Smart Agricultural Technology 14 (2026) 101953

Contents lists available at ScienceDirect

Smart Agricultural Technology

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

Effect of cover crops on soil quality, yield, and prediction using machine learning in papaya (*Carica papaya* L.)

Pedro A. Torres-Herrera [a,*], Marielita Arce-Inga [a], Ever Tarrillo [a], Elton Ocupa [b], Nilton Atalaya-Marin [a], Héctor Cabrera-Hoyos [c], Juancarlos Cruz-Luis [d], Victor H. Taboada-Mitma [a], Darwin Gómez-Fernández [a], Daniel Tineo [a], Malluri Goñas [a]

[a] Centro Experimental Yanayacu, Dirección de Servicios estratégicos agrarios (DSEA), Instituto Nacional de Innovación Agraria (INIA), Carretera Jaén San Ignacio KM 23.7, Jaén 06801, Cajamarca, Perú
[b] Laboratorio de Suelos Aguas y Foliares, Dirección de Servicios Estratégicos Agrarios (DSEA), Instituto Nacional de Innovación Agraria (INIA), Carretera Jaén San Ignacio KM 23.7, Jaén 06801, Cajamarca, Perú
[c] Estación Experimental Agraria Baños del Inca, Dirección de Servicios Estratégicos Agrarios (DSEA), Instituto Nacional de Innovación Agraria (INIA), Jirón Wiracocha S/N, Cajamarca 06001, Cajamarca, Perú
[d] Dirección de Servicios Estratégicos Agrarios (DSEA), Instituto Nacional de Innovación Agraria (INIA), Av . Molina 1981, Lima 15024, Perú

ARTICLE INFO

Keywords:
Machine learning
Papaya yield prediction
Vegetation indices
Soil physicochemical attributes
UAV multispectral imagery

ABSTRACT

The integration of vegetative cover crops and machine learning-based predictive models represents an innovative strategy to enhance the sustainability and productivity of tropical fruit production systems. This study evaluated the effects of four soil cover treatments, spontaneous vegetation, *Arachis pintoi, Canavalia ensiformis*, and *Centrosema macrocarpum*, in addition to a no-cover control, on yield performance and soil quality in papaya (*Carica papaya* L.) cultivation. Agronomic variables, vegetation indices derived from multispectral imagery, and meteorological factors were integrated to develop yield prediction models using Random Forest, K-Nearest Neighbors, and Extreme Gradient Boosting algorithms. Analysis of variance revealed significant differences among treatments ($p < 0.05$), with *Centrosema macrocarpum* achieving the highest yield (102.22 t ha^{-1}), representing a 37% increase compared to spontaneous vegetation. Furthermore, cover treatments improved soil pH, suggesting reduced acidity and a positive contribution to the long-term sustainability of the production system. Among the evaluated models, Extreme Gradient Boosting demonstrated the best predictive performance ($R^2 = 0.85$; RMSE = 11.56 t ha^{-1}). These findings indicate that the combined use of vegetative cover strategies and precision agriculture tools can optimize decision-making, enhance resource-use efficiency, and strengthen the resilience of papaya production systems.

1. Introduction

Papaya (*Carica papaya* L.) is a tropical crop of great economic and nutritional importance in hot-climate regions, characterized by a short production cycle and high susceptibility to phytopathogenic diseases, which can reduce yield by up to 40 % [1,2]. Global papaya production exceeds 13.8 million tons, with major producers including India, Brazil, Indonesia, and the Dominican Republic [3]. In Peru, papaya production reaches approximately 175 thousand tons [4], establishing it as an important fruit crop and an alternative to traditional agricultural systems [5].

Agricultural production in the province of Jaén is primarily based on crops such as coffee, rice, and cocoa, which constitute the pillars of the local economy [6]. In contrast, papaya records an average annual production of 489.79 tons, highlighting its relatively limited participation in the regional agricultural landscape [5]. Nonetheless, this crop represents a strategic alternative to diversify the agricultural base and strengthen the sustainability of local production systems [7]. Its establishment requires continuous management practices such as weeding, fertilization, and pest control applications, which are essential for optimal phenological development and yield maximization [8,9]. However, approximately 50 % of cultivated areas are affected by the misuse of agronomic practices, particularly those leading to soil degradation [1].

* Corresponding author.
E-mail address: path201098@gmail.com (P.A. Torres-Herrera).

https://doi.org/10.1016/j.atech.2026.101953
Received 26 November 2025; Received in revised form 6 March 2026; Accepted 6 March 2026
Available online 10 March 2026

Phytosanitary issues represent one of the main constraints in papaya production, with the papaya ringspot virus (PRSV) capable of reducing yield by up to 50 % [2]; anthracnose (*Colletotrichum gloeosporioides*), responsible for postharvest losses of about 30 % [10]; neck and root rot (*Phytophthora* spp.), associated with wilting and mortality in poorly drained soils [11]; and pests such as fruit flies (*Anastrepha* spp.), which affect both yield and commercial quality [12]. In response to these challenges, the implementation of cover crops such as *Centrosema macrocarpum, Arachis pintoi*, and *Canavalia ensiformis* has emerged as a sustainable strategy for papaya cultivation. These species enhance soil fertility and structure through nitrogen fixation and contribute to greater availability of essential nutrients such as potassium (K), phosphorus (P), and calcium (Ca) [13,14].

These cover crop species have demonstrated positive effects in tropical systems, particularly when associated with maize and banana crops [15,16], where their incorporation has promoted soil recovery and ecological fertility management. Additionally, by enhancing microbial diversity and improving soil conditions, cover crops can reduce the incidence of soil pathogens such as *Phytophthora* spp. and limit pest pressure and alternative virus hosts [17], thereby contributing to disease mitigation and the development of more resilient and sustainable production systems.

Precision agriculture, on the other hand, has become a fundamental tool for early diagnosis of crop status and for guiding management decisions. In particular, unmanned aerial vehicles (UAVs) have shown great potential for generating high-resolution geospatial data, enabling the estimation of vegetation indices that reflect crop vigor and canopy status [18–20]. When combined with agronomic and meteorological data, these indices can be used to develop predictive models with strong capacity to estimate crop yield [21,22]. In the case of papaya, monitoring cover crops using multispectral sensors mounted on UAVs allows linking the edaphic and physiological benefits of species such as *Centrosema macrocarpum, Arachis pintoi*, and *Canavalia ensiformis* with objective productivity indicators, offering a comprehensive strategy to improve the sustainability and stability of tropical fruit systems [23,24].

In Peru, studies on papaya cultivation remain limited, and the systematic use of precision agriculture tools for yield prediction has not yet been established. Moreover, there is no prior research integrating the use of cover crops with multispectral monitoring and predictive analytics. Therefore, this study aimed to evaluate the effects of different cover crops associated with papaya cultivation on agronomic performance and soil physicochemical properties, and to predict yield using advanced remote sensing and modeling methodologies. The ultimate goal was to generate robust scientific evidence to improve understanding of cover crop-papaya interactions and to support the design of more resilient and sustainable agricultural systems.

2. Methods

2.1. Study area

The experimental plot was established at the Centro Experimental Yanayacu of the Instituto Nacional de Innovación Agraria (INIA) (Fig. 1), located in the district and province of Jaén, Cajamarca region, Peru. The site is situated at an altitude of 535 m a.s.l. and is characterized by a warm, humid climate throughout the year, classified as semi-dry with abundant moisture (C(r)A). Maximum temperatures range from 29 °C to 33 °C, while minimum temperatures vary between 19 °C and 23 °C. Relative humidity remains between 75 % and 90 % from January to December, and the mean annual rainfall fluctuates between 900 mm and 1200 mm. The dry season extends from May to October, whereas the wet season occurs from October to April [25].

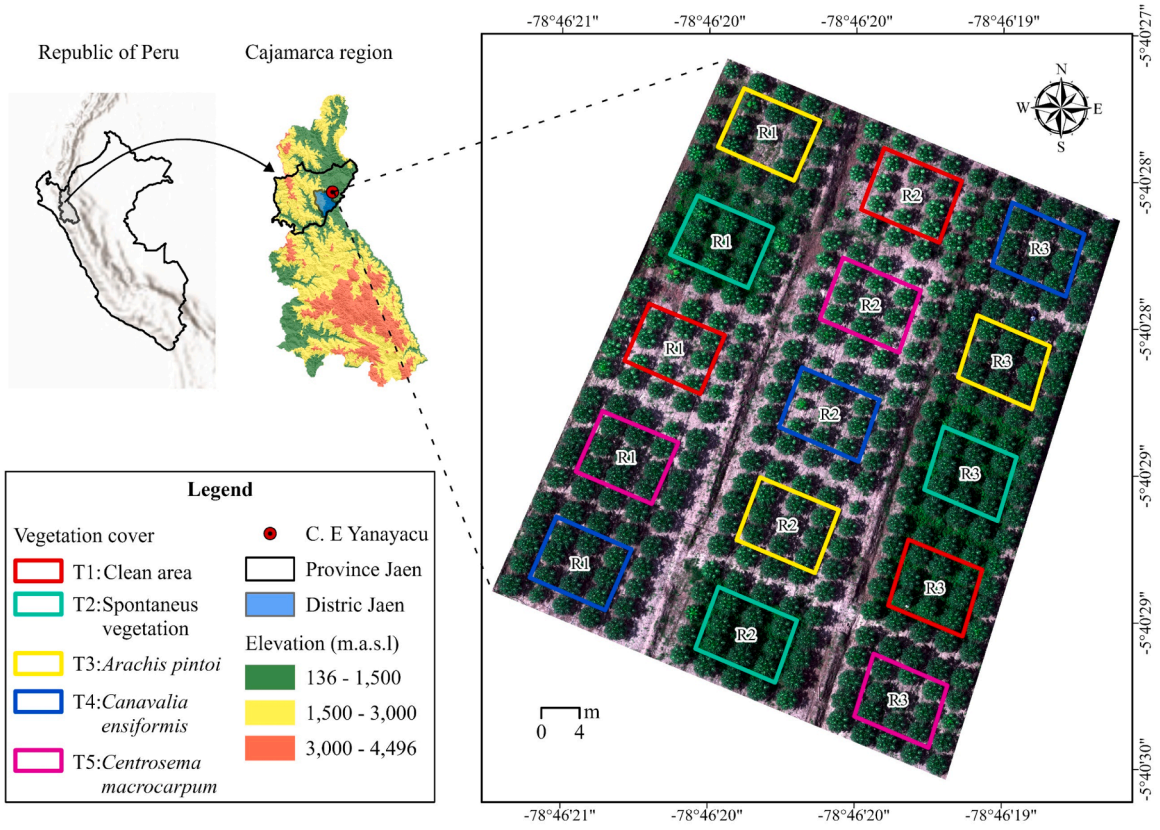

Fig. 1. Location of the papaya experimental plot at the Centro Experimental Yanayacu, Jaen province, Cajamarca region, Peru. Created by the authors and geospatially represented using ArcGIS Pro v3.1.0 software (https://pro.arcgis.com/en/pro-app/latest/get-started/install-and-sign-in-to-arcgis-pro.htm). The elevation map was downloaded from ASF Data Search Vertex (https://search.asf.alaska.edu/#/).

2.2. Experimental design

An experimental plot of papaya (*Carica papaya* L.) cv. 'Sinta F1' was established under a completely randomized design (CRD) in an area of 3554 m^2. Five treatments were evaluated, consisting of four vegetation covers and one bare soil treatment associated with *C. papaya*: clean area, spontaneous vegetation, *Arachis pintoi, Canavalia ensiformis*, and *Centrosema macrocarpum*. In total, 15 experimental units of 7.5 × 9 m were established. Nine central plants per unit were evaluated, excluding border plants to avoid edge effects, resulting in a total of 135 plants assessed across the entire experiment.

The experimental plot was managed under uniform fertilization and irrigation conditions. Fertilization consisted of a total application of 300 g of nitrogen (N), 300 g of phosphorus pentoxide (P_2O_5), and 360 g of potassium oxide (K_2O), distributed over three crop development stages: 10 % during establishment, 40 % from flowering to fruit set, and 50 % at the onset of the production phase. Nutrients were supplied using urea, diammonium phosphate, and potassium sulfate as sources.

Irrigation was applied through a surface drip system to maintain adequate soil moisture within the active root zone (20–40 cm depth) of papaya plants and their associated covers. During the establishment stage, irrigation was performed every two days for two hours per day, divided into two sessions (morning and afternoon). In subsequent stages of vegetative growth, flowering, and fruiting, irrigation frequency was adjusted to every three days while maintaining the same total daily duration and split sessions. This approach reduced direct soil evaporation and promoted more efficient water use.

2.3. Methodological outline

The present study was structured in three consecutive phases aimed at characterizing the agronomic, edaphic, and predictive performance of papaya cultivation under different plant covers. In the first phase, systematic monitoring of agronomic variables (AG), vegetation indices (VIs), and meteorological parameters (ME) was carried out throughout six phenological stages of the crop. This phase sought to describe the dynamic responses of *Carica papaya* L. cv. 'Sinta F1' under the influence of each vegetation cover treatment. The second phase focused on the development of yield estimation models using machine learning techniques. RIDGE regression was implemented to ensure statistical robustness and minimize multicollinearity, while the Random Forest (RF), XGBoost (XG), and K-nearest neighbors (KNN) algorithms were applied to enhance predictive accuracy. Model performance was evaluated using the coefficient of determination (R^2) and root mean square error (RMSE) as adjustment indicators. Finally, the third phase involved the evaluation of soil physicochemical properties at two critical crop stages: flowering (S1) and production (S2). Soil samples were collected directly from the experimental plots and analyzed at the Laboratorio de Suelos, Aguas y Foliares (LABSAF). This analysis aimed to establish correlations between soil nutrient dynamics and the agronomic performance of the papaya crop, as summarized in Fig. 2.

2.4. Agronomic and meteorological data collection

The following agronomic variables were evaluated: stem diameter (measured with a Vernier 6 Truper digital caliper), plant height (measured with a Gripper 5 M Truper flexometer winch), chlorophyll index-Soil Plant Analysis Development (measured with a MINOLTA SPAD 502 Plus meter-2900P), number of fruits, fruit set, and flower buds during six phenological stages, as well as fruit weight only in the final production phase, as detailed in Fig. 3.

Yield was estimated at the end of the first harvest, 205 days after planting *C. papaya* L., following the methodology established for tropical soils. During the production stage, ten fruits per experimental unit were selected as valid samples, collected, and individually weighed for each treatment with three replications. From these data, the average fruit weight was calculated and multiplied by the total number of fruits per plant in each experimental unit. The results were expressed in tons per hectare (t ha^{-1}) and were analyzed for normality and homogeneity of variances.

Meteorological data (temperature and relative humidity) were obtained from the Automatic Meteorological Station (EMA) of SENAMHI, located 0.8 km from the experimental plot.

2.4.1. Multispectral image collection and processing

For the processing of multispectral images, Pix4D Mapper v4.5 software was used, incorporating four ground control points (GCPs) georeferenced in UTM Zone 17S coordinates using a GNSS South Galaxy G7 receiver. From these data, a point cloud was generated with a spatial resolution of 1.2 cm per pixel. Subsequently, multispectral orthomosaics were constructed, and using the Pairwise Buffer geoprocessing tool with radii adjusted according to the crop's phenological stage, information was extracted through the Zonal Statistics as Table tool in ArcGIS Pro v3.1.0. Finally, the vegetation indices were organized and computed in Microsoft Excel 2016 using the equations listed in Table 1.

2.5. Evaluation of the physicochemical properties of the soil

A total of 30 soil samples were collected, 15 samples at 97 days and 15 samples at 205 days after the treatments were established. The samples were homogenized, air-dried, sieved, and analyzed at the Laboratorio de Aguas, Suelos y Foliares (LABSAF) of the Centro

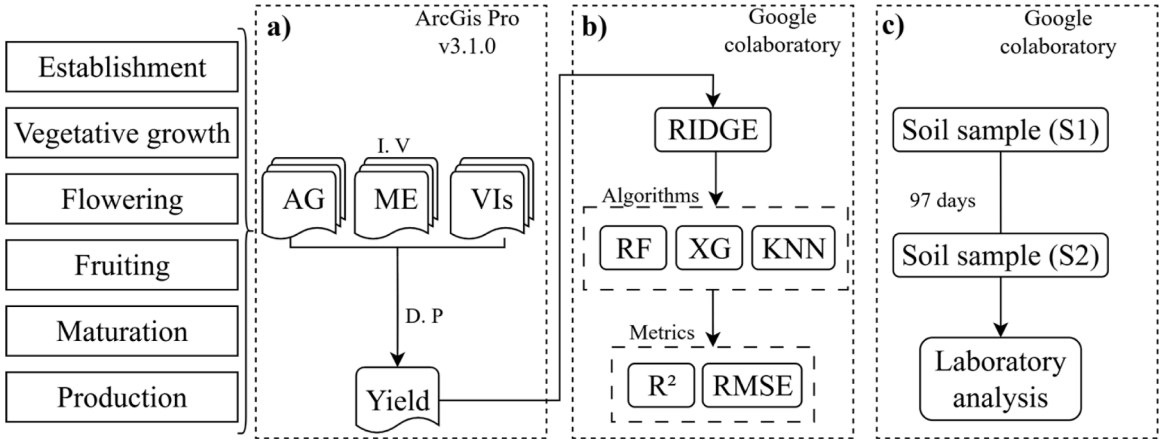

Fig. 2. Methodological scheme followed to estimate the yield of papaya cultivation under four vegetation cover treatments considering a) Acquisition of monitoring information, b) Multitemporal analysis and c) Evaluation of the physicochemical properties of the soil.

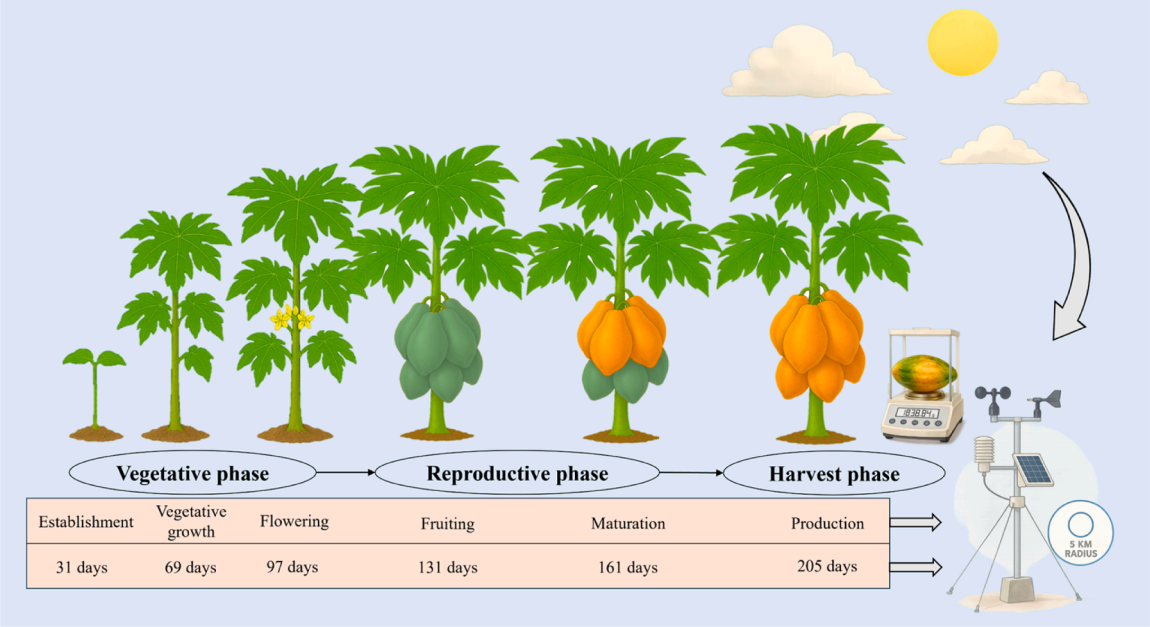

Fig. 3. Evaluation of agronomic variables and meteorological monitoring across the six phenological stages of papaya crop development: establishment, vegetative growth, flowering, fruiting, maturation, and production.

Table 1
Vegetation indices calculated in the study.

Indices	Formula	Reference
Normalized Difference Vegetation Index (NDVI)	$\frac{NIR - Red}{NIR + Red}$	[26]
Modified Soil Adjusted Vegetation Index (MSAVI)	$\frac{1}{2} * [2(\mathrm{NIR} + 1) - \sqrt{(2 \cdot NIR + 1)2 - 8(NIR - Red)}]$	[27]
Enhanced Vegetation Index (EVI)	$2.5 * \frac{NIR - Red}{(NIR + 6\ Red - 7.5Blue) + 1}$	[28]
Green Normalized Difference Vegetation Index (GNDVI)	$\frac{NIR - Green}{NIR + Green}$	[29]
Normalized Difference Red-Edge (NDRE)	$\frac{NIR - RedEdge}{NIR + RedEdge}$	[30]

Experimental Yanayacu, part of the Instituto Nacional de Innovación Agraria (INIA). The physicochemical analyses detailed in Table 2 were performed following the procedures established by the International Organization for Standardization (ISO, 2006) and the Environmental Protection Agency (EPA).

Table 2
Methods used to determine the physicochemical characteristics of soils.

Parameter	Unit	Method	Reference
pH	Unit.	EPA method 9045D	[31]
Electrical conductivity	mS m^{-1}	EPA method 9045D	[31]
Calcium Carbonate Equivalent	%	Walkley & Black.	[32]
Organic matter	%	Gravimetric method	[33]
Phosphorus Available	mg kg^{-1}	Kjeldahl Method	[34]
Kjeldahl Total Nitrogen	mg kg^{-1}	Bray and Kurtz Method	[35]
Available Potassium	mg kg^{-1}	Kjeldahl Method	[34]
Total Organic Carbon	mg kg^{-1}	Kjeldahl Method	[34]
Sand	%	Bouyoucos Hydrometer Method	[36]
Clay	%	Bouyoucos Hydrometer Method	[36]
Silt	%	Bouyoucos Hydrometer Method	[36]

2.6. Modeling the yield of papaya crops

The development of the predictive model for estimating papaya yield was based on the integration of agronomic (AG), vegetation index (VIs), and meteorological (ME) variables. Seven modeling scenarios were proposed (Table 3), in which the RIDGE algorithm [37] was implemented. A conservative regularization parameter ($\alpha = 1$) was applied to penalize predictors with low contribution. Following variable standardization, only those with absolute coefficients greater than 1×10^{-4} were retained, thereby reducing the risk of overfitting.

First, variable selection was performed to identify the most relevant predictors and reduce the risk of overfitting. The dataset was then randomly partitioned into 80 % training and 20 % independent testing subsets. Subsequently, Random Forest (RF) regression [38] was implemented using the following parameters: n_estimators = 100, max_depth = 3, max_features = 'sqrt', and random_state = 42. In parallel, the K-Nearest Neighbors (KNN) algorithm [39] was configured with n_neighbors = 3, and the Extreme Gradient Boosting (XGBoost) algorithm [40] was set with n_estimators = 100, learning_rate = 0.1, max_depth = 5, and random_state = 42. All modeling procedures were implemented in Python 3.10 using the Scikit-learn [41], NumPy, and Pandas [42] libraries, which enabled the identification of the variable set with the highest predictive capacity for yield. This methodology allowed for a comparative evaluation of model accuracy across scenarios and the establishment of a robust framework for papaya yield

Table 3
Proposed scenarios for estimating yield.

N	Scenario
1	Agronomic variables (AG)
2	Vegetation indices (VIs)
3	Meteorological variables (ME)
4	AG + Vis
5	AG + ME
6	VIs + ME
7	AG + VIs + ME

estimation.

2.6.1. RF, KNN, and XGBoost accuracy estimation and validation

After applying the Random Forest (RF), K-nearest neighbors (KNN), and Extreme Gradient Boosting (XG) regression algorithms, the models were validated using the coefficient of determination (R^2) and the root mean square error (RMSE) [43]. To assess the accuracy of these estimation models, the corresponding equations available in the Scikit-learn library were applied, and the models with the highest goodness of fit were selected [41]:

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

Where $\hat{y}_i$ represents the predicted value of the i th sample, and y_i denotes its corresponding observed value in a dataset of n samples. The values of this goodness-of-fit index range from $-\infty$ to 1, where a value of 1 indicates a perfect fit, while negative values suggest that the model performs arbitrarily worse than the mean prediction.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{(II)}$$

Additionally, k-fold cross-validation ($k = 8$) was applied to assess the performance of the best predictive model. Subsequently, the standard deviation (σ) was calculated to quantify model variability and statistical reliability.

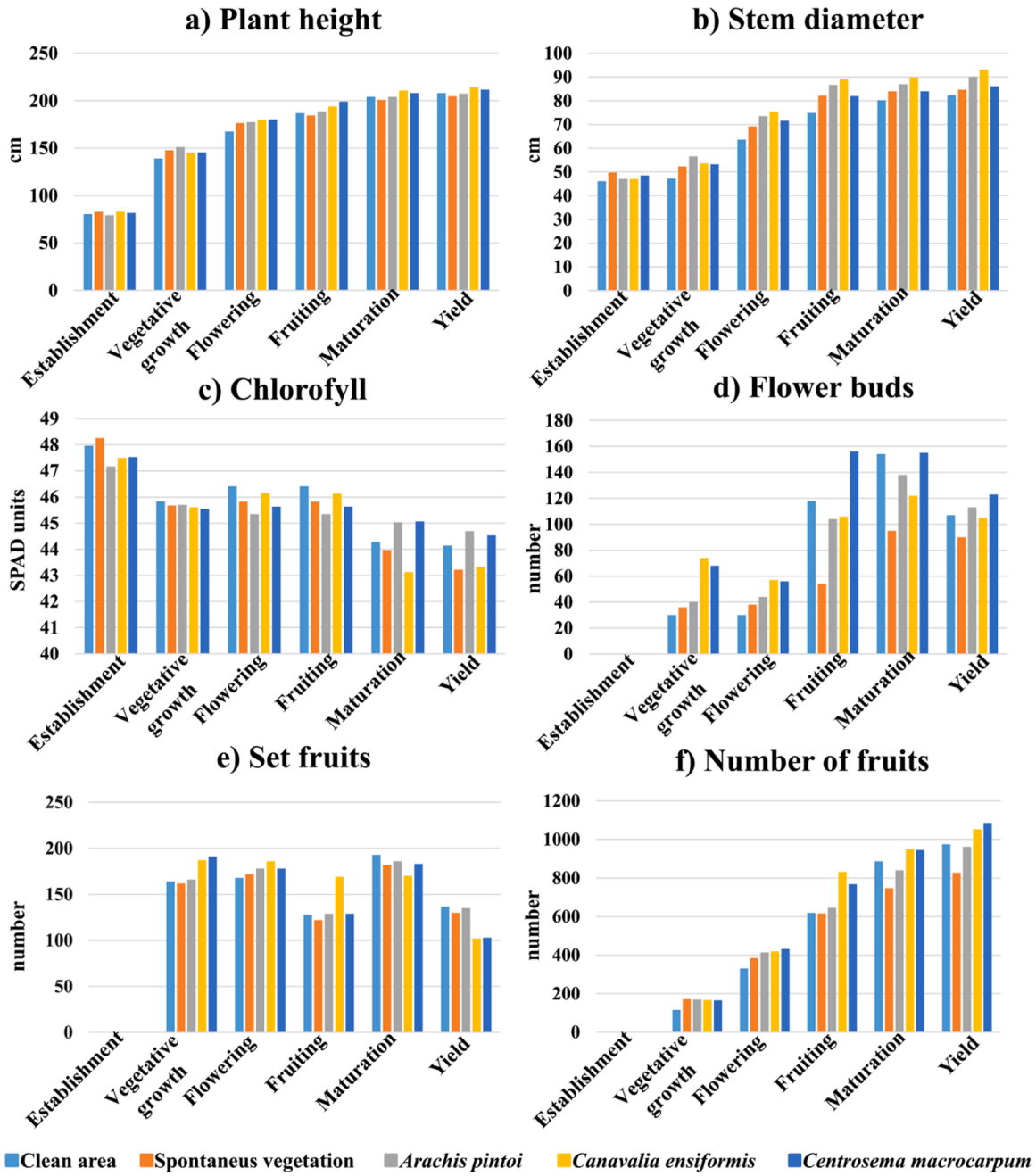

Fig. 4. Agronomic variables evaluated: a) plant height, b) stem diameter, c) chlorophyll content (SPAD), d) number of fruits, e) fruit set, and f) flower buds across six phenological phases, with five treatments: T1, clean area (light blue); T2, spontaneous vegetation (orange); T3, A. pintoi (gray); T4, C. ensiformis (yellow); and T5, C. macrocarpum (blue).

2.6.2. Statistical analysis of data

The estimation of papaya crop yield was analyzed using analysis of variance (ANOVA), followed by Tukey's multiple comparison test to identify significant differences among treatment means. In addition, a Student's *t*-test was performed to detect significant differences in physicochemical characteristics, and Pearson's correlation analysis was applied to evaluate the relationships between agronomic variables at two phenological stages, before (S1) and after (S2) the establishment of plant covers.

3. Results

3.1. Monitoring of variables

3.1.1. Agronomic variables

During the production phase of the papaya crop, the highest average values were recorded in both the control and the treatments with plant covers for the variables of plant height and stem diameter (Fig. 4). Regarding plant height, the maximum value was obtained with *C. ensiformis* (214.26 cm), while spontaneous vegetation showed the minimum (204.81 cm) (Fig. 4a). A similar trend was observed for stem diameter, where the highest value (93.12 cm) was recorded with *C. ensiformis*, and the lowest (82.37 cm) with the control treatment (Fig. 4b).

For chlorophyll leaf content, the highest value was observed with spontaneous vegetation (48.25 SPAD units), while *A. pintoi* presented the lowest (47.17 SPAD units) during the establishment stage (Fig. 4c). In the case of flower buds, the maximum number was reached with *C. macrocarpum* (156 units), while spontaneous vegetation showed the minimum (54 units) during the fruiting stage (Fig. 4d).

In terms of fruit set, the highest value was recorded with 193 units, while *C. ensiformis* showed the lowest (170 units) during the maturation stage (Fig. 4e). Finally, the greatest total number of fruits was obtained with *C. macrocarpum* (1086 units), whereas spontaneous vegetation presented the lowest value (827 units) during the production stage (Fig. 4f).

3.1.2. Meteorological variables

In the evaluated period from January to June 2025, temperature and relative humidity showed variations closely linked to crop development (Fig. 5). The average temperature remained relatively stable between 24.9 °C and 25.7 °C, although with maximums ranging from 32.0 °C in flowering to 36.2 °C in establishment, reflecting the warmer conditions at the beginning of the cycle. Minimums remained constant at around 20 °C, with slight decreases towards the production phase, where they reached 19.9 °C (Fig. 5a).

In contrast, relative humidity showed a more variable dynamic. During the establishment, it registered an average of 62.4 %, increasing in the vegetative growth stage to 67.2 %. In flowering, the wettest condition was reached with 70.1 %, and then gradually decreased in fruiting to 64.4 %, in maturation to 60.6 % and finally reached its lowest value in the production phase with 59.5 %. These variations reflect a greater environmental amplitude throughout the phenological cycle of papaya (Fig. 5b).

During the phenological cycle of *C. papaya* L., temperature (a) and relative humidity (b) were evaluated using the Mann-Kendall trend test. The mean temperature ranged from 25.2 °C to 25.7 °C, with a reference line of 25.5 °C; the test indicated $p = 0.369$, with no significant trend throughout the six phenological phases (Fig. 6a).

The mean relative humidity ranged from 62.2 % to 65.0 %, with a baseline of 64.2 %; the test yielded $p = 0.428$, confirming the absence of a significant trend during the phenological cycle (Fig. 6b).

3.1.3. Vegetation indices (VIs)

Regarding crop monitoring by vegetation indices, Fig. 7 shows the values obtained in the phenological production phase, where clear differences in the response of each index are evident. These contrasts reflect the different sensitivity of the indices to the physiological state of the plant at this stage. The NDVI reached 0.98 (Fig. 7b) and the MSAVI registered 0.96 (Fig. 7c), showing high sensitivity to vigor and

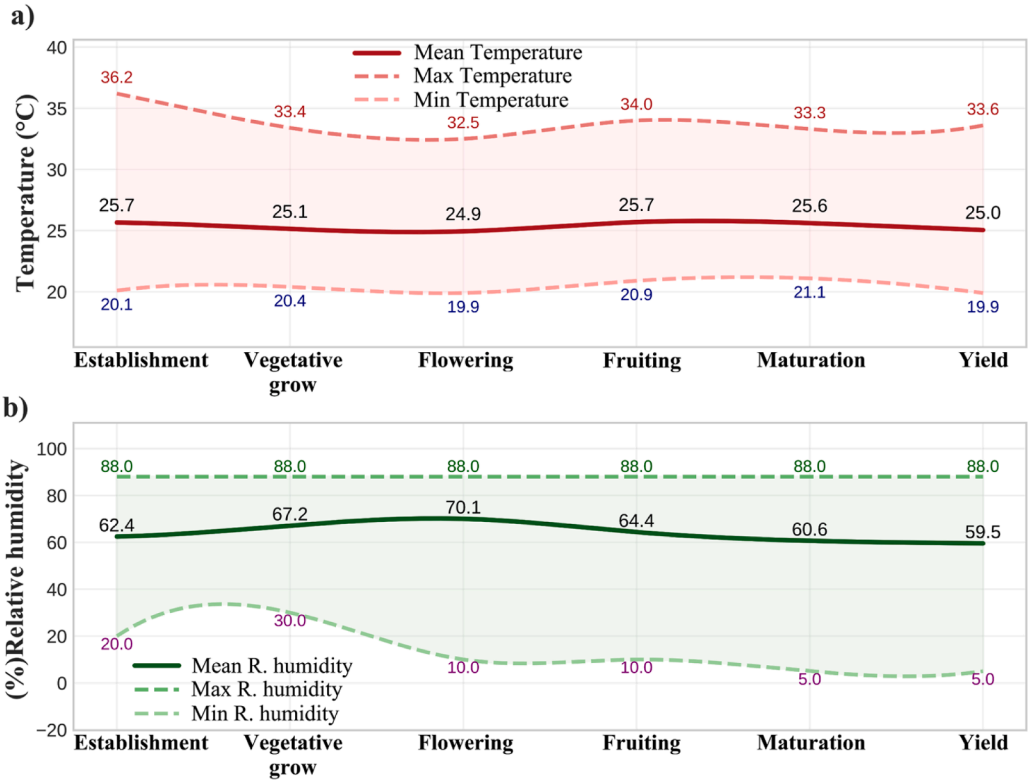

Fig. 5. Behavior of a) temperature (°C) and b) relative humidity (%) during the phenological phases from establishment to production.

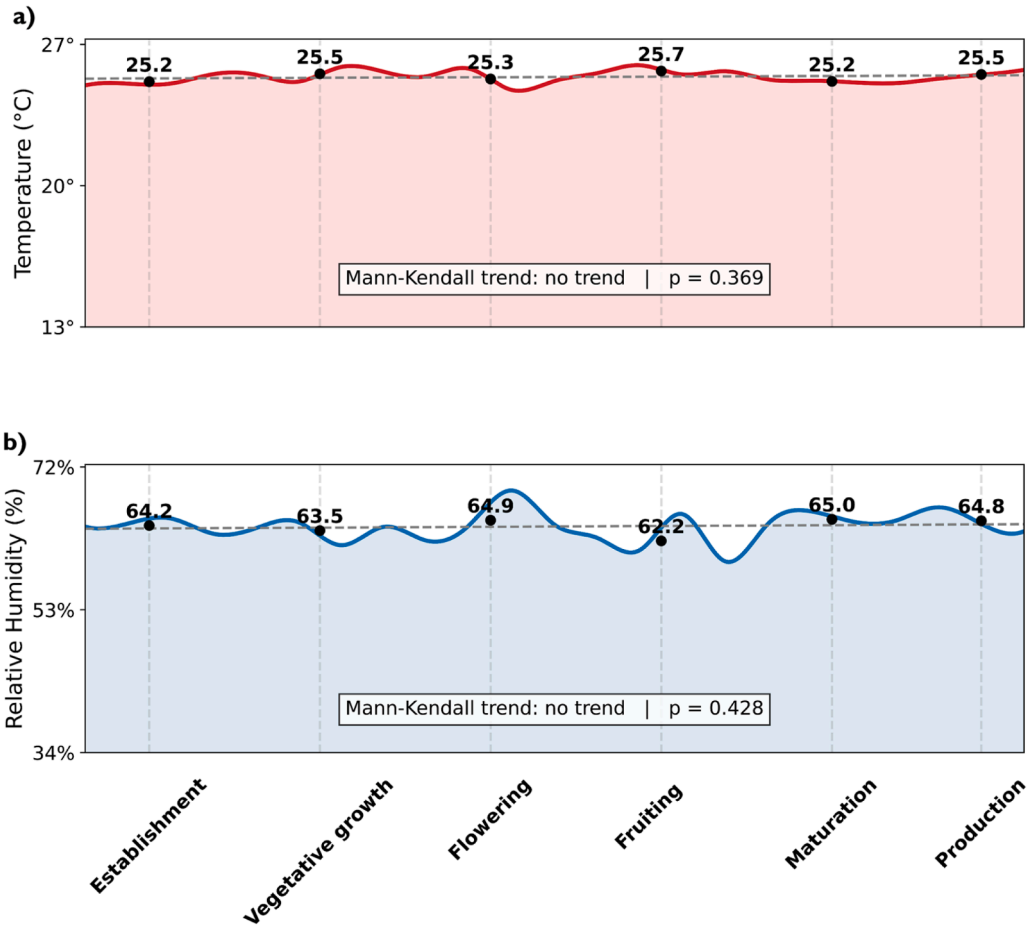

Fig. 6. Daily meteorological trends during six phenological stages of papaya cultivation, figure a) shows the evolution of the mean temperature, while b) represents the variation of relative humidity, each graph a trend line and Mann-Kendall test to evaluate the presence of significant trends throughout the papaya phenological cycle.

vegetation cover, with less influence from the soil.

EVI reached 2.2 (Fig. 7d), reinforcing its ability to discriminate variations in leaf vigor. On the other hand, the GNDVI obtained 0.96 (Fig. 7e), associated with chlorophyll content, while the NDRE presented 0.92 (Fig. 7f), showing better ability to characterize variations in chlorophyll and nitrogen content. To understand more fully the dynamics of these indices throughout the phenological cycle, the results for the other phases are presented in the supplementary figures (Figures S1 to S5). This behavior confirms that the indices have different capacities to capture variations in the structure and physiological condition of the crop.

On the other hand, Fig. 8 shows the evolution of NDVI during the six phenological phases from establishment to production. It is observed that the NDVI has a progressive dynamic in the vegetation cover.

From the first phenological stage of establishment to production, NDVI values showed an increase. This behavior was homogeneous, obtaining higher NDVI values of up to 0.98 recorded in the phenological phases of maturation (Fig. 8e) and production (Fig. 8f), respectively, while vegetative growth (Fig. 8b) presented the lowest value of −0.27.

The dynamics of vegetation indices showed a progressive increase from the establishment stage to production, with notable variations between treatments (Fig. 9). EVI and NDVI showed high sensitivity to leaf cover and canopy structure, with EVI standing out for its accuracy in capturing variations in foliage architecture and upper canopy definition. The NDRE remained stable, reflecting its usefulness in characterizing chlorophyll content and canopy integrity.

In the treatments of clean area (Fig. 9a) and spontaneous vegetation (Fig. 9b), the indices increased steadily until fruiting and then stabilized. *A. pintoi* (Fig. 9c) maintained homogeneous values, with a GNDVI close to 0.6, while *C. ensiformis* (Fig. 9d) showed the most pronounced EVI, with peaks greater than 1.1 in flowering and fruiting. *C. macrocarpum* (Fig. 9e) showed greater spectral variability, with peaks of EVI in flowering and stability of the MSAVI throughout the cycle.

During the phenological cycle of papaya cultivation, the spectral trends associated with its development were analyzed. Although the EVI and MSAVI indices showed values close to significance, none of the vegetation indices showed statistically significant trends according to the Mann-Kendall test ($p > 0.05$). Fig. 10 shows the behavior of the four treatments and the control, each with its respective trend line.

3.2. Analysis of physicochemical variables

3.2.1. Analysis of changes in physicochemical variables

After the implementation of the plant covers, Fig. 11 shows the changes in the physicochemical variables evaluated, where significant differences are indicated by the letters (ab) and non-significant differences by the letter (a). Soil pH showed a significant reduction from alkaline conditions (8.17–8.33) to ranges close to neutrality (7.83–7.93), with statistical differences between treatments. The lowest values were recorded in *C. ensiformis, A. pintoi* and spontaneous vegetation. Organic matter showed a slight increase, reaching 2.40 % in *C. ensiformis* and 2.03 % in spontaneous vegetation, although without significant differences with respect to the cleaned area (1.53 %).

The available phosphorus showed its highest concentration in *C. ensiformis* (17.50 mg kg-1), but the variability between replicates prevented statistical differences from being established. Total nitrogen

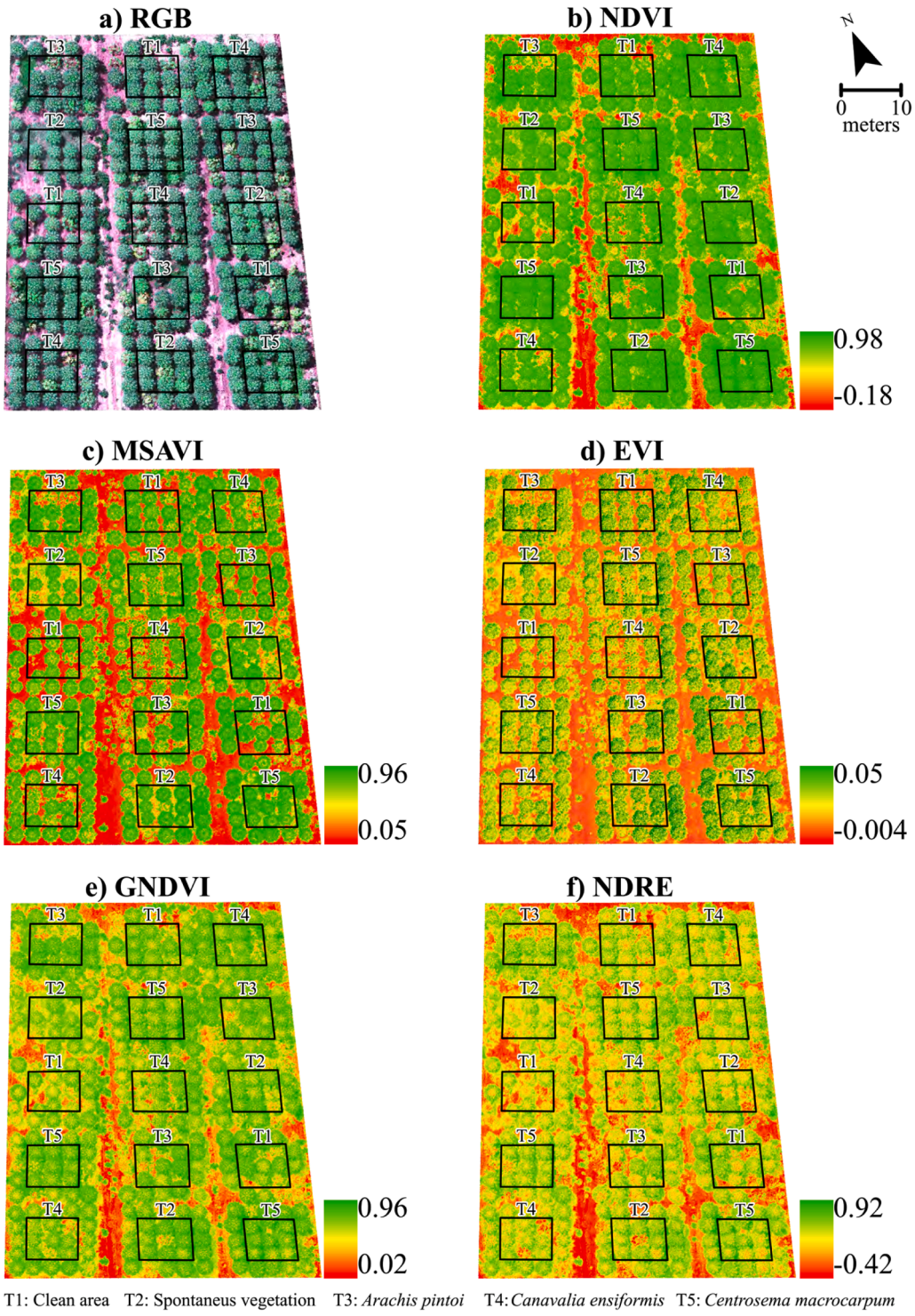

Fig. 7. Calculated vegetation indices (b-f) and contrast in natural color (a) of the papaya crop in the phenological phase of production.

and available potassium ranged from 1.20–1.40 mg kg-1 to 167.00–213.50 mg kg-1, respectively, while total organic carbon ranged from 0.83 to 1.89 mg kg-1, with no significant differences. The soil texture remained constant in all treatments, preserving the sandy-loam classification.

3.2.2. Analysis of the correlation of physicochemical variables with agronomic variables

The correlation diagram showed that plant height was the main indicator of productivity, as it was closely associated with stem diameter (coefficient = 0.90), number of fruits (coefficient = 0.98) and flower buds (coefficient = 0.89). Likewise, fruit sets were significantly correlated with plant height (coefficient = 0.80) and stem diameter (coefficient = 0.94) (Fig. 12a). At the soil level, negative correlations of pH with clay (coefficient = −0.64) and total nitrogen with sand (coefficient = −0.81) stood out.

After the installation of plant covers, the correlations were reorganized, reflecting a greater integration between soil fertility and agronomic performance (Fig. 12b). Organic matter was associated with available phosphorus (coefficient = 0.78) and organic carbon with silt

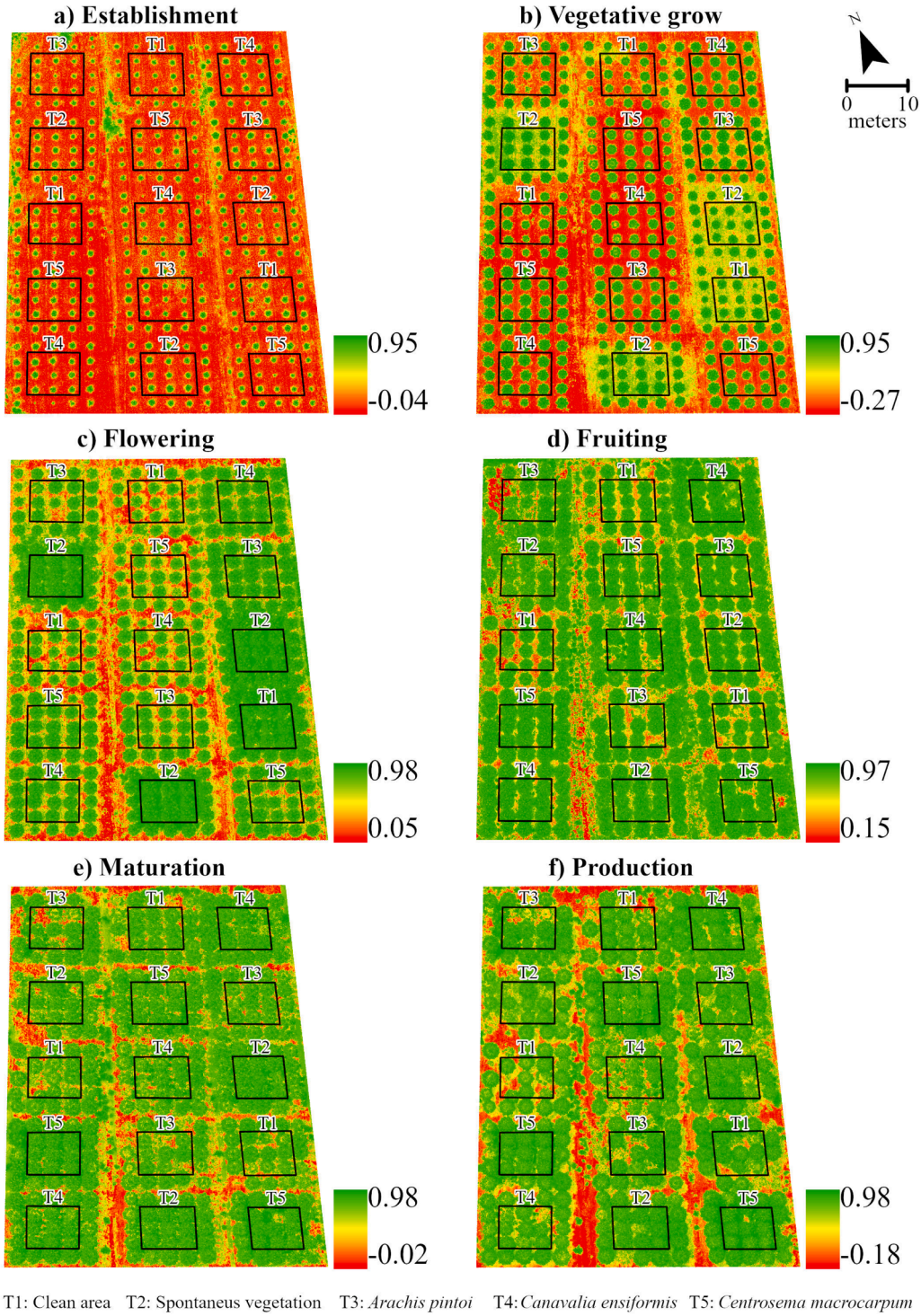

Fig. 8. The NDVI of the six phenological phases is observed: a) establishment, b) vegetative growth, c) flowering, d) fruiting, e) maturation and f) production.

(coefficient = 0.73). In the agronomic field, yield was linked to the number of fruits (coefficient = 0.57), flower buds (coefficient = 0.56) and calcium carbonate (coefficient = 0.49), while notable negative associations were recorded between fruit set and fruit number and plant height (coefficient = −0.71 and −0.86, respectively).

3.3. Yield modeling

3.3.1. Determination of yield

All experimental units were harvested and yield per hectare was projected. The highest average was obtained with *C. macrocarpum* (T5), reaching 102.22 ± 13.65 t ha^{-1}. On the other hand, spontaneous vegetation (T2) presented the lowest yield with 74.50 ± 9.65 t ha^{-1} (Fig. 13). The Shapiro-Wilk test confirmed that the residuals conformed to a normal distribution ($W = 0.955$, $p = 0.608$), validating the statistical consistency of the data. The analysis of variance showed a significant effect of the treatments on yield, and Tukey's HSD post-hoc test indicated that *C. macrocarpum* differed significantly from spontaneous vegetation, with an average difference of >20 t ha^{-1}. The other comparisons showed no significant differences ($p > 0.05$), suggesting a

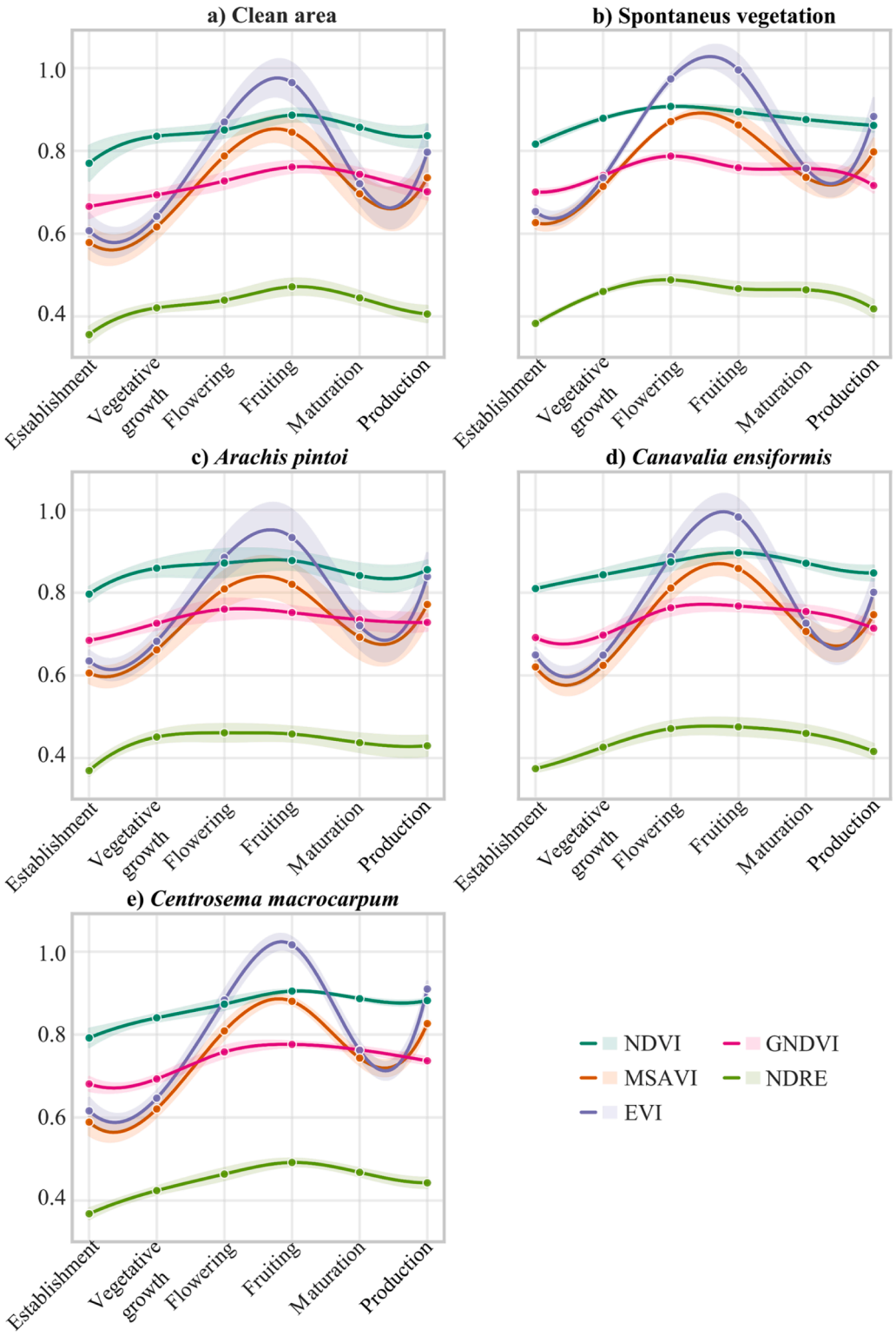

Fig. 9. Behavior of vegetation indices in each phenological phase of papaya cultivation.

statistically homogeneous behavior between these treatments.

3.3.2. *Determining R^2 from estimation models*

Fig. 14 shows the predictive performance recorded by the XGBoost, Random Forest and KNN algorithms. In the clean area treatment, the best fit was obtained with the XGBoost algorithm and the combination of agronomic, vegetation and meteorological variables, reaching an R^2 of 0.59, while the Random Forest and KNN algorithms recorded notably lower values (Fig. 14a). In spontaneous vegetation, XGBoost with agronomic variables achieved an R^2 of 0.54, surpassing Random Forest, which barely reached an R^2 of 0.17 (Fig. 14b).

In *A. pintoi*, the most robust performance of the study was obtained, with an R^2 of 0.85 with the XGBoost regression algorithm and the combination of agronomic and meteorological variables, while KNN and Random Forest did not exceed values of 0.68 (Fig. 14c). In *C. ensiformis*, the KNN algorithm achieved the highest fit with an R^2 of 0.75 when

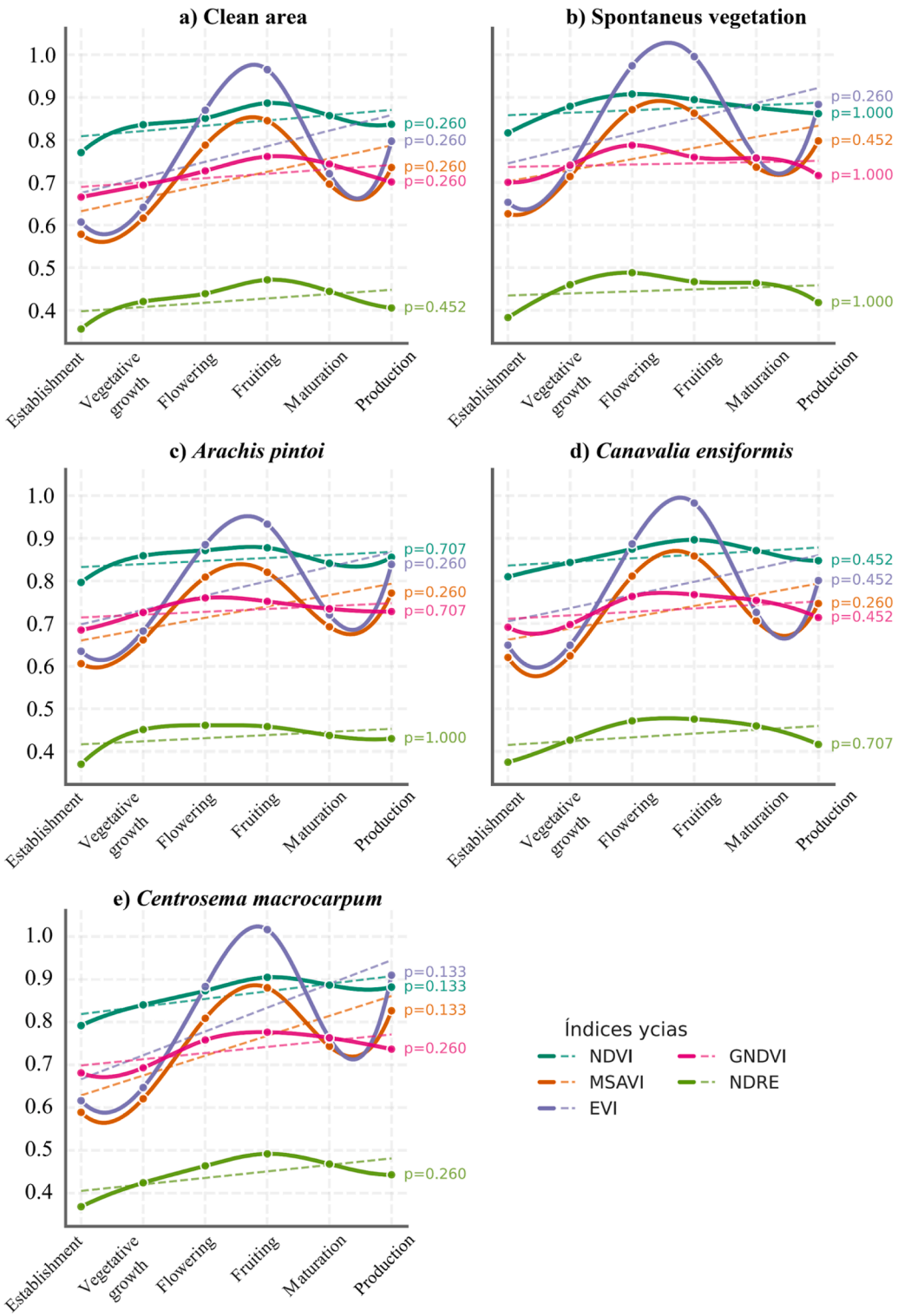

Fig. 10. Man Kendall test ($p < 0.05$) with trend lines of the five vegetation indices along six phenological stages of papaya cultivation.

using agronomic variables, in contrast to Random Forest, which decreased to 0.26 (Fig. 14d). Finally, in *C. macrocarpum*, the lowest values of the entire experiment were recorded, where XGBoost with the integration of agronomic variables, vegetation and meteorological indices reached only an R^2 of 0.48, slightly surpassing KNN and Random Forest, whose values ranged between 0.37 and 0.45 (Fig. 14e).

Yield estimation was performed in the phenological production phase, identified as the most suitable for predictive modeling by generating adjustment values closer to unity. In the clean area, the XGBoost algorithm achieved an R^2 of 0.59 using the combination of agronomic, vegetation and meteorological variables, while in spontaneous vegetation, the same algorithm recorded an R^2 of 0.54 using only agronomic variables. For *A. pintoi*, the best overall performance was obtained, reaching an R^2 of 0.85 with XGBoost and the combination of agronomic and meteorological variables. In *C. ensiformis*, the K-NN algorithm achieved an R^2 of 0.75 using agronomic variables, and finally, in *C. macrocarpum*, XGBoost reported the lowest adjustment with an R^2 of 0.48 when using the combination of agronomic variables, vegetation

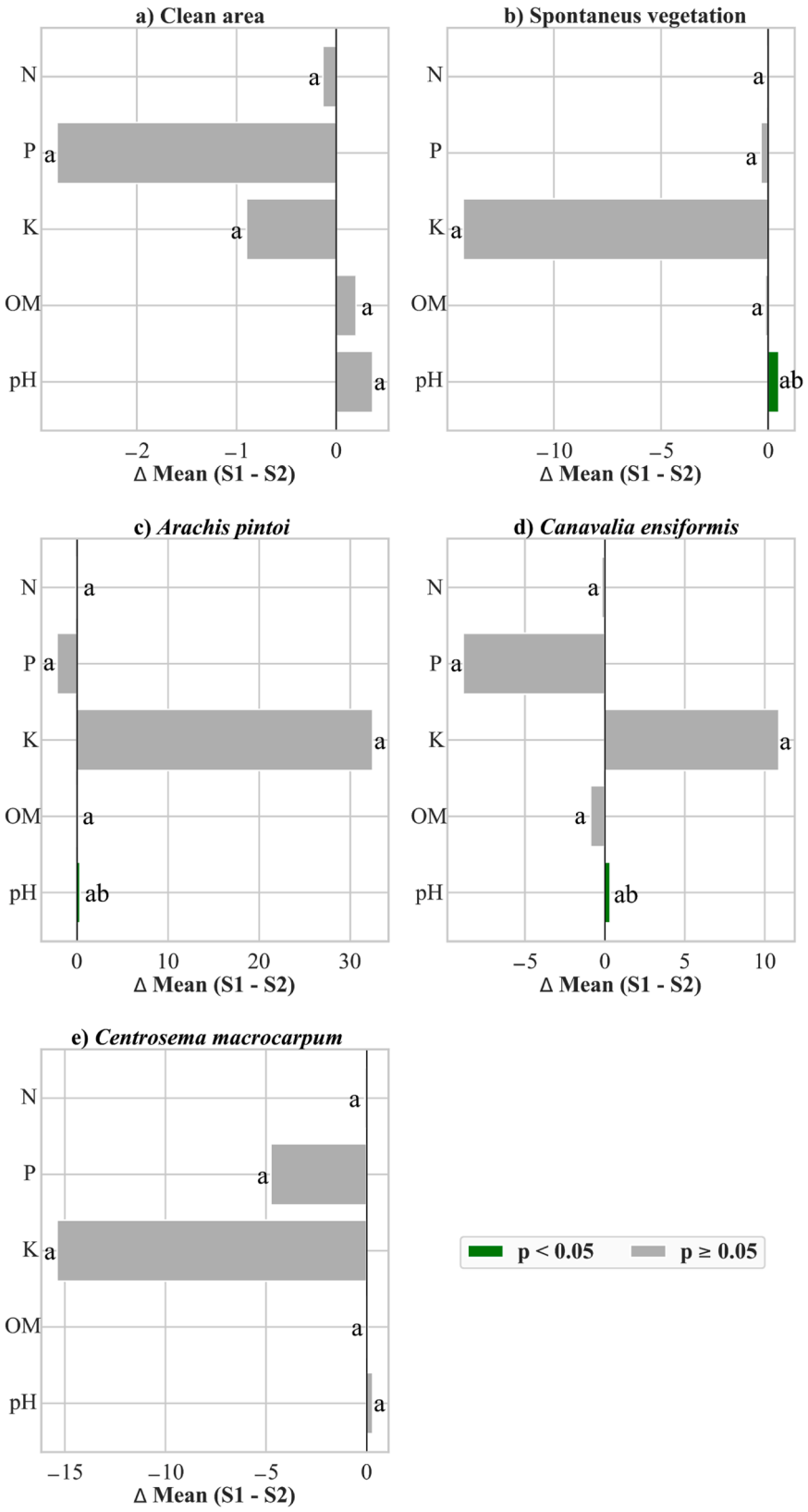

Fig. 11. Physicochemical variables of the soil before (S1) and after (S2) the installation of plant covers in the four treatments and a control, the means with letters in common are not significantly different, according to the t-student test ($p < 0.05$).

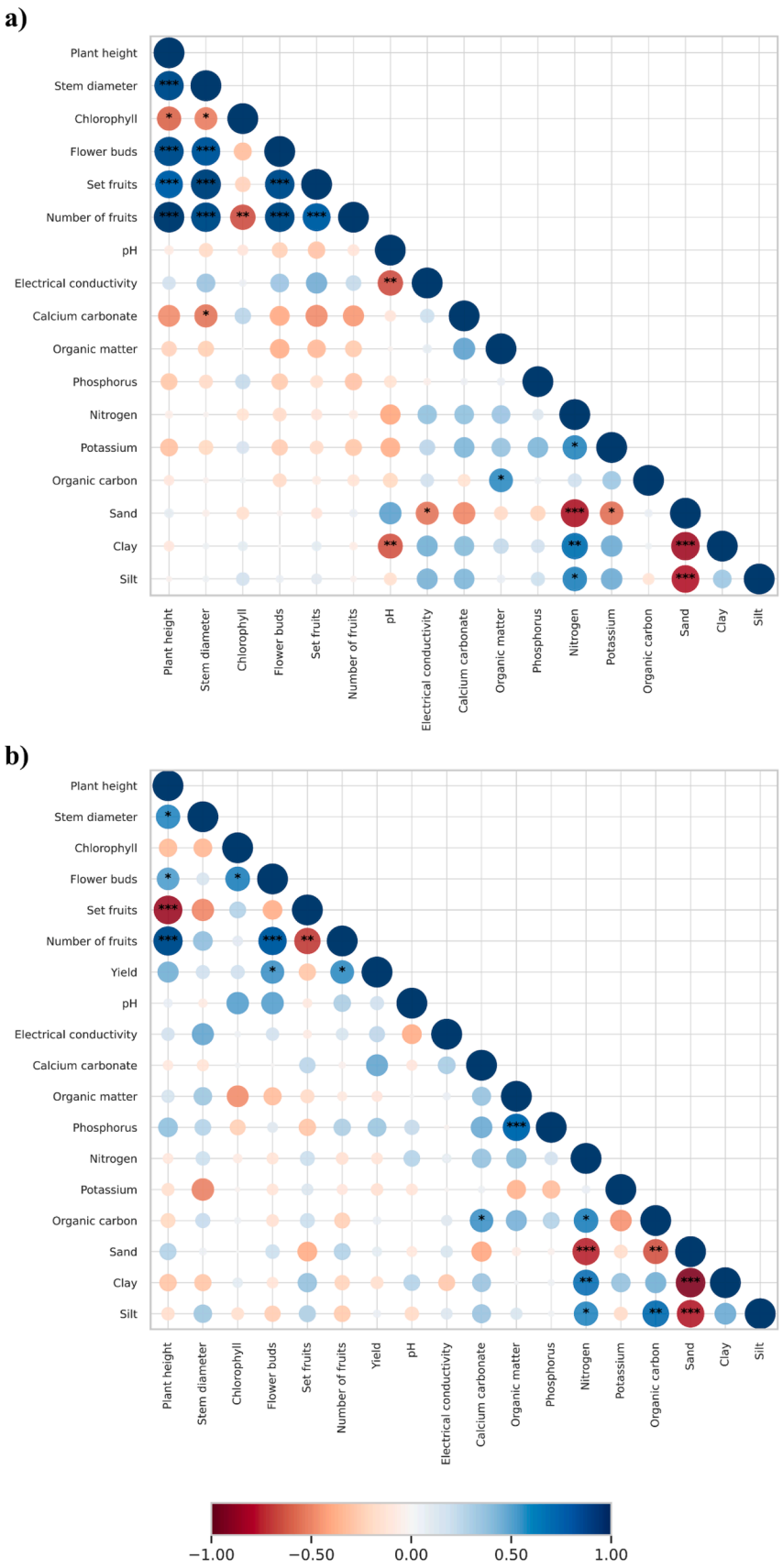

Fig. 12. Pearson's correlation matrix ($p < 0.05$) *** highly significant correlation values between agronomic variables and physicochemical variables a) before and b) after planting of plant covers, with a scale categorized from −1 to 1 in RdBu_r, where blue denotes the highest correlation values from 0 to 1, and red denotes the lowest correlation values from 0 to −1.

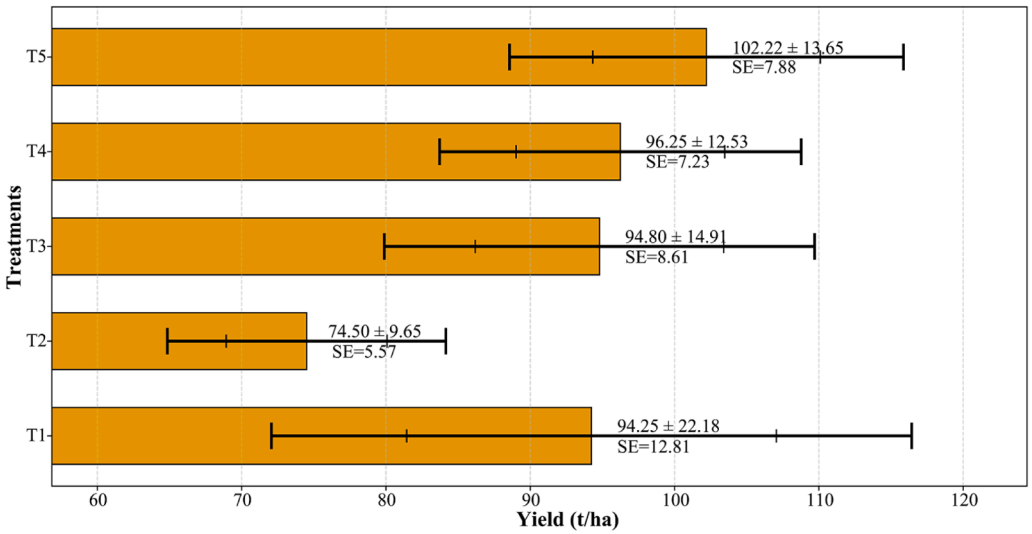

Fig. 13. Papaya crop yield (t ha^{-1}), expressed as mean ± standard deviation and standard error (SE) per treatment.

and meteorological indices (Table 4).

To better understand the most influential variables in yield estimation, Fig. 15 presents a SHAP summary plot illustrating the importance and effect of the three most relevant predictors in the model with the highest R^2 using the Extreme Gradient Boosting (XGBoost) algorithm for *Arachis pintoi* and *Centrosema macrocarpum*, respectively.

For Arachis pintoi, the most influential variable was fruit number, followed by plant height and chlorophyll content. These input variables exhibited greater dispersion in SHAP values (Fig. 15a). In contrast, for *Centrosema macrocarpum*, fruit number was also the most influential predictor, followed by plant height and fruit set (Fig. 15b). However, both the magnitude and dispersion of SHAP values were lower compared to *Arachis pintoi*, which may be associated with the comparatively reduced predictive performance of the model.

3.3.3. *Determination of RMSE from prediction algorithms*

Subsequently, the predictive performance of the regression algorithms was evaluated using the root mean square error (RMSE). Fig. 16 shows the correspondence between the observed values (blue line) and the values estimated by the three models, Random Forest (black line), K-Nearest Neighbors (KNN, green line) and XGBoost (orange line), applied to the four experimental treatments and the control. Among the models evaluated, XGBoost exhibited the highest adjustability, achieving the lowest prediction errors in clean area (16.19 t ha^{-1}), spontaneous vegetation (21.35 t ha^{-1}), *A. pintoi* (11.56 t ha^{-1}) and *C. ensiformis* (32.45 t ha^{-1}), while the treatment with *C. macrocarpum* performed best with the Random Forest model, with an RMSE of 9.51 t ha^{-1}. Taken together, these results indicate that XGBoost achieved greater accuracy and predictive stability than the other algorithms in most treatments, reflecting its superior ability to capture yield variability under different vegetation cover conditions.

4. Discussion

Plant height and stem diameter increased progressively throughout the phenological cycle of papaya cultivation (Fig. 4). The highest plant height values were observed in *C. ensiformis* during the production stage, reaching 214.26 cm and 93.11 cm, respectively. Similarly, the number of fruits was highest in *C. macrocarpum*, averaging 40 fruits per plant, reflecting a strong relationship between vegetative vigor and productivity [44,45]. This positive effect is associated with the greater nitrogen availability provided by legume species such as *C. macrocarpum*, which can fix between 145 and 311 kg N ha^{-1} year^{-1} [46], enhancing protein synthesis and enzymatic activity in the soil. Additionally, its high aboveground and root biomass contributes to increased soil organic matter and organic carbon content, thereby promoting crop development [47]. In contrast, *A. pintoi*, with a limited superficial root system mainly oriented toward weed control, and *C. ensiformis*, whose contribution is rapid but less persistent, exert lower effects on soil fertility and crop yield, whereas the deeper root system and resilience of *C. macrocarpum* to poor or water-stressed soils explain its superior performance [48].

Moreover, *C. macrocarpum* contributed to maintaining more stable chlorophyll levels between 45 and 46 SPAD units compared with the control (Fig. 4). This behavior is linked to its nitrogen-fixing capacity, ensuring a consistent foliar nitrogen supply, essential for the synthesis of chlorophyll and photosynthetic proteins [49,50]. Furthermore, the microclimate generated by its abundant aboveground biomass acts as a regulator of thermal and water stress, reducing photoinhibition and preserving the integrity of photosynthetic pigments. This mechanism also contributes to decreased flower and fruit drop by maintaining a stable water balance during reproductive phases, as previously reported in tropical crops [51,52]. These ecophysiological advantages translated into productivity, with *C. macrocarpum* achieving the highest yield (102.22 t ha^{-1}), surpassing spontaneous vegetation (74.50 t ha^{-1}) (Figs. 4 and 13).

Although temperature and relative humidity remained within suitable ranges for papaya cultivation during the study period (Fig. 5), the Mann-Kendall trend analysis (Fig. 6) indicated a gradual decrease in relative humidity. This trend aligns with the characteristic seasonal regime of northeastern Peru, where the dry quarter from June to August is associated with increased temperature and reduced atmospheric humidity. These subtle variations reflect growing microclimatic instability in tropical regions, attributed to climate change [53]. Consequently, tropical agriculture faces increasingly challenging conditions to maintain productivity and resilience, emphasizing the importance of adaptive strategies such as plant covers and crop associations, which regulate the soil microclimate, reduce moisture loss, and enhance thermal stability and physiological efficiency [54,55].

Among the evaluated physicochemical properties, soil pH showed significant differences ($p < 0.05$), decreasing from alkaline values (8.1–8.3) to near neutrality (7.8–7.9) in several treatments (Fig. 11). This finding is particularly relevant, as pH regulates soil fertility and microbial activity. The shift toward neutral values indicates a buffering effect from the organic matter contributed by the cover biomass. In alkaline soils, essential nutrients such as phosphorus, zinc, and manganese tend to precipitate, limiting their availability, whereas near-neutral conditions enhance solubilization and absorption, reflected in a modest

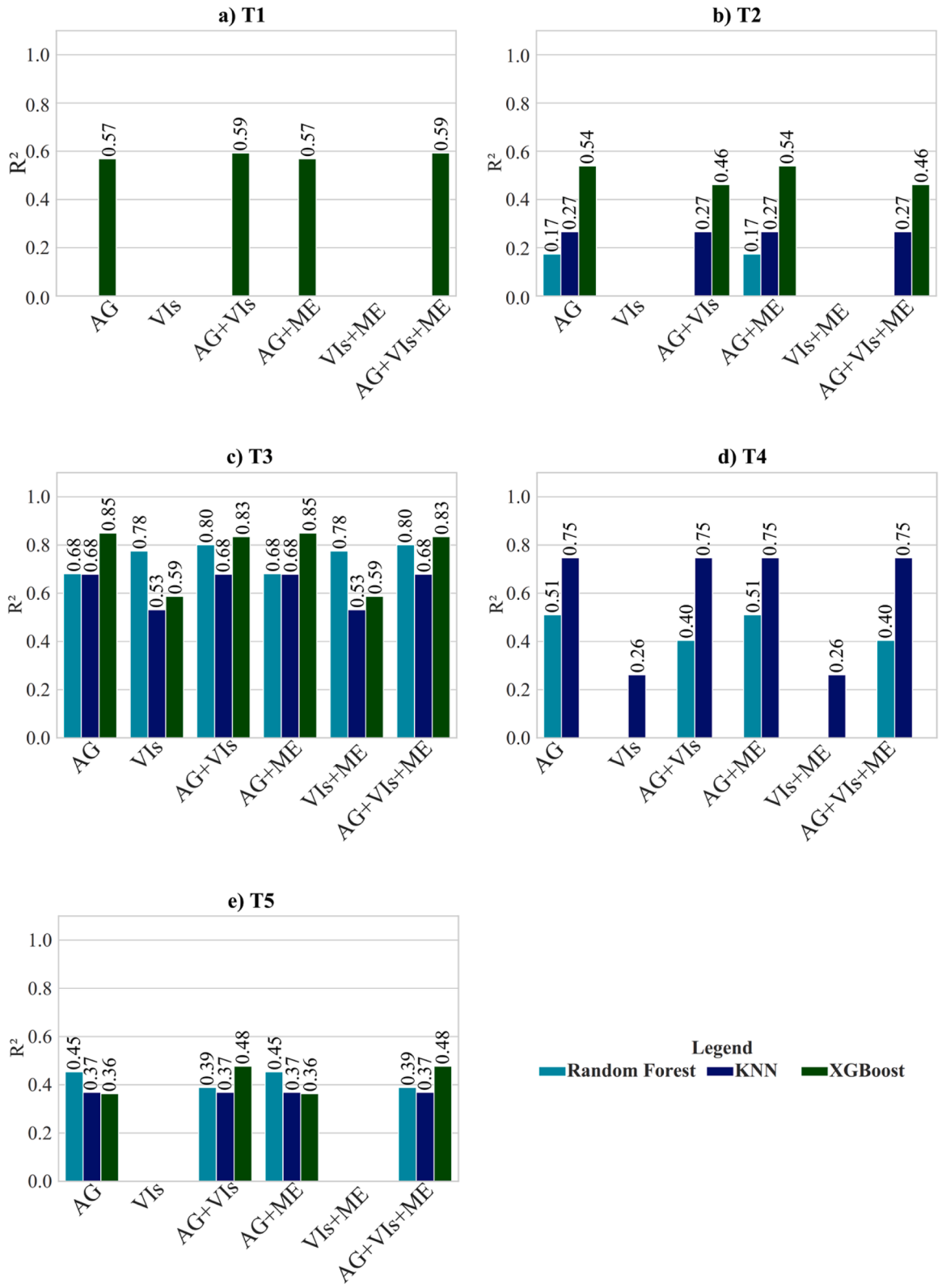

Fig. 14. Maximum R^2 values recorded from performance estimation models constructed from Random Forest (RF), K-Nearest Neighbors (KNN), and Extreme Gradient Boosting (XGBoost) algorithms using the feature combinations.

increase in phosphorus and its association with growth and productivity [56–59].

Although other properties such as organic matter, organic K, C, and P did not show significant differences, they remained stable, corroborating research indicating that root decomposition and cover biomass improve organic matter quality and regulate soil fertility [60–62]. This stability likely explains why *C. macrocarpum* excelled in yield, promoting more consistent soil regulation due to its nitrogen fixation and high biomass production.

Pearson correlations further support this effect (Fig. 12), showing

Tabla 4
Prediction models with better accuracy by treatments according to the highest value of coefficient of determination (R^2). The algorithm, phenological stage, combination of variables, inputs and the corresponding R^2 value are presented.

Tratamient	Algoritm	Combination	Inputs	R^2
Clean area	XGBoost	AG + VIs + ME	High, diameter, chlorophyll, Fruits, Set Fruits, Flower Buds, NDVI, MSAVI, EVI, GNDVI y NDRE	0.59
Spontaneous vegetation	XGBoost	AG	High, diameter, chlorophyll, Fruits, Set Fruits, Flower Buds	0.54
A. pintoi	XGBoost	AG + ME	High, diameter, chlorophyll, Fruits, Set Fruits, Flower Buds	0.85
C. ensiformis	KNN	AG	High, diameter, chlorophyll, Fruits, Set Fruits, Flower Buds	0.75
C. macrocarpum	XGBoost	AG + VIs + ME	High, diameter, chlorophyll, Fruits, Set Fruits, Flower Buds, NDVI, MSAVI, EVI, GNDVI y NDRE	0.48

that before installation (S1), relationships were primarily morphological (height-diameter-fruits), whereas after installation (S2), stronger correlations emerged between organic matter, total C and N, and productivity indicators (fruits, flower buds, SPAD). This demonstrates that improvements in soil organic quality strengthened the integration between soil fertility and agronomic performance, consistent with findings in tropical soils where organic matter accumulation enhances nutrient use efficiency [63].

In terms of remote sensing, vegetation indices accurately captured phenological dynamics, with EVI reaching 2.2 and NDVI 0.98, while MSAVI and GNDVI reached 0.96, and NDRE 0.92 (Fig. 7). EVI showed a marked increase between flowering (V2) and production (V4), maintaining sensitivity under high biomass conditions and exceeding the saturation observed in NDVI (Fig. 8). This behavior is consistent with previous studies that recognize NDVI as a general indicator of vegetation vigor, although it presents limitations in dense canopies due to signal saturation [64,65], whereas EVI enhances the detection of chlorophyll variations under conditions of high canopy cover [28], likewise, MSAVI reduces the influence of soil background effects, making it more suitable during early growth stages [27], GNDVI and NDRE exhibit a stronger association with chlorophyll content and nitrogen status, enabling early detection of nutritional stress and providing a more robust assessment of the crop's physiological variability throughout the phenological cycle [30,66]. In addition, Mann-Kendall analysis (Fig. 10) revealed no statistically significant trends, though a gradual increase was observed across all indices (Fig. 9), reflecting steady phenological development and increasing photosynthetic activity.

In predictive terms (Fig. 14), although *C. macrocarpum* achieved the highest yield (102.22 t ha^{-1}), it exhibited the lowest model fit (R^2 = 0.48), suggesting a production system characterized by greater ecophysiological complexity and nonlinear dynamics. High-biomass cover crops modify the microenvironment and canopy architecture, thereby generating nonlinear relationships between productivity and spectral indices. In particular, NDVI tends to saturate under dense canopy conditions, losing sensitivity at high ranges of leaf area index and biomass [28,67], which reduces the explanatory variance available for predictive modeling. Moreover, structurally heterogeneous systems exhibit hierarchical interactions and ecological thresholds that may constrain the generalization capacity of machine learning algorithms when phenological variables or explicit spatial validation are not incorporated [68]. Although XGBoost demonstrated overall superiority over Random Forest and K-Nearest Neighbors due to its sequential ensemble structure and regularization capacity [40,69], its performance may decline when predictors exhibit saturation effects or high collinearity. Therefore, the lower predictive performance observed in *C. macrocarpum* reflects greater structural complexity rather than reduced productive stability, highlighting the need to integrate less saturation-prone indices such as EVI and crop-specific phenological metrics.

The results were interpreted considering that the evaluations were conducted during a single growing season under spatially stable climatic conditions. Furthermore, the experiment was carried out in a tropical

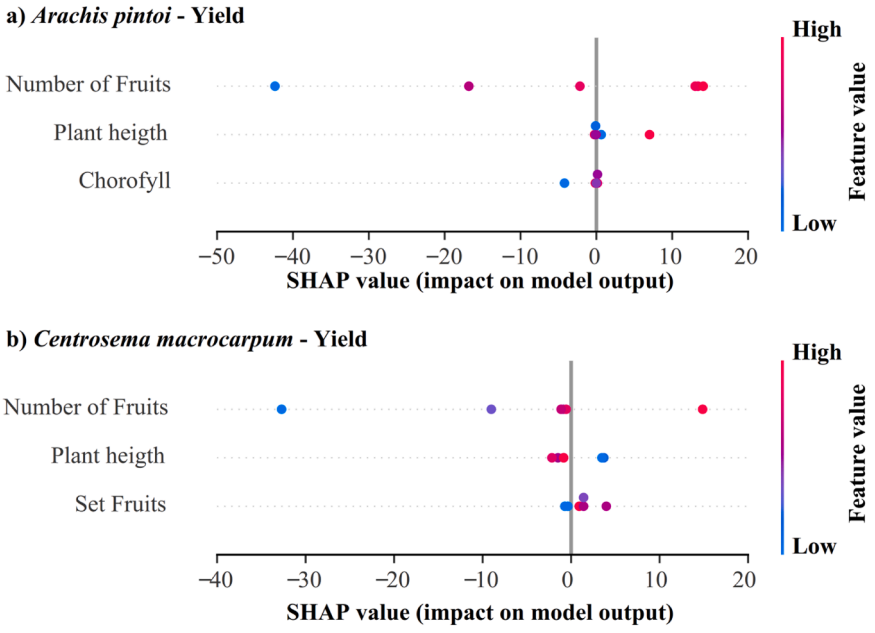

Fig. 15. SHAP summary plot showing the magnitude of influence of agronomic variables with the greatest contribution to yield prediction during the yield phenological stage under two treatments: a) *Arachis pintoi* and b) *Centrosema macrocarpum*. The x-axis represents SHAP magnitude values, indicating the direction and extent of the impact of each input feature on the model output. Each point corresponds to an individual observation and is color-coded according to feature value (blue = low; pink = high).

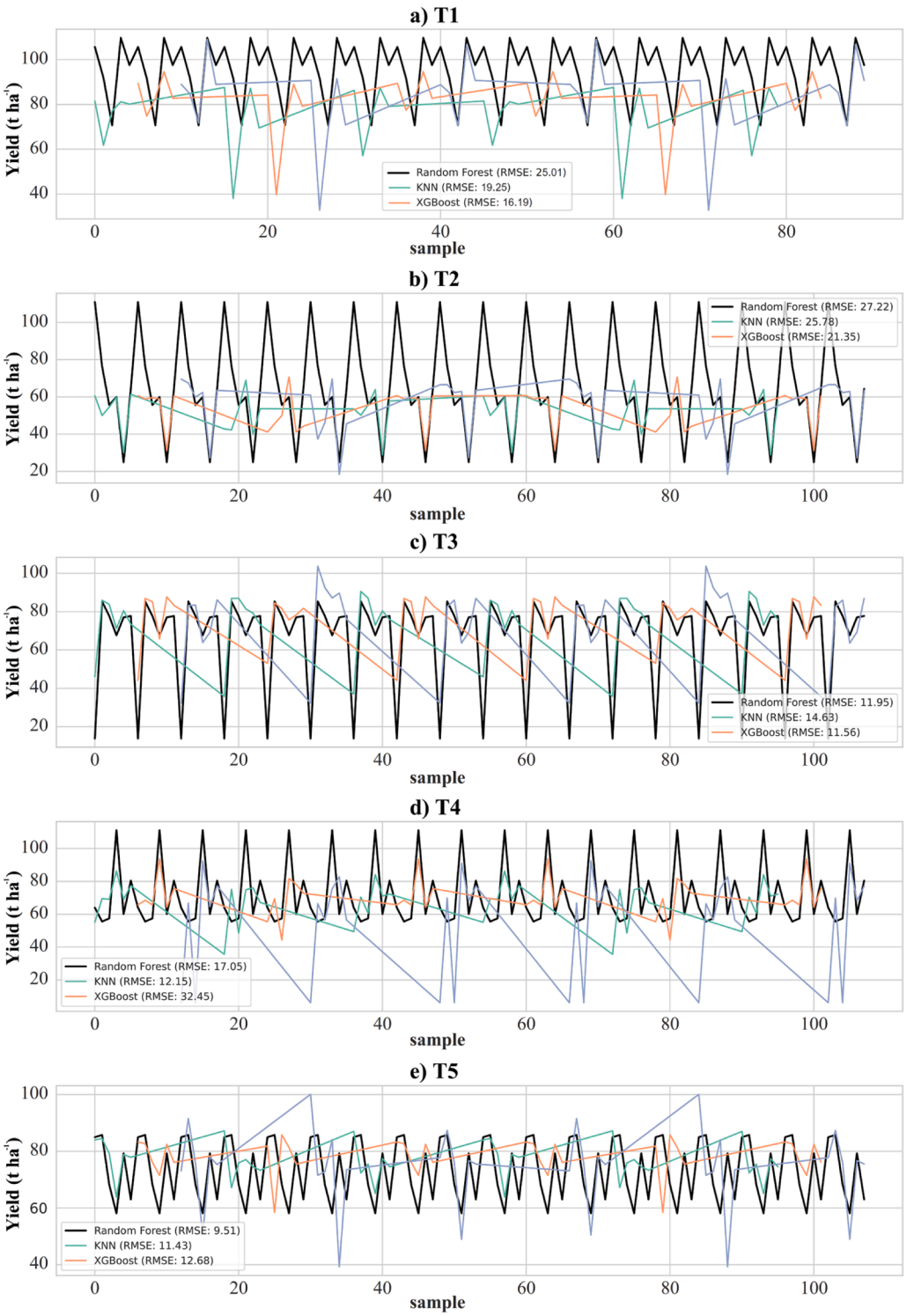

Fig. 16. Determination of the root mean square error (RMSE) of Random Forest regression (black line), K-Nearest Neighbors (KNN) (green line), and XGBoost (orange line). The X-axis represents the sample number evaluated by the prediction algorithms and the Y-axis indicates the yield in kilograms (t ha^{-1}). In each treatment, the mean square error (RMSE) for each model is reported as a measure of accuracy.

region, which limits the extrapolation of these findings to scenarios with different climate variability. Dependence on spectral indices susceptible to saturation also restricts detection of physiological variations in advanced reproductive stages. These limitations underscore the need for multi-year evaluations across diverse agroecological contexts, incorporating hyperspectral sensors and advanced modeling approaches that integrate XGBoost with deep neural networks and explainability tools.

Overall, this study demonstrates that *C. macrocarpum* is a highly effective plant cover for improving papaya productivity through nitrogen fixation, abundant biomass production, and microclimate regulation, enhancing soil-plant interactions and soil physicochemical properties. Furthermore, integration of spectral indices and machine learning algorithms provides a robust framework for real-time yield prediction, supporting decision-making and enabling the development

of more resilient and sustainable tropical fruit systems.

5. Conclusions

This study demonstrates that plant covers are an effective agronomic practice to increase papaya yield under tropical conditions. Among the treatments evaluated, *C. macrocarpum* achieved the highest yield with 102.22 t ha^{-1}, surpassing spontaneous vegetation by >27 t ha^{-1}, which registered 74.50 t ha^{-1}, demonstrating the direct impact of cover on productivity. Likewise, soil analyses revealed modifications in pH, confirming the influence of the covers on soil acidity. Regarding yield prediction, the XGBoost model showed the highest accuracy ($R^2 = 0.85$ with an RMSE error = 11.56) in the treatment with *A. pintoi*, highlighting the potential of machine learning algorithms in precision agriculture. Taken together, these results not only validate the role of mulches in improving productivity and soil health, but also demonstrate that the integration of multispectral, meteorological and agronomic data with predictive models constitutes a robust and applicable strategy to optimize the sustainable management of papaya crops. In this context, extension services and local governments should promote the adoption of vegetative cover crops as productive allies through agroecological programs and climate-smart agriculture strategies.

Ethical statement

The present study did not involve human participants, animals, or the use of any personal or sensitive data. Therefore, ethical approval was not required. All experimental procedures were conducted in accordance with institutional, national, and international guidelines for responsible conduct of research. The authors confirm that the study complies with the ethical standards of the journal and relevant regulations.

Disclosure statement

The authors declare no potential conflict of interest.

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

During the preparation of this work the author used Chat GPT in order to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Pedro A. Torres-Herrera: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Marielita Arce-Inga:** Writing – review & editing, Software, Methodology, Investigation. **Ever Tarrillo:** Formal analysis, Data curation. **Elton Ocupa:** Formal analysis, Data curation. **Nilton Atalaya-Marin:** Writing – review & editing, Software, Investigation. **Héctor Cabrera-Hoyos:** Visualization, Validation, Supervision, Data curation. **Juancarlos Cruz-Luis:** Validation, Supervision, Resources, Project administration. **Victor H. Taboada-Mitma:** Visualization, Validation, Supervision. **Darwin Gómez-Fernández:** Writing – review & editing, Validation, Formal analysis, Data curation. **Daniel Tineo:** Writing – review & editing, Validation, Formal analysis, Data curation. **Malluri Goñas:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Data curation.

Declaration of competing interest

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

Acknowledgements

The authors acknowledge the support of the National Institute of Agrarian Innovation (INIA) through the Investment Project CUI No. 2472675, "Improvement of Research and Agricultural Technology Transfer Services at the Baños del Inca Experimental Agricultural Station," located in the district of Baños del Inca, province of Cajamarca, department of Cajamarca. The authors also wish to thank Yolmer Leonardo Dávila Hernández, Brigith Guadalupe Díaz Zelada, and Johana Marisol Coronado Burga for their valuable contribution to the implementation of this project.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.atech.2026.101953.

Data availability

Data will be made available on request.

References

[1] J.C. Noa-Carrazana, D. González-De-León, B.S. Ruiz-Castro, D. Piñero, L. Silva-Rosales, Distribution of Papaya ringspot virus and Papaya mosaic virus in papaya plants (Carica papaya) in Mexico, Plant Dis. 90 (2006) 1004–1011, https://doi.org/10.1094/PD-90-1004.
[2] B. Patil, S. Tripathi, Differential expression of microRNAs in response to Papaya ringspot virus infection in differentially responding genotypes of papaya (Carica papaya L.) and its wild relative, Front. Plant Sci. 15 (2024), https://doi.org/10.3389/fpls.2024.1398437.
[3] A. Kumar, R. Singh, A.P. Singh, Production Technology of Papaya. Production Technology of Fruit Crops, 2025. https://www.researchgate.net/publication/391369585.
[4] O.A. Galagarza, A. Ramirez-Hernandez, H.F. Oliver, M.V. Álvarez Rodríguez, M.D. C. Valdez Ortiz, E. Pachari Vera, Y. Cereceda, Y.K. Diaz-Valencia, A.J. Deering, Occurrence of chemical contaminants in peruvian produce: a food-safety perspective, Foods 10 (2021), https://doi.org/10.3390/foods10071461.
[5] Ministerio de Desarrollo Agrario y Riego [MIDAGRI], Valor Bruto De La Producción Agropecuaria, Lima, 2025. www.gob.pe/midagri.
[6] Servicio Nacional de Sanidad [SENASA], Guía para la implementación de Buenas Prácticas Agrícolas (BPA) para el cultivo de Papaya, 2023. www.gob.pe/senasa.
[7] Instituto Nacional de Innovación Agraria [INIA], Proyecto de mejoramiento de la red de servicios de innovación, Transferencia Tecnológica y Extensión Tecnológica Agraria En Las Seis Estaciones Experimentales Agrarias Del INIA, Lima, 2024. www.gob.pe/inia.
[8] J.A. Mono, S.E. Ndongo, O.T. Adegono Assiene, A. Mewoli, R.A. Nguefack Assona, R.H. Bitete, G.U. Defo Tatchum, C. Takoumbe, Impact of extraction methods on the properties of Carica papaya pseudostem fibers from Cameroon used as reinforcement in biocomposites, Heliyon 11 (2025), https://doi.org/10.1016/j.heliyon.2024.e41093.
[9] A.M. Sánchez-Moreiras, Y. Vieites-Álvarez, D. Fernández Calviño, C. Campillo, P. Gonzalez-de-Santos, T. Rodríguez Silva, T. Fedoniuk, A. Synowiec, M. Ângelo Almeida Pinheiro de Carvalho, A. Vityi, F. Araniti, J. Nacher, İ. Bolat, R. Zornoza, M. Verdeguer, M. Shanskiy, H. Reinhardt-Weik, H. Jóhannesson, L. Piron, Estrategias agroecológicas para el manejo sostenible de malas hierbas en cultivos europeos de relevancia económica (AGROSUS) agroecological strategies for sustainable weed management in key European crops, Revista de Ciências Agrárias 2024 (2024) 107–112, https://doi.org/10.19084/rca.34857.
[10] C. Maeda, S. Nelson, Anthracnose of Papaya in Hawai'i (2014). www.ctahr.hawaii.edu/freepubs.
[11] S. Nelson, Cooperative Extension Service Phytophthora Blight of Papaya (2008). http://www.ctahr.hawaii.edu/freepubs.
[12] M. Aluja, A. Jiménez, M. Camino, J. Piñero, L. Aldana, V. Castrejón, M.E. Valdés, Habitat manipulation to reduce papaya fruit fly (Diptera: tephritidae) damage: orchard design, use of trap crops and border trapping, J. Econ. Entomol. 90 (1997) 1567–1576, https://doi.org/10.1093/jee/90.6.1567.
[13] J.A. Achicanoy, R. Rojas-Robles, J.E. Sánchez, Análisis y proyección de las coberturas vegetales mediante el uso de sensores remotos y Sistemas de Información Geográfica en la localidad de Suba, Gestión y Ambiente 21 (2018) 41–58, https://doi.org/10.15446/ga.v21n1.68285.
[14] P. Sharma, T. Reitz, S.P. Singh, A. Worrich, E.M. Muehe, Going beyond improving soil health: cover plants as contaminant removers in agriculture, Trends Plant Sci (2025), https://doi.org/10.1016/j.tplants.2025.01.009.

[15] F. Puertas, E. Arévalo, L. Zúñiga, J. Alegre, O. Loli, H. Soplin, V. Baligar, Establishment of cover crops and their growth and nutrient uptake in a humid tropical soil of the Peruvian Amazon, Ecología Aplicada 7 (2008).
[16] S.S. Quispe, K.M. Dávalos, S. Sangay-Tucto, R.C.C. de la Cruz, Use of cover crops for sustainable soil management associated with corn (Zea mays L.) cultivation, Scientia Agropecuaria 12 (2021) 329–336, https://doi.org/10.17268/SCI.AGROPECU.2021.036.
[17] P. Singh, J. Prakash, A.K. Goswami, K. Singh, Z. Hussain, A.K. Singh, Genetic variability and correlation studies for vegetative, reproductive and yield attributing traits in papaya, Indian J. Hortic. 75 (2018) 1–7, https://doi.org/10.5958/0974-0112.2018.00001.4.
[18] L. Kouadio, M. El Jarroudi, Z. Belabess, S.E. Laasli, M.Z.K. Roni, I.D.I. Amine, N. Mokhtari, F. Mokrini, J. Junk, R. Lahlali, A review on UAV-based applications for plant disease detection and monitoring, Remote Sens. (Basel) (2023) 15, https://doi.org/10.3390/rs15174273.
[19] E. Salamí, C. Barrado, E. Pastor, UAV flight experiments applied to the remote sensing of vegetated areas, Remote Sens. (Basel). 6 (2014) 11051–11081, https://doi.org/10.3390/rs61111051.
[20] M. Der Yang, Y.C. Hsu, Y.H. Chen, C.Y. Yang, K.Y. Li, Precision monitoring of rice nitrogen fertilizer levels based on machine learning and UAV multispectral imagery, Comput. Electron. Agric. 237 (2025), https://doi.org/10.1016/j.compag.2025.110523.
[21] J.A. Fernandez-Jibaja, N. Atalaya-Marin, V.H. Taboada-Mitma, J. Cruz-Luis, D. Tineo, Y.A. Álvarez-Robledo, M. Goñas, D. Gómez-Fernández, Integration of agronomic information, vegetation indices (VIs), and meteorological data for phenological monitoring and yield estimation of rice (Oryza sativa L.), Smart Agric. Technol. (2025), https://doi.org/10.1016/j.atech.2025.101203.
[22] B.N. Puelles Condori, J. Cárdenas Rodríguez, A.C. Estrada Zuñiga, Estimación de biomasa y carga animal en humedales ribereños utilizando ortofotografías multiespectrales adquiridas con microsensores transportados en vehículos aéreos no tripulados "Drone, Revista de Investigaciones Altoandinas - J. High Andean Res. 24 (2022) 248–256, https://doi.org/10.18271/ria.2022.442.
[23] D. Ponce de Leon, M. Garcia Rubido, R. Rivera, D. Mancero-Castillo, Y. Garcia, Benefits of canavalia ensiformis, arbuscular mycorrhizal fungi, and mineral fertilizer management in tobacco production, Front. Agron. 6 (2024), https://doi.org/10.3389/fagro.2024.1386656.
[24] D. Venkata Sai Chakradhar Reddy, R.N. Sahoo, T. Kondraju, R.G. Rejith, R. Ranjan, A. Bhandari, A. Moursy, S.C. Tripathi, N. Kumar, Drone-based multispectral imaging for precision monitoring of crop growth variables, Biol. Life Sci. Forum (2025) 10, https://doi.org/10.3390/blsf2025041010.
[25] SENAMHI, Climas del Perú - Mapa de Clasificación Climática Nacional, Lima (2021). www.senamhi.gob.pe.
[26] R.W.H. Rouse, J.A.W. Haas, D.W. Deering, Monitoring vegetation systems in the great plains with erts, Remote Sens. Center (1974) ntrs.nasa.gov/citations/19740022614.
[27] J. Qi, A. Chehbouni, A.R. Huete, Y.H. Kerr, S. Sorooshian, A modified soil adjusted vegetation index, Remote Sens. Environ. 48 (1994) 119–126, https://doi.org/10.1016/0034-4257(94)90134-1.
[28] A. Huete, K. Didan, T. Miura, E.P. Rodriguez, X. Gao, L.G. Ferreira, Overview of the radiometric and biophysical performance of the MODIS vegetation indices, Remote Sens. Environ. (2002). www.elsevier.com/locate/rse.
[29] A.A. Gitelson, M.N. Merzlyak, Y. Grits, Novel algorithms for remote sensing of chlorophyll content in higher plant leaves, Geosci. Remote Sens. Sympos. 4 (1996) 2355–2357, https://doi.org/10.1109/igarss.1996.516985.
[30] E. Barnes, T. Clarke, S. Richards, P. Colaizzi, J. Haberland, M. Kostrzewski, P. Waller, C. Choi, E. Riley, T. Thompson, R.J. Lascano, H. Li, M. Moran, Coincident Detection of Crop Water stress, Nitrogen Status and Canopy Density Using Ground-Based Multiespectral Data (Paper #1356), American Society of Agronomy., 2000. www.cabidigitallibrary.org.
[31] Environmental Protection Agency [EPA], SW-846 Test Method 9045D — Soil and Waste pH (electrometric method) (2004). www.epa.gov.
[32] A. Walkley, I. Black, Estimation of soil organic carbon by the chromic acid titration method, Soil Sci (1934), https://doi.org/10.1097/00010694-193401000-00003.
[33] G.W. Gee, J.W. Bauder, Particle size analysis. Methods of Soil Analysis, Soil Science Society of America, 1986, https://doi.org/10.2136/sssabookser5.1.2ed.c15.
[34] J.M. Bremner, Determination of nitrogen in soil by the Kjeldahl method, J. Agric. Sci. (1960), https://doi.org/10.1017/S0021859600021572.
[35] R. Bray, L. Kurtz, Determination of total organic and available forms of phosphorus in soils, Soil Sci (1945), https://doi.org/10.1097/00010694-194501000-00006.
[36] G.J. Bouyoucos, Hydrometer method improved for making particle size analysis of soils, Agron. J. (1962), https://doi.org/10.2134/agronj1962.00021962005400050028x.
[37] A.E. Hoerl, R.W. Kennard, Ridge regression: biased estimation for nonorthogonal problems, Technometrics 12 (1970), https://doi.org/10.2307/1267351.
[38] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[39] T.M. Cover, P.E. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 24 (1967) 335–342, https://doi.org/10.1109/TIT.1967.1053964.
[40] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, 2016, pp. 785–794, https://doi.org/10.1145/2939672.2939785.
[41] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: Mach. Learn. Python (2018). http://arxiv.org/abs/1201.0490.
[42] The pandas development team, pandas-dev/pandas: pandas (Version 2.2.0) [Software], Zenodo (2024) zenodo.org.
[43] F. Imtiaz, A.A. Farooque, G.S. Randhawa, X. Wang, T.J. Esau, S.E. Hashemi Garmdareh, B. Acharya, Optimizing potato yield mapping and prediction: integrating satellite-based remote sensing and machine learning for sustainable agriculture, Comput. Electron. Agric. 237 (2025), https://doi.org/10.1016/j.compag.2025.110636.
[44] M. de S. Silva, S. Leonel, J.M.A. Souza, R.B. Ferreira, A.C.B. Bolfarini, M.A. de O. Júnior, Correlations between agronomic traits in papaya tree (Carica papaya L.) grown under subtropical climate of Brazil, Aust. J. Crop Sci. 12 (2018) 886–891, https://doi.org/10.21475/ajcs.18.12.06.PNE840.
[45] S. Atmaca, H.İ. Yolcu, G. Erdoğan, H. Gübbük, H. Sert, Comparison of the morphological characteristics, yield, and quality traits of fruits of two Papaya cultivars grown under protected cultivation, Horticulturae 10 (2024), https://doi.org/10.3390/horticulturae10111192.
[46] E.B. Moore, V.R. Radovanovíc, S. Fahad, E.M. Silva, D. Bruce, J.C. Dawson, Cover crop-based reduced tillage management impacts organic squash yield, pest pressure, and management time, 2022. https://doi.org/10.3389/fsufs.2022.991463.
[47] M. Matías-Ramos, C.I. Hidalgo-Moreno, M. Fuentes-Ponce, J. Delgadillo-Martínez, J.D. Etchevers, Potential of legume species as soil fertility enhancers in tropical regions, Revista Mexicana Ciencias Agrícolas 14 (2023) 531–541.
[48] O. Guenni, E. Romero, Y. Guédez, M.P. Macías, D. Infante, Survival strategies of Centrosema molle and C. macrocarpum in response to drought, Trop. Grasslands-Forrajes Tropicales 5 (2017) 1–18, https://doi.org/10.17138/TGFT(5)1-18.
[49] C. Li, Z. Yang, C. Zhang, J. Luo, N. Jiang, F. Zhang, W. Zhu, Heat stress recovery of chlorophyll fluorescence in tomato (Lycopersicon esculentum Mill.) leaves through nitrogen levels, Agronomy 13 (2023), https://doi.org/10.3390/agronomy13122858.
[50] Y. chun Lin, Y. gao Hu, C. zhong Ren, L. chun Guo, C. long Wang, Y. Jiang, X. jiao Wang, H. Phendukani, Z. hai Zeng, Effects of nitrogen application on chlorophyll fluorescence parameters and leaf gas exchange in naked oat, J. Integr. Agric. 12 (2013) 2164–2171, https://doi.org/10.1016/S2095-3119(13)60346-9.
[51] S. Conti Taguali, R. Pöter, F. Aloi, C. Fernández-Trujillo, A. Acedo, F. La Spada, M. G. Li Destri Nicosia, A. Pane, L. Schena, S.O. Cacciola, Influence of environmental and agronomic variables on soil microbiome in citrus orchards: a comparative analysis of organic and conventional farming system, Microbiol. Res. 299 (2025), https://doi.org/10.1016/j.micres.2025.128260.
[52] M. Madin, D. Goodin, L. Moley, K. Nelson, Environmental factors related to biophysical suitability and agronomic effects of biodegradable mulch applications: benchmarking key variables using machine learning, Environ. Chall. 18 (2025), https://doi.org/10.1016/j.envc.2025.101105.
[53] W.S. Lavado Casimiro, J. Ronchail, D. Labat, J.C. Espinoza, J.L. Guyot, Analyse de la pluie et de l'écoulement au Pérou (1969–2004) : bassins versants du Pacifique, du Lac Titicaca et de l'Amazone, Hydrol. Sci. J. 57 (2012) 625–642, https://doi.org/10.1080/02626667.2012.672985.
[54] R. Iqbal, M.A.S. Raza, M. Valipour, M.F. Saleem, M.S. Zaheer, S. Ahmad, M. Toleikiene, I. Haider, M.U. Aslam, M.A. Nazar, Potential agricultural and environmental benefits of mulches—A review, Bull. Natl. Res. Cent. 44 (2020) 75, https://doi.org/10.1186/s42269-020-00290-3.
[55] K. Choudhary, J. Singh, N.K. Meena, N. Al-Ansari, S. Choudhary, R.K. Tiwari, M. Choudhary, D.K. Vishwakarma, S. El-Hendawy, M.A. Mattar, Water volumes and mulches affect plant growth, leaf nutrient status and orchard soil mineral content of sweet orange cv, Mosambi, Sci. Rep. 14 (2024), https://doi.org/10.1038/s41598-024-73262-6.
[56] G. Agegnehu, T. Amede, T. Erkossa, C. Yirga, C. Henry, R. Tyler, M.G. Nosworthy, S. Beyene, G.W. Sileshi, Extent and management of acid soils for sustainable crop production system in the tropical agroecosystems: a review, Acta Agric. Scand. B Soil Plant Sci. 71 (2021) 852–869, https://doi.org/10.1080/09064710.2021.1954239.
[57] R. Dong, W. Hu, L. Bu, H. Cheng, G. Liu, Legume cover crops alter soil phosphorus availability and microbial community composition in mango orchards in karst areas, Agric. Ecosyst. Environ. 364 (2024) 108906, https://doi.org/10.1016/j.agee.2024.108906.
[58] J. Rousk, E. Bååth, P.C. Brookes, C.L. Lauber, C. Lozupone, J.G. Caporaso, R. Knight, N. Fierer, Soil bacterial and fungal communities across a pH gradient in an arable soil, ISME J. 4 (2010) 1340–1351, https://doi.org/10.1038/ismej.2010.58.
[59] S. Zhang, Q. Zhu, W. de Vries, G.H. Ros, X. Chen, M.A. Muneer, F. Zhang, L. Wu, Effects of soil amendments on soil acidity and crop yields in acidic soils: a worldwide meta-analysis, J. Environ. Manage. 345 (2023) 118531, https://doi.org/10.1016/j.jenvman.2023.118531.
[60] M. Abdalla, A. Hastings, K. Cheng, Q. Yue, D. Chadwick, M. Espenberg, J. Truu, R. M. Rees, P. Smith, A critical review of the impacts of cover crops on nitrogen leaching, net greenhouse gas balance and crop productivity, Glob. Chang. Biol. 25 (2019) 2530–2543, https://doi.org/10.1111/gcb.14644.
[61] M.B. Villamil, G.A. Bollero, R.G. Darmody, F.W. Simmons, D.G. Bullock, No-till corn/soybean systems including winter cover crops, Soil Sci. Soc. Am. J. 70 (2006) 1936–1944, https://doi.org/10.2136/sssaj2005.0350.
[62] H. Gao, G. Tian, M. Khashi u Rahman, F. Wu, Cover crop species composition alters the soil bacterial community in a continuous pepper cropping system, Front. Microbiol. 12 (2022), https://doi.org/10.3389/fmicb.2021.789034.
[63] A. Kumar, S. Vallabhbhai Patel, Y. Prasad, P. Chaudhary, N. Kumar, C. Ankur Kumar, Studies on genetic variability, character association and path analysis

among yield and yield contributing traits in papaya (carica papaya L.), J. Pharmacogn. Phytochem. (2018) 1.
[64] M. Palchaudhuri, S. Biswas, Application of LISS III and MODIS-derived vegetation indices for assessment of micro-level agricultural drought, Egypt. J. Remote Sens. Space Sci. 23 (2020) 221–229, https://doi.org/10.1016/j.ejrs.2019.12.004.
[65] B. Matsushita, W. Yang, J. Chen, Y. Onda, G. Qiu, Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: a case study in high-density cypress forest, Sensors 7 (2007) 2636–2651. www.mdpi.org/sensors.
[66] A.A. Gitelson, Y.J. Kaufman, M.N. Merzlyak, J. Blaustein, Use of a green channel in remote sensing of global vegetation from EOS-MODIS, Remote Sens. Environ. 58 (1995) 10010.
[67] A.A. Gitelson, Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation (2004). http://www.elsevier-deutschland.de/jplhp.
[68] D.R. Roberts, V. Bahn, S. Ciuti, M.S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J.J. Lahoz-Monfort, B. Schröder, W. Thuiller, D.I. Warton, B. A. Wintle, F. Hartig, C.F. Dormann, C.F.D. David R. Roberts, Volker Bahn, Simone Ciuti, Mark S. Boyce, Jane Elith, Gurutzeta Guillera-Arroita, Severin Hauenstein, José J. Lahoz-Monfort, Boris Schröder, Wilfried Thuiller, David I. Warton, Brendan A. Wintle, Florian Hartig, Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure, Ecography 40 (2017) 913–929, https://doi.org/10.1111/ecog.02881.
[69] F. Huber, A. Yushchenko, B. Stratmann, V. Steinhage, Extreme gradient boosting for yield estimation compared with Deep Learning approaches, Comput. Electron. Agric. 202 (2022) 107346, https://doi.org/10.1016/j.compag.2022.107346.