agronomy

MDPI

Article

Estimation of Strawberry Crop Productivity by Machine Learning Algorithms Using Data from Multispectral Images

Larissa Silva de Oliveira [1,*], Renata Castoldi [2], George Deroco Martins [3] and Matheus Henrique Medeiros [1]

[1] Institute of Agricultural Science, Universidade Federal de Uberlândia, Campus Glória, BR 050, km 78, Uberlândia 38410-337, MG, Brazil; matheusmedeiros@ufu.br
[2] Institute of Agricultural Science, Universidade Federal de Uberlândia, Campus Monte Carmelo, LMG-746, km 1, Monte Carmelo 38500-000, MG, Brazil; rcastoldi@ufu.br
[3] Institute of Geography, Universidade Federal de Uberlândia, Campus Monte Carmelo, LMG-746, km 1, Monte Carmelo 38500-000, MG, Brazil; deroco@ufu.br
* Correspondence: larissaagroliveira@outlook.com

Abstract: Currently, estimations of strawberry productivity are conducted manually, which is a laborious and subjective process. The use of more efficient and precise estimation methods would result in better crop management. The objective of this study was to assess the performance of two regression algorithms-Linear Regression and Support Vector Machine—in estimating the average weight and number of fruits and the number of leaves on strawberry plants, using multispectral images obtained by a remotely piloted aircraft (RPA). The experiment, which was conducted in the experimental area of the Botany Laboratory at the Federal University of Uberlândia-Monte Carmelo Campus (Universidade Federal de Uberlândia, Campus Monte Carmelo), was carried out using a randomized block design with six treatments and four replications. The treatments comprised six commercial strawberry varieties: San Andreas, Albion, PR, Festival, Oso Grande, and Guarani. Images were acquired on a weekly basis and then preprocessed to extract radiometric values for each plant in the experimental area. These values were then used to train the production prediction algorithms. During the same period, data on the average fruit weight, number of fruits per plant, and number of leaves were collected. The total fruit weight in the field was 48.08 kg, while the linear regression (LR) and Support Vector Machine (SVM) estimates were 48.04 and 43.09 kg, respectively. The number of fruits obtained in the field was 4585, and the number estimated by LR and SVM algorithms was 4564 and 3863, respectively. The number of leaves obtained in the field was 10,366, and LR and SVM estimated 10,360 and 10,171, respectively. It was concluded that LR and SVM can estimate strawberry production and the number of fruits and leaves using multispectral unmanned aerial vehicle (UAV) images. The LR algorithm was the most efficient in estimating production, with 99.91% accuracy for average fruit weight, 99.55% for the number of fruits and 99.94% for the number of leaves. SVM exhibited 89.62% accuracy for average fruit weight, 84.26% for the number of fruits, and 98.12% for the number of leaves.

Keywords: prediction models; regression algorithms; strawberry productivity; vegetation indices

check for updates

Citation: Oliveira, L.S.d.; Castoldi, R.; Martins, G.D.; Medeiros, M.H. Estimation of Strawberry Crop Productivity by Machine Learning Algorithms Using Data from Multispectral Images. *Agronomy* **2023**, *13*, 1229. https://doi.org/10.3390/agronomy13051229

Academic Editors: Yanbo Huang, Ali C. Gurbuz, Xin Zhang and Gniewko Niedbała

Received: 26 February 2023
Revised: 17 April 2023
Accepted: 19 April 2023
Published: 27 April 2023

1. Introduction

Despite its high production, strawberry cultivation requires a high level of technological knowledge, numerous crop management practices, and modern phytosanitary management methods, in addition to continuous monitoring, which is normally performed manually in the field [1]. Thus, the use of tools such as remote sensing may enhance crop management [2–5]. Remote sensing (RS) provides information on objects and surfaces by recording the electromagnetic energy absorbed or reflected by these targets using sensors installed on platforms. According to [6], using platforms such as unmanned aerial vehicles (UAVs) produces low-cost, high-resolution images [7,8]. In this respect, remote sensing techniques based on multispectral images with more than one electromagnetic image band have been increasingly used to create estimating and yield mapping models [9].

In addition to remote sensing, vegetation indices (VIs) and machine learning (ML) have performed well in quantifying and understanding data-intensive processes in operational agricultural environments [10]. According to [11], vegetation indices are the most used resource for forecasting crop yields. ML and VIs have been applied in different agricultural contexts, such as strawberry phenotyping and crop management [12], in order to predict leaf color [13], assess fruit quality [14], detect pests and diseases on the leaves [15], estimate harvest time [16], forecast yield and price [17], recognize ripe strawberries [18], and carry out automatized harvesting [19], among other applications. The authors of [20] show that vegetation indices are good substitutes for monitoring the temporal dynamics of sweet potato crop growth and for differentiating phenological stages.

Real-time detection of powdery mildew on strawberry plants was performed by [21] with 87% accuracy, while [22] found Eleusine indica in strawberry and tomato crops with 93% and 77% accuracy, respectively.

The number of plants can be obtained with aerial images in order to estimate their production [23]. According to [24], the number of strawberry plants obtained from unmanned aerial vehicle images exhibits less than 4% prediction errors, demonstrating the feasibility of automatically estimating strawberry crops. The authors of [25] managed to predict tomato productivity 6 weeks before harvest, with a percentage error of 9.28% in relation to the average production volume (154.5 tons).

The strawberry yield was predicted by [2] using high-resolution images, whereby flower and fruit counts showed prediction errors of 26.3% and 25.7%, respectively. The strawberry yield based on a deep neural network was predicted by [3], using high-resolution aerial orthoimages, with an average accuracy of 83% and 72% for all the objects detected at a height of 2 and 3 m, respectively.

Several other studies have applied RS in a cultivated area for different purposes, such as predicting the presence of metals in the soil [26]; estimating soil moisture content [27]; spatial monitoring of agricultural drought [28]; estimating crops [29]; detecting biotic stress in plants [30]; detecting plant nutritional status [31]; certifying organic crops [32]; monitoring pests and diseases [33]; and estimating biomass [34], among others.

Although there are a number of studies on the use of RS in strawberry cultivation [2–5,35,36], this technique may be limited to estimating yield because the plants are relatively small with indeterminate growth patterns, that is, they exhibit leaves, stolons, inflorescences, and fleshy fruits, with several harvests during the cycle [4]. RS has been widely applied to estimate the production of annual crops, which have a specific ripening period and a single harvest during the cycle [37–40].

Remote sensing makes it possible to count strawberry flowers and accurately estimate yield [3]. Efficient and accurate strawberry production estimates would enable resource planning for the harvest, transport, and commercialization of the product. The estimation process is normally conducted manually, a lengthy, laborious, and subjective endeavor [3]. As such, the objective of the present study was to assess the performance of linear regression and support vector machine algorithms in estimating the number and average weight of fruits and the number of strawberry leaves using multispectral UAV images.

2. Materials and Methods

2.1. Field Experiment

The experiment was conducted in the experimental area of the Botany Laboratory (LABOT) of the Federal University of Uberlândia-Monte Carmelo Campus (Universidade Federal de Uberlândia, Campus Monte Carmelo) (located at 18°43′36.05″ South latitude, 47°31′31.77″ West longitude at an average altitude of 902 m) (Figure 1).

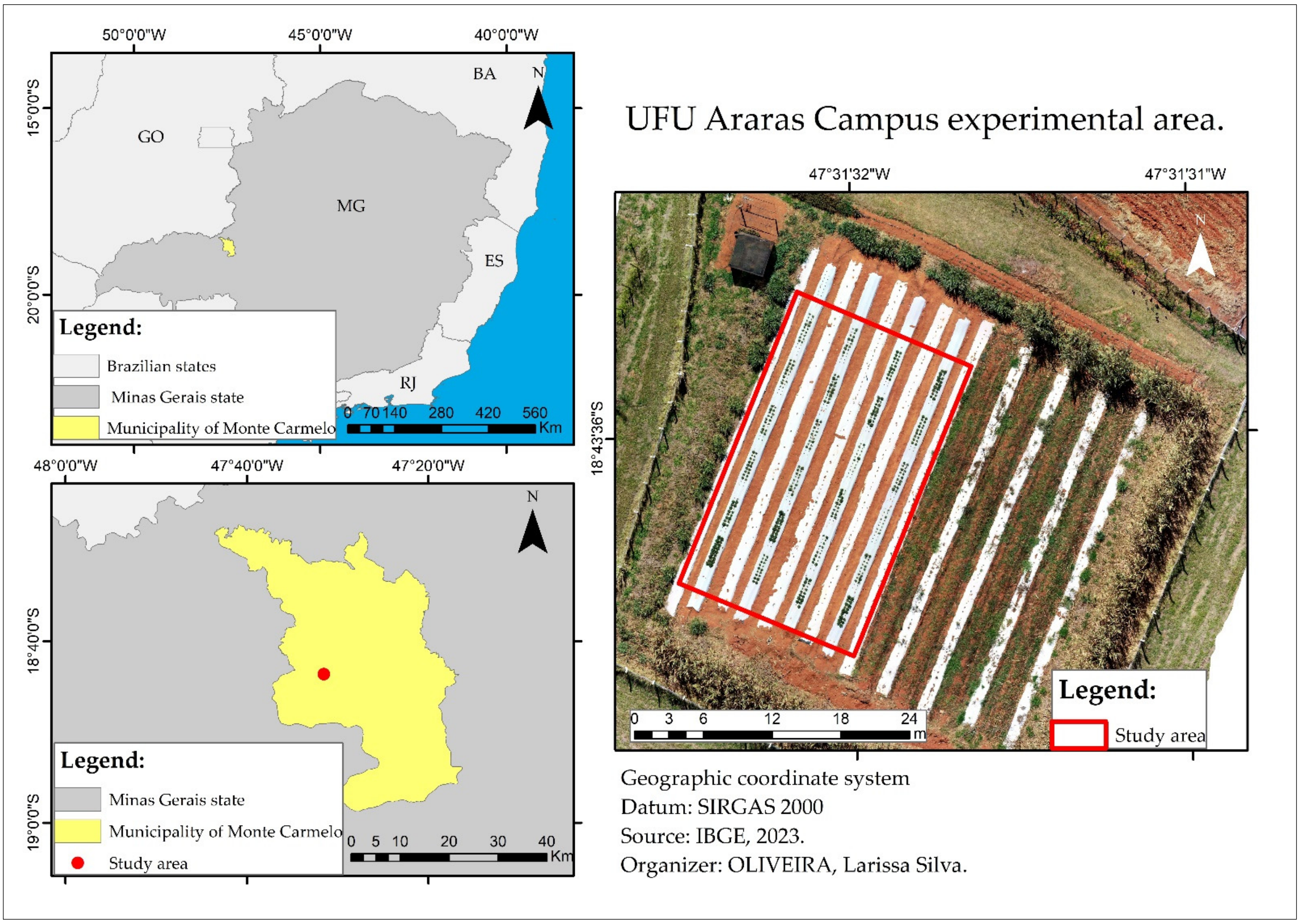


Figure 1. Location map of the experimental area.

A randomized block design, comprising six treatments and four replications (totaling 24 experimental plots), was used. The treatments consisted of six commercial strawberry cultivars: San Andreas, Albion, PR, Festival, Oso Grande, and Guarani. The use of different cultivars was for generating genetic variability to test the estimation accuracy of the algorithms.

Each experimental plot consisted of 18 plants, arranged into two rows spaced 0.3 m apart, with a 0.3 m space between plants. The 10 central plants of each plot were considered for evaluation. Before the installation of the experiment, the soil was sampled at a depth of 0–0.2 m and taken to the soil analysis laboratory for the determination of its chemical and physical properties. The results obtained are shown in Table 1.

Table 1. Soil chemical characterization.

Deep (cm)	pH H_2O (1:2.5)	P Meh mg dm^{-3}	K^+	$H^+ + Al^{3+}$	Al^{3+}	Ca^{+2} cmolc dm^{-3}	Mg^{+2}	SB	T	t	V %
0–20	6.2 ^{L}	9.8 ^{L}	0.34 ^{L}	1.90 ^{M}	0.00 ^{VL}	2.65 ^{L}	0.58 ^{M}	3.57 ^{M}	5.47 ^{M}	3.57 ^{L}	65 ^{M}

^{L}: Low; ^{M}: Medium; ^{VL}: Very low; pH H_2O = hydrogen potential in water; P meh = phosphor; K^+ = potassium; $H^+ + Al^{3+}$ = hydrogen + aluminum; Al^{3+} = aluminum; Ca^{+2} = calcium; Mg^{+2} = magnesium; SB = Sum of Bases; T = CTC pH 7.0; t = effective CTC; V = saturation of bases; pH in H_2O, Ca, Mg, Al = KCl solution (1 mol L^{-1}); P, K = 0.05 mol L^{-1} HCl + H_2SO_4 0.0125 mol L^{-1}; H + Al = SMP buffer solution (pH 7.5). Methodology source [41].

To increase the V% to 80%, liming was performed 60 days before transplanting the seedlings. In total, 1.7 tons of limestone per hectare was spread evenly over the entire area. The limestone used contained 30% CaO and 8% MgO and had a total neutralization relative power of 80%.

Fifty days after liming, four planting beds (each 21.4 m long and 1.2 m wide) were created using a 1.20 m-wide rotary cultivator attached to a tractor. Then, planting fertilizers were distributed in the beds. Following the guidelines of [41], a fertilizer composition of 220 kg ha^{-1} N, 200 kg ha^{-1} P_2O_5, and 80 kg ha^{-1} K_2O was applied for the entire crop cycle, accounting for 16% of N, 100% of P_2O_5, and 70% of K_2O, which were derived from urea, single superphosphate, and potassium nitrate, respectively. Top-dressing fertilizer was applied weekly through the irrigation system to distribute the remaining 84% of urea-derived N and 30% of potassium nitrate-derived K_2O, totaling 23 fertilizer applications via irrigation over the 164 days of the experiment.

On 16 March 2020, the seedlings were transplanted to the beds, which were then covered with a black- and white-sided plastic film (mulching), with the black side facing the soil and the white side facing outward. The seedlings were planted in holes of 75 mm diameter, which were drilled along the entire length of the plastic film following the spatial distribution of the plants in the experimental plots.

A drip irrigation system consisting of two dripping tubes per bed was used, with a daily watering schedule applied. The schedule was based on recommendations for the strawberry crop coefficient (Kc) and the daily evapotranspiration data from a weather station located 200 m from the experiment site. The drippers were self-compensating, with a diameter of 25 mm and flow rate of 1.6 L h^{-1}.

To control insect pests, the insecticide Abamectin Nortox® 400 WG, Nortox, Arapongas, Brazil was used at a concentration of 75 mL per 100 L of solution, which was applied using a 20 L manual backpack pump. To control microbial diseases, Metiltiofan®, Sipcam Nichino Brazil S.A., Uberaba, Brazil (70 g per 100 L of solution) and Folicur® 200 EC, Bayer S.A., São Paulo, Brazil (75 g per 100 L of solution) were applied on a weekly basis using a manual backpack pump.

Harvesting—defined as the point when the fruits had reached 75% of their full size and displayed a red coloration—was carried out weekly, starting 40 days after planting (DAP) the seedlings and ending 164 DAP, which resulted in a total of 16 harvests. After harvesting, the fruits were transported to the Botany Laboratory at the Federal University of Uberlândia for crop productivity evaluation; namely, determining the number of fruits per plant and the average weight of the fruits, the latter of which was calculated by dividing the total fruit weight per plant by the number of fruits per plant. The values were expressed in grams (g).

Additionally, the number of leaves on each plant in the useful area of all plots was evaluated every two weeks. This was performed 11 times to allow enough time for new leaves to develop and be counted.

2.2. Image Acquisition

The image acquisition flights were carried out weekly at peak radiation time (between 12 PM and 1 PM) for a total of four months (from April to August) during the entire crop production cycle. Of these flights, 16 were conducted to obtain data on the number of fruits per plant and average fruit weight, and 11 flights were conducted to obtain data on the number of leaves per plant.

The images were captured using a Phantom 4 Pro RPA developed by Da-Jiang Innovations Science and Technology Co., Ltd. (DJI), Shenzhen, China. The RPA was also equipped with a Mapir Survey3W camera, which has a resolution of 12 megapixels and green (550 nm), red (660 nm), and near-infrared (850 nm) channels with FWHM values of 40, 60, and 80 nm, respectively.

The flight plan was created using the free Drone Deploy© app, California, USA, with a front overlap of 80% and lateral overlap of 80%. Five stripes were established, with an altitude of 30 m, a speed of 3 m/s, a flight time of approximately 4 min, and a photo capture interval of 3 s. For the Mapir Survey3W camera, atmospheric correction was performed for all flights.

2.3. Image Preprocessing

2.3.1. Mosaic Creation

For each flight, mosaics of the images captured by the Mapir sensor were created using Agisoft PhotoScan Professional software. The program aligned the images by determining the camera position and finding common points between them, resulting in a sparse point cloud. A dense point cloud was then created from the sparse point cloud, and a 3D polygonal mesh model was used to generate a surface, producing a total of 16 mosaics for each flight and sensor.

2.3.2. Atmospheric Correction

The purpose of atmospheric correction is to reduce the impact of atmospheric and topographical disturbances on images. In this study, this was achieved prior to each flight by using the Mapir Survey3W camera to take a photograph of the atmospheric correction target (Figure 2), which consisted of four pallets of known reflectance values in different colors (white, black, light gray, and dark gray). The correction process was then carried out using Mapir Camera Control software.

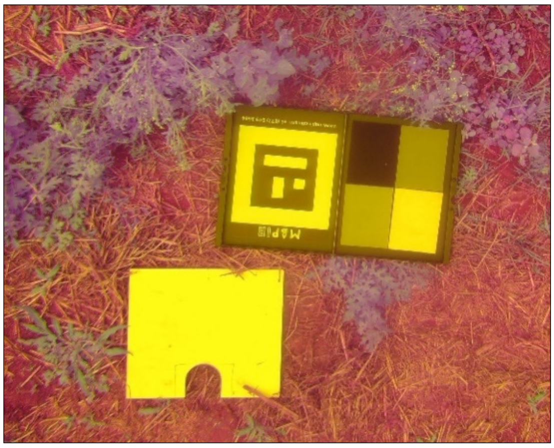


Figure 2. Atmospheric correction target for the Mapir Survey3W camera.

Atmospheric correction was achieved by inserting the mosaics into the software program and specifying the lens (87° HFOV, 19 mm, f/2.8 aperture, −1% distortion, glass lens) and camera filter used. The image containing the target for the flight day was then selected. Using this information, the software program automatically performed the atmospheric correction, producing a corrected orthomosaic with surface reflectance values as the final output.

2.3.3. Radiometric Evaluation

The normalization process was carried out using the Environment for Visualizing Images (ENVI) 5.0 program. First, a reference image was selected, followed by the image to be normalized. Then, the program selected the brightest and darkest areas on the image and extracted reflectance values from the Mapir camera's bands (R for red, G for green, and N for near-infrared) for each flight. Finally, for each spectral band, these values were applied to Equation (1) as follows:

$$\mathrm{Ti} = \mathrm{mi} \times \mathrm{xi} + \mathrm{bi} \quad (1)$$

where

mi = (Bri × Dri)/(Bsi − Dsi)
bi = (Dri × Bsi − Dsi × Bri)/(Bsi − Dsi)
Ti = normalized image reflectance
xi = reflectance of the original image to be normalized
Bri = average of the clear reference set
Dri = mean of the dark reference set
Bsi = mean of the clear set to be normalized
Dsi = mean of the dark set to be normalized
i = bands of the sensor under study

The equation was processed through the Band Math function in ENVI 5.0, which generated separate normalized images for each spectral band. These images were then combined using the Build Layer Stack tool in ENVI 5.0 to create a single file.

2.4. Extraction of Radiometric Data and Calculation of the Normalized Difference Vegetation Index

Radiometric data (i.e., the reflectance values captured by the Mapir camera) were obtained from 240 plants. This process was conducted for each of the 16 flights. The extraction was performed using the corrected orthoimages and the Region of Interest (ROI) function in ENVI 5.0. A polygon was created over the ROI of each plant to extract the radiometric values, which were then tabulated to form the classification data set.

The radiometric data set from 240 plants were processed to remove outliers (data points that deviated from the pattern) in order to reduce the errors produced by classifiers. Outliers were defined as plants that were assigned null production values. However, this does not mean that these plants did not produce fruits; it simply indicates that no fruits were present at the time of data collection. After removal of the outliers, a total of 150 sample data points remained.

The normalized difference vegetation index (NDVI) was calculated using the average reflectance values from the Mapir camera (Equation (2)), as it is a popular index and is known for its sensitivity to vegetation.

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{RED}}{\mathrm{NIR} + \mathrm{RED}}, \tag{2}$$

where

NIR = the reflectance value in the near-infrared spectral range
RED = the reflectance value in the red spectral range

2.5. Supervised Regression

The tabulated reflectance values were used to train the classifier algorithms (LR and SVM) available in the Waikato Environment for Knowledge Analysis (Weka) 3.9.4 software program. The algorithms were applied using standard Weka settings. For the LR algorithm, these were: attribute selection method: M5 method; batch size: 100; debug: false; do not check capabilities: false; eliminated colinear attributes: true; minimal: false; num decimal places: 4; output additional stats: false; ridge: 1.0×10^{-8}; and use QR decomposition: false. For the SVM algorithm, the settings were batch size: 100; c: 1.0; debug: false; do not check capabilities: false; filter type: normalize raining data; kernel: Polykernel–E1.0-C250007; num decimal places: 2; and reg optimizer: RegSMOimproved. The performance was evaluated using the Cross Validation function, with the number of folds being equal to the number of samples in the inserted data set, which was a total of 150 samples per flight.

The regression data sets were created using data collected in the field. These data included information about the average weight of the fruits, the number of fruits per plant, and the number of leaves per plant as well as values from radiometric data from the sensor bands of the Mapir camera. The LR classifier and SVM algorithms of Weka were used to analyze the average fruit weight, number of fruits, and number of leaves.

2.6. Analysis of the Algorithms

The performances of the LR and SVM algorithms were evaluated on the basis of the correlation coefficient (CC) and root mean squared error (RMSE) values. The CC measures the relationship of the average fruit weight, number of fruits, and number of plant leaves with the NDVI (near-infrared and red) spectral bands. By contrast, the RMSE indicates how much the algorithm's estimate deviates from the actual field measurement.

After obtaining the CC and RMSE values, the Student *t*-test was applied to the data that met the required assumptions for the test (normal distribution, normality of residuals, homogeneity of variances, and additivity) to analyze if there was a difference between the two algorithms. Because only the "number of leaves" variable evaluated in flight 5 met these assumptions, the Student *t*-test was applied for that variable and flight only.

For flights that did not meet the criteria for the Student *t*-test, the Mann–Whitney–Wilcoxon test was used instead. This test has the same purpose as the Student *t*-test, but instead of examining the means of two groups for a specific characteristic, it analyzes the medians of the groups to determine if there are any differences between them for that variable. The tests were conducted using R Studio v.4.1.2 [42].

Analysis of the 16 flights was performed both collectively and individually. In the overall flight analysis, the CC and RMSE values generated by the LR and SVM algorithms were analyzed using the Mann–Whitney–Wilcoxon test to determine if there was a difference between the two algorithms in determining these two sets of values. The individual flights were also analyzed to determine which algorithm showed the best performance in estimating the average fruit weight, number of fruits, and number of plant leaves for each flight.

3. Results

Figure 3A shows the experimental area immediately after strawberry seedling transplanting in the vegetative development phase, characterized by very small plants with a reduced leaf area, the absence of fruits and a little apparent vegetative canopy in the image. Figure 3B illustrates well-developed strawberry plants 40 days after transplanting, with an apparent and more voluminous canopy in relation to Figure 3A. At this time, harvesting and assessments were initiated, obtaining 9.28 leaves per plant, 1.57 fruits per plant and average fruit weight of 13.40 g in the first assessment.

(**A**) (**B**)

Figure 3. Experimental area immediately after strawberry seedling transplanting (**A**), 40 days after and at the onset of assessments (**B**).

When comparing the CCs, which reflect the degree of association between two or more variables, no significant difference was found between the LR and SVM algorithms in terms of the correlation of the average fruit weight, number of fruits, or number of plant leaves with the near-infrared and red spectral bands. With the LR algorithm, the highest average CC value obtained was 0.82 for the number of leaves, whereas the values were 0.51 and 0.42 for the average fruit weight and number of fruits, respectively; that is, both had a lower correlation with the spectral bands as compared with the number of leaves (Table 2).

Table 2. Mean correlation coefficients reflecting the relationship of the average fruit weight, number of fruits, and number of plant leaves with the spectral bands, and root mean squared error values of the Linear Regression and Support Vector Machine algorithms for estimating the average fruit weight, number of fruit, and number of plant leaves from the data of the 16 flights.

	Average Fruit Weight		**Number of Fruits**		**Number of Leaves**	
	CC	**RMSE**	**CC**	**RMSE**	**CC**	**RMSE**
M (*p*-value)	120 (0.7804 ^{ns})	131(0.9260 ^{ns})	121 (0.8091 ^{ns})	121 (0.8065 ^{ns})	27 (0.1797 ^{ns})	10 (0.2403 ^{ns})
MLR	0.5180	14.2046	0.4290	1.5022	0.8213	2.8497
MSVM	0.5427	13.9049	0.4393	1.5268	0.7032	10.2891

CC: correlation coefficient. RMSE: root mean squared error. MLR: mean of the Linear Regression algorithm. MSVM: mean of the Support Vector Machine algorithm. M: Mann–Whitney–Wilcoxon test value at 5% significance. *p*-value: probability F of Snedecor. ^{ns}: not significant.

Of the three variables analyzed, the number of leaves was found to be the one that correlated the most with the spectral bands (CC of 0.8213 for LR and 0.7032 for SVM; Table 2). This meant that this variable camouflaged the other two variables with regard to the spectral reflectance of the plant. The high correlation can be attributed to the fact that the vegetative canopy absorbs a low amount of energy, especially in the near-infrared region, and therefore reflects a high amount of light. When the number of leaves was correlated with the NDVI bands, it stood out as the most significant agronomic variable compared with the average fruit weight and number of fruits.

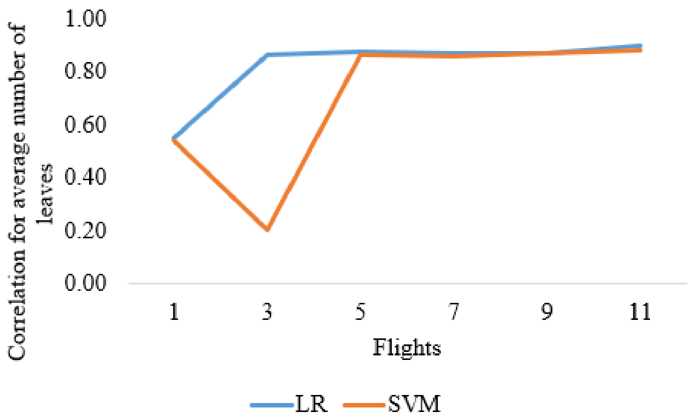
During the strawberry production cycle, the strongest correlations between the red/near-infrared spectral bands and the average fruit weight (Figure 4A) and number of fruits (Figure 4B) were found in flights 1–5 and 10–16, whereas the weakest correlations were found in flights 6–9. During the same period, the CC values for the number of leaves (Figure 4C) were greater than 0.8, indicating that the reflectance associated with this variable is more closely linked to the leaf features than to the spectral behaviors of average fruit weight and number of fruits.

This is confirmed by the fact that during the same time window, the correlations between the red/near-infrared spectral bands and the number of leaves were higher (LR: 0.82; SVM: 0.70), while those between the spectral bands and the average fruit weight (LR: 0.51; SVM: 0.54) and the number of fruits (LR: 0.42; SVM: 0.43) were lower. These results are due to the fact that during flights 6–9, there was an increase in the number of leaves per plant and hence the leaf area index, which affected the reflectance of the plants in the near-infrared region.

(**A**)

Figure 4. *Cont*.

(**B**)



(**C**)

Figure 4. Correlation coefficient values for the average fruit weight (**A**), number of fruits (**B**), and number of leaves (**C**) estimated using the Linear Regression (LR) and Support Vector Machine (SVM) algorithms for six strawberry cultivars during all flights.

With regard to the RMSE, which reflects the difference between the field-measured values and the algorithm-estimated ones, there was no significant difference found between the LR and SVM algorithms for the variables studied (Table 2). For the average fruit weight, the RMSE of the estimates generated with the LR and SVM algorithms were 14.2 and 13.9 g plant^{-1}, respectively. By contrast, the RMSE values of the LR- and SVM-determined estimates were, respectively, 1.5 and 1.53 fruits plant^{-1} for the number of fruits and 2.8 and 10.28 leaves plant^{-1} for the number of leaves.

The error between the field-measured number of leaves and the SVM-determined estimate was high (45.83 leaves plant^{-1}) for the flight 3 data (Figure 5C). This could be due to the presence of an outlier in the data set, which caused an increase in the RMSE value for this variable. However, the LR algorithm estimated the number of leaves effectively, generating a low average RMSE value of 2.84 leaves plant^{-1} for the same flight. Considering that the SVM algorithm was able to estimate the number of leaves with low error values for flights 1, 2, and 4–11 and that the LR algorithm was also able to estimate this variable with low errors using the same data set, we can conclude that the increase in RMSE in flight 3 was caused by the presence of an outlier.

Flights 1–6, 10, and 13–16 showed lower RMSE values for estimating the average fruit weight. There were peaks in flights 7, 8, 9, 11, and 12 where the estimation error of the actual fruit weight was higher (Figure 5A). During this time, the correlation between the average fruit weight and the spectral bands was lower (Figure 4A) and the number of leaves per plant increased (Figure 4C). This indicates that the plants were going through a period of increased energy investment in vegetative development, with more energy and nutrients being directed toward leaf production.

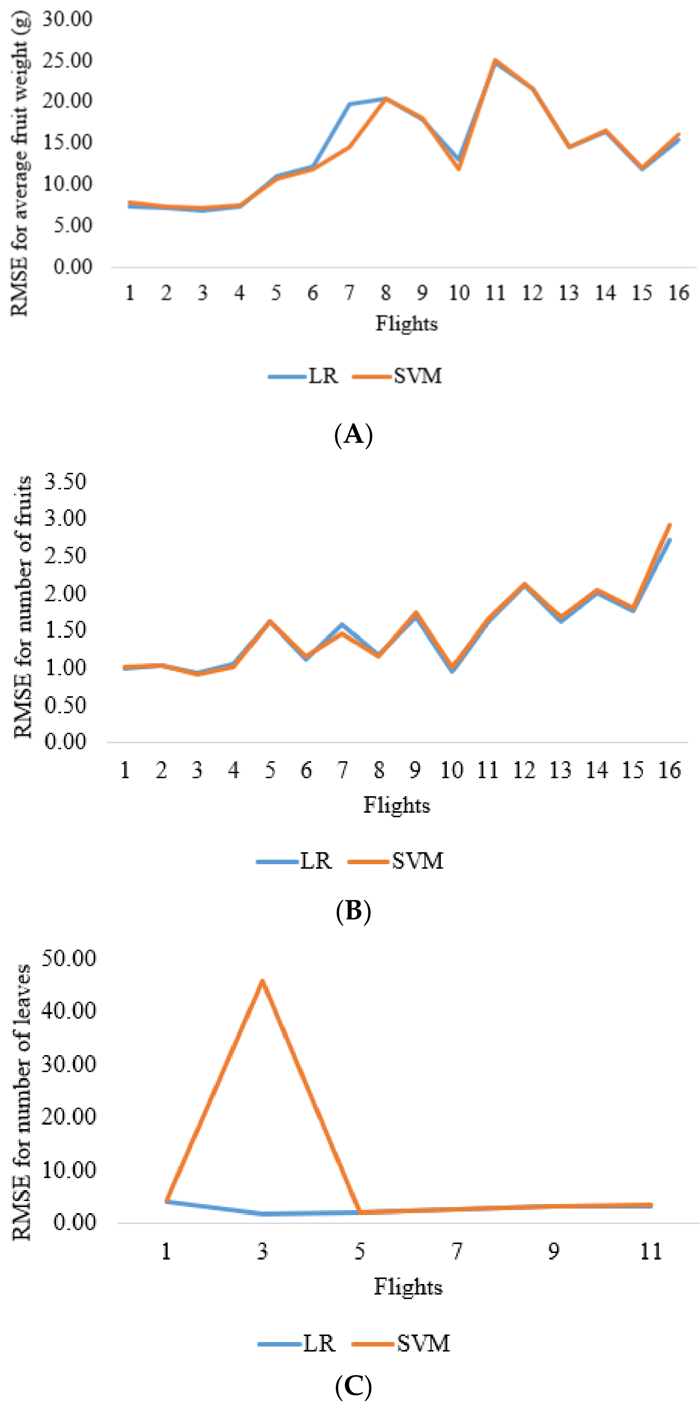


Figure 5. Root mean squared error (RMSE) values for the average fruit weight (**A**), number of fruits (**B**), and number of leaves (**C**) estimated using Linear Regression (LR) and Support Vector Machine (SVM) algorithms for six strawberry cultivars during all flights.

In the near-infrared region, the reflectance will increase if the leaf area index increases. This is because most of the energy reflected in this part of the electromagnetic spectrum is attributed to the internal structure of the leaves. This may have reduced the plant's spectral behavior in relation to the average fruit weight.

With regard to the variable numbers of fruits per plant, the two algorithms generated similar estimates, with errors ranging from 1 to 2 fruits plant^{-1} for most flights, except for flights 12 and 16, which had errors above 2 fruits plant^{-1} (Figure 5B). By contrast, the RMSE values for number of leaves were below 5 leaves plant^{-1} throughout the analyzed period, except for flight 3, where the SVM algorithm had a peak error of 45.83 leaves plant^{-1} (Figure 5C).

By comparing the average fruit weight measured in the field with the mean values estimated using the LR and SVM methods, it was found that both algorithms were effective in estimating this variable, generating results that were very close to those obtained in the field. With regard to the performance of the algorithms per flight, the LR algorithm was found to show greater accuracy in estimating the average fruit weight for most flights. For flights 1, 2, 5, 7, 12, 14, and 15, the LR and SVM algorithms did not differ significantly. For this variable, the LR and SVM estimates differed the most significantly for flights 9, 10, and 13, whereas the values were not significantly different for flights 1, 2, 5, 7, 12, 14, and 15 (Figure 6).

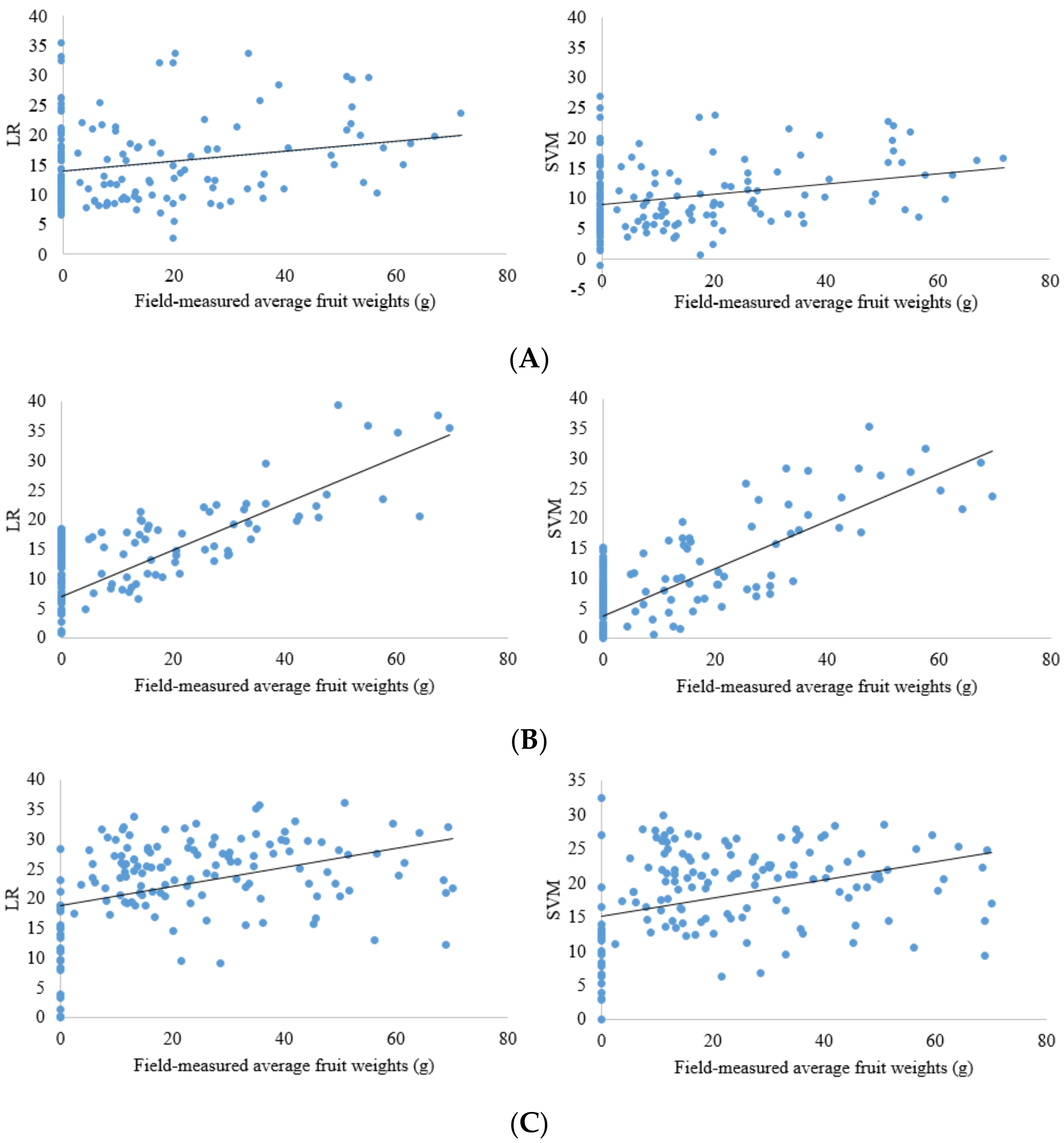
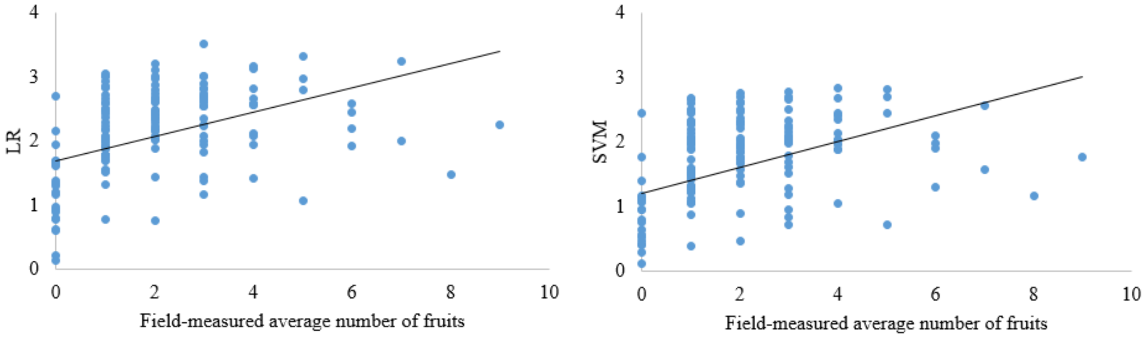


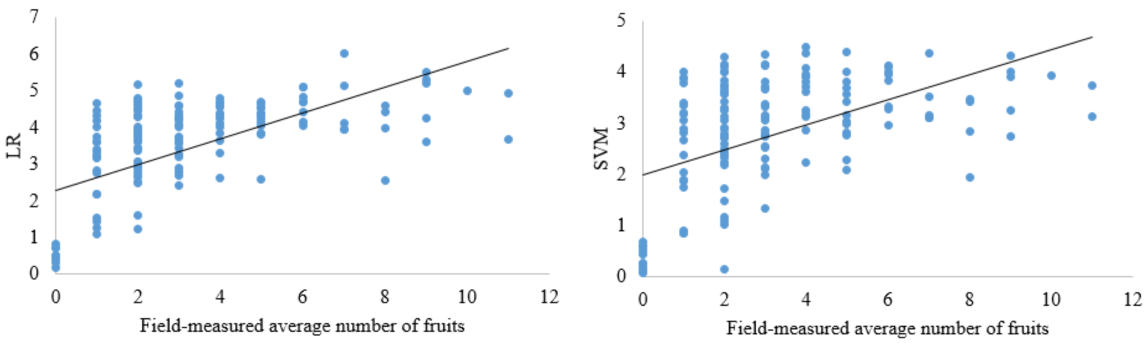
Figure 6. Average fruit weights estimated by the Linear Regression (LR) and Support Vector Machine (SVM) algorithms for flights 9 (**A**), 10 (**B**), and 13 (**C**), as a function of the field-measured values.

The LR algorithm was found to be more effective than the SVM algorithm in determining the number of fruits per plant, generating estimates that were closer to the actual measured values. Despite this, both algorithms performed similarly when estimating this variable from the data sets of flights 1, 2, 3, 8, and 10.

For both algorithms, the most accurate estimates for the number of fruits per plant were generated using the data of flights 4–7, 9, and 11–16. However, only the estimates for flights 13 and 16 are shown herein (Figure 7).



(**A**)



(**B**)

Figure 7. Average number of fruits estimated by the Linear Regression (LR) and Support Vector Machine (SVM) algorithms for flights 13 (**A**) and 16 (**B**), as a function of the field-measured values.

Both the LR and SVM algorithms effectively estimated the amount of leaves per plant, with no significant statistical difference between them. However, for the flight 1 data set, the LR algorithm outperformed the SVM algorithm in estimating this variable (Figure 8).

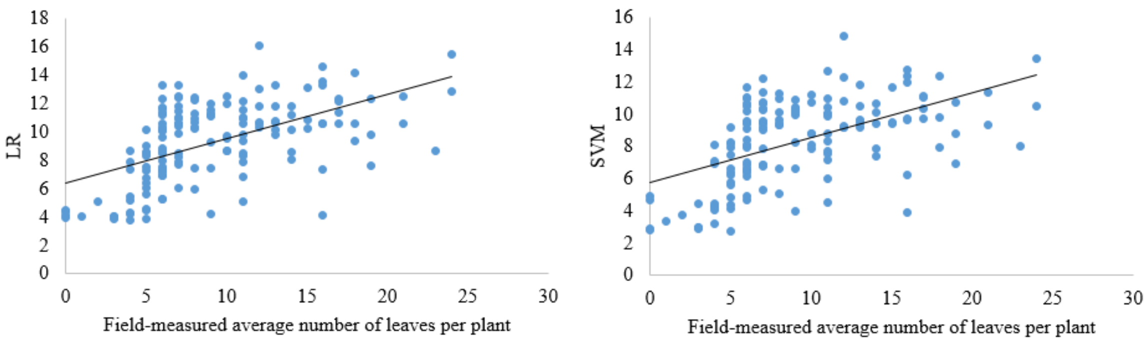


Figure 8. Average number of leaves estimated by the Linear Regression (LR) and Support Vector Machine (SVM) algorithms for flight 1, as a function of the field-measured values.

The comparison of the total estimate with the field values showed that the LR algorithm was more accurate in estimating the fruit weight. The total fruit weight measured in the field was 48.08 kg in an area of 21.6 m^2, which is equivalent to 22,259.26 kg ha^{-1}. The LR algorithm estimated the fruit weight to be 48.04 kg, or approximately 22,240.74 kg ha^{-1}, which was 99.91% of the actual value. By contrast, the SVM algorithm estimated the fruit weight to be 43.09 kg, or approximately 19,949.07 kg ha^{-1}, which was only 89.62% of the actual value.

In total, 4585 fruits were obtained from the field. The LR algorithm generated the best estimate for this variable, with a result of 99.55% of the actual number of fruits. By contrast, the SVM algorithm estimated the number of fruits to be approximately 84.26% of the actual total. Both algorithms were able to estimate the number of leaves with a high degree of accuracy. Compared with the 10,366 leaves measured in the field, the LR algorithm estimated 10,360.31 leaves (99.4% of the actual count), whereas the SVM algorithm estimated 10,171.89 leaves (98.12% of the actual number).

In Figure 9, the NDVI images from the first and last flights are displayed to show the difference in the spectral properties between the two time periods analyzed. When the plants appear green in color, it indicates a low fruit yield at the time of image acquisition. By contrast, a more intense red coloring of the plants indicates higher levels of leaf and fruit area yield. This allows us to differentiate the more productive from the less productive areas as well as the spectral changes in the plants during different stages of the strawberry production cycle. Figure 9A shows the results from the first flight that took place during the early developmental stage of the crop, which was characterized by growth and vegetative development. Figure 9B represents the high fruit yield phase.

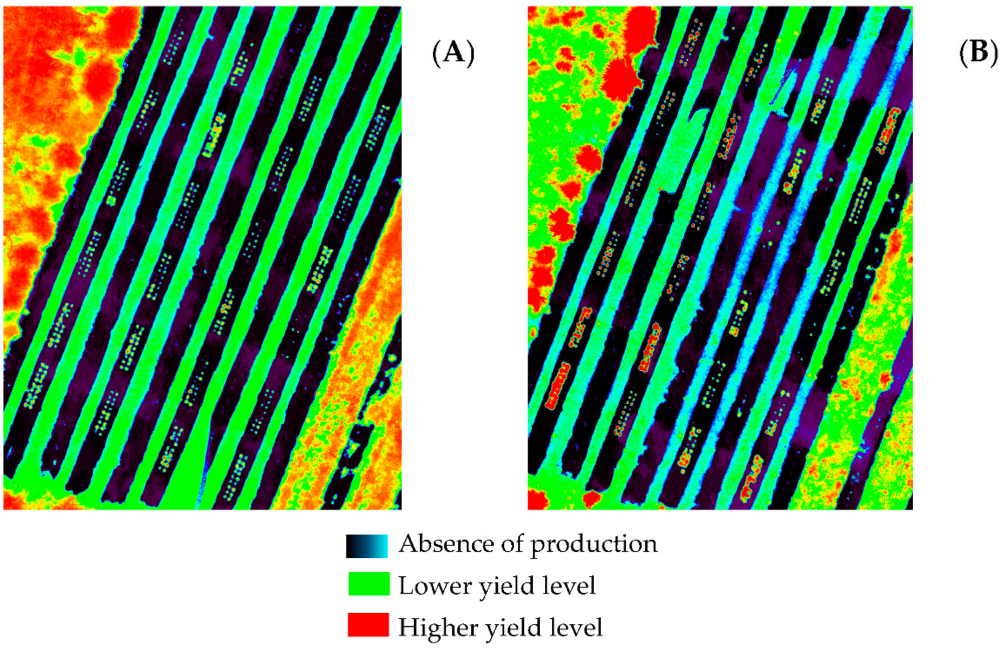


Figure 9. Normalized difference vegetation index image of the experimental area for flights 1 (**A**) and 16 (**B**).

The use of tools that efficiently and accurately estimate strawberry production enables resource planning for the harvesting, transport, and commercialization of the product. It also lowers labor costs, since this estimation process is carried out manually, which is lengthy, laborious, and subjective.

4. Discussion

The RL and SVM algorithms estimated the variables analyzed, exhibiting no differences between the correlation coefficients and RMSE after analysis of all the flights. However, when the flights were analyzed separately, the RL algorithm more accurately estimated average weight and number of fruits, and for the number of leaves variable, there was no significant difference between the algorithms for most of the flights. Similarly, ref. [43] estimated potato biomass and yield using machine learning and remote sensing, obtaining estimated biomass (RMSE = 5.8) and yield (RMSE < 20%), and the better analysis method was RL, which was more accurate in estimating the tubercle formation phase.

The strawberry fruit weight was estimated by [44] using machine learning. They found that the RL prediction model was slightly more accurate than support vector regression (SVR) models in estimating fruit weight, indicating a linear relationship between the number of pixels and fruit weight. The authors report that this method has advantages since it is not destructive, in addition to being economical for the regular monitoring of strawberry weight. According to [44], more strawberry samples from several crops should be studied to improve model performance in estimating fruit weight. The author of [44] also inferred that the shape of strawberry cultivars might affect the number of pixels in the captured images, which are directly related to fruit weight.

In addition, ref. [45] predicted the total soluble solid (TSS) content and pH of strawberries using machine learning, where the results indicated that the SVM models performed better than their LR counterparts, explaining 84.1% of TSS and 78.8% of pH. Similarly, ref. [46] estimated strawberry plant water content, obtaining a correlation coefficient of 0.82 and RMSE of 0.0092 g^{-1}, SVM maturation classification with 99.7% accuracy for ripe fruits and 94.9% for their green counterparts [46]. Analysis of the studies showed that RL algorithms estimated fruit weight and leaf area better, while their SVM counterparts were superior in estimating physiological parameters such as TSS, pH and water content.

Among the parameters assessed, number of leaves displayed the highest correlation with the spectral bands (SVM: 0.70 and RL: 0.82), possibly because the vegetation canopy of plants, mainly in the near infrared region, absorbs a low amount of energy, causing high reflectance. When the number of leaves is correlated with the bands applied by NDVI, this parameter stands out in relation to average weight and number of fruits. According to the results obtained by [4], the VIs of the infrared region are good biomass predictors. This is likely because these VIs may be sensitive to gaps in the plant canopy and changes in senescence and may therefore effectively reflect the differences in canopy density or in leaf weight between strawberry plants [4].

The correlation values for the average weight and number of fruits show a low correlation with the spectral bands in flights 6 to 9, and in this same time interval, a high correlation was observed with the number of leaves. This likely occurred due to the high potential to reflect energy in the spectral bands in question, with greater energy expression for this variable when compared to the spectral response to average mass and number of fruits. This may be because the strawberry plant undergoes several stages simultaneously (presence of leaves, fruits and inflorescences), hampering characterization of the plant's spectral response to certain agronomic parameters.

The RMSE values for the average weight and number of fruits is low, demonstrating that both algorithms can estimate these variables. For the number of leaves, the error between the average field value and that estimated by the SVM algorithm was high, because the presence of an outlier in the database resulted in a high error in the number of leaves (45.83 leaves $plant^{-1}$) in flight number 3 (Figure 5C), thereby raising the RMSE of this variable. For this same variable and flight, the RL algorithm estimated an average error of 2.84 leaves $plant^{-1}$. Given that the SVM algorithm estimated the number of leaves, with low error values for flight 1 to 11, except flight 3, and that the RL algorithm, using the same database as SVM, estimated the number of leaves with a low error rate, it can be inferred that the increase in RMSE caused by flight 3 is due to the presence of an outlier that was not identified in the database.

Strawberry yield was predicted by [3], using a deep neural network and high-resolution aerial orthoimages, obtaining an average accuracy of 91.0% in detecting ripe Sensation fruits, on a 2 m-high image, and 84.1% for fruit count. This demonstrates the feasibility of using high temporal resolution images to extract information on canopy size, number of fruits and flowers to predict strawberry production throughout the crop cycle. The strawberry yield was also predicted by [2], using images, with errors of 26.3 and 25.7% for flower and fruit counts, respectively.

Analysis of the flights separately showed that the RL algorithm is more efficient than SVM for average weight and number of fruits, while both algorithms are efficient for the number of leaves.

Assessing machine learning algorithms in predicting corn yield, ref [47] concluded that linear algorithms (Linear discriminant analysis—LDA and Logistic Regression—LR) predicted the corn yield closest to those observed when compared to their nonlinear counterparts Naive Bayes (NB), K-nearest neighbor (KNN), Classification and regression trees (CART), and Support Vector Machine (SVM). In general, the LDA algorithm was the best tool and SVM the worst in predicting corn yield.

5. Conclusions

This study underscored the ability of remote sensing techniques to accurately predict the yield of strawberries that exhibit high phenological dynamics, that is, vegetative development occurring at the same time as phenological changes, given that phenological changes caused in different amounts of cover and green biomass on the soil surface throughout the crop cycle influence spectral behavior in the images.

Linear regression and support vector machine algorithms can estimate strawberry production and the number of fruits and leaves, using multispectral UAV images.

Linear regression was the most efficient in estimating crop yields, with 99.91% accuracy for the average fruit weight, 99.55% for the number of fruits and 99.94% for the number of leaves.

SVM was 89.62% accurate for the average fruit weight, 84.26% for the number of fruits and 98.12% for the number of leaves.

Studies similar to the present investigation showed that yield prediction models are constructed from images collected in a single crop development phase, but the models presented here can accurately predict total strawberry yield, since images were taken throughout the crop development phases. Thus, the models presented were more effective in predicting when constructed with a historical series of images, that is, those taken between the first and last harvest.

One of the study limitations is that the prediction models were constructed based on a single experimental area. For future investigations, a larger number of cultivated areas under different biotic and abiotic conditions should be used, making it possible to create more robust and less biased models.

It is important to note the following advantages and disadvantages of using this methodology: the use of images to monitor yield may provide an important source of georeferenced monitoring of the strawberry crop, allowing specific management in different crop development phases; the prediction models consist of a limited number of plant samples, mitigating the need for in loco monitoring of all the plants; and the use of wavelengths such as the near infrared, in the composition of prediction models may provide information that is not collected in classic yield control and crop quality techniques.

Author Contributions: Conceptualization, R.C. and G.D.M.; methodology, L.S.d.O., R.C. and G.D.M.; software, L.S.d.O. and G.D.M.; validation, L.S.d.O., R.C., M.H.M. and G.D.M.; formal analysis, L.S.d.O.; investigation, L.S.d.O., R.C. and M.H.M.; resources, R.C.; data curation, L.S.d.O. and M.H.M.; writing—original draft preparation, L.S.d.O.; writing—review and editing, L.S.d.O., R.C. and G.D.M.; visualization, L.S.d.O.; supervision, R.C. and G.D.M.; project administration, L.S.d.O., R.C., M.H.M. and G.D.M.; funding acquisition, R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data are contained within the manuscript and available on request from the corresponding author.

Acknowledgments: We thank the Federal University of Uberlândia (*Universidade Federal de Uberlândia*—UFU), the Institute of Agricultural Sciences (*Instituto de Ciências Agrárias*—ICIAG), the Center for Research in Olericulture (*Núcleo de Pesquisa em Olericultura*—NUPOL), and the Coordination for the Improvement of Higher Education Personnel (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*—CAPES) for granting us a master's scholarship.

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

References

1. Castro, A.; Spalding, L.E.S.; Rieder, R. *Plataforma Embarcada para Monitoramento da Cultura do Morangueiro*; Universidade de Passo Fundo: Passo Fundo, Brazil, 2017.
2. Abd-Elrahman, A.; Wu, F.; Agehara, S.; Britt, K. Improving Strawberry Yield Prediction by Integrating Ground-Based Canopy Images in Modeling Approaches. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 239. [CrossRef]
3. Chen, Y.; Lee, W.S.; Gan, H.; Peres, N.; Fraisse, C.; Zhang, Y.; He, Y. Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages. *Remote Sens.* **2019**, *11*, 1584. [CrossRef]
4. Zheng, C.; Abd-Elrahman, A.; Whitaker, V.; Dalid, C. Prediction of Strawberry Dry Biomass from UAV Multispectral Imagery Using Multiple Machine Learning Methods. *Remote Sens.* **2022**, *14*, 4511. [CrossRef]
5. Guan, Z.; Abd-Elrahman, A.; Fan, Z.; Whitaker, V.M.; Wilkinson, B. Modeling Strawberry Biomass and Leaf Area Using Object-Based Analysis of High-Resolution Images. *ISPRS J. Photogramm. Remote Sens.* **2020**, *163*, 171–186. [CrossRef]
6. Franchini, J.C.; Balbinot, J.; Alvadi, A.; Jorge, L.A.C.; Debiasi, H.; Dias, W.P.; Godoy, C.V.; Oliveira Junior, A.; Corrêa, F.B.; Oliveira, M.C.N. *Uso de Imagens Aéreas Obtidas com Drones em Sistemas de Produção de Soja*, 1st ed.; Embrapa Soja; Londrina–PR, Documentos (INFOTECA-E); 2018; 39p. Available online: https://ainfo.cnptia.embrapa.br/digital/bitstream/item/190247/1/Doc-408-OL-alta.pdf (accessed on 13 March 2022).
7. Elkhrachy, I. Accuracy Assessment of Low-Cost Unmanned Aerial Vehicle (UAV) Photogrammetry. *Alex. Eng. J.* **2021**, *60*, 5579–5590. [CrossRef]
8. Yao, H.; Qin, R.; Chen, X. Unmanned Aerial Vehicle for Remote Sensing Applications—A Review. *Remote Sens.* **2019**, *11*, 1443. [CrossRef]
9. de Oliveira Bertolin, N.; Filgueiras, R.; Venancio, L.P.; Mantovani, E.C. Predição da produtividade de milho irrigado com auxílio de imagens de satélite. *Rbai* **2017**, *11*, 1627–1638. [CrossRef]
10. Liakos, K.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*, 2674. [CrossRef]
11. Muruganantham, P.; Wibowo, S.; Grandhi, S.; Samrat, N.H.; Islam, N. A Systematic Literature Review on Crop Yield Prediction with Deep Learning and Remote Sensing. *Remote Sens.* **2022**, *14*, 1990. [CrossRef]
12. Zheng, C.; Abd-Elrahman, A.; Whitaker, V. Remote Sensing and Machine Learning in Crop Phenotyping and Management, with an Emphasis on Applications in Strawberry Farming. *Remote Sens.* **2021**, *13*, 531. [CrossRef]
13. Madhavi, B.G.K.; Basak, J.K.; Paudel, B.; Kim, N.E.; Choi, G.M.; Kim, H.T. Prediction of Strawberry Leaf Color Using RGB Mean Values Based on Soil Physicochemical Parameters Using Machine Learning Models. *Agronomy* **2022**, *12*, 981. [CrossRef]
14. Li, X.; Cai, C.; Zheng, H.; Zhu, H. Recognizing Strawberry Appearance Quality Using Different Combinations of Deep Feature and Classifiers. *J. Food Process Eng.* **2022**, *45*, 10. [CrossRef]
15. Shin, J.; Chang, Y.K.; Heung, B.; Nguyen-Quang, T.; Price, G.W.; Al-Mallahi, A. A Deep Learning Approach for RGB Image-Based Powdery Mildew Disease Detection on Strawberry Leaves. *Comput. Electron. Agric.* **2021**, *183*, 106042. [CrossRef]
16. Yang, M.-H.; Nam, W.-H.; Kim, T.; Lee, K.; Kim, Y. Machine Learning Application for Predicting the Strawberry Harvesting Time. *Agric. Sci. Res.* **2019**, *46*, 381–393. [CrossRef]
17. Nassar, L.; Okwuchi, I.E.; Saad, M.; Karray, F.; Ponnambalam, K.; Agrawal, P. Prediction of Strawberry Yield and Farm Price Utilizing Deep Learning. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–7.
18. Li, X.; Li, J.; Tang, J. A Deep Learning Method for Recognizing Elevated Mature Strawberries. In Proceedings of the 2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC), Nanjing, China, 18–20 May 2018; pp. 1072–1077.
19. Huang, Z.; Sklar, E.; Parsons, S. Design of Automatic Strawberry Harvest Robot Suitable in Complex Environments. In Proceedings of the Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, Cambridge, UK, 23 March 2020; pp. 567–569.

20. Tedesco, D.; de Oliveira, M.F.; dos Santos, A.F.; Costa Silva, E.H.; de Souza Rolim, G.; da Silva, R.P. Use of Remote Sensing to Characterize the Phenological Development and to Predict Sweet Potato Yield in Two Growing Seasons. *Eur. J. Agron.* **2021**, *129*, 126337. [CrossRef]
21. Mahmud, M.S.; Zaman, Q.U.; Esau, T.J.; Chang, Y.K.; Price, G.W.; Prithiviraj, B. Real-Time Detection of Strawberry Powdery Mildew Disease Using a Mobile Machine Vision System. *Agronomy* **2020**, *10*, 1027. [CrossRef]
22. Sharpe, S.M.; Schumann, A.W.; Boyd, N.S. Goosegrass Detection in Strawberry and Tomato Using a Convolutional Neural Network. *Sci. Rep.* **2020**, *10*, 9548. [CrossRef]
23. Serejo, G.L.; Santos, V.A.; Gomes, A.C.S.; Silva, A.F.B. Estimativa Da Produção de Uma Lavoura de Agricultura Familiar Através de Imagens Digitais Capturadas Por Veículo Aéreo Não Tripulado (VANT). In Proceedings of the Anais do XII Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais (WCAMA 2021); Sociedade Brasileira de Computação: Tucuruí, Brazil, 2021; pp. 79–86.
24. Barreto, A.; Lottes, P.; Ispizua Yamati, F.R.; Baumgarten, S.; Wolf, N.A.; Stachniss, C.; Mahlein, A.-K.; Paulus, S. Automatic UAV-Based Counting of Seedlings in Sugar-Beet Field and Extension to Maize and Strawberry. *Comput. Electron. Agric.* **2021**, *191*, 106493. [CrossRef]
25. Lillo-Saavedra, M.; Espinoza-Salgado, A.; García-Pedrero, A.; Souto, C.; Holzapfel, E.; Gonzalo-Martín, C.; Somos-Valenzuela, M.; Rivera, D. Early Estimation of Tomato Yield by Decision Tree Ensembles. *Agriculture* **2022**, *12*, 1655. [CrossRef]
26. Tu, Y.-H.; Johansen, K.; Phinn, S.; Robson, A. Measuring Canopy Structure and Condition Using Multi-Spectral UAS Imagery in a Horticultural Environment. *Remote Sens.* **2019**, *11*, 269. [CrossRef]
27. Greifeneder, F.; Notarnicola, C.; Wagner, W. A Machine Learning-Based Approach for Surface Soil Moisture Estimations with Google Earth Engine. *Remote Sens.* **2021**, *13*, 2099. [CrossRef]
28. Shahzaman, M.; Zhu, W.; Bilal, M.; Habtemicheal, B.A.; Mustafa, F.; Arshad, M.; Ullah, I.; Ishfaq, S.; Iqbal, R. Remote Sensing Indices for Spatial Monitoring of Agricultural Drought in South Asian Countries. *Remote Sens.* **2021**, *13*, 2059. [CrossRef]
29. Hara, P.; Piekutowska, M.; Niedbała, G. Selection of Independent Variables for Crop Yield Prediction Using Artificial Neural Network Models with Remote Sensing Data. *Land* **2021**, *10*, 609. [CrossRef]
30. Kashyap, B.; Kumar, R. Sensing Methodologies in Agriculture for Monitoring Biotic Stress in Plants Due to Pathogens and Pests. *Inventions* **2021**, *6*, 29. [CrossRef]
31. Fiorentini, M.; Zenobi, S.; Orsini, R. Remote and Proximal Sensing Applications for Durum Wheat Nutritional Status Detection in Mediterranean Area. *Agriculture* **2021**, *11*, 39. [CrossRef]
32. Denis, A.; Desclee, B.; Migdall, S.; Hansen, H.; Bach, H.; Ott, P.; Kouadio, A.L.; Tychon, B. Multispectral Remote Sensing as a Tool to Support Organic Crop Certification: Assessment of the Discrimination Level between Organic and Conventional Maize. *Remote Sens.* **2020**, *13*, 117. [CrossRef]
33. Gao, D.; Sun, Q.; Hu, B.; Zhang, S. A Framework for Agricultural Pest and Disease Monitoring Based on Internet-of-Things and Unmanned Aerial Vehicles. *Sensors* **2020**, *20*, 1487. [CrossRef]
34. Acorsi, M.G.; das Dores Abati Miranda, F.; Martello, M.; Smaniotto, D.A.; Sartor, L.R. Estimating Biomass of Black Oat Using UAV-Based RGB Imaging. *Agronomy* **2019**, *9*, 344. [CrossRef]
35. Gastli, M.S.; Nassar, L.; Karray, F. Deep Learning Models for Strawberry Yield and Price Forecasting Using Satellite Images. In Proceedings of the 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Melbourne, Australia, 17 October 2021; pp. 1790–1796.
36. Heylen, R.; van Mulders, P.; Gallace, N. Counting Strawberry Flowers on Drone Imagery with a Sequential Convolutional Neural Network. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11 July 2021; pp. 4880–4883.
37. Khaki, S.; Pham, H.; Wang, L. Simultaneous Corn and Soybean Yield Prediction from Remote Sensing Data Using Deep Transfer Learning. *Sci. Rep.* **2021**, *11*, 11132. [CrossRef]
38. Sun, J.; Lai, Z.; Di, L.; Sun, Z.; Tao, J.; Shen, Y. Multilevel Deep Learning Network for County-Level Corn Yield Estimation in the U.S. Corn Belt. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 5048–5060. [CrossRef]
39. Stepanov, A.; Dubrovin, K.; Sorokin, A.; Aseeva, T. Predicting Soybean Yield at the Regional Scale Using Remote Sensing and Climatic Data. *Remote Sens.* **2020**, *12*, 1936. [CrossRef]
40. Xu, W.; Chen, P.; Zhan, Y.; Chen, S.; Zhang, L.; Lan, Y. Cotton Yield Estimation Model Based on Machine Learning Using Time Series UAV Remote Sensing Data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102511. [CrossRef]
41. Ribeiro, A.C.; Guimarães, P.T.G.; Alvarez, V.H. *Recomendações para o uso de Corretivos e Fertilizantes em Minas Gerais: 5ª Aproximação*; Comissão de Fertilidade do Solo do Estado de Minas Gerais: Viçosa, Brazil, 1999; 359p.
42. R Core Team. *R: A language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021. Available online: https://www.R-project.org/ (accessed on 15 March 2022).
43. Li, C.; Ma, C.; Pei, H.; Feng, H.; Shi, J.; Wang, Y.; Chen, W.; Li, Y.; Feng, X.; Shi, Y. Estimation of Potato Biomass and Yield Based on Machine Learning from Hyperspectral Remote Sensing Data. *J. Agric. Sci. Technol. B* **2020**, *10*, 195–213. [CrossRef]
44. Basak, J.K.; Paudel, B.; Kim, N.E.; Deb, N.C.; Kaushalya Madhavi, B.G.; Kim, H.T. Non-Destructive Estimation of Fruit Weight of Strawberry Using Machine Learning Models. *Agronomy* **2022**, *12*, 2487. [CrossRef]

45. Basak, J.K.; Madhavi, B.G.K.; Paudel, B.; Kim, N.E.; Kim, H.T. Prediction of Total Soluble Solids and PH of Strawberry Fruits Using RGB, HSV and HSL Colour Spaces and Machine Learning Models. *Foods* **2022**, *11*, 2086. [CrossRef] [PubMed]
46. Raj, R.; Cosgun, A.; Kulić, D. Strawberry Water Content Estimation and Ripeness Classification Using Hyperspectral Sensing. *Agronomy* **2022**, *12*, 425. [CrossRef]
47. Mupangwa, W.; Chipindu, L.; Nyagumbo, I.; Mkuhlani, S.; Sisito, G. Evaluating Machine Learning Algorithms for Predicting Maize Yield under Conservation Agriculture in Eastern and Southern Africa. *SN Appl. Sci.* **2020**, *2*, 952. [CrossRef]