# Estimation of biometric, physiological, and nutritional variables in lettuce seedlings using multispectral images[1]

## Estimativa de variáveis biométricas, fisiológicos e nutricionais em mudas de alface usando câmara multiespectral

George D. Martins[2], Onésio F. da Silva Neto[3], Glecia J. dos S. Carmo[4],
Renata Castoldi[4]*, Ludymilla C. S. Santos[3] & Hamilton C. de O. Charlo[3]

[1] Research developed at Uberaba, MG, Brazil
[2] Universidade Federal de Uberlândia/ Departamento de Engenharia Civil, Monte Carmelo, MG, Brazil
[3] Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro/Departamento de Agronomia, Uberaba, MG, Brazil
[4] Universidade Federal de Uberlândia/ Instituto de Ciências Agrárias, Monte Carmelo, MG, Brazil

*HIGHLIGHTS:*

*Multispectral imaging, using a low-cost camera, makes it possible to assess the agronomic characteristics of lettuce.*
*Predictive models can be obtained using parametric and non-parametric algorithms based on machine-learning approaches.*
*Specific wavelengths of the reflective spectrum of lettuce are important plant quality indicators.*

**ABSTRACT:** The formation of seedlings is one of the most important phases of lettuce cultivation. Therefore, any strategy that aims to obtain high-quality seedlings can increase productivity. One of these strategies is the prediction of morphophysiological attributes based on optical properties. The objective of this study was to quantitatively estimate the biometric variables of lettuce from parametric and non-parametric models based on the response of multispectral camera images. The experiment was conducted in a greenhouse in the municipality of Uberaba, Minas Gerais State, Brazil. Twenty days after sowing, multispectral images of the plants were captured using a MAPIR Survey 3 camera. To compose the estimation models, along with the original bands of the camera, the multispectral vegetation indices were calculated using the calibrated original camera bands. Bands $B_{550}$, $B_{660}$, and $B_{850}$ and the near-infrared indices contributed significantly to estimating the physiological variable models, with $B_{850}$ contributing the most to the biometric and nutritional variables. From the near-infrared band ($B_{850}$) and derived indices, it was possible to estimate all the agronomic variables from the models generated by the M5 algorithm, with an accuracy of up to 1.6% for the maximum quantum yield. Thus, it is possible to quantify the biometric, physiological, and nutritional variables of lettuce using a multispectral camera. Among the Mapir camera bands, $B_{660}$ exhibited the greatest variability, showing that the red range was the most sensitive.

Key words: *Lactuca sativa*, morphophysiological variables, prediction models

**RESUMO:** A formação de mudas é umas das fases mais importantes da cultura da alface. Desta forma, qualquer estratégia que vise obtenção de mudas de alta qualidade pode representar aumento de produtividade. Uma dessas formas tem sido a estimativa de atributos morfofisiológicos baseados nas propriedades ópticas. Com este estudo, o objetivo foi estimar quantitativamente variáveis biométricas da alface a partir de modelos paramétricos e não paramétricos baseados na resposta de imagens multiespectrais tomadas por câmara multiespectral. O experimento foi conduzido em casa de vegetação, no município de Uberaba, MG, Brasil. Aos 20 dias após a semeadura foram capturadas imagens multiespectrais das plantas, utilizando-se para isto a câmara Survey 3 da MAPIR. Para compor os modelos de estimativa, juntamente com as bandas originais da câmara, foram calculados os índices de vegetação multiespectrais, a partir das bandas originais calibradas da câmara. As bandas $B_{550}$, $B_{660}$, and $B_{850}$ e os índices derivados do infravermelho próximo contribuiram ao máximo para estimar os modelos das variáveis fisiológicas. Já a banda $B_{850}$ foi a que mais contribuiu para as variáveis biométricas e nutricionais. A partir da banda do infravermelho próximo ($B_{850}$) e índices derivados, foi possível estimar quase todas as variaveis agronômicas a partir de modelos gerados pelo algoritmo M5, com precisão de até 1,6% para rendimento quântico máximo. Conclui-se que é possível quantificar as variaveis biométricas, fisiológicas e nutricionais da alface por meio de câmara multiespectral. Dentre as bandas de câmeras Mapir, a $B_{660}$ exibiu a maior variabilidade, mostrando que a faixa do vermelho foi a mais sensível.

Palavras-chave: *Lactuca sativa*, variáveis morfofisiológicas, modelos de predição

## INTRODUCTION

The development of seedlings is one of the most critical phases of the crop cycle because it directly influences final plant performance from both a nutritional and productive perspective (Gusatt et al., 2019; Desai et al., 2020).

The purchase price of lettuce seeds represents 8% of the input cost and 4.15% of the total production cost of this vegetable (Oliva et al., 2016). Thus, any strategy aimed at obtaining high-quality seedlings may represent an increase in productivity and, consequently, higher profitability for the producer.

However, one of the limitations is the lack of rapid and efficient technologies to select high-vigor seedlings that result in more productive plants.

Remote sensing has several agricultural applications (Weiss et al., 2020) and uses leaf or canopy reflectance to calculate vegetation indices to assess agronomic variables such as nutritional status, biomass, leaf area, and drought resistance (Gogoi et al., 2018).

The use of optical sensors has been found to be efficient for the assessment of several plant species (Picoli et al., 2013; Makanza et al., 2018; Maciel et al., 2019); however, there are few reports on *Lactuca sativa* seedlings. These sensors can cover an extensive area in a very short time, providing rapid assessment of all the seedlings and minimizing the effect of rapidly changing environmental conditions, such as wind speed, cloud cover, and solar radiation.

Thus, the present study aimed to quantitatively estimate the biometric, physiological, and nutritional variables of lettuce, using parametric and non-parametric models (machine learning) based on images taken using a multispectral camera.

## MATERIAL AND METHODS

The study was conducted in the municipality of Uberaba, MG, Brazil, which is in the mesoregion of Triângulo Mineiro and Alto Paranaíba, during December 8 and 28, 2019, with a cultivar adapted to tropical conditions in late spring and early summer, when there is an increasing market demand for lettuce and seedling production.

The arch-type greenhouse where the study was conducted was conducted in double-span plastic and measured 14 m wide and 51 m long, with eaves and arc heights of 3.00 and 1.5 m, respectively, in the east-west direction and lies at coordinates 19° 39' 44" S and 47° 58' 02" W, at an altitude of 790 m. It was covered with a 150-μm light-diffusing plastic film with closed sides and 50% shade cloth (Figure 1).

Temperature and relative air humidity data were collected daily in the greenhouse using a thermohygrometer (Testo 174 H data logger) installed 1.20 m above ground level to assess the environmental conditions just above the plants.

Data variability was achieved by producing lettuce seedlings with variations in the fertigation start time and application intervals. Two experiments were performed. For both experiments, a randomized block design was used with six repetitions. Each experimental unit consisted of 140 plants.

The treatments in the first experiment consisted of six fertigation starting times ($T_1 = 0$, $T_2 = 3$, $T_3 = 6$, $T_4 = 9$, $T_5 = 12$,



**Figure 1.** Greenhouse view where the experiments were conducted

and $T_6 = 15$ days after emergence). After the first application of each treatment, fertigation was repeated at five-day intervals. Thus, depending on the starting time of the first application, the number of applications differed between treatments, with four applications for $T_1$, three for $T_2$ and $T_3$, two for $T_4$ and $T_5$, and one for $T_6$.

The treatments in the second experiment consisted of five fertigation application intervals ($T_1 = 3$, $T_2 = 4$, $T_3 = 5$, $T_4 = 6$, and $T_5 = 7$ days). In this experiment, the first application occurred three days after plant emergence, with five applications in $T_1$, four in $T_2$, three in $T_3$ and $T_4$, and three in $T_5$.

Fertigation depth varied as a function of seedling development stage. More advanced stages required increased fertigation depths. Nutrient solution was applied as follows: 0.4 L from emergence to 3 days after emergence (DAE); 0.5 L from 4 to 7 DAE; 0.6 L from 8 to 11 DAE; 0.7 L from 12 to 14 DAE and 0.8 L from 15 to 18 DAE.

The lettuce cultivar used to produce seedlings was Vanda (from Sakata®). The seeds were sown in polyethylene trays, with 12.5 cm³ cells filled with Bioplant Plus® substrate. The experiment was conducted on wire benches, 90 cm above the ground.

The seedlings were irrigated four times a day (8:00 a.m., 12:00 p.m., 2:00 p.m., and 4:30 p.m.) using an automated sprinkler irrigation system without drainage to maintain the substrate in the containers. From sowing to emergence (3 days after sowing), all treatments were irrigated with water only. Emergence was considered to be established when at least 90% of the tray cells contained emerged seedlings. Fertigation was then initiated within the time frames proposed in the experimental design and performed only once a day (8:00 a.m.).

The sprinkler irrigation system installed inside the greenhouse provides an average depth of 3.98 mm h⁻¹, which corresponds to 3.98 L of water for each m² of greenhouse in 1 h. Water depth was measured every 3 days by weighing the trays. The trays were irrigated until the water began to drain, that is, until the container capacity was reached, when they were weighed. They were weighed again after 2.5 h and the procedure was repeated four times throughout the day. The difference between weight measurements was attributed to crop evapotranspiration (ETc), which was subsequently converted

into the amount of water to be applied daily via irrigation or fertigation.

The irrigation depth varied as a function of the seedling development stage. More advanced stages required greater irrigation depth. The depth applied up to 3 days after emergence was 4.98 mm day$^{-1}$, from 4 to 7 DAE, 5.24 mm day$^{-1}$, from 8 to 11 DAE, 5.44 mm day$^{-1}$, from 12 to 14 DAE, 5.70 mm day$^{-1}$ and from 15 to 18 DAE, 5.97 mm day$^{-1}$.

Each day, all treatments received the same amount of water. However, when fertigation was applied, the water was replaced with the same volume of nutrient solution for the first irrigation of the day (8:00 a.m.). Fertigation was applied using a backpack sprayer with a capacity of 20 L and flow rate of 0.8 L h$^{-1}$. After this application, a small volume of water was sprayed on the seedlings to prevent the nutrient solution from accumulating on the leaves. The other irrigations (12:00, 2:00, and 4:30 p.m.) were performed with water only for all treatments.

At 15 DAE, the number of leaves (NF15) was evaluated by counting the fully developed leaves of 10 plants from each plot.

At 17 DAE, the physiological variables of the seedlings were evaluated through the OJIP transient fluorescence of chlorophyll a, including the initial fluorescence ($F_0$), variable fluorescence ($F_v$), maximum fluorescence ($F_m$), maximum quantum yield ($F_v/F_m$), number of photons absorbed by the antenna complex (ABS/RC), amount of energy flowing through the antenna complex and captured by the PSII reaction center (TRo/RC), forward electron flow from the reaction center (ETo/RC), and amount of energy dissipated by non-photochemical dissipation (DIo/RC). These variables were measured using PSI's (Photon Systems Instruments) Fluorometer FP100 model, performing readings on the second true leaf of five plants in each plot between 12:00 and 3:00 a.m. so that the plants were adapted to the dark. The chlorophyll a index was evaluated on the same day using Falker's CFL1030 Clorofilog model, taking five readings per plot on the second true leaf from 11:00 a.m. to 3:00 p.m.

At 21 days after sowing, or 18 days after the beginning of the treatments, when the seedlings reached the commercial point of transplanting, the following variables were assessed in 40 plants from each experimental plot: leaf area (Area), expressed in cm$^2$ plant$^{-1}$; shoot dry mass (SDM), expressed in g plant$^{-1}$; and root dry mass (RDM), expressed in g plant$^{-1}$.

Shoot N, P, K, Ca, Mg, and S (g kg$^{-1}$) concentrations were also determined in 40 seedlings from each experimental plot.

After obtaining the data, dispersion analysis was performed, and the variables were subjected to descriptive statistical analysis to obtain the mean, standard deviation, and coefficient of variation.

At 20 days after sowing, multispectral images of the plants were captured using a MAPIR Survey 3 camera, which has a 12-bit resolution, 19 mm focal length, 2.3 cm ground sample distance (GSD), and green (550 nm), red (660 nm), and near-infrared (850 nm) bands.

Images of six plants from each experimental plot were captured for each experiment, totaling 66 images. The camera was installed on a horizontally and vertically leveled support (1.20 m) from the ground in the nadir position. The plants were placed under a black nonwoven fabric to mitigate the effects of reflectance from neighboring targets.

To ensure maximum absorption and reflection conditions of solar electromagnetic radiation, the images were taken from 11:30 a.m. to 12:30 p.m. without shading, which could be caused by clouds or anthropic features near the frame of the capture area.

Radiometric calibration of the images was performed using the Mapir Camera Control software. This process was possible because the images were calibrated with the reflected radiance of the calibration plate, provided by the Survey 3 camera manufacturer, at the same time as the images were taken.

After calibration, radiometric normalization of all the images was performed to compensate for the lighting effects from the first to the last shot, where the image taken at 12:00 p.m. was set as the reference. Normalization was performed using ENVI 5.1 software, according to the methodology proposed by Jensen (2009).

During the normalization process, in the reference image, as well as the images to be normalized, the radiance values were manually extracted from the set of light and dark pixels in all the bands of the camera. In all cases, the pixels were extracted from the calibration plate.

The following equation was used to determine the coefficients of linear transformation:

$$Ti = mi\ xi + bi \qquad (1)$$

where:

| | |
|---|---|
| mi | - (Bri - Dri)/(Bsi - Dsi); |
| bi | - (Dri × Bsi - Dsi × Bri)/(Bsi - Dsi); |
| Ti | - radiance of the reference image |
| Xi | - radiance of the image to be normalized. |
| Bri | - mean of the light reference set; |
| Dri | - mean of the dark reference set; |
| Bsi | - mean of the light set to be normalized; |
| Dsi | - mean of the dark set to be normalized; and, |
| i | - bands of the sensor under study. |

The multispectral vegetation indices were calculated using Eqs. 2, 3, and 4 to compose the estimation models, along with the original camera bands, from the original calibrated Survey 3 camera bands (Table 1), using ENVI 5.1 software.

The individual mean brightness values of the original bands and the derived multispectral indices were extracted for all plants. The average brightness values were extracted from the irregular polygons surrounding the plant crowns using the regions of interest tool created in the ENVI 5.1 software. The mean zonal method was used for data extraction.

**Table 1.** Equations and references for calculations of vegetation indices derived from the original bands of the Survey 3 Camera

| Index | Equation | Eq. | Reference |
|---|---|---|---|
| Simple ratio | SR = (B$_{850}$)/(B$_{660}$) | (2) | Birth & McVey (1968) |
| Normalized difference vegetation index | NDVI = (B$_{850}$ − B$_{660}$)/(B$_{850}$ + B$_{660}$) | (3) | Rouse et al. (1973) |
| Green normalized difference vegetation index | GNDVI = (B$_{850}$ − B$_{550}$)/(B$_{850}$ + B$_{550}$) | (4) | Gitelson et al. (1996) |

The estimation process used in this study is based on parametric and non-parametric regression methods. The non-parametric multilayer perceptron (neural network) and parametric multiple linear regression (MLR) algorithms were used as estimators, and the results were compared.

Neural network implementation was used to estimate and validate the lettuce variables (biometric, physiological, and nutritional), in Weka software (Weka 3.9.4, University of Waikato).

The neural network variables were set to their default values (two neurons and one layer). It is important to emphasize that several tests were conducted to optimize the network variables, but the standard architecture of the software always exhibited more accurate estimates.

This study used the MLR method, which attempts to estimate the variables of the model and describe the relationship between two or more independent variables and a response variable by fitting a linear equation to the observed data, often using the least squares method in Weka (Weka 3.9.4, University of Waikato). In Weka, the selection of the features model was completed using a backward elimination method called "M5," wherein the attribute with the smallest standardized coefficient is removed until no improvements are observed in the Akaike information criterion (Akaike, 1974).

All estimation models were created from the combination of the original bands and the derived multispectral indices. For model training, the radiometric values of 53 randomly defined plants (80% of the sample set) were considered. The root mean squared error (RMSE) (Eq. 5), and the normalized RMSE (RMSE%) (Eq. 6) were calculated to validate the accuracy of the models, considering the residue of the difference between the estimated and measured agronomic variables for 13 plants (20% of the sample set).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(x_o - x_e\right)^2}{n}} \qquad (5)$$

$$RMSE\left(\%\right) = \sqrt{\frac{\sum_{i=1}^{n}\left(x_o - x_e\right)^2}{n}} \frac{100n}{\sum_{i=1}^{n} x_e} \qquad (6)$$

where:

$x_o$ - observed value;

$x_e$ - estimated value; and,

n - number of samples.

## RESULTS AND DISCUSSION

The average temperature and relative air humidity during the experiment were 27.51 °C and 67.16%, respectively. Maximum and minimum temperatures of 46.10 and 18.10 °C were recorded on December 22 and 24, 2019, respectively, and maximum (94.6%) and minimum relative air humidity (27.9%) on December 19 and 24, 2019, respectively.

The standard deviation varied from 0.006 (RDM) to 4,184.01 (Fm). This shows that the higher the value, the greater is the data dispersion (Table 2).

The coefficients of variation (CV) ranged from 0.81% ($B_{550}$) to 62.19% (ETo/RC), which represents low data dispersion for all the evaluated variables (Cantelli et al., 2016), except for ETo/RC and Area (Table 2).

For the data extracted from the image, the highest CV was for $B_{660}$ (5.35%), indicating that the red spectrum is the best region for discriminating between the different types of treatment. For this experiment, the greater spectral variability in this range is associated with the high variability of the biometric and physiological variables of lettuce, given that reflected energy in the 630-680 nm spectral range is influenced by photosynthesis-related variables, such as NF15 and chlorophyll a (Jensen, 2009).

The low CV values for the $B_{550}$ (0.81%) and $B_{850}$ (2.26%) bands show spectral regions with less potential for discriminating between the different treatments applied to the lettuce samples. The low variability of the $B_{550}$ band is associated with the condition that all the samples exhibited the typical appearance of healthy green vegetation, that is, the same phycocyanin concentrations.

The variability of the $B_{850}$ band, which was slightly higher than that of $B_{550}$, was directly influenced by the low variability of the nutritional variables of lettuce (Jensen, 2009), which had a maximum CV for N (17.47%). In this case, in addition to the influence of nutritional variables, a low CV was associated with

**Table 2.** Means, standard deviations, and coefficients of variation (CV) of biometric, physiological, and nutritional variables evaluated in lettuce seedlings

| Variables | Means | Standard deviations | CV (%) |
|---|---|---|---|
| $B_{850}$ | 252.89 | 6.59 | 2.26 |
| $B_{550}$ | 244.06 | 1.96 | 0.81 |
| $B_{660}$ | 225.10 | 12.05 | 5.35 |
| Biometric | | | |
| NF15 | 3.57 | 0.41 | 11.45 |
| SDM | 0.08 | 0.02 | 25.30 |
| RDM | 0.02 | 0.006 | 27.06 |
| Area | 32.35 | 10.41 | 32.17 |
| Physiological | | | |
| $F_0$ | 7,281.91 | 696.98 | 9.57 |
| $F_m$ | 39,588.33 | 4,184.01 | 10.57 |
| $F_v$ | 32,306.42 | 3,603.00 | 11.15 |
| $F_v/F_m$ | 0.82 | 0.01 | 1.25 |
| ABS/RC | 2,309.99 | 148.94 | 6.45 |
| TRo/RC | 1,880.77 | 110.03 | 5.85 |
| ETo/RC | 446.30 | 277.55 | 62.19 |
| Dlo/RC | 0.43 | 0.05 | 10.62 |
| Chlorophyll a | 17.36 | 1.55 | 8.95 |
| Nutritional | | | |
| N | 13.39 | 2.34 | 17.47 |
| P | 6.49 | 0.80 | 12.37 |
| K | 31.54 | 3.39 | 10.75 |
| Ca | 9.23 | 1.05 | 11.37 |
| Mg | 2.35 | 0.16 | 7.01 |
| S | 1.20 | 0.18 | 15.07 |

NF15 - Number of leaves at 15 DAE; SDM - Shoot dry mass (g plant$^{-1}$); RDM - Root dry mass (g·plant$^{-1}$); Chlorophyll a - Chlorophyll a index; Area - Leaf area (cm$^2$ plant$^{-1}$); $F_0$ - Initial fluorescence; $F_m$ - Maximum fluorescence; $F_v$ - Variable fluorescence; $F_v/F_m$ - Maximum quantum yield; ABS/RC - Absorption flow/center of reaction; TRo/RC - Aaptured energy flow/center of reaction; ETo/RC - Electron transport flow/center of reaction; Dio/RC - Non-photochemical energy flow/center of reaction; N, P, K, Ca, Mg, and S - Leaf concentrations (g kg$^{-1}$); CV - Coefficient of variation, n = 66

the calibration process and radiometric normalization, which contributed significantly to the lower range and adherence of the values as a function of the mean.

The multiple linear models presented higher coefficients of determination ($R^2$) because they incorporated more index bands in the regression equations, as verified for the variables $F_v/F_m$ (43%), DIo/RC (34%), and P (58%) (Table 3). However, it is evident that the greater number of predictor variables did not guarantee that these models were the most accurate (Table 4).

As in Oliveira et al. (2020), the models that estimate agricultural variables, when composed of a single predictive variable, tend to have a low coefficient of determination, as shown in Table 3. However, this does not necessarily result in an inaccurate estimate of the variable because if the validation data are normally distributed, the errors in data variability around the mean will be null and decrease the RMSE.

The $B_{550}$, $B_{660}$, and $B_{850}$ bands and the indices derived from the near-infrared region contribute mostly to the estimation models of the physiological variables (Table 3). Band $B_{850}$ contributed the most to the biometric and nutritional variables (Table 3). The variables NF15, SDM, RDM, and Area represent the vegetation biomass, which may explain these results because multispectral indices such as SR and GNDVI are highly correlated with the biomass of canopies of several crops (Jensen, 2009).

For the physiological variables, bands $B_{550}$ and $B_{850}$ were present in most prediction models, except for the variables chlorophyll a and ETo/RC (Table 3). For chlorophyll a, the $B_{660}$ red spectrum band was the predictor variable because this specific portion of the reflective spectrum is causally related to chlorophyll a and b absorption (Jensen, 2009).

For the nutritional variables, the predominance of the $B_{850}$ band in the prediction models reflects the direct relationship between the near-infrared and lettuce nutrients, as observed by Mao et al. (2015).

**Table 4.** Performance of algorithms in estimating biometric, physiological, and nutritional variables

| Variables | Neural network | | M5 | |
|---|---|---|---|---|
| | RMSE | RMSE% | RMSE | RMSE% |
| Biometric | | | | |
| NF15 | 0.421 | 11.79 | 0.382 | 10.7 |
| SDM | 0.022 | 27 | 0.025 | 32 |
| RDM | 0.005 | 15.6 | 0.007 | 22 |
| Area | 11.94 | 43 | 12.2 | 44 |
| Physiological | | | | |
| $F_0$ | 742 | 10.2 | 826 | 11.38 |
| $F_m$ | 4700 | 12.5 | 5253 | 14.26 |
| $F_v$ | 4087 | 13.37 | 4538 | 14.85 |
| $F_v/F_m$ | 0.013 | 1.7 | 0.012 | 1.6 |
| ABS/RC | 154 | 6.74 | 148.32 | 6.5 |
| TRo/RC | 112.38 | 6.06 | 107 | 5.8 |
| ETo/RC | 300 | 77 | 265 | 63 |
| DIo/RC | 0.050 | 12 | 0.052 | 12.1 |
| Chlorophyll a | 1.305 | 5.9 | 1.31 | 7.5 |
| Nutritional | | | | |
| N | 3.4 | 23.5 | 3.19 | 21.9 |
| P | 1.11 | 16.2 | 1.07 | 15.6 |
| K | 3.11 | 9.2 | 3.32 | 10.7 |
| Ca | 1.09 | 11.2 | 1.34 | 13.7 |
| Mg | 0.21 | 8.6 | 0.18 | 7.5 |
| S | 0.15 | 12.4 | 0.14 | 11.5 |

M5 - Backward elimination method; NF15 - Number of leaves at 15 DAE; SDM - Shoot dry mass (g plant⁻¹); RDM - Rroot dry mass (g·plant⁻¹); Chlorophyll a - Chlorophyll a index; Area - Leaf area (cm² plant⁻¹); $F_0$ - Initial fluorescence; $F_m$ - Maximum fluorescence; $F_v$ - Variable fluorescence; $F_v/F_m$ - Maximum quantum yield; ABS/RC - Absorption flow/center of reaction; TRo/RC - Captured energy flow/center of reaction; ETo/RC - Electron transport flow/center of reaction; Dio/RC -Non-photochemical energy flow/center of reaction; N, P, K, Ca, Mg, and S - Leaf concentrations (g kg⁻¹)

Table 4 shows the performance of the algorithms (RMSE and RMSE%) for estimating the biometric, physiological, and nutritional variables.

Regarding the biometric variables, the M5 model had higher precision in the NF15 estimate (10.7%), while the neural networks provided more precise estimates of SDM (27%) and RDM (15.6%) (Table 4).

**Table 3.** Linear models and coefficient of determination ($R^2$) for the estimation of biometric, physiological, and nutritional variables

| Variables | Model | $R^2$ (%) |
|---|---|---|
| Biometric | | |
| Number of leaves at 15 DAE | $NF15 = -3.44 \times GNDVI + 3.7$ | 22 |
| Shoot dry mass (g plant⁻¹) | $SDM = -0.01 \times B_{850} + 2.7$ | 18 |
| Root dry mass (g plant⁻¹) | $RDM = 0.00002 \times B_{850} + 0.2$ | 20 |
| Leaf area (cm² plant⁻¹) | $Area = -5.7 \times B_{850} + 1341.6$ | 2 |
| Physiological | | |
| Initial fluorescence | $F_0 = -99 + 7 \times B_{850} + 32647.2$ | 19 |
| Maximum fluorescence | $F_m = 328.48 \times B_{550} - 39469.7$ | 26 |
| Variable fluorescence | $F_v = 315.06 \times B_{550} - 43630.2$ | 15 |
| Maximum quantum yield | $F_v/F_m = 0.0095 \times B_{850} - 0.16 \times GNDVI - 1.58$ | 43 |
| Absorption flow/center of reaction | $ABS/RC = -47.51 \times B_{850} + 14330.4$ | 7.8 |
| Captured energy flow/center of reaction | $TRo/RC = 33.67 \times B_{550} - 7735.06$ | 12 |
| Electron transport flow/center of reaction | $ETo/RC = 33.67 \times B_{660} - 7735.06$ | 15 |
| Non-photochemical energy flow/center of reaction | $Dio/RC = -0.0336 \times B_{850} + 0.5251 \times GNDVI + 8.909$ | 34 |
| Chlorophyll a index | $Chlorophyll\ a = -0.054 \times B_{660} + 29.4$ | 15 |
| Nutritional | | |
| Nitrogen leaf concentrations (g kg⁻¹) | $N = 1.33 \times B_{850} - 323.42$ | 11 |
| Phosphorus leaf concentrations (g kg⁻¹) | $P = 0.46 \times B_{850} - 41.1 \times GNDVI - 21.31 \times SR - 89.07$ | 58 |
| Potassium leaf concentrations (g kg⁻¹) | $K = 1.18 \times B_{850} - 269.55$ | 14 |
| Calcium leaf concentrations (g kg⁻¹) | $Ca = 0.18 \times B_{850} - 37.37$ | 10 |
| Magnesium leaf concentrations (g kg⁻¹) | $Mg = 0.06 \times B_{850} - 12.28$ | 16 |
| Sulfur leaf concentrations (g kg⁻¹) | $S = 0.06 \times B_{850} - 14.62$ | 12.35 |

SR - Simple ratio; GNDVI - Green normalized difference vegetation index

The estimated leaf area was not satisfactory because the errors were 43% and 44% for the neural networks and M5 models, respectively (Table 4). In contrast, when monitoring lettuce health, Ren et al. (2017) verified that the estimation of leaf area by remotely piloted aircraft cameras was possible and determinant for the development of a methodology capable of detecting leaf deterioration.

Burmgarder et al. (2012) also highlighted the possibility of monitoring the leaf area index (LAI) using RGB images and machine learning algorithms. The authors were able to estimate the LAI of lettuce samples under different treatments with an accuracy ranging from 71% to 95%.

As shown in Table 3, the performance of the $B_{660}$ chlorophyll estimate regression model stands out. The composition of the red spectrum response may be associated with the fact that this range is sensitive to chlorophyll a and b (Jensen, 2009). Hyperspectral data can also be used in similar studies. Simko et al. (2015) determined that lettuce leaf deterioration can be detected using hyperspectral images. In the same study, using hyperspectral indices, chlorophyll was classified according to leaf health with an accuracy of 97%.

Hyperspectral curves obtained by field spectroradiometers can also provide spectral models capable of estimating biometric variables. In Kizil et al. (2012), chlorophyll was estimated with an accuracy of 97%, and the authors built models using neural networks from hyperspectral indices derived from field radiometric measurements.

With regard to the physiological variables, the M5 model showed higher precision in the estimates of $F_v/F_m$ (1.66%), ABS/RC (6.5%), TRo/RC (5.8%), and Dio/RC (12%), while the neural networks had higher precision in the estimates of $F_0$ (10.2%), $F_m$ (12.5%), $F_v$ (13.37%), and Chlorophyll a (5.9%). The estimate of the ETo/RC parameter was not satisfactory because the errors were 77% and 63% for the neural networks and M5 models, respectively (Table 4).

Among the nutritional variables, the M5 model was more accurate in estimating N (21.9%), P (15.6%), Mg (7.5%), and S (11.5%), while the neural networks had higher precision in the estimation of K (9.2%) and Ca (11.2%) (Table 4).

With respect to N, Mao et al. (2015) discussed the potential of composing N prediction models in lettuce from hyperspectral images, where it was possible to characterize in detail the spectral response of the plant and detect nutritional characteristics omitted by multispectral cameras.

Furthermore, although it is possible to estimate Ca by M5, the architecture created by the neural networks proved to be more efficient. Various studies aimed at monitoring Ca in lettuce have already been reported. Story et al. (2010) determined that it was possible to detect Ca deficiency using computational vision algorithms by considering the integration between images composed by visible RGB bands, hue-saturation-luminance (HSL), and morphological features.

It is important to emphasize that, based on the results obtained, except for chlorophyll a, TRo/RC, and ETo/RC contents, it is possible to estimate the biometric, physiological, and nutritional variables of lettuce in absolute values, using multispectral cameras to monitor seedling yield without the need for destructive analyses, in the shortest possible time and

without complex architectures, given that the simplest models generated by the M5 parametric algorithm were efficient in estimating the variables assessed.

Despite the possibility of accurately estimating these lettuce variables, it is important to emphasize that the methodology exhibits implementation limitations in situ. The challenges lie mainly in the radiometric data acquisition protocol, because under ideal electromagnetic radiation incidence, the interval between images should be restricted to a short timeframe, between 11 a.m. and 1 p.m.

Another point to highlight is the possibility of continuous calibration of coefficients contained in the regression models and the configuration of neural network architecture, since the models presented here are calibrated for specific seasonal conditions at the time and place where the images were taken. Thus, with the confirmed possibility of estimating quality variables in lettuce using the Mapir camera, it is recommended that models be recalibrated for every field campaign.

## Conclusions

1. In this study, it was possible to quantify the estimated biometric, physiological, and nutritional variables of lettuce using multispectral cameras.

2. Among the Mapir camera bands, $B_{660}$ exhibited the greatest variability, showing that the red range was the most sensitive to different treatments. Except for the variables Chlorophyll a, TRo/RC, and ETo/RC from the $B_{850}$ band and derived indices, it was possible to estimate all the agronomic variables using the models generated by the M5 algorithm.

## Literature Cited

Akaike, H. A new look at the statistical model identification. IEEE Transactions on Automatic Control, v.19, p.716-723, 1974. https://doi.org/10.1109/TAC.1974.1100705

Birth, G. S.; Mcvey, G. R. Measuring the color of growing turf with a reflectance spectrophotometer. Agronomy Journal, v.60, p.640-643, 1968. https://doi.org/10.2134/agronj1968.00021962006000060016x

Bumgarner, N. R.; Miller, W. S.; Kleinhenz, M. D. Digital image analysis to supplement direct measures of lettuce biomass. Horticulture and Crop , v.22, p.547-555, 2012. https://doi.org/10.21273/HORTTECH.22.4.547

Cantelli, D. A. V.; Hamawaki, O. T.; Rocha, M. R.; Nogueira, A. P. O.; Hamawaki, R. L.; Sousa, L. B.; Hamawaki, C. D. L. Analysis of the genetic divergence of soybean lines through hierarchical and Tocher optimization methods. Genetics and Molecular Biology, v.15, p.1-13, 2016. https://doi.org/10.4238/gmr.15048836

Desai, S.; Bagyaraj, D. J.; Ashwin, R. Inoculation with microbial consortium promotes growth of tomato and capsicum seedlings raised in pro trays. Proceedings of the National Academy of Sciences, India Section B: Biological Sciences, v.90, p.21-28, 2020. https://doi.org/10.1007/s40011-019-01078-w

Gitelson, A. A.; Kaufman, Y. J.; Merzlyak, M. N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment, v.58, p.289-298, 1996. https://doi.org/10.1016/S0034-4257(96)00072-7

Gogoi, N. K.; Deka, B.; Bora, L. C. Remote sensing and its use in detection and monitoring plant diseases: A review. Agricultural Reviews, v.39, p.307-313, 2018. https://doi.org/10.18805/ag.R-1835

Gusatti, M.; Zanuzo, M. R.; Machado, R. A. F.; Vieira, C. V.; Cavalli, E. Performance of agricultural substrates in the production of lettuce seedlings (*Lactuca sativa* L.). Scientific Electronic Archives, v.12, p.40-46, 2019. https://doi.org/10.36560/1252019807

Jensen, J. R. Remote sensing of the environment: A perspective on terrestrial resources. São José dos Campos: Parenthesis, 2009. 672p.

Kizil, U.; Genç, L.; Inalpulat, M.; Sapolyo, D.; Mirik, M. Lettuce (*Lactuca sativa* L.) yield prediction under water stress using artificial neural network (ANN) model and vegetation indices. Zemdirbyste, v.99, p.409-418, 2012.

Maciel, G. M.; Gallis, R. B. A.; Barbosa, R. L.; Pereira, L. M.; Siquieroli, A. C. S.; Peixoto, J. V. M. Image phenotyping of inbred red lettuce lines with genetic diversity regarding carotenoid levels. International Journal of Applied and Earth Observation Geoinformation, v.81, p.154-160, 2019. https://doi.org/10.1016/j.jag.2019.05.016

Makanza, R.; Zaman-Allah, M.; Cairns, J. E.; Magorokosho, C.; Tarekegne, A.; Olsen, M.; Prasanna, B. M. High-throughput phenotyping of canopy cover and senescence in maize field trials using aerial digital canopy imaging. Remote Sensing, v.10, p.1-13, 2018. https://doi.org/10.3390/rs10020330

Mao, H.; Gao, H.; Zhang, X.; Kumi, F. Nondestructive measurement of total nitrogen in lettuce by integrating spectroscopy and computer vision. Scientia Horticulturae, v.184, p.1-7, 2015. https://doi.org/10.1016/j.scienta.2014.12.027

Oliva, F. A.; Vieira, M. A.; Baldoto, P. V.; Pocaya, A. P. Production cost and sustainability of lettuce. Colloquium Agrariae, v.12, p.30-35, 2016. https://doi.org/10.5747/ca.2016.v12.nesp.000167

Oliveira, R. A.; Näsi, R.; Niemeläinen, O.; Nyholm, L.; Alhonoja, K.; Kaivosoja, J.; Jauhiainen, L.; Viljanen, N.; Nezami, S.; Markelin, L. Machine learning estimators for the quantity and quality of grass swards used for silage production using drone-based imaging spectrometry and photogrammetry. Remote Sensing of Environment, v.246, p.1-20, 2020. https://doi.org/10.1016/j.rse.2020.111830

Picoli, M. C. A.; Lamparelli, R. A. C.; Sano, E.; Mello, J. R. B.; Rocha, J. V. Effect of sugar cane planting row directions on ALOS/PALSAR satellite images. Remote Sensing, v.50, p.349-357, 2013. https://doi.org/10.1080/15481603.2013.808457

Ren, D. D. W.; Tripathi, S.; Li, L. K. B. Low-cost multispectral imaging for remote sensing of lettuce health. Journal of Applied Remote Sensing, v.11, 016006-1, 2017. https://doi.org/10.1117/1.JRS.11.016006

Rouse, J. W.; Haas, R. H.; Well, J. A.; Deering, D. W. Monitoring vegetation systems in the great plains with ERTS. Proceedings, v.1, p.309-317, 1973.

Simko, I.; Jimenez-Berni, J. A.; Furbank, R. T. Detection of decay in fresh-cut lettuce using hyperspectral imaging and chlorophyll fluorescence imaging. Postharvest Biology and Technology, v.106, p.44-52, 2015. https://doi.org/10.1016/j.postharvbio.2015.04.007

Story, D.; Kacira, M.; Kubota, C.; Akoglu, A.; An, L. Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. Computers and Electronics in Agriculture, v.74, p.238-243, 2010. https://doi.org/10.1016/j.compag.2010.08.010

Weiss, M.; Jacob, F.; Duveillerc, G. Remote sensing for agricultural applications: A meta-review. Remote Sensing of Environment, v.236, 111402, 2020. https://doi.org/10.1016/j.rse.2019.111402