in comparison with the other algorithms. It was able to classify with almost complete accuracy the spectral wavelengths of the “pre-stress” and “during-stress” phases and returned approximately 98% accuracy in the “post-stress” phase (Figure 2). It also demonstrated almost no false-positive rates for these classes in any momentum, unlike the other algorithms. The ANN algorithm, based on a multi-layer perceptron, proved to be sufficient to model the spectral response, not requiring deeper networks. Recently, deep neural networks have been highly capable of performing improved estimates than shallow networks. However, the amount of labelled data required far surpasses the needs of a machine learning model.

The remaining algorithms also returned satisfactory classification metrics, being, alongside the ANN, CN2 rule, DT, and SVM, the most prominent ones (Figure 2). As for the other algorithms, the performance varied. While this may not be a common outcome with traditional shallow learners, it was observed in previous studies conducted with hyperspectral data for different types of vegetation-related classification tasks [16,45–47]. As aforementioned in the Method section, we conducted a PCA as a pre-processing step in an experimental phase, and while it did help return similar performances for the LR, NB,

and RF models, the impact was not sufficient to surpass the top four algorithms (ANN, CN2 rule, DT, and SVM). Another important observation is that SVM, which is considered a powerful method in shallow learning comparison, should theoretically benefit from a dimension reduction analysis like PCA. However, the performance was not impacted by its usage, maintaining the overall values observed with the original reflectance data.

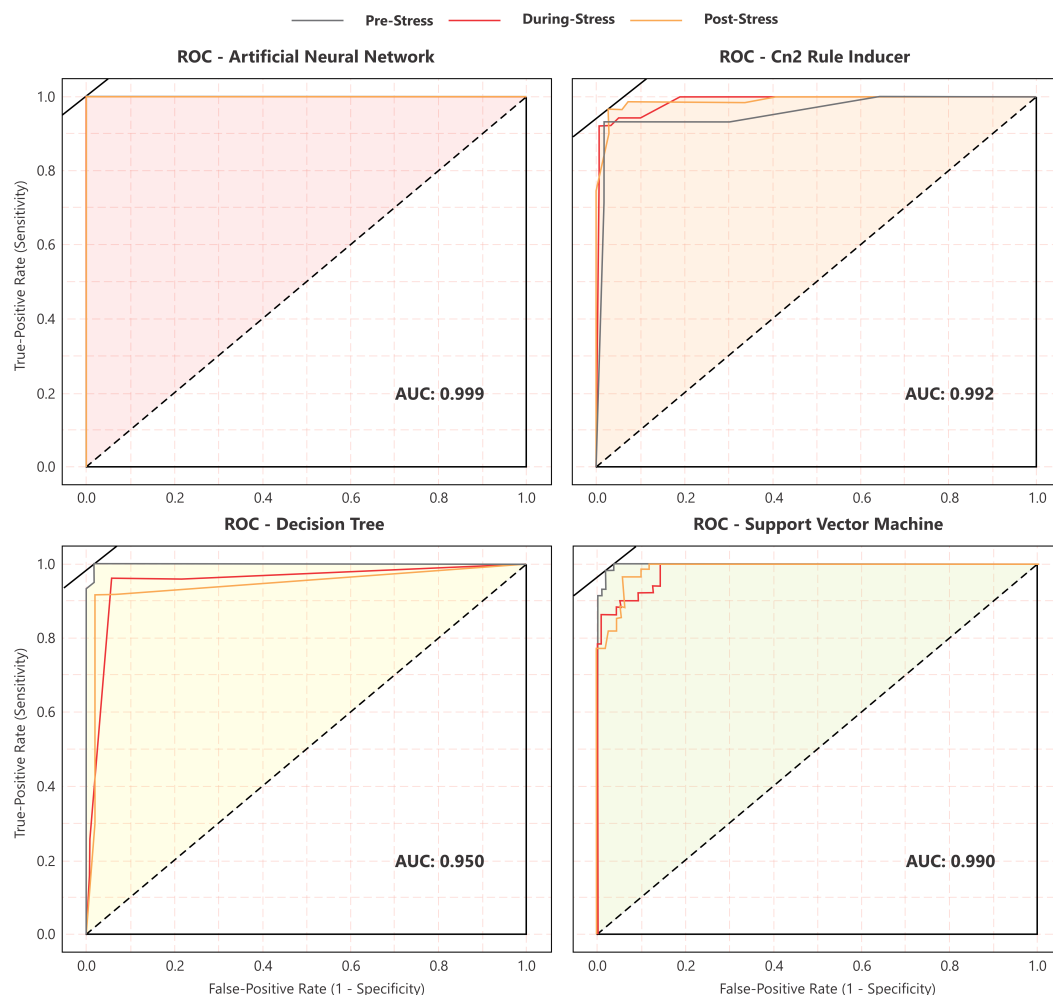


Figure 3. ROC and AUC curves of the most suitable algorithms indicating their performance in each class.

In this analysis, the four best learners (Figure 2) were considered as the most prominent methods to model the hyperspectral response of the investigated thermal stress. It is also important to note the difference between the accuracies of the methods since some of them were not able to properly understand the patterns demonstrated by the different classes. The ANN, which achieved the highest evaluation metrics, has already presented potential in previous research for evaluating water-stress effects on hyperspectral measurements [16]. Regardless, as mentioned, some of the remaining algorithms were also feasible for the given task. In this sense, it is still early to determine which algorithm may perform better given a more complete and generative dataset, and future research could focus on observing those issues.

3.2. Ranking the Contribution of Spectral Wavelengths

Another investigation of our framework proposal regards the identification of spectral regions and wavelengths that had the highest contribution to model the thermal-stress response. Overall, the induced stress increased the reflectance, especially in the green,

red-edge, and NIR regions (Figure 4). This pattern continued even after returning the temperature to the pre-existing conditions. This demonstrates how the effects of thermal stress can occur in the spectral behaviour of the plant even after the stress period. The green region is often linked to foliar pigmentation [4,48,49], while red-edge changes are associated with plant stress [49,50]. The NIR region is related to the structuring and organisation of molecules in the mesophyll [49,51]. During the induced thermal stress, the plants showed wilting signals and chlorosis symptoms, which may explain the differences observed here.

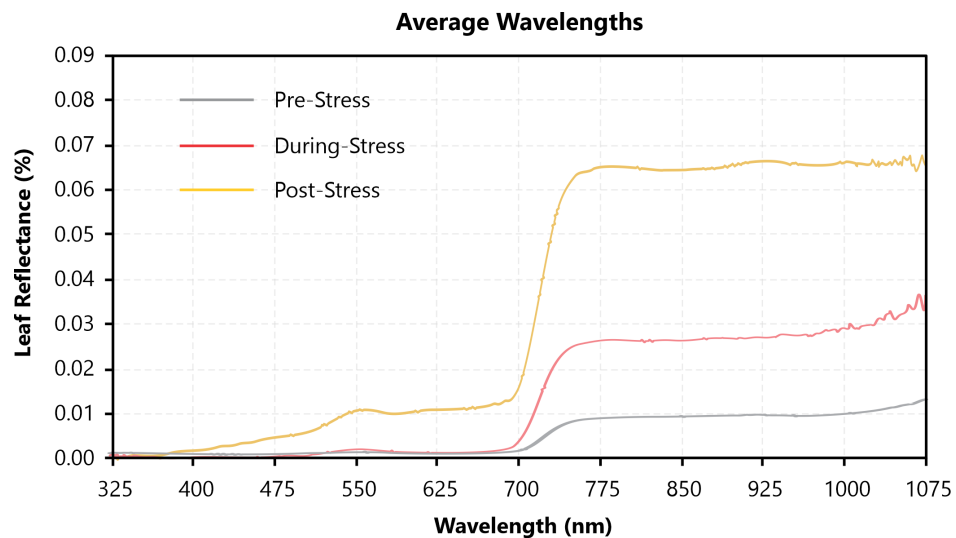


Figure 4. Average spectral wavelengths of each condition: pre-stress, during stress, and post-stress.

To indicate the most constitutive spectral wavelengths to the overall best-chosen models' performance, the gain ratio metric was used. The gain ratio is calculated concerning the accuracy returned by the spectral wavelength alone and in combination with others when used as input for the given model. The proposed approach resulted in different gain ratios for the analysed wavelengths (Figure 5). In the visible spectral region, the major contribution to the algorithms came from wavelengths in the green region, between 530 and 570 nm. In the red edge, a smaller region, ranging from 700 to 710 nm, had a similar contribution. This indicates how important both spectral regions were in differentiating the measurements before, during, and after the thermal stress for both cultivar types. The blue (and ultra-blue) and red regions (375–500 nm and 600–700 nm, respectively) presented a lower and noisier gain ratio than the others.

Furthermore, a qualitative evaluation of the mean reflectance value of each thermal-stress period highlighted differences in the leaf reflectance measurements (Figure 6). The amplitude between spectral wavelengths in the green region shows an increase in the reflectance factor as the thermal stress was applied. This effect continues to persist even afterwards in the stress phase. In the red-edge region, the amplitude between wavelengths was also high. However, it also increases even after the 710 nm range. This region, despite not achieving the same contribution as a range of 700–710 nm, still presented a more stable behaviour than other wavelengths, forming a “plateau” in terms of gain ratio (Figure 6).

The best contribution to the algorithm's classification originated from the spectral regions of 530 to 570 nm and 700 to 710 nm. These regions presented a high amplitude among themselves when their mean curves were evaluated (Figure 6). Machine learning possesses the advantage of working non-linearly and non-parametrically. Also, the models are capable of performing various calculations and combinations in a matter of seconds [37,40]. These combinations indicated that the green and red-edge regions showed here can explain the phenomenon observed with better sensitivity than others, constituting a more correlated indicator (Figure 5). This information is important as it can be used for

the creation of simpler models or spectral indices able to model the response of thermal stress by focusing on these wavelengths.

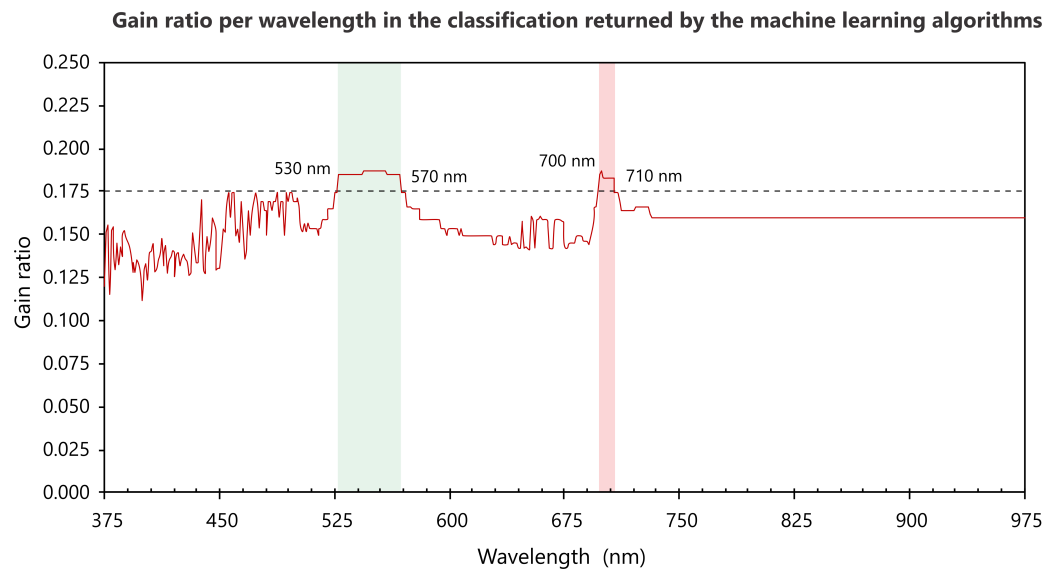


Figure 5. Gain ratio per wavelength in the classification returned by the machine learning algorithms. The shaded areas highlight the most contributory spectral regions: green (530–570 nm) and red-edge (700–710 nm).

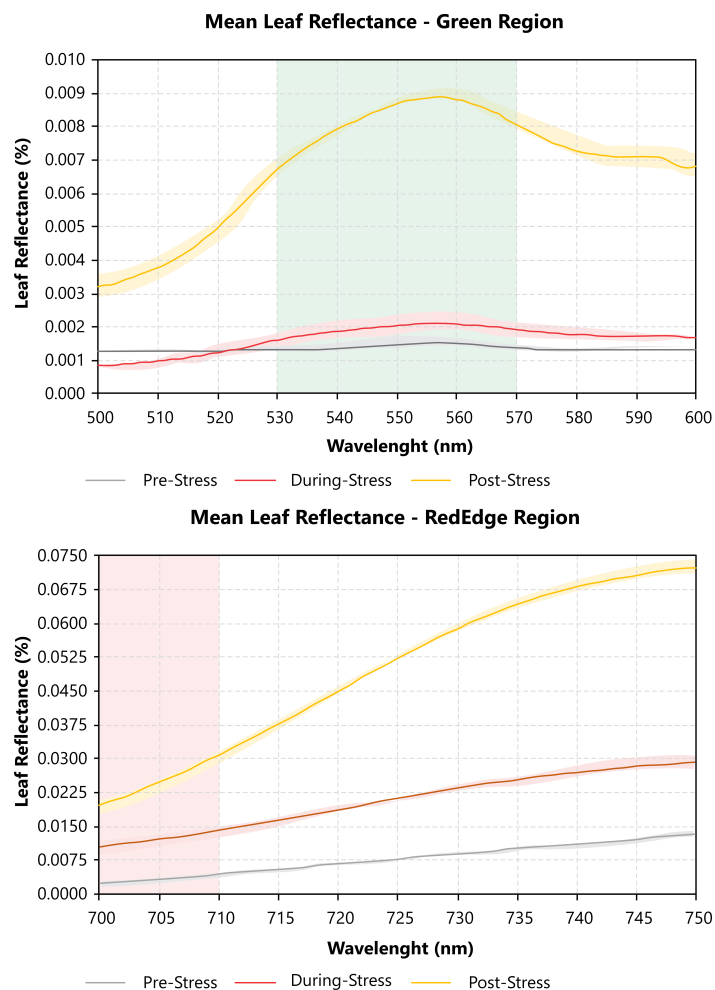


Figure 6. Mean leaf reflectance in the green- and red-edge regions between the analysed periods.

Studies have shown that, under certain stresses, leaf reflectance can be increased in the visible and NIR regions [11,19,28,29,31]. This shows how these spectral regions are linked with stress response in plants. The present study indicated how much of this contribution is attainable since the proposed approach to multiple machine learning algorithms returned overall accuracies up to 99.4%. This is important as it allows future research to model the spectral behaviour of plants that have experienced this type of stress. This behaviour continued as the amount of energy reflected by the leaf increased in the last phases, exceeding values measured before and during stress induction.

Machine learning algorithms are important tools in the analysis of plant spectral behaviour. From the dataset constructed here, it was possible to estimate with feasible overall accuracy the spectral behaviour of the measurements. Although limited to the visible to NIR regions, this study demonstrates the potential of these algorithms to classify whether the plant is under stress or has suffered thermal stress. However, other studies may use the framework proposed here for different plant species. In bean plants, it is recommended that future research focus on the green (530–570 nm) and red-edge (700–710 nm) wavelengths or vegetation indices that adopt both spectral regions to improve methods and operations that model the thermal-stress response in plants.

4. Discussion

The high-performance results of our machine learning framework for detecting thermal stress in bean plants offer significant insights into the capabilities of these models and the physiological changes captured by hyperspectral data. The superior performance of the Artificial Neural Network (ANN), CN2 Rule Induction, Decision Tree (DT), and Support Vector Machine (SVM) algorithms, with overall accuracies ranging from 90.1% to 99.4%, stands as a crucial finding. The ANN's accuracy in separating the different stress periods underscores its exceptional ability to identify complex, non-linear patterns within high-dimensional hyperspectral data. A key practical implication of this is that for this specific application, a sophisticated, multi-layer perceptron architecture is sufficient, negating the need for more computationally intensive deep learning networks that typically require much larger datasets.

The strong performance of the rule-based (CN2) and tree-based (DT) algorithms is also noteworthy. Their ability to achieve high accuracies (94.7% and 93.6%, respectively) suggests that the relationship between thermal stress and spectral response can be effectively modelled through logical rules and hierarchical decisions. This characteristic makes these models highly interpretable, a significant advantage in agricultural applications where understanding the underlying factors driving the classification is important for decision-making. However, the ANN's slight edge over these well-performing models likely stems from its capacity to model highly complex, non-linear relationships inherent in hyperspectral data. Unlike rule-based (CN2) or tree-based (DT) approaches that rely on hierarchical partitioning or explicit rule generation based on specific feature thresholds, the ANN's multi-layer architecture allows it to learn subtle patterns and interactions across a wide range of wavelengths simultaneously through weighted connections. This implicit feature weighting mechanism could have been particularly advantageous for our data, where the relevant stress signal could be encoded in nuanced variations across multiple correlated bands rather than being dependent on a few discrete wavelengths, allowing the ANN to capture the subtle signature of thermal stress with slightly higher fidelity. In contrast, the SVM, while generally a powerful shallow learner, showed a specific weakness in discriminating the "post-stress" phase, resulting in higher false-positive rates. This may be attributed to the more subtle and variable spectral changes that occur after the

stress period has ended, which could make it difficult for the SVM's hyperplane-based classification to consistently separate this class from the others.

A noteworthy result was the contrast in performance between the ANN (99.4% accuracy) and Random Forest (RF) (65.7% accuracy). While RF is often highly effective for high-dimensional, collinear data like hyperspectral measurements, precisely because its ensemble nature and feature subsampling can mitigate overfitting and handle redundancy, its suboptimal performance in this specific instance warrants consideration. One strong possibility lies in hyperparameter sensitivity. Achieving optimal RF performance with over 1700 spectral features often requires extensive tuning of parameters like the number of features considered at each split, tree depth, and the number of estimators. While our tuning was systematic, it is possible that the chosen parameters, representing a relatively standard configuration, were not ideal for capturing the specific nuances differentiating the thermal stress classes in this particular dataset. Given that our goal was primarily to compare the baseline effectiveness of various algorithms and other models that had already achieved high accuracy, we did not pursue more computationally intensive hyperparameter optimisation techniques, such as grid search or random search, which might have potentially improved the RF results.

The ranking of spectral wavelengths via the gain-ratio metric provided a key contribution to our study. The consistent importance of the green- (530–570 nm) and red-edge (700–710 nm) regions confirms their physiological significance as indicators of thermal stress. The observed increase in leaf reflectance within the green region during and after stress induction is directly linked to the breakdown of chlorophyll pigments, a common physiological response to high temperatures. Similarly, the changes in the red-edge region, a well-established indicator of plant stress, are consistent with alterations in the plants' photosynthetic efficiency and overall health. Our findings provide a much more specific wavelength range (700–710 nm) than previous studies, which often broadly identified the VIS and NIR regions. This precision is invaluable for future research aimed at developing targeted spectral indices or simpler models for thermal stress detection.

A critical question arising from these findings is whether the response in the green- and red-edge regions is specific to thermal stress or represents a more generalised plant stress response. Physiologically, the increased reflectance in the green region is directly linked to chlorophyll degradation, while shifts in the red-edge are tied to changes in both chlorophyll content and leaf cellular structure. These are common physiological reactions to a wide array of stressors, including water deficit, nutrient deficiency, and disease.

While the underlying biological response may be general, the exceptional accuracy of our framework suggests that the machine learning models were able to capture subtle, high-dimensional patterns within these bands that are characteristic of thermal stress. It is plausible that the specific magnitude, rate of change, and the combined pattern across the green- and red-edge regions provided a unique signature for the thermal stress conditions in our experiment. For example, the rapid onset of thermal stress might induce a spectral shift with a different dynamic than the slower progression of nutrient deficiency. To definitively isolate the unique spectral signature of thermal stress, future research should include controlled experiments that directly compare it against other common agricultural stresses on the same crop.

Our results demonstrate a marked improvement over traditional methods for thermal stress detection, which have shown low to moderate coefficients of determination (R^2) of 0.54 and 0.71. The ability of our machine learning framework to achieve accuracies up to 99.4% highlights its feasibility and potential for practical applications in precision agriculture. The framework can model and classify the spectral behaviour of plants that

have undergone thermal stress, even after the stress period has ended, when the reflected energy continues to increase.

Despite the promising results, this study was conducted under controlled laboratory conditions, and scaling this approach to real-world agricultural fields presents several significant challenges. Beyond the influence of soil background and mixed canopy signals, transitioning to a field environment introduces critical complexities. First, the variability in illumination is a primary obstacle. Our controlled setup used a stable, fixed light source, whereas field measurements are subject to changes in sun angle, cloud cover, and atmospheric conditions like haze, all of which alter the spectral signal. Second, the target itself is far more complex; a sensor captures an integrated signal from a 3D canopy with varying leaf angles, wind-induced movement, and intricate self-shadowing. Finally, biological factors are rarely isolated in the field, as plants may experience a combination of thermal, water, and nutrient stress simultaneously, creating confounding spectral signatures.

Therefore, a critical next step is to validate this framework using data from Unmanned Aerial Vehicles (UAVs). UAVs offer a practical balance, providing the high spatial resolution needed to mitigate some of these challenges while still allowing for field-scale analysis. The integration of UAV-based hyperspectral imaging with machine learning, as demonstrated for applications like early yield prediction [52] and disease/pest detection using deep learning [53], highlights the potential and trajectory of this technology in intelligent agriculture. However, for monitoring very large and potentially arid regions, where UAV deployment can be inefficient, satellite remote sensing becomes necessary. This approach is not without its own limitations, such as lower image resolution and frequent obstruction by cloud cover. In such large-scale contexts, integrating multiple data sources may be required. For example, methodologies using radar data, such as those explored by [54], offer a robust alternative for crop monitoring because they are unaffected by clouds. Future research could explore a multi-sensor approach, combining the detailed biochemical information from UAV-based hyperspectral data with the broad, all-weather coverage of satellite radar for a comprehensive stress monitoring system.

5. Conclusions

This study demonstrates a framework using machine learning to model the hyperspectral leaf reflectance of thermal-stressed bean plants. Data from two species of bean plants, at two plant stages, and three stress conditions (before, during, and after said stress), were used to compose our dataset. The Artificial Neural Network (ANN) proved to be the most effective algorithm, achieving an overall accuracy of over 99%. Furthermore, the high performance of other models such as CN2 rule induction, decision tree, and support vector machine, confirms the viability of various machine learning approaches for this critical task. A key outcome of our study is the definitive identification of the green- (530–570 nm) and red-edge (700–710 nm) spectral regions as the most suitable ranges for discriminating thermal-induced stress in bean plants.

In summary, our work provides a robust data-driven methodology for the early and accurate detection of thermal stress in bean plants. The specific spectral ranges identified can serve as a foundation for developing new, targeted spectral indices and portable sensing devices. Specifically, future work should focus on formulating and testing a novel vegetation index based on the 530–570 nm (green-edge) and 700–710 nm (red-edge) regions, designed for enhanced sensitivity to thermal stress compared to existing broadband indices. Validating such an index in field conditions could provide a cost-effective tool for practical stress monitoring, bridging the gap between high-dimensional hyperspectral analysis and more accessible multispectral sensors. We recommend that this framework be applied to

other crop species and validated in real-world field conditions to fully assess its potential for large-scale agricultural monitoring and management.

Author Contributions: Conceptualisation, L.P.O., A.P.M.R. and F.F.d.A.; methodology, L.P.O. and É.A.S.M.; software, L.P.O., J.M.J. and W.N.G.; validation, L.A.d.C.J. and V.L.; investigation, É.A.S.M. and B.C.d.L.; resources, N.N.I. and F.F.d.A.; data curation, É.A.S.M.; writing—original draft preparation, L.P.O., É.A.S.M. and B.C.d.L.; writing—review and editing, A.P.M.R., J.M.J., W.N.G., L.A.d.C.J., V.L., J.L., A.S.F.d.A., N.N.I. and F.F.d.A.; visualisation, L.P.O.; supervision, A.P.M.R., L.A.d.C.J., V.L., J.L. and A.S.F.d.A.; project administration, J.M.J., W.N.G., N.N.I. and F.F.d.A.; funding acquisition, N.N.I. and F.F.d.A. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by the Coordination for the Improvement of Higher Education Personnel (CAPES; Finance Code 001), FAPESP [grant number 2017/16202-2], FUNDUNESP/Print (grant number 3030/2019), FAPESC (grant number 2017TR1762, 2019TR816), and CNPq (grant number 310128/2018-8, 310517/2020-6, 310145/2020-1, 308747/2021-6, 305296/2022-1, 308481/2022-4, 305814/2023-0, 403213/2023-1, and 312816/2025-1).

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to institutional data sharing policies.

Conflicts of Interest: Author Lucio André de Castro Jorge is an employee at the Brazilian Agricultural Research Corporation (EMBRAPA). The remaining authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Letcher, T.M. Why do we have global warming? In *Managing Global Warming*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 3–15.
2. Moriarty, P.; Honnery, D. Global renewable energy resources and use in 2050. In *Managing Global Warming*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 221–235.
3. Badzmierowski, M.J.; McCall, D.S.; Evanylo, G. Using Hyperspectral and Multispectral Indices to Detect Water Stress for an Urban Turfgrass System. *Agronomy* **2019**, *9*, 439. [[CrossRef](#)]
4. Cao, Z.; Yao, X.; Liu, H.; Liu, B.; Cheng, T.; Tian, Y.; Cao, W.; Zhu, Y. Comparison of the abilities of vegetation indices and photosynthetic parameters to detect heat stress in wheat. *Agric. For. Meteorol.* **2019**, *265*, 121–136. [[CrossRef](#)]
5. Gerhards, M.; Schlerf, M.; Mallick, K.; Udelhoven, T. Challenges and Future Perspectives of Multi-/Hyperspectral Thermal Infrared Remote Sensing for Crop Water-Stress Detection: A Review. *Remote Sens.* **2019**, *11*, 1240. [[CrossRef](#)]
6. Krishna, G.; Sahoo, R.N.; Singh, P.; Bajpai, V.; Patra, H.; Kumar, S.; Dandapani, R.; Gupta, V.K.; Viswanathan, C.; Ahmad, T.; et al. Comparison of various modelling approaches for water deficit stress monitoring in rice crop through hyperspectral remote sensing. *Agric. Water Manag.* **2019**, *213*, 231–244. [[CrossRef](#)]
7. Egerer, M.H.; Lin, B.B.; Threlfall, C.G.; Kendal, D. Temperature variability influences urban garden plant richness and gardener water use behavior, but not planting decisions. *Sci. Total Environ.* **2019**, *646*, 111–120. [[CrossRef](#)] [[PubMed](#)]
8. Khanna, R.; Schmid, L.; Walter, A.; Nieto, J.; Siegwart, R.; Liebisch, F. A spatio temporal spectral framework for plant stress phenotyping. *Plant Methods* **2019**, *15*, 1–18. [[CrossRef](#)] [[PubMed](#)]
9. Maimaitiyiming, M.; Ghulam, A.; Bozzolo, A.; Wilkins, J.L.; Kwasniewski, M.T. Early Detection of Plant Physiological Responses to Different Levels of Water Stress Using Reflectance Spectroscopy. *Remote Sens.* **2017**, *9*, 745. [[CrossRef](#)]
10. Karadağ, K.; Tenekeci, M.E.; Taştaltın, R.; Bilgili, A. Detection of pepper fusarium disease using machine learning algorithms based on spectral reflectance. *Sustain. Comput. Inform. Syst.* **2020**, *28*, 100299. [[CrossRef](#)]
11. Abdullah, H.; Skidmore, A.K.; Darvishzadeh, R.; Heurich, M. Timing of red-edge and shortwave infrared reflectance critical for early stress detection induced by bark beetle (*Ips typographus*, L.) attack. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101900. [[CrossRef](#)]
12. Chen, J.; Li, F.; Wang, R.; Fan, Y.; Raza, M.A.; Liu, Q.; Wang, Z.; Cheng, Y.; Wu, X.; Yang, F.; et al. Estimation of nitrogen and carbon content from soybean leaf reflectance spectra using wavelet analysis under shade stress. *Comput. Electron. Agric.* **2019**, *156*, 482–489. [[CrossRef](#)]
13. Sonobe, R.; Yamashita, H.; Mihara, H.; Morita, A.; Ikka, T. Hyperspectral reflectance sensing for quantifying leaf chlorophyll content in wasabi leaves using spectral pre-processing techniques and machine learning algorithms. *Int. J. Remote Sens.* **2021**, *42*, 1311–1329. [[CrossRef](#)]

14. Zhou, W.; Zhang, J.; Zou, M.; Liu, X.; Du, X.; Wang, Q.; Liu, Y.; Li, J. Prediction of cadmium concentration in brown rice before harvest by hyperspectral remote sensing. *Environ. Sci. Pollut. Res.* **2018**, *26*, 1848–1856. [[CrossRef](#)] [[PubMed](#)]
15. Rocha, A.D.; Groen, T.A.; Skidmore, A.K. Spatially-explicit modelling with support of hyperspectral data can improve prediction of plant traits. *Remote Sens. Environ.* **2019**, *231*, 111200. [[CrossRef](#)]
16. Osco, L.P.; Ramos, A.P.M.; Moriya, É.A.S.; Bavaresco, L.G.; de Lima, B.C.; Estrabis, N.; Pereira, D.R.; Creste, J.E.; Júnior, J.M.; Gonçalves, W.N.; et al. Modeling Hyperspectral Response of Water-Stress Induced Lettuce Plants Using Artificial Neural Networks. *Remote Sens.* **2019**, *11*, 2797. [[CrossRef](#)]
17. Virnodkar, S.S.; Pachghare, V.K.; Patil, V.C.; Jha, S.K. Remote sensing and machine learning for crop water stress determination in various crops: A critical review. *Precis. Agric.* **2020**, *21*, 1121–1155. [[CrossRef](#)]
18. Feng, X.; Zhan, Y.; Wang, Q.; Yang, X.; Yu, C.; Wang, H.; Tang, Z.Y.; Jiang, D.; Peng, C.; He, Y. Hyperspectral imaging combined with machine learning as a tool to obtain high-throughput plant salt-stress phenotyping. *Plant J.* **2020**, *101*, 1448–1461. [[CrossRef](#)]
19. Zhai, Y.; Cui, L.; Zhou, X.; Gao, Y.; Fei, T.; Gao, W. Estimation of nitrogen, phosphorus, and potassium contents in the leaves of different plants using laboratory-based visible and near-infrared reflectance spectroscopy: Comparison of partial least-square regression and support vector machine regression methods. *Int. J. Remote Sens.* **2012**, *34*, 2502–2518.
20. Dmitriev, P.A.; Dmitrieva, A.A.; Kozlovsky, B.L. Random Reflectance: A new hyperspectral data preprocessing method for improving the accuracy of machine learning algorithms. *AgriEngineering* **2025**, *7*, 90. [[CrossRef](#)]
21. Weinstein, B.G.; Marconi, S.; Bohlman, S.; Zare, A.; White, E. Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks. *Remote Sens.* **2019**, *11*, 1309. [[CrossRef](#)]
22. Ahmed, N.; Mahmud, S.; Elahi, M.M.L.; Ahmed, S.; Sujaudhin, M. Forecasting river sediment deposition through satellite image driven unsupervised machine learning techniques. *Remote Sens. Appl. Soc. Environ.* **2019**, *13*, 435–444. [[CrossRef](#)]
23. Zhang, J.; Huang, Y.; Reddy, K.N.; Wang, B. Assessing crop damage from dicamba on non-dicamba-tolerant soybean by hyperspectral imaging through machine learning. *Pest Manag. Sci.* **2019**, *75*, 3260–3272. [[CrossRef](#)] [[PubMed](#)]
24. Gao, J.; Meng, B.; Liang, T.; Feng, Q.; Ge, J.; Yin, J.; Wu, C.; Cui, X.; Hou, M.; Liu, J.; et al. Modeling alpine grassland forage phosphorus based on hyperspectral remote sensing and a multi-factor machine learning algorithm in the east of Tibetan Plateau, China. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 104–117. [[CrossRef](#)]
25. Oliveira, A.K.d.S.; Rizzo, R.; Silva, C.A.A.C.; Ré, N.C.; Caron, M.L.; Fiorio, P.R. Prediction of corn leaf nitrogen content in a tropical region using Vis-NIR-SWIR spectroscopy. *AgriEngineering* **2024**, *6*, 4135–4153. [[CrossRef](#)]
26. Osco, L.P.; Ramos, A.P.M.; Pereira, D.R.; Moriya, É.A.S.; Imai, N.N.; Matsubara, E.T.; Estrabis, N.; de Souza, M.; Junior, J.M.; Gonçalves, W.N.; et al. Predicting Canopy Nitrogen Content in Citrus-Trees Using Random Forest Algorithm Associated to Spectral Vegetation Indices from UAV-Imagery. *Remote Sens.* **2019**, *11*, 2925. [[CrossRef](#)]
27. Loggenberg, K.; Strever, A.; Greyling, B.; Poona, N. Modelling Water Stress in a Shiraz Vineyard Using Hyperspectral Imaging and Machine Learning. *Remote Sens.* **2018**, *10*, 202. [[CrossRef](#)]
28. Asaari, M.S.M.; Mertens, S.; Dhondt, S.; Inzé, D.; Wuyts, N.; Scheunders, P. Analysis of hyperspectral images for detection of drought stress and recovery in maize plants in a high-throughput phenotyping platform. *Comput. Electron. Agric.* **2019**, *162*, 749–758. [[CrossRef](#)]
29. Elvanidi, A.; Katsoulas, N.; Ferentinos, K.P.; Bartzanas, T.; Kittas, C. Hyperspectral machine vision as a tool for water stress severity assessment in soilless tomato crop. *Biosyst. Eng.* **2018**, *165*, 25–35. [[CrossRef](#)]
30. Gerhards, M.; Schlerf, M.; Rascher, U.; Udelhoven, T.; Juszczak, R.; Alberti, G.; Miglietta, F.; Inoue, Y. Analysis of Airborne Optical and Thermal Imagery for Detection of Water Stress Symptoms. *Remote Sens.* **2018**, *10*, 1139. [[CrossRef](#)]
31. Huo, H.; Li, Z.-L.; Xing, Z. Temperature/emissivity separation using hyperspectral thermal infrared imagery and its potential for detecting the water content of plants. *Int. J. Remote Sens.* **2018**, *40*, 1672–1692. [[CrossRef](#)]
32. Alves, V.C.D.; de Lima, S.F.; Santana, D.C.; Barreto, R.F.; da Cunha, R.A.; da Silva Cândido Seron, A.C.; Teodoro, L.P.R.; Teodoro, P.E.; de Cássia Félix Alvarez, R.; Campos, C.N.S.; et al. Predicting pineapple quality from hyperspectral data of plant parts applied to machine learning. *AgriEngineering* **2025**, *7*, 170. [[CrossRef](#)]
33. Mahmud, M.S.; Zahid, A.; Das, A.K.; Muzammil, M.; Khan, M.U.; Tran, D.H.; Chang, Y.K. Trends and Prospect of Machine Vision Technology for Stresses and Diseases Detection in Precision Agriculture. *AgriEngineering* **2023**, *5*, 20–39.
34. Marques Ramos, A.P.; Prado Osco, L.; Elis Garcia Furuya, D.; Nunes Gonçalves, W.; Cordeiro Santana, D.; Pereira Ribeiro Teodoro, L.; Antonio da Silva Junior, C.; Fernando Capristo-Silva, G.; Li, J.; Henrique Rojo Baio, F.; 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](#)]
35. Anderson, K.; Rossini, M.; Pacheco-Labrador, J.; Balzarolo, M.; Mac Arthur, A.; Fava, F.; Julitta, T.; Vescovo, L. Inter-comparison of hemispherical conical reflectance factors (HCRF) measured with four fibre-based spectrometers. *Opt. Express* **2013**, *21*, 605. [[CrossRef](#)] [[PubMed](#)]
36. Lavrac, N.; Flach, P.A.; Kasek, B.; Todorovski, L. Rule induction for subgroup discovery with CN2-SD. In Proceedings of the ECML/PKDD'02 Workshop on Integration and Collaboration Aspects of Data Mining, Decision Support and Meta-Learning, Helsinki, Finland, 19 August 2002; pp. 77–87.

37. Marrs, J.; Ni-Meister, W. Machine Learning Techniques for Tree Species Classification Using Co-Registered LiDAR and Hyperspectral Data. *Remote Sens.* **2019**, *11*, 819. [\[CrossRef\]](#)
38. Abdulridha, J.; Batuman, O.; Ampatzidis, Y. UAV-Based Remote Sensing Technique to Detect Citrus Canker Disease Utilizing Hyperspectral Imaging and Machine Learning. *Remote Sens.* **2019**, *11*, 1373. [\[CrossRef\]](#)
39. Pourghasemi, H.; Gayen, A.; Park, S.; Lee, C.-W.; Lee, S. Assessment of Landslide-Prone Areas and Their Zonation Using Logistic Regression, LogitBoost, and NaïveBayes Machine-Learning Algorithms. *Sustainability* **2018**, *10*, 3697. [\[CrossRef\]](#)
40. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.; Tiede, D.; Aryal, J. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection. *Remote Sens.* **2019**, *11*, 196. [\[CrossRef\]](#)
41. Jamil, A.; Bayram, B. The delineation of tea gardens from high resolution digital orthoimages using mean-shift and supervised machine learning methods. *Geocarto Int.* **2019**, *36*, 758–772. [\[CrossRef\]](#)
42. Feng, P.; Wang, B.; Liu, D.L.; Yu, Q. Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. *Agric. Syst.* **2019**, *173*, 303–316. [\[CrossRef\]](#)
43. Liang, L.; Di, L.; Huang, T.; Wang, J.; Lin, L.; Wang, L.; Yang, M. Estimation of Leaf Nitrogen Content in Wheat Using New Hyperspectral Indices and a Random Forest Regression Algorithm. *Remote Sens.* **2018**, *10*, 1940. [\[CrossRef\]](#)
44. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [\[CrossRef\]](#)
45. Furuya, D.E.G.; Ma, L.; Pinheiro, M.M.F.; Gomes, F.D.G.; Gonçalves, W.N.; Junior, J.M.; Rodrigues, D.d.C.; Blassioli-Moraes, M.C.; Michereff, M.F.F.; Borges, M.; et al. Prediction of insect-herbivory-damage and insect-type attack in maize plants using hyperspectral data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102608.
46. Osco, L.P.; Ramos, A.P.M.; Pinheiro, M.M.F.; Moriya, É.A.S.; Imai, N.N.; Estrabis, N.; de Araújo, F.F.; Liesenberg, V.; Jorge, L.A.d.C.; Li, J.; et al. A Machine Learning Framework to Predict Nutrient Content in Valencia-Orange Leaf Hyperspectral Measurements. *Remote Sens.* **2020**, *12*, 906. [\[CrossRef\]](#)
47. Ramos, A.P.M.; Gomes, F.D.G.; Pinheiro, M.M.F.; Furuya, D.E.G.; Gonçalves, W.N.; Junior, J.M.; Michereff, M.F.F.; Blassioli-Moraes, M.C.; Borges, M.; Alaumann, R.A.; et al. Detecting the attack of the fall armyworm (*Spodoptera frugiperda*) in cotton plants with machine learning and spectral measurements. *Precis. Agric.* **2022**, *23*, 470–491. [\[CrossRef\]](#)
48. Ballester, C.; Brinkhoff, J.; Quayle, W.C.; Hornbuckle, J. Monitoring the Effects of Water Stress in Cotton using the Green Red Vegetation Index and Red Edge Ratio. *Remote Sens.* **2019**, *11*, 873. [\[CrossRef\]](#)
49. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*, 2nd ed.; Prentice Hall: Harlow, UK, 2014; pp. 333–378.
50. Jiang, C.; Chen, Y.; Wu, H.; Li, W.; Zhou, H.; Bo, Y.; Shao, H.; Song, S.; Puttonen, E.; Hyypä, J. Study of a High Spectral Resolution Hyperspectral LiDAR in Vegetation Red Edge Parameters Extraction. *Remote Sens.* **2019**, *11*, 2007. [\[CrossRef\]](#)
51. Poona, N.K.; Ismail, R. Developing optimized spectral indices using machine learning to model *Fusarium circinatum* stress in *Pinus radiata* seedlings. *J. Appl. Remote Sens.* **2019**, *13*, 1. [\[CrossRef\]](#)
52. Zhu, H.; Lin, C.; Dong, Z.; Xu, J.L.; He, Y. Early Yield Prediction of Oilseed Rape Using UAV-Based Hyperspectral Imaging Combined with Machine Learning Algorithms. *Agriculture* **2025**, *15*, 1100. [\[CrossRef\]](#)
53. Zhu, H.; Lin, C.; Liu, G.; Wang, D.; Qin, S.; Li, A.; He, Y. Intelligent agriculture: Deep learning in UAV-based remote sensing imagery for crop diseases and pests detection. *Front. Plant Sci.* **2024**, *15*, 1435016. [\[CrossRef\]](#)
54. Verkhoturov, A.L.; Stepanov, A.S.; Illarionova, L.V. Using radar data to monitor the condition of agricultural crops in the southern Russian Far East. *Inform. Autom.* **2024**, *23*, 1221–1245.

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.