Article

Prediction of Soybean Plant Density Using a Machine Learning Model and Vegetation Indices Extracted from RGB Images Taken with a UAV

Predrag Ranđelović *, Vuk Đorđević, Stanko Milić, Svetlana Balešević-Tubić, Kristina Petrović[ID]**, Jegor Miladinović**[ID] **and Vojin Đukić**

Institute of Field and Vegetable Crops, Maksima Gorkog 30, Novi Sad 21000, Serbia;
vuk.djordjevic@ifvcns.ns.ac.rs (V.Đ.); stanko.milic@ifvcns.ns.ac.rs (S.M.); svetlana.tubic@ifvcns.ns.ac.rs (S.B.-T.);
kristina.petrovic@ifvcns.ns.ac.rs (K.P.); jegor.miladinovic@ifvcns.ns.ac.rs (J.M.);
vojin.djukic@ifvcns.ns.ac.rs (V.Đ.)

* Correspondence: predrag.randjelovic@ifvcns.ns.ac.rs; Tel.: +381-21-489-8480; Fax: +381-21-489-8131

check for updates

Abstract: Soybean plant density is an important factor of successful agricultural production. Due to the high number of plants per unit area, early plant overlapping and eventual plant loss, the estimation of soybean plant density in the later stages of development should enable the determination of the final plant number and reflect the state of the harvest. In order to assess soybean plant density in a digital, nondestructive, and less intense way, analysis was performed on RGB images (containing three channels: RED, GREEN, and BLUE) taken with a UAV (Unmanned Aerial Vehicle) on 66 experimental plots in 2018, and 200 experimental plots in 2019. Mean values of the R, G, and B channels were extracted for each plot, then vegetation indices (VIs) were calculated and used as predictors for the machine learning model (MLM). The model was calibrated in 2018 and validated in 2019. For validation purposes, the predicted values for the 200 experimental plots were compared with the real number of plants per unit area (m^2). Model validation resulted in the correlation coefficient—R = 0.87, mean absolute error (MAE) = 6.24, and root mean square error (RMSE) = 7.47. The results of the research indicate the possibility of using the MLM, based on simple values of VIs, for the prediction of plant density in agriculture without using human labor.

Keywords: soybean; machine learning; vegetation indices; UAV; RGB images

1. Introduction

Field-based phenotyping can be promising for gathering information about the number of plants per plot, which is important for the description of plant traits in agricultural practices [1]. Soybean plant architecture is highly influenced by plant density. An insufficient number of plants per unit area results in more branches and more pods per plant [2]. The number of plants per unit area at emergence is often different from the number of plants in the harvest. Plant losses are usually caused by technological operations (machinery, inter-row cultivation, mechanical weed control and pesticide application), severe weather conditions (hay, flooding, and frost) and biotic factors (pests and diseases). Due to plant loss, knowledge of the number of plants in the later stages of canopy development should enable a precise approximation of the final number of plants in the harvest. The number of plants per unit area will provide information about emergence and potential losses in plant density, which is important for both agricultural science and production. Furthermore, the later stages of crop development provide an extended timeframe for plant density estimation, which is not limited to the early period of crop emergence. The usual way of obtaining data about plant density on experimental plots can be tiring

and implies a lot of manual work. This can be avoided through the implementation of remote sensing techniques and tools.

Unmanned Aerial Vehicle (UAV) are new instruments useful for expanding our knowledge about precision farming and phenotyping [1]. The use of a UAV equipped with a suitable multispectral or RGB camera (containing three channels: RED, GREEN, and BLUE) in precision agriculture has been increasing due to the reduced use of human labor and the speed of data collection. Also, a UAV is a cheaper and more precise alternative to satellite and airborne vehicles [3]. A lot of information can be obtained using these machines, such as high-resolution digital elevation models (DEMs), maps of vegetation height, or the calculation of different vegetation indices (VIs) [4,5].

Many VIs can be obtained from a multispectral camera that has five channels—RED (R), GREEN (G), BLUE (B), NEER INFRARED (NIR), and RED EDGE (RE) or from RGB cameras. VIs derived from multispectral camera like the Normalized Difference Vegetation Index (NDVI), Soil-Adjusted Vegetative Index (SAVI), Enhanced Vegetative Index (EVI), or Normalized Pigment Chlorophyll Ratio Index (NPCI) were used to provide significant information on how to improve the validation of numerous agronomic traits like Leaf Area Index (LAI), leaf chlorophyll content, and plant senescence [6].

NDVI is based on the information derived from two spectral channels (R and NIR), which enables assessment of different information about the crops, like determination of nitrogen use efficiency [7], and yield estimation in wheat [8]. In previous studies, multispectral VIs like NDVI, Green Ratio Vegetation Index (GRVI), and Wide Dynamic Range Vegetation Index (WDRVI) were implemented in the models used for crop height and development analysis during the vegetation season [9]. Although multispectral cameras and different VIs can provide wider information about crop development they are a relatively expensive tool, so in this study we used a cheaper but still good alternative: a UAV with a simple RGB camera. The RGB camera is a good alternative to a multispectral one not just because the lower price but also because of the possibility of calculating many VIs using appropriate equations from RGB images. VIs like Excess Green (ExG), Excess Green Red (ExGR), or Color Index of Vegetation Extraction (CIVE), obtained from RGB images were used to distinguish vegetation from the soil which is important for classification and extraction of the plants [10]. Also in the previously conducted studies, VIs like CIVE, ExGR, ExG, Triangular Greenness Index (TGI), and other RGB VIs were used for early prediction of yield, lodging, and other important soybean traits [11]. RGB and/or multispectral imagery was used for the estimation of important indicators of crop development like biomass [12] and temperature [13]. In research on sunflowers and maize, the ExG index obtained from RGB images taken with UAV was used for distinguishing crops from weeds [14]. VIs calculated from RGB digital imagery were also used for obtaining the information about LAI and biomass in cereal breeding program [15]. The extraction of green pixels from UAV images in order to detect plant ground cover is considered a good alternative to the traditional methods, which are destructive and require a lot of human labor [16]. Because of that and based on the many above listed examples, in research on soybean plant density prediction we chose the following VIs: TGI, Green Leaf Index (GLI), Normalized Green (NG), ExGR, Red Green Difference (RGD), Normalized Green Red Difference (NGRD), Modified Normalized Green Red difference (MNGRD), and Modified Excess Green (MExG).

The results of plant ground cover detection are improved when machine learning models (MLMs) are also used in the detection process. The use of MLMs for image classification has been increasing [17–21]. With the information extracted from UAV images, these models can be a strong and effective tool for the prediction of different crop parameters. One of the MLMs used for classification and calculation is Random Forest (RF). The model is based on binary trees and can be used for correlation and regression [22]. RF uses many trees to classify a set of data, after which it calculates the predictions using data from all the trees [23]. In previous studies, RF was used for estimation of leaf coverage in maize [24], soybean yield prediction [25], and determination of leaf chlorophyll content in wheat [26].

The main objective of this study was to develop an MLM based on values of simple VIs obtained from RGB images for the prediction of soybean plant density in mid-development stages. An additional objective of this study was to validate the model in an independent environment to test its robustness.

2. Materials and Methods

2.1. Trial Description

The trial was conducted in 2018 and 2019 on the experimental fields of the Institute of Field and Vegetable Crops in Rimski Šančevi, Serbia. The trial was performed in both years on chernozem soil with homogeneous texture across the entire experimental site, and these sites were deep and well drained. In both years, standard cultivation practices were applied to the experimental field with the sowing dates, row and seed spacing in soybean planting recorded (Table 1).

Table 1. Year and location (latitude and longitude) of the experimental trials along with the sowing date and row and seed spacing, for the experimental trials in 2018 and 2019.

Year	Site	Latitude	Longitude	Sowing Date	Row Spacing (cm)	Seed Spacing (cm)
2018	Rimski Šančevi	45°19′33″ N	19°50′10″ E	6.4.2018	45	5
2019	Rimski Šančevi	45°20′03″ N	19°50′13″ E	19.4.2019	45	5

In 2018, 66 soybean genotypes, each sown on its own 8 m^2 plot were used for calibration of the MLM, while in 2019, 200 soybean genotypes, each sown on its own 10 m^2 plot, were used for model validation. In both years, trials were planted on different fields. In total, the analysis included 266 different soybean genotypes, which were experimental lines from the soybean breeding programs. The genotypes included in the trial represented different maturity groups and different plant architecture.

2.2. UAV Description

The UAV platform used for collecting the RGB photos was Drone Phantom 4 (DJI, Shenzhen, China), powered by four propellers and operated with a remote controller that runs on 2.4 GHz. Soybean plots were filmed using an integrated camera with the following characteristics: 1/2.3″ CMOS (Complementary Metal Oxide Semiconductor) sensor with 12.4 megapixels, focal length of 8.8 mm and 1.84 cm/pixel resolution. The maximum wind speed that allows for image acquisition with the UAV is 10 m/s. To determine the geographic position of each image, the UAV used the global positioning systems GPS/GLONASS (Global Positioning System/Global Navigation Satellite System).

2.3. Field Based and Remote Data Collection

In 2018, the number of plants was first counted manually for each of the 66 experimental plots and then, by dividing the total number of plants per plot by the plot area (8 m^2) we calculated the number of plants per unit area (m^2) for each plot. After that, RGB photos were taken with a UAV in the soybean development phases of four unfolded trifoliolate leaves (V4) (Figure 1a) and beginning pod (R3) (Figure 1b). In 2019, the same data collection procedure was repeated for 200 plots for validation of the model developed in 2018.

In both trial years, photos were taken on a sunny day, at a wind speed equal to or less than 10 m/s, and between 10:00 and 14:00; more details about the UAV photo acquisition are given in Table 2.

In 2018, after acquisition of all the individual images in phase V4, an orthophoto of the entire trial was created, and the same process was repeated with the images taken in the R3 phase. The same procedure was repeated in 2019. The creation of the orthophoto involved a process of stitching together all the individual photos, which was carried out using the open-source software WebODM [27].

The next step was the analysis of the individual plots on the RGB orthophoto through open-source software for image analysis called Fiji [28]. First, in Fiji, the region of interest (ROI) was created on

RGB images for 66 plots from the 2018 trial, and 200 experimental plots from the 2019 trial. After the creation of ROI for each plot, with the Fiji's function stack to images, RGB images were separated into the individual channels R, G and B (Figure 2).
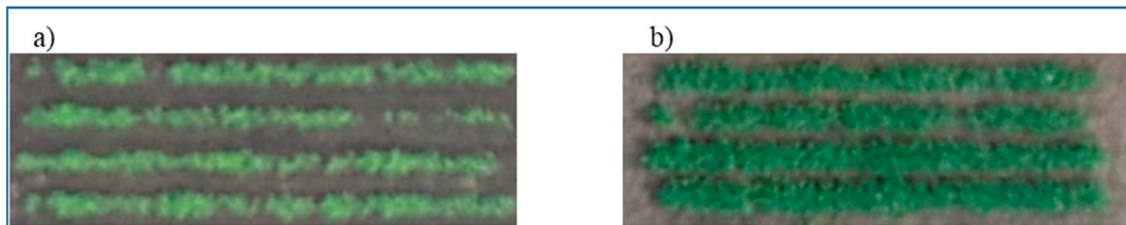
Figure 1. Plot images of soybean in the V4 phase: (**a**) four unfolded trifoliolate leaves and the (**b**) R3 phase—beginning pod; images were taken with the UAV (Unmanned Aerial Vehicle) in 2018.

Table 2. Basic details about acquisition of photos.

Year	Flight Date	Flight Altitude (m)	Soybean Development Stage	Ground Resolution (cm/pix)
2018	25.5.2018	100	Four unfolded trifoliolate leaves (V4)	1.84
	20.6.2018	100	Beginning pod (R3)	1.84
2019	30.5.2019	100	Four unfolded trifoliolate leaves (V4)	1.84
	14.6.2019	100	Beginning pod (R3)	1.84
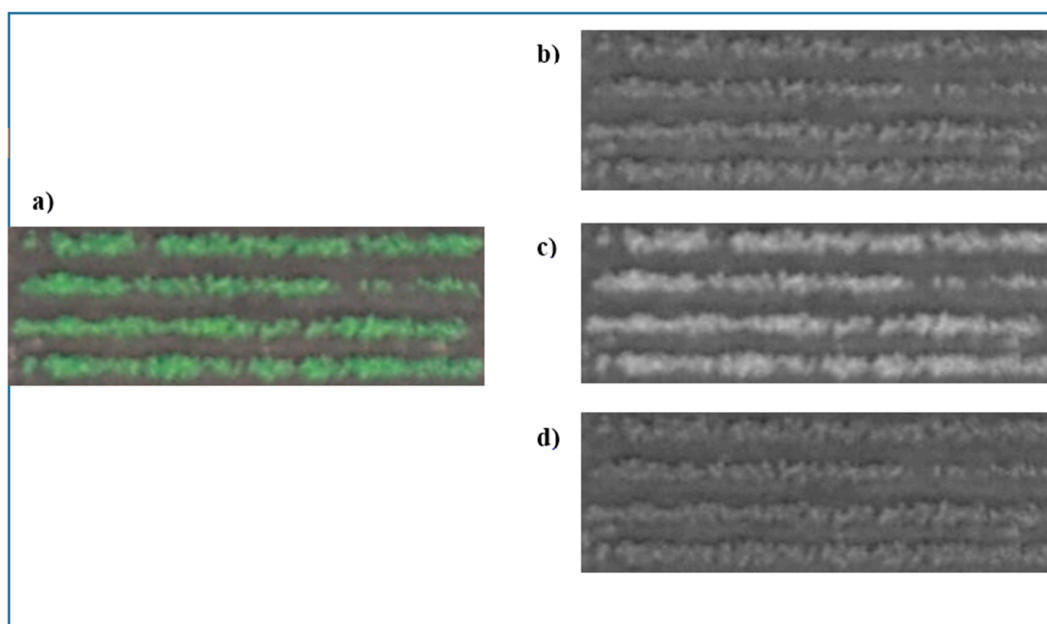
Figure 2. Example of the plot image in the RGB (containing three channels: RED, GREEN, and BLUE) color space (**a**) and individual channels of the same plot: R (**b**), G (**c**), and B (**d**) after separation of the RGB image in Fiji.

The further step implied the extraction of the mean values from each individual channel (R, G, and B) for every plot (ROI) in both years (2018 and 2019) and for both development phases (V4 and R3). That was accomplished using Fiji's measure tool, which was applied to each of the three individual channels. The individual values obtained for all experimental plots were used for the calculation of the simple VIs, which were predictors for the MLM.

2.4. Vegetation Indices Calculated from UAV Images

Eight VIs were used for the prediction model of the number of soybean plants/m^2 (Table 3).

Table 3. Vegetation indices (VIs) used as predictors for the machine learning model (MLM). G—mean value of the GREEN channel, R—mean value of the RED channel, and B—mean value of the BLUE channel.

Vegetation Index	Name	Formula	References
TGI	Triangular greenness index	$G - 0.39 \times R - 0.61 \times B$	[29]
GLI	Green leaf index	$(2 \times G - R - B)/(2 \times G + R + B)$	[30]
NG	Normalized green	$G/(R + G + B)$	[31]
ExGR	Excess green red	$(3 \times G - 2.4 \times R - B)/(R + G + B)$	[32]
RGD	Red green difference	$R - G$	[33]
NGRD	Normalized green red difference	$(G - R)/(G + R)$	[34]
MNGRD	Modified normalized green red difference	$(G^2 - R^2)/(G^2 + R^2)$	[12]
MExG	Modified excess green	$1.262 \times G - 0.884 \times R - 0.311 \times B$	[35]

2.5. Prediction Model

R package RF was used for the prediction of soybean plant density, including the following settings (maxnodes = 50 and ntree = 100) [36,37]. To work properly, the machine learning algorithm needs training and test data sets as input parameters.

In the first trial year, the prediction model was based on the number of plants/m^2 and the VIs calculated on the 66 experimental plots. In this case, VIs and the number of plants counted from 80% of the randomly selected plots were used as the training set, while VIs for the remaining 20% of the plots were used as the test set. This 80/20 data partition is done by the code that is included in the RF model. After running the model and obtaining the predicted values of the number of plants/m^2 for 20% of the plots, the obtained values were compared to the manually counted values, so as to assess the quality of the model.

In the second trial year, VIs and the number of plants, manually counted on 66 experimental plots in 2018 were used as training sets for model validation, while the VIs collected on 200 plots in 2019 were used for the test set to predict the number of plants/m^2. For the validation of the model, the predicted values for 200 plots in 2019 were also compared with those counted manually. The results obtained after the comparison were shown through the correlation coefficient (R), coefficient of determination (R^2), mean of the absolute error (MAE), and root mean square error (RMSE).

3. Results

RGB images that were collected with the UAV in both soybean trial years were of good quality because they provided sharp orthophoto of the experimental trials in both years after a stitching procedure and were usable for the calculation of the simple VIs. The results showed a high correlation between the real and predicted number of plants/m^2, with a relatively low mean absolute error and root mean squared error (Table 4).

Table 4. Statistical parameters for comparison of the real and predicted number of plants/m^2 for each plot in model calibration. R—correlation coefficient, R^2—correlation of determination, MAE—mean absolute error, and RMSE—root mean squared error.

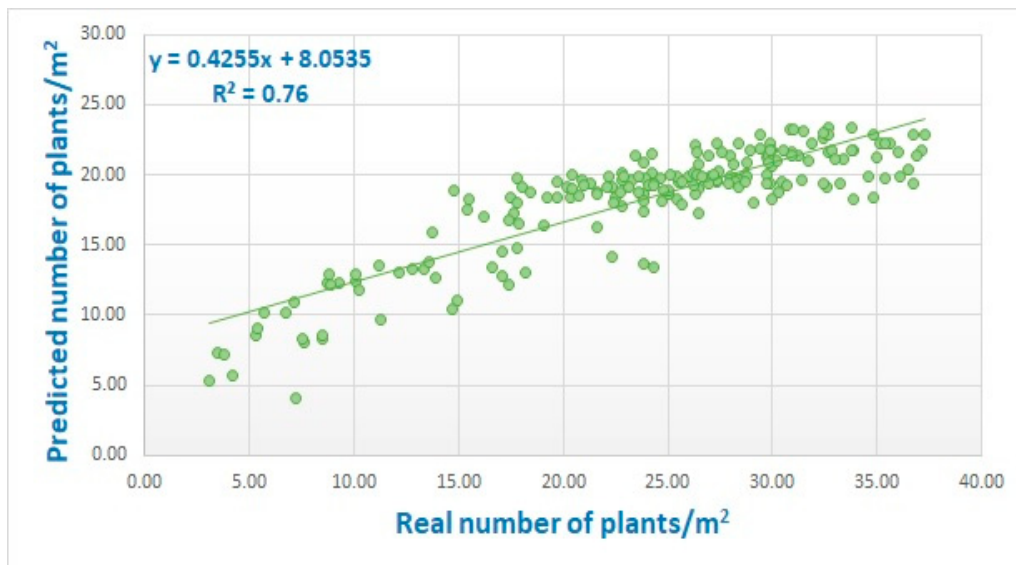| Model Calibration in 2018	Predicted Number of Plants/m^2 for Each Plot			
	R	R^2	MAE	RMSE
Real number of plants/m^2 for each plot	0.90	0.80	3.07	3.91

The model in 2018 showed good results, however, for accepting it as a tool for prediction of the number of plants per unit area, it is important that the model provides a similar quality in different years. The model from 2018 was therefore evaluated the following year on the 200 experimental plots (Table 5).

Comparing the predicted and real values of number of plants/m^2, it can be concluded that the high correlation coefficient and R^2 between these two variables indicates the possibility of using this model

as a tool for digital counting of plants on experimental plots (Figure 3). Lower R^2 and higher error for model validation in 2019 were observed compared to the results obtained in the model calibration in 2018. This was expected and illustrates the effect of uncontrolled factors.

Table 5. Statistical parameters for comparison of real and predicted number of plants/m^2 for each plot in model calibration. R—correlation coefficient, R^2—coefficient of determination, MAE—mean absolute error, and RMSE—root mean squared error.

Model Validation in 2019	Predicted Number of Plants/m^2 for Each Plot			
	R	R^2	MAE	RMSE
Real number of plants/m^2 for each plot	0.87	0.76	6.24	7.47

Figure 3. Correlation between the real and predicted number of plants/m^2 for each of the 200 experimental plots for model validation in 2019. R^2—value of coefficient of determination.

After obtaining the predicted values for 200 experimental plots in 2019, descriptive statistics were calculated for the real and predicted number of plants/m^2 (Table 6).

Table 6. Descriptive statistics for 200 plots for the real and predicted number of plants/m^2.

Descriptive Statistics of Number of Plants/m^2 for 200 Plots	Real Values	Predicted Values
Mean	23.94	18.24
Standard Error	0.58	0.28
Standard Deviation	8.24	4.03
Sample Variance	67.92	16.20
Minimum	3.10	4.04
Maximum	37.30	23.33

The results show a higher variability between the individual plots for the real values than between the same plots, but with the predicted values of number of plants/m^2 for each plot. This is indicated by a higher value of standard deviation, which resulted in the higher standard error for real values compared to the predicted ones. Also, the range between the maximum and minimum number of plants/m^2 per plot is narrower for the predicted values than for the real ones. This is mainly because of the lower value of maximum number of plants/m^2 per plot that was calculated by the prediction model.

The number of plants/m^2 manually counted on 200 plots in 2019 and the number of plants/m^2 obtained from the same plots using the prediction model were subtracted; the difference is shown in the box plot (Figure 4).
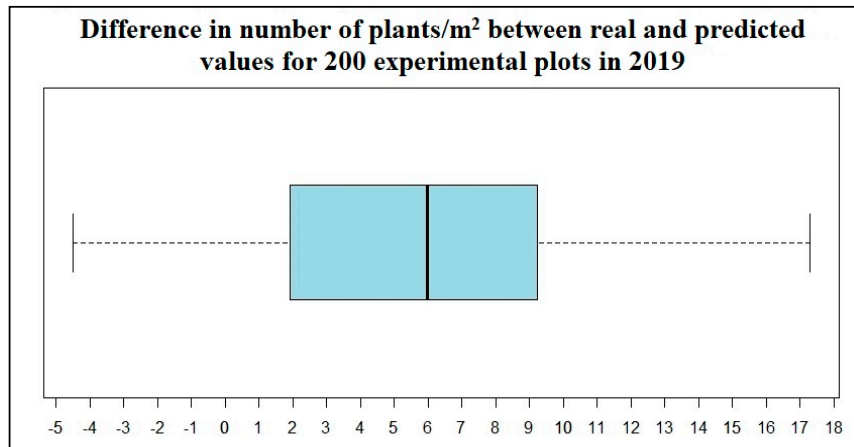
Figure 4. Box plot of difference in the number of plants/m^2 between the real and predicted values for 200 experimental plots in 2019.

The difference between the real and predicted number of plants/m^2 for all plots varied between −5 and +18. This indicates that the prediction model underestimated the real number of plants/m^2 for the plots, which led to a positive difference between the real and predicted values, while overestimating the real number of plants/m^2 for the plots, which led to a negative difference between the real and predicted values.

Predictions of the number of plants per unit area for the 200 plots were based on the simple VIs and the RF machine learning algorithm. Values of the individual VIs used in the RF model derived from the images taken in two middle phases of soybean development (V4 and R3), had a different impact on the final result of the prediction (Figure 5). This is represented through the Increase in Node Purity (IncNodePurity).
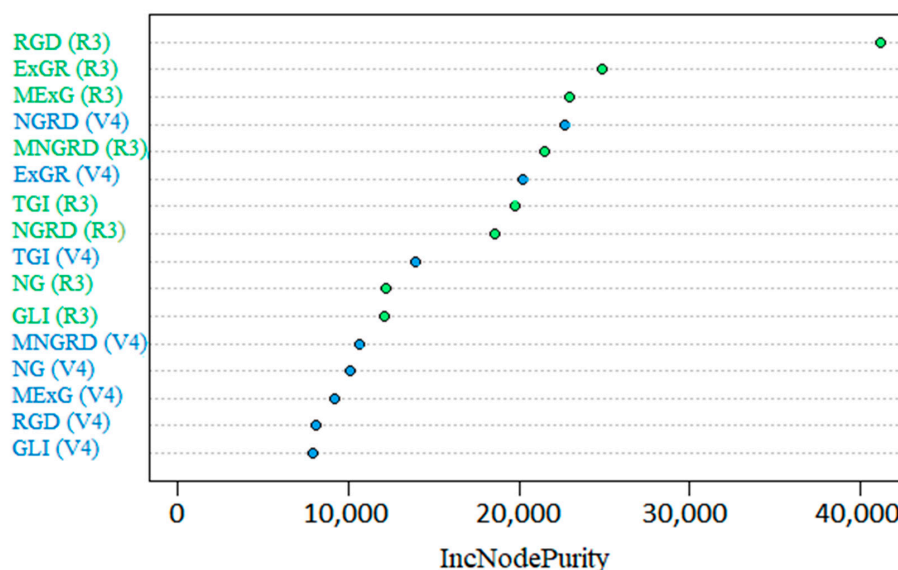
Figure 5. The impact of vegetation indices (Vis) on the final results of plant density prediction, shown through the Increase in Node Purity (IncNodePurity). (V4) and (R3) next to the names of the VIs represent two middle phases of soybean development from which the indices were obtained.

4. Discussion

The main reason why the model underestimated the real number of plants/m^2 lies in the overlap between the plants on the denser plots, so the model could not distinguish every soybean plant individually. For those plots that had a higher number of plants/m^2 in reality, the model predicted lower values.

On the other hand, one possible reason why the model overestimated the real values for some plots was the presence of weeds on those plots so the model counted them as soybean plants, which resulted in a mismatch between the real and predicted values. This indicates that, for higher precision of the results, plots must be free of weeds so as not to disturb the prediction model.

The higher value of IncNodePurity indicates a greater influences of the variable on the final results of the prediction [21]. On average, the values of indices extracted from the later (R3) phase of soybean development had a greater influence as predictors than the values of the same indices extracted from the earlier (V4) phase. This is because the plants were more robust in the R3 phase, with more leaves, which increased the number of green pixels per plot and thus improved the index efficiency. Still, the values of VIs calculated from phase V4 also had an important role on the final results of the prediction, especially for indices NGRD and ExGR which were verified by their high IncNodePurity values.

In the study, the middle phases (V4 and R3) of soybean development were used for plant density prediction. However, in a two-year experiment on safflower, a density estimation model was proposed, analyzing the pixel green ratio of plants that were in the of 2–4 leaves stage [38]. The results of the study showed that better values of prediction (for 2017: $R^2 = 0.88$; for 2018 $R^2 = 0.86$) were obtained during the early growth stages because of less overlapping between the plants. This indicates the higher accuracy of the soybean prediction model if the VIs are calculated from the images of plants in the earlier growth stages than in stages V4 and R3 when the plants overlap. The reason for using the middle soybean development stages in soybean trials was to avoid any errors in calculating the predictions which could happen for two reasons. One mistake could occur if an analysis is done too early is that, due to the uneven emergence of all individual plants in the plots, some plants are not taken into account. A second mistake can be potential plant losses occurring as a consequence of inter-row cultivation.

A study on maize has pointed out that the simple amount of green pixel extracted from the images is not enough for digital counting of the number of plants [1]. This is reflected in the low value of $R^2 = 0.023$ calculated between the number of green pixels and the number of plants counted manually. Only after conducting the additional image transformation did the results improve significantly ($R^2 = 0.89$). Excessive and often complicated image transformations were avoided during the soybean research. This was achieved by using only the values of pixels to calculate simple VIs and importing them into the MLM, which resulted in a precise prediction of the number of plants per unit area ($R^2 = 0.80$ in 2018 and $R^2 = 0.76$ in 2019).

5. Conclusions

Plant density is one of the most important factors in successful agricultural production, not only for soybeans but for all crops. It is of key importance in obtaining high yields. Therefore, information about the number of plants per unit area is necessary. This information can be obtained by means of a nondestructive method, using the MLMs and VIs derived from RGB images. A favorable time for the extraction of VIs is the middle phase of plant development, because all plants that emerge earlier or later are taken into account for the calculation. Although more research is needed, the predictions calculated using RF in the soybean trials gave good results and showed that the method can be used as new tool for gathering significant information about crops.

References

1. Gnädinger, F.; Schmidhalter, U. Digital counts of maize plants by Unmanned Aerial Vehicles (UAVs). *Remote Sens.* **2017**, *9*, 544. [CrossRef]

2. Rigsby, B.; Board, J.E. Identification of soybean cultivars that yield well at low plant populations. *Crop Sci.* **2003**, *43*, 234–239. [CrossRef]

3. Jannoura, R.; Brinkmann, K.; Uteau, D.; Bruns, C.; Joergensen, R.G. Monitoring of crop biomass using true colour aerial photographs taken from from a remote controlled hexacopter. *Biosyst. Eng.* **2015**, *129*, 341–351. [CrossRef]

4. Hatfield, J.L.; Prueger, J.H. Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. *Remote Sens.* **2010**, *2*, 562–578. [CrossRef]

5. Maes, W.H.; Steppe, K. Perspectives for remote sensing with Unmanned Aerial Vehicles in precision agriculture. *Trends Plant Sci.* **2018**, *24*, 152–164. [CrossRef]

6. Tao, H.; Feng, H.; Xu, L.; Miao, M.; Long, H.; Yue, J.; Li, Z.; Yang, G.; Yang, X.; Fan, L. Estimation of crop growth parameters using UAV-based hyperspectral remote sensing data. *Sensors* **2020**, *20*, 1296. [CrossRef]

7. Naser, M.A.; Khosla, R.; Longchamps, L.; Dahal, S. Characterizing variation in nitrogen use efficiency in wheat genotypes using proximal canopy sensing for sustainable wheat production. *Agronomy* **2020**, *10*, 773. [CrossRef]

8. Naser, M.A.; Khosla, R.; Longchamps, L.; Dahal, S. Using NDVI to differentiate wheat genotypes productivity under dryland and irrigated conditions. *Remote Sens.* **2020**, *12*, 824. [CrossRef]

9. Maresma, A.; Ariza, M.; Martinez, E.; Lloveras, J.; Martinez-Casanovas, J.A. Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (zea mays L.) from a standard UAV service. *Remote Sens.* **2016**, *8*, 973. [CrossRef]

10. Romeo, J.; Pajares, G.; Montalvo, M.; Guerrero, J.M.; Guijaro, M.; de la Cruz, J.M. A new expert system for greenness identification in agricultural images. *Expert. Syst. Appl.* **2013**, *40*, 2275–2286. [CrossRef]

11. Yuan, W.; Wijewardane, N.K.; Jenkins, S.; Bai, G.; Ge, Y.; Graef, L.G. Early prediction of soybean traits through color and texture features of canopy RGB imagery. *Sci. Rep.* **2019**, *9*, 14089. [CrossRef] [PubMed]

12. Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M.L.; Bareth, G. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs.* **2015**, *39*, 79–87. [CrossRef]

13. Zarco-Tejada, P.J.; Gonzalez-dugo, V.; Berni, J.A. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sens. Environ.* **2012**, *117*, 322–337. [CrossRef]

14. Perez-Ortiz, M.; Pena-Barragan, J.M.; Gutierez, P.A.; Torres-Sanchez, J.; Hervas Martinez, C.; Lopez-Granados, F. A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method. *Appl. Soft Comput.* **2015**, *37*, 533–544. [CrossRef]

15. Casadesús, J.; Villegas, D. Conventional digital cameras as a tool for assessing leaf area index and biomass for cereal breeding. *J. Integr. Plant Biol.* **2014**, *56*, 7–14. [CrossRef] [PubMed]

16. Kipp, S.; Mistele, B.; Baresel, P.; Schmidhalter, U. High-throughput phenotyping early plant vigour of winter wheat. *Eur. J. Agron.* **2014**, *52*, 271–278. [CrossRef]

17. Guo, W.; Fukatsu, T.; Ninomiya, S. Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images. *Plant Methods* **2015**, *11*, 7. [CrossRef]

18. Peters, J.; de Baets, B.; Verhoest, N.E.C.; Samson, R.; Degroeve, S.; de Becker, P.; Huybrechts, W. Random forests as a tool for ecohydrological distribution modeling. *Ecol. Model.* **2007**, *207*, 304–318. [CrossRef]

19. Van Beijma, S.; Comber, A.; Lamb, A. Random forest classification of salt marsh vegetation habitats using quad-polarimetric airborne SAR, elevation and optical RS data. *Remote Sens. Environ.* **2014**, *149*, 118–129. [CrossRef]

20. Wiesmeier, M.; Barthold, F.; Blank, B.; Koegel-Knabner, I. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. *Plant Soil* **2011**, *340*, 7–24. [CrossRef]

21. Khun, S.; Neumann, S.; Egert, B.; Steinbeck, C. Building blocks for automated elucidation of metabolites: Machine learning methods for NMR prediction. *BMC Bioinform.* **2008**, *9*, 400.

22. Jeong, J.; Resop, J.P.; Mueller, N.D.; Fleisher, D.H.; Kyungdahm, Y.; Butler, E.E.; Timlin, D.; Kyo-Moon, S.; Gerber, J.; Vangimalla Ramakrishna, R.; et al. Random Forests for global and regional crop yield predictions. *PLoS ONE* **2016**, *11*, e0156571. [CrossRef] [PubMed]

23. Cutler, D.R.; Edwards, T.C.; Bear, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random Forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [CrossRef]

24. Zhou, C.; Ye, H.; Xu, Z.; Hu, J.; Shi, X.; Hua, S.; Yue, J.; Yang, G. Estimating maize-leaf coverage in field conditions by applying a machine learning algorithm to UAV remote sensing images. *Appl. Sci.* **2019**, *9*, 2389. [CrossRef]

25. Parmley, K.A.; Higgins, R.H.; Ganapathysubramanian, B.; Sarkar, S.; Singh, A.K. Machine learning approach for prescriptive plant breeding. *Sci. Rep.* **2019**, *9*, 17132. [CrossRef]

26. Shah, S.H.; Angel, Y.; Houborg, R.; Ali, S.; McCabe, M.F. A Random Forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Remote Sens.* **2019**, *11*, 920. [CrossRef]

27. WebODM. Available online: https://www.opendronemap.org/webodm/ (accessed on 14 July 2020).

28. Fiji is JustImageJ. Available online: http://fiji.sc/Fiji (accessed on 14 July 2020).

29. Hunt, E.R.J.; Daughtry, C.S.T.; Eitel, J.U.H.; Long, D.S. Remote sensing leaf chlorophyll content using a visible band index. *Agron. J.* **2011**, *103*, 1090–1099. [CrossRef]

30. Hunt, E.R.J.; Doraiswamy, P.C.; McMurtrey, J.E.; Daughtry, C.S.T.; Perry, E.M.; Akhmedov, B. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *Int. J. Appl. Earth Obs.* **2013**, *21*, 103–112. [CrossRef]

31. Woebbecke, D.M.; Meyer, G.E.; Bargen, K.V.; Mortensen, D.A. Color indices for weed identification under various soil, residue, and lighting conditions. *Trans. ASAE* **1995**, *38*, 259–269. [CrossRef]

32. Meyer, G.E.; Neto, J.C. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agr.* **2008**, *63*, 282–293. [CrossRef]

33. Sanjerehei, M.M. Assessment of spectral vegetation indices for estimating vegetation cover in arid and semiarid shrublands. *Range Manag. Agrofor.* **2014**, *35*, 91–100.

34. Hamuda, E.; Glavin, M.; Jones, E. A survey of image processing techniques for plant extraction and segmentation in the field. *Comput. Electron. Agr.* **2016**, *125*, 184–199. [CrossRef]

35. Burgos-Artizzu, X.P.; Ribeiro, A.; Guijarro, M.; Pajares, G. Real-time image processing for crop/weed discrimination in maize fields. *Comput. Electron. Agr.* **2011**, *75*, 337–346. [CrossRef]

36. R Core Team. R development core team. *RA Lang. Environ. Stat. Comput.* **2013**, *55*, 275–286.

37. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.

38. Koh, C.O.J.; Hayden, M.; Daetwyler, H.; Kant, S. Estimation of crop plant density at early mixed growth stages using UAV imagery. *Plant Methods* **2019**, *15*, 64. [CrossRef]