

# Total Nitrogen Estimation in Agricultural Soils via Aerial Multispectral Imaging and Libs

**Md Abir Hossen**

South Dakota School of Mines and Technology

**Prasoon K Diwakar**

South Dakota School of Mines and Technology

**Shankarachary Ragi** (✉ [Shankarachary.Ragi@sdsmt.edu](mailto:Shankarachary.Ragi@sdsmt.edu))

South Dakota School of Mines and Technology

---

## Research Article

**Keywords:** Measuring soil health indicators, real-time unmanned aerial vehicle (UAV), computed vegetation indices, laser-induced breakdown spectroscopy (LIBS)

**Posted Date:** February 15th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-228145/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Total Nitrogen Estimation in Agricultural Soils via Aerial Multispectral Imaging and LIBS

Md Abir Hossen<sup>1</sup>, Prasoon K Diwakar<sup>2</sup>, and Shankarachary Ragi<sup>3,\*</sup>

<sup>1</sup>Graduate Research Assistant, Department of Electrical Engineering

<sup>2</sup>Assistant Professor, Department of Mechanical Engineering

<sup>3</sup>Assistant Professor, Department of Electrical Engineering

<sup>1, 2, 3</sup>South Dakota School of Mines and Technology, Rapid City, SD-57701, USA

\*Shankarachary.Ragi@sdsmt.edu

## ABSTRACT

Measuring soil health indicators is an important and challenging task that affects farmers' decisions on timing, placement, and quantity of fertilizers applied in the farms. Most existing methods to measure soil health indicators (SHIs) are in-lab wet chemistry or spectroscopy-based methods, which require significant human input and effort, time-consuming, costly, and are low-throughput in nature. To address this challenge, we develop an artificial intelligence (AI)-driven near real-time unmanned aerial vehicle (UAV)-based multispectral sensing (UMS) solution to estimate total nitrogen (TN) of the soil, an important macro-nutrient or SHI that directly affects the crop health. Accurate prediction of soil TN can significantly increase crop yield through informed decision making on the timing of seed planting, and fertilizer quantity and timing. We train two machine learning models including multi-layer perceptron and support vector machine to predict the soil nitrogen using a suite of data classes including multispectral characteristics of the soil and crops in red, near-infrared, and green spectral bands, computed vegetation indices, and environmental variables including air temperature and relative humidity. To generate the ground-truth data or the training data for the machine learning models, we measure the total nitrogen of the soil samples (collected from a farm) using laser-induced breakdown spectroscopy (LIBS).

## Introduction

Soil health indicators are a composite set of measurable physical, chemical and biological properties which can be used to determine soil health status. Among the chemical indicators, we particularly focus on nitrogen (N) because N is the most limiting nutrient in many of the world's agricultural areas<sup>1</sup>. Insufficient use of N causes economic loss, in contrast, excessive use of N implies wasting fertilizer, causes nitrate pollution and increases the cost<sup>2,3</sup>. Nitrogen treatment can account for up to 30% from the total production cost<sup>4</sup>.

Chlorophyll meter (CM) measures the chlorophyll content of crops to estimate their N nutrition status. In recent years, the use of CM has increased among researchers and farmers<sup>5,6</sup>. For N management, determination of soil spatial patterns is necessary but requires collection and analysis of a large number of samples which is labor intensive and time-consuming<sup>6,7</sup>.

Satellite-based remote sensing is one alternative to ground-based measurements. Satellite-based techniques utilize images at the spectral level for crop growth monitoring and real-time management<sup>8-10</sup>. For instance, vegetation indices (VIs), evaluated using the data obtained from satellite-based multispectral sensors, have been used to detect the N stress at V4-V7 (4-7 leaves with visible leaf collar) stages<sup>11-13</sup>. However, satellite-based sensing suffers from lower spatial and temporal resolution, and sensing disruption may occur during image acquisition in some areas because of cloud cover and/or sprinkler irrigation<sup>14</sup>. Farmers' adaption of the system is still limited. Additionally, the high cost of obtaining these images for relatively small areas is a significant drawback<sup>15</sup>. Multispectral cameras mounted on unmanned aerial vehicles (UAVs) have enormous potential to resolve this problem. For image acquisition, UAVs can be deployed rapidly and frequently, resulting in reduced costs, greater flexibility in terms of data resolution and mission timing<sup>16-18</sup>.

The purpose of the present study is to develop a machine learning (ML)-based predictive model to estimate soil TN using soil spectral characteristics measured from the multispectral images captured from a UAV. Specifically, we train a multi-layer perceptron (MLP) and support vector machine (SVM) model for TN prediction. We use root mean square error (RMSE) and computational time (CT) as performance metrics to measure the performance of the above predictive model. To reduce the RMSE and lower CT in the machine learning models, we perform hyper-parameter optimization (HPO). The HPO tuning process depends on the ML model used for prediction<sup>19</sup>. Several HPO techniques<sup>20</sup> have been used for different application such as grid search (GS), random search (RS), Bayesian optimization, and particle swarm optimization. In this study, we implement GS, RS and GA for hyper-parameter optimization.

Laser induced breakdown spectroscopy (LIBS) is an analytical method for qualitative and quantitative elemental detection. LIBS can be readily applied to soil samples, providing rapid chemical analysis of soil samples and their constituents (e.g, Nitrogen, Potassium, Phosphorus, Calcium). The combination of an autonomous drone, LIBS, and machine learning can be used to achieve in-field measurement which provides instant results for deficient nutrient analysis and fertilization planning. With appropriate calibration, the LIBS analysis can provide quantitative measurement for most elements in soil including carbon, nitrogen, potassium, sulfur, and phosphorus<sup>21,22</sup>. There have been some applications of standalone LIBS systems in precision agriculture<sup>21–24</sup>, however there has been no detailed research of LIBS application in combination with ML and drones. Some studies have found it challenging to measure nitrogen using LIBS due to environmental factors; Earth's atmosphere is almost 80% nitrogen which will interfere with the sample measurement result since soil is less than 1% nitrogen. Testing in a vacuum or in low-pressure conditions has been suggested to improve measurement accuracy<sup>22</sup>. In this study, we conducted LIBS analysis on soil samples under a normal atmosphere for observation. Low laser pulse energies were used to minimize breakdown of air and thereby minimize the influence of atmospheric nitrogen.

## Methods

In supervised learning, the goal is to obtain an optimal predictive model function  $f^*$  based on the input  $x$  and the output  $y$  to minimize the cost function  $L(f(x), y)$ . In this study, we particularly use MLP and SVM which can be used for both classification and regression problems. We applied HPO techniques to determine the best set of hyper-parameters from the ML models and train the ML models using those hyper-parameters on the training dataset.

### Multi-layer Perceptron (MLP)

Multi-layer perceptron is a supervised learning algorithm that learns a function  $f(\cdot) : R^x \rightarrow R^o$  by training on a dataset<sup>25</sup>, where  $x$  is the number of input dimension and  $o$  is the number of output dimension. We designed the MLP (Figure 1a) with multiple organized layers consisting of various neuron-like processing units. Each node in the layer was connected with the nodes in the previous layer. Each node may have symmetrical or differing strengths and weights. The data in the network enters with the input layer and gradually runs through each layer to reach the output layer. For a given a set of features  $x = \{R, NIR, G, NDVI, Air\ temperature, RH\}$  and target  $y = TN$ ,  $f(\cdot) : R^6 \rightarrow R^1$ . To train the MLP from a given set of input-output pairs  $X = \{(\vec{x}_1, y_1), \dots, (\vec{x}_N, y_N)\}$ , learning consists of iteratively updating the values of weight and bias of the perceptron to minimize RMSE. The hyper-parameter configuration (Table 1) was created using the solver type<sup>26</sup>, activation function<sup>27</sup>, learning rate and hidden layer sizes.

### Support Vector Machines (SVM)

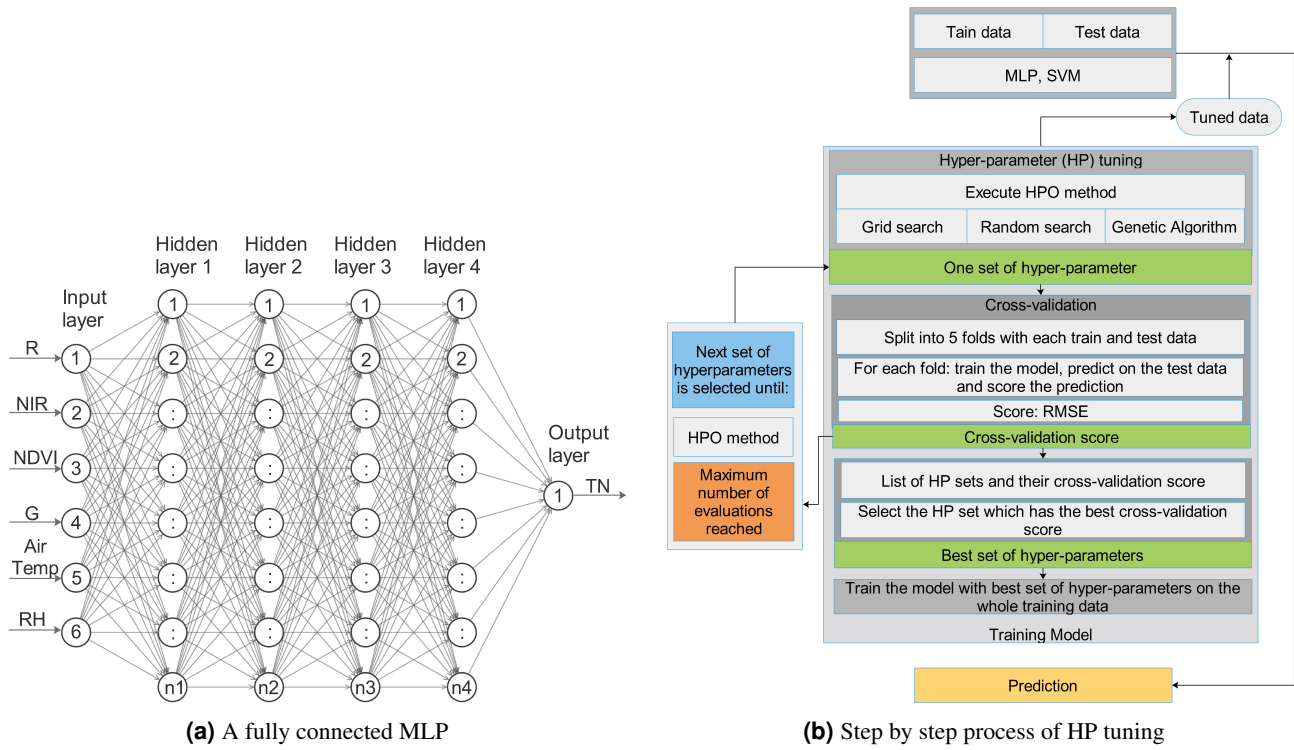
SVM makes data points linearly separable by mapping them from low-dimensional to high-dimensional space. The classification boundary creates a partition between the data points by generating a hyperplane<sup>28</sup>. SVM uses different types of kernel functions to measure the similarities between two data points<sup>29</sup>. Linear kernels, radial basis function (RBF), polynomial kernels, and sigmoid kernels are common kernel types in SVM. We created the hyper-parameter configuration (Table 1) using the kernel types, penalty parameter ( $C$ )<sup>30</sup> and distance error (*epsilon*) of the loss function<sup>27</sup>.

### Hyper-parameter Optimization (HPO)

GS, RS and GA HPO techniques were executed within their respective hyper-parameters to train the model. We performed cross-validation by splitting the train and test data into 5-folds. After obtaining the RMSE from the cross-validation score, we selected the hyper-parameters which yielded the lowest RMSE. Finally, using the best set of hyper-parameters we trained the MLP and SVM models for each HPO technique. Figure 1b shows the step-by-step process of HPO, training the dataset, and prediction of test data.

GS exhaustively evaluates all the combinations in the hyper-parameter configuration space specified by the user in the form of a grid configuration<sup>31</sup>. The user must identify the global optimums manually since GS cannot utilize the well-performing regions<sup>20</sup>. However, in RS, the user defines a budget (i.e., time) as well as the upper and lower bounds of the hyper-parameter values. RS randomly selects the values from the pre-defined boundary and trains until the budget is exhausted<sup>20</sup>. If the configuration space is wide enough, RS can detect the global optima. Assuming a model has  $k$  parameters and each of them has  $n$  distinct values, the GS computational complexity increases exponentially at a rate of  $O(n^k)$ <sup>32</sup>. Therefore, the effectiveness of GS depends on the size of the hyper-parameter configuration space. For RS, the computational complexity is defined as  $O(n)$ , where  $n$  is specified by the user before the optimization process starts<sup>20</sup>.

GA<sup>33</sup> randomly initializes the population and chromosomes. Genes represents the entire search space, hyper-parameters, and hyper-parameter values. GA uses a fitness function to evaluate the performance of each individual in the current generation similarly to the objective function of a ML model. To produce a new generation, GA performs selection, crossover, and mutation



**Figure 1.** (a) HPO for GS, RS and GA with cross-validation, training the ML models with the tuned HP and prediction. (b) MLP with four hidden network having different weights, where input layer  $\in \mathbb{R}^6$ , hidden layer  $\in \mathbb{R}^4$ , output layer  $\in \mathbb{R}^1$  and n1, n2, n3, and n4 represent the number of perceptronw in each hidden layer, respectively.

operations on the chromosomes involving the next hyper-parameter configurations to be evaluated. The cycle continues until the algorithm reaches the global optimum.

## Experimental Design

An aerial survey was carried out with Mavic 2 Pro UAV. We obtained multispectral images using the Sentera high-precision NDVI single sensor which was mounted on the UAV (Figure 2a). The sensor is 1.2 MP CMOS with a 60° horizontal FOV and a 47° vertical FOV and works with two wide spectral bands: red (625 nm CWL x 100 nm width) and NIR (850 nm CWL x 40 nm width) with a pixel count of 1248 horizontal / 950 vertical. The green band is typically unused. The sensor has a total weight of 30 g and size of 25.4 x 33.8 x 37.3 mm.

### Data Collection: Multispectral Images and Soil Samples

The farm used for data collection is located at Sturgis, South Dakota, USA (44° 25' 27"N; 103° 22' 34"W). We created an autonomous UAV flight plan for minimal passes similar to a raster scan pattern using the coordinates of the four corner of the filed (44.25.39 N, 103.22.60 W; 44.25.28 N, 103.22.60 W; 44.25.39 N, 103.23.16 W; 44.25.39 N, 103.23.16 W). We captured 865 multispectral images while the drone was following the raster scan pattern using the following parameters and experimental setup,

#### (i) Parameters:

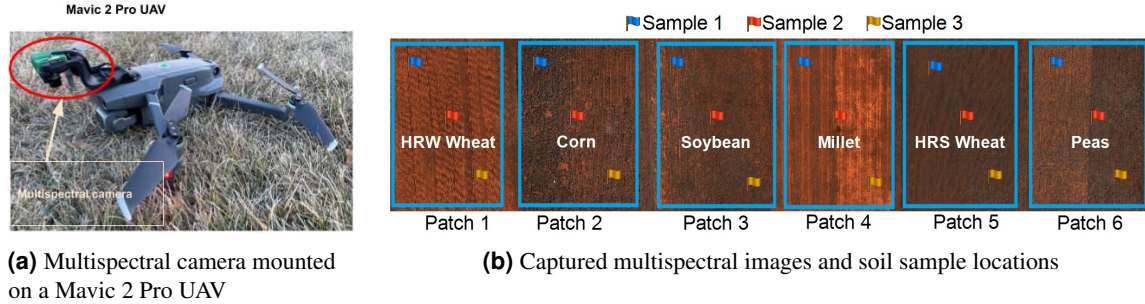
- Flight Type: *QuickTile*
- Overlap Setting: 75%
- Altitude: 200ft.
- Speed: 15mph

#### (ii) Experimental setup:

- Desired resolution: The Ground Sample Distance (GSD)/pixel of the multispectral camera was set to 2.1 inches for a 200 ft. altitude.



- Cloud cover and time of day: The drone was flown when the sun was highest in the sky for more accurate data. Data is best when sky conditions are consistent, ideally 100% sunny or 100% cloudy. Flying with a mix of sun and clouds causes inconsistency in brightness and contrast while stitching images. Therefore, the stitched image will provide an inaccurate NDVI value.

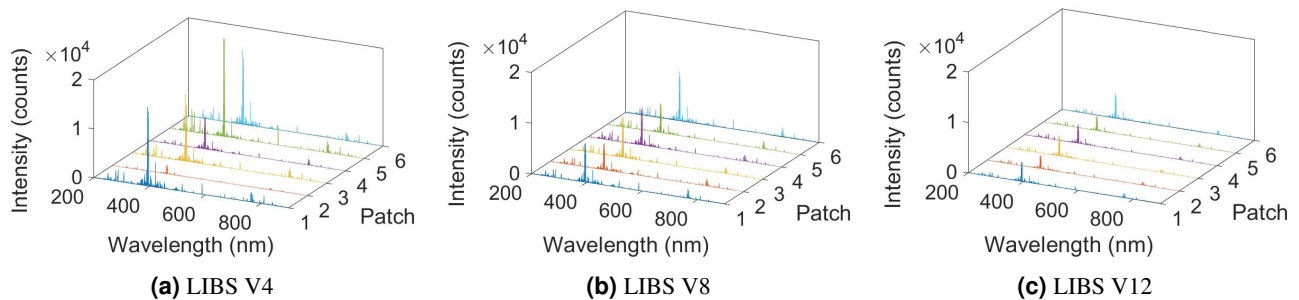


**Figure 2.** (a) Mavic 2 pro UAV with multispectral camera mounted. (b) The the flags shows the sample locations of the corresponding crops. The patches have crops including Peas, HRS Wheat, Millet, Soybean, Corn and HRW Wheat respectively.

Soil samples were collected at an 8-inch depth in each of the 6 patches at the V4, V8 and V12 stages using a hydraulic probe. We avoided sampling from the areas where conditions were different from the rest of the field (e.g., former manure piles, fertilizer bands, or fence lines). Figure 2b shows the sample locations across the patches.

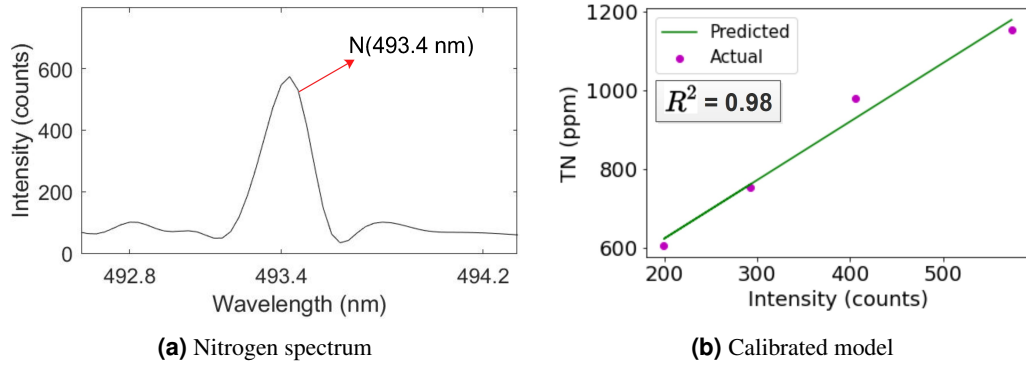
### Calibration

LIBS utilizes a high energy pulsed laser which generates high a temperature ranging from  $10^5$ -  $20,000^\circ\text{K}$  resulting in plasma formation when focused on a sample. This, in turn, leads to ablation of a minuscule amount of sample leading to excitation of the sample's constituent elements. As the plasma cools, these excited atoms and electrons emit photons which correspond to specific elements present in the sample. These photons are collected by a spectrometer and result in quantitative and qualitative analysis of samples. The SciAps Z-300 handheld LIBS analyzer was used for these measurements. This device has an extended spectrometer wavelength range from 190 nm to 950 nm. The extended range allows emission lines from elements H, F, N, O, Br, Cl, Rb and S to be measured. The LIBS instrument is equipped with a Q-switched Nd:YAG laser, 5-6 mJ per pulse at 1064 nm. Ten laser pulses are shot on the soil samples to obtain averaged data on each measurement. The focused laser on the soil surface forms a  $\mu\text{m}$  size of a sample into  $>10,000^\circ\text{K}$  plasma. The unique emission spectrum is collected by the spectrometer as the plasma cools. We used NIST database<sup>34</sup> to determine the N lines from the emission spectrum (Figure 3) and found N lines at 493.4 nm (Figure 4a). From the soil samples, we obtain actual TN in ppm and analyze the samples in LIBS to determine the maximum intensity from the N spectrum for all the samples (Figure 5) at V4, V8 and V12 stages.

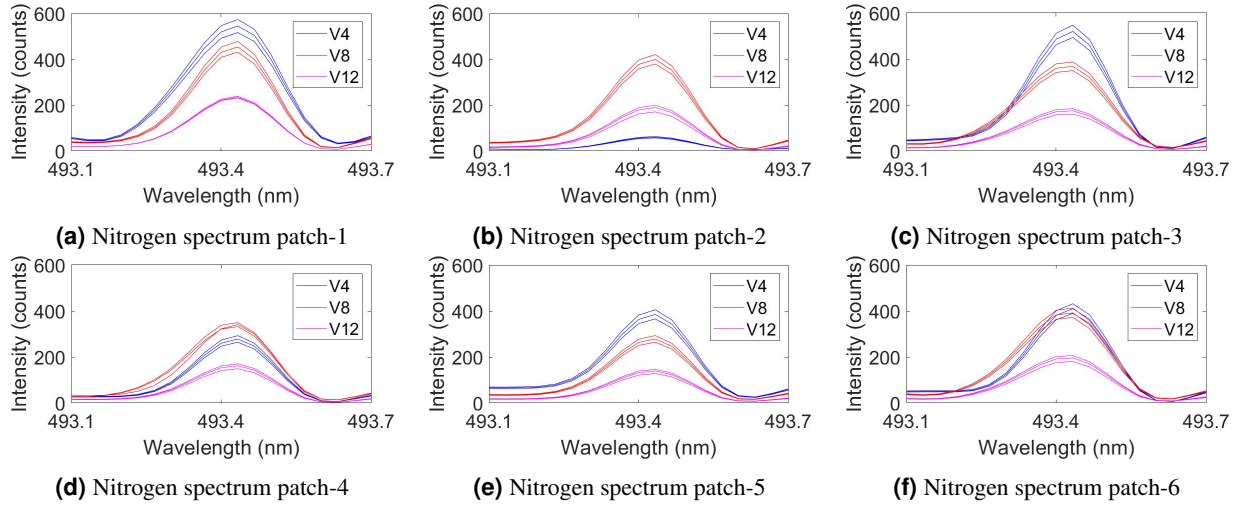


**Figure 3.** Emission lines of soil samples at the V4, V8, and V12 stages for six patches.

Using the correlation between actual TN and the maximum intensity of N spectrum, we generate a calibration plot through linear regression (Figure 4b). We use  $R^2$  as our calibration metric and find  $R^2 = 0.98$  showing a strong correlation between the actual soil TN and the peak intensity of the N spectrum. Using the calibrated model, we converted the peak intensity of the N spectrum to TN (ppm) for all the samples.



**Figure 4.** Calibration plot for computing soil TN using the peak intensity of the nitrogen spectrum at 493.4 nm



**Figure 5.** Nitrogen spectrum of the soil samples for six patches at the V4, V8 and V12 stages.

### Feature Extraction and Dataset

The multispectral images are composed of three channels, channel-1: R, channel-2: G and channel-3: NIR. The multispectral sensor's datasheet<sup>35</sup> shows that channel-1 contains both R and NIR light. Therefore, the NIR light needed to be removed to isolate R and compute NDVI. The equation for R and NIR light are,

$$R = 1.0 * DN_{ch1} - 1.012 * DN_{ch3} \quad (1)$$

$$NIR = 9.605 * DN_{ch3} - 0.6182 * DN_{ch1} \quad (2)$$

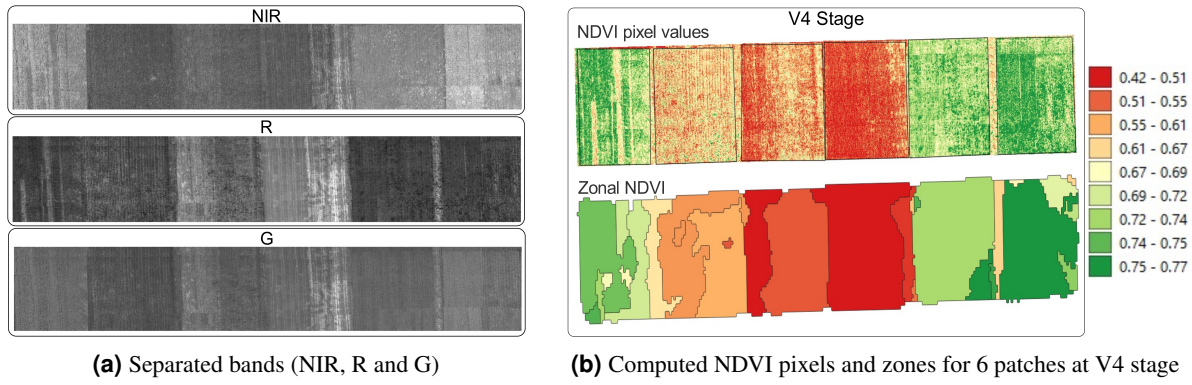
where  $DN_{ch1}$  is the Digital Number (pixel value) of channel one, and  $DN_{ch3}$  is the Digital Number (pixel value) of channel three. The coefficients of  $DN$  were provided in the datasheet<sup>35</sup>.

Using equations 1 and 2, band separation (Figure 6a) was performed to compute NDVI (Figure 6b) and extract the pixel values from each of the bands. The dataset was created using the mean NDVI and the mean pixel values of each of the bands from individual zones at the V4, V8 and V12 stages. The equation for computing NDVI,

$$NDVI = \frac{1.236 * DN_{ch3} - 0.188 * DN_{ch1}}{1.000 * DN_{ch3} + 0.044 * DN_{ch1}} \quad (3)$$

### Results and Discussion

To evaluate the HPO methods, we implemented 5-fold cross validation and used RMSE as the performance metric. Additionally, we measured CT as a model efficiency metric. CT is the total time required to complete an HPO process. The optimal



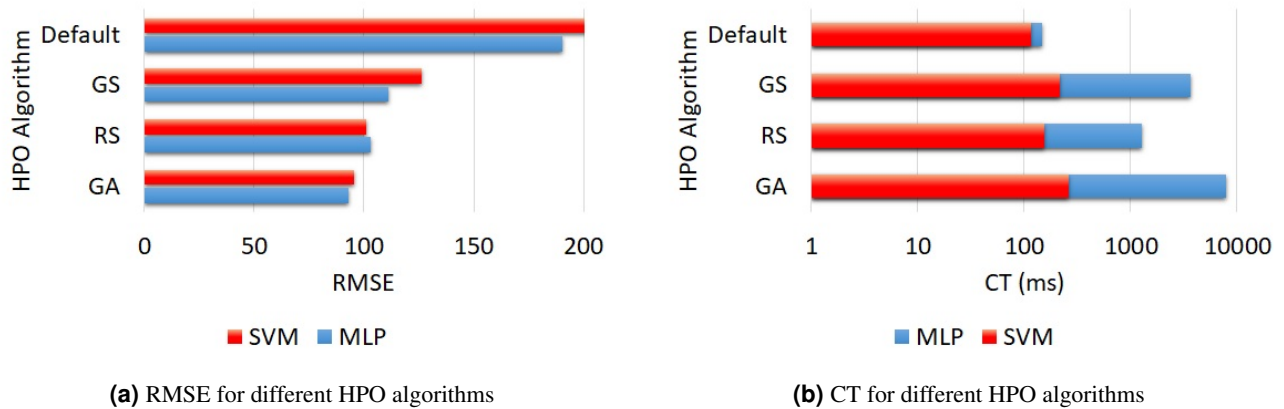
**Figure 6.** Band separation, and computed NDVI pixels and zonal NDVI

**Table 1.** Specifics of the configuration space for the hyper-parameters

ML model	Hyper-parameter	Type	Search Space
MLP Regressor	activation	Categorical	['relu', 'tanh', 'logistic', 'identity']
	solver	Categorical	['adam', 'lbfgs']
	learning_rate	Categorical	['constant', 'adaptive', 'invscaling']
	hidden_layer_sizes	Discrete	[50,150]
SVM Regressor	kernel	Categorical	['poly', 'rbf', 'sigmoid']
	C	Discrete	[1,10000]
	epsilon	Discrete	[0.0001,1]

hyper-parameter configuration was determined based on the lowest RMSE. We specify the same hyper-parameter configuration space (Table 1) for all HPO methods to fairly compare the GS, RS and GA HPO methods.

We tuned the models on a machine with an 8 Core i7-9700K processor and 16 gigabytes (GB) of memory. We used Python 3.5, multiple open-source Python libraries, and open-source Python frameworks, including *sklearn*<sup>27</sup>. Figure 7 shows that for both MLP and SVM, RS produces much faster results than GS while maintaining lower RMSE for the same search space size. In general, GA offers lower RMSE for both models but has a higher CT compared to GS and RS. Overall, MLP outperforms SVM in terms of performance. However, we achieved better efficiency with SVM in our dataset.



**Figure 7.** Performance comparison for different HPO algorithms

We introduced the machine learning approach to estimate the soil TN using NDVI and multispectral characteristics (R, NIR and G) of the images. We also consider the environmental factors such as air temperature and RH. The performance of MLP and SVM models were tested on a fixed configuration space for the hyper-parameters under various hyper-parameter optimization techniques. For both MLP and SVM, the default HP configuration do not yield the lowest RMSE, this demonstrates the significance of utilizing HPO. Table 2 shows that the estimation error of predicting soil TN is lowest in GA compared to GS

and RS for both MLP and SVM, where  $\mu$  is the mean and  $\sigma$  is the standard deviation. While training the models, we split our dataset into train and test, where we use 80% of the data for training and 20% for testing.

**Table 2.** Estimation error for predicting soil TN

ML model	HPO method	Error ( $\mu \pm \sigma$ )
MLP	Default	141.53 $\pm$ 84.52
	GS	76.37 $\pm$ 61.51
	RS	65.21 $\pm$ 59.91
	GA	51.11 $\pm$ 43.33
SVM	Default	171.14 $\pm$ 99.96
	GS	80.99 $\pm$ 53.78
	RS	48.47 $\pm$ 49.38
	GA	44.46 $\pm$ 33.60

## Conclusions

In this paper, we have demonstrated the ability of a UAV-based multispectral sensing solution to estimate soil total nitrogen. Specifically, we implemented two machine learning models *multilayer perceptron* and *support vector machine* to predict soil total nitrogen using a suite of data classes including UAV-based imaging data in red, near infrared, and green spectral bands, normalized difference vegetation indices (computed using the multispectral images), air temperature, and relative humidity. We performed hyperparameter optimization methods to tune the models for prediction performance. Overall, our numerical studies confirm that our machine learning-based predictive models can estimate total nitrogen of the soil with a root mean square percent error (RMSPE) of 10.8%.

## References

1. Fageria, N. & Baligar, V. Enhancing nitrogen use efficiency in crop plants. *Adv. agronomy* **88**, 97–185 (2005).
2. Bausch, W. C. & Duke, H. Remote sensing of plant nitrogen status in corn. *Transactions ASAE* **39**, 1869–1875 (1996).
3. Khan, S., Mulvaney, R. L. & Hoelt, R. A simple soil test for detecting sites that are nonresponsive to nitrogen fertilization. *Soil Sci. Soc. Am. J.* **65**, 1751–1760 (2001).
4. Lloveras Vilamanyà, J., Cabasés i Piqué, M. *et al.* Costes de producció de cultivos extensivos en secano y regadio. *Vida Rural*. 2015, núm. 401, p. 38-47 (2015).
5. Bagheri, N., Ahmadi, H., Alavipanah, S. K. & Omid, M. Multispectral remote sensing for site-specific nitrogen fertilizer management. *Pesquisa Agropecuária Brasileira* **48**, 1394–1401 (2013).
6. Bausch, W. & Khosla, R. Quickbird satellite versus ground-based multi-spectral data for estimating nitrogen status of irrigated maize. *Precis. Agric.* **11**, 274–290 (2010).
7. Bausch, W. C. & Duke, H. Remote sensing of plant nitrogen status in corn. *Transactions ASAE* **39**, 1869–1875 (1996).
8. Daughtry, C., Walthall, C., Kim, M., De Colstoun, E. B. & McMurtrey Iii, J. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote. sensing Environ.* **74**, 229–239 (2000).
9. ZHANG, D.-y., HUANG, W.-j. *et al.* A field-based pushbroom imaging spectrometer for estimating chlorophyll content of maize. *Spectrosc. Spectr. Analysis* **31**, 771–775 (2011).
10. Zarco-Tejada, P. J., Catalina, A., González, M. & Martín, P. Relationships between net photosynthesis and steady-state chlorophyll fluorescence retrieved from airborne hyperspectral imagery. *Remote. Sens. Environ.* **136**, 247–258 (2013).
11. Sripada, R. P., Heiniger, R. W., White, J. G. & Meijer, A. D. Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agron. J.* **98**, 968–977 (2006).
12. Ma, B.-L., Wu, T.-Y. & Shang, J. On-farm comparison of variable rates of nitrogen with uniform application to maize on canopy reflectance, soil nitrate, and grain yield. *J. Plant Nutr. Soil Sci.* **177**, 216–226 (2014).
13. Jones, J. *et al.* Influence of soil, crop residue, and sensor orientations on ndvi readings. *Precis. agriculture* **16**, 690–704 (2015).



14. Hunt, E. R., Cavigelli, M., Daughtry, C. S., McMurtrey, J. E. & Walthall, C. L. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* **6**, 359–378 (2005).
15. Robert, P. C. Precision agriculture: a challenge for crop nutrition management. In *Progress in Plant Nutrition: Plenary Lectures of the XIV International Plant Nutrition Colloquium*, 143–149 (Springer, 2002).
16. Strachan, I. B., Pattey, E. & Boisvert, J. B. Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. *Remote. sensing Environ.* **80**, 213–224 (2002).
17. Quemada, M., Gabriel, J. L. & Zarco-Tejada, P. Airborne hyperspectral images and ground-level optical sensors as assessment tools for maize nitrogen fertilization. *Remote. sensing* **6**, 2940–2962 (2014).
18. Papadopoulos, A. *et al.* Preliminary results for standardization of ndvi using soil nitrates in corn growing. *Fresen. Environ. Bull* **23**, 348–354 (2014).
19. DeCastro-García, N., Muñoz Castañeda, Á. L., Escudero García, D. & Carriegos, M. V. Effect of the sampling of a dataset in the hyperparameter optimization phase over the efficiency of a machine learning algorithm. *Complexity* **2019** (2019).
20. Yang, L. & Shami, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing* **415**, 295–316 (2020).
21. Erler, A., Riebe, D., Beitz, T., Löhmansröben, H.-G. & Gebbers, R. Soil nutrient detection for precision agriculture using handheld laser-induced breakdown spectroscopy (libs) and multivariate regression methods (plsr, lasso and gpr). *Sensors* **20**, 418 (2020).
22. Harris, R. D., Cremers, D. A., Ebinger, M. H. & Bluhm, B. K. Determination of nitrogen in sand using laser-induced breakdown spectroscopy. *Appl. spectroscopy* **58**, 770–775 (2004).
23. Tran, M., Sun, Q., Smith, B. W. & Winefordner, J. D. Determination of c: H: O: N ratios in solid organic compounds by laser-induced plasma spectroscopy. *J. Anal. At. Spectrom.* **16**, 628–632 (2001).
24. Yu, K., Ren, J. & Zhao, Y. Principles, developments and applications of laser-induced breakdown spectroscopy in agriculture: A review. *Artif. Intell. Agric.* **4**, 127 – 139, DOI: <https://doi.org/10.1016/j.aiia.2020.07.001> (2020).
25. Gardner, M. W. & Dorling, S. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment* **32**, 2627–2636 (1998).
26. Fine, T. L. *Feedforward neural network methodology* (Springer Science & Business Media, 2006).
27. Pedregosa, F. *et al.* Scikit-learn: Machine learning in python. *J. machine Learn. research* **12**, 2825–2830 (2011).
28. Yang, L., Muresan, R., Al-Dweik, A. & Hadjileontiadis, L. J. Image-based visibility estimation algorithm for intelligent transportation systems. *IEEE Access* **6**, 76728–76740 (2018).
29. Soliman, O. S. & Mahmoud, A. S. A classification system for remote sensing satellite images using support vector machine with non-linear kernel functions. In *2012 8th International Conference on Informatics and Systems (INFOS)*, BIO–181 (IEEE, 2012).
30. Zhang, J., Jin, R., Yang, Y. & Hauptmann, A. Modified logistic regression: An approximation to svm and its applications in large-scale text categorization. In *Proceedings of the Twentieth International Conference on Machine Learning (ICML, 2003)*.
31. Hutter, F., Kotthoff, L. & Vanschoren, J. *Automated machine learning: methods, systems, challenges* (Springer Nature, 2019).
32. Lorenzo, P., Nalepa, J., Kawulok, M., Ramos, L. & Ranilla, J. Particle swarm optimization for hyper-parameter selection in deep neural networks. *Proc. Genet. Evol. Comput. Conf.* (2017).
33. Gogna, A. & Tayal, A. Metaheuristics: review and application. *J. Exp. & Theor. Artif. Intell.* **25**, 503–526 (2013).
34. Kramida, A. & Ralchenko, Y. NIST Atomic Spectra Database, NIST Standard Reference Database 78, DOI: [10.18434/T4W30F](https://doi.org/10.18434/T4W30F) (1999). Type: dataset.
35. False color to ndvi conversion precision ndvi single sensor. *Sentera, LLC*.

## Acknowledgement

This work was supported in part by South Dakota GOED i6 program through the Proof of Concept grant. We thank Christopher Graham for support in data collection.

## Author contributions statement

M.A.H. conducted the data collection, experiments and prepared the results. The LIBS experiments were conducted by P.K.D. and M.A.H. Analysis was done by all authors. The manuscript was prepared by M.A.H. and S.R. S. R. served as the principal investigator for this project. All authors reviewed the manuscript.

## Source code

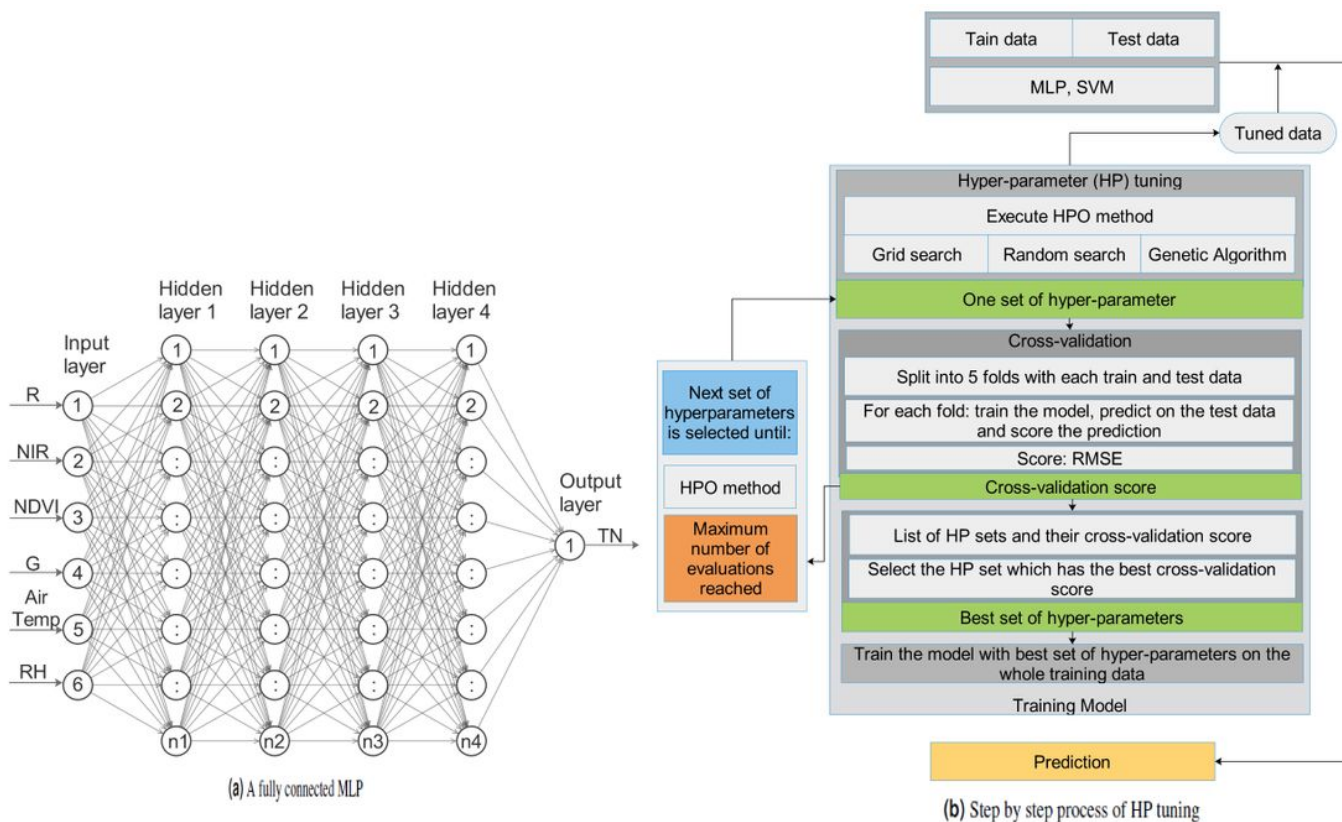
The source code, and the training data are available at <https://github.com/abirhossen786/Nitrogen-prediction-using-multispectral-images>.

## Additional information

**Competing interests:** The authors declare that they have no competing interests.



# Figures

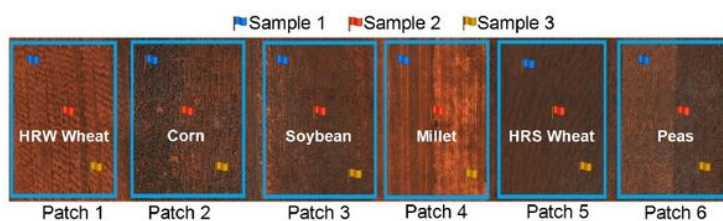


**Figure 1**

(a) HPO for GS, RS and GA with cross-validation, training the ML models with the tuned HP and prediction. (b) MLP with four hidden network having different weights, where input layer 2 R6, hidden layer 2 R4, output layer 2 R1 and n1, n2, n3, and n4 represent the number of perceptronw in each hidden layer, respectively



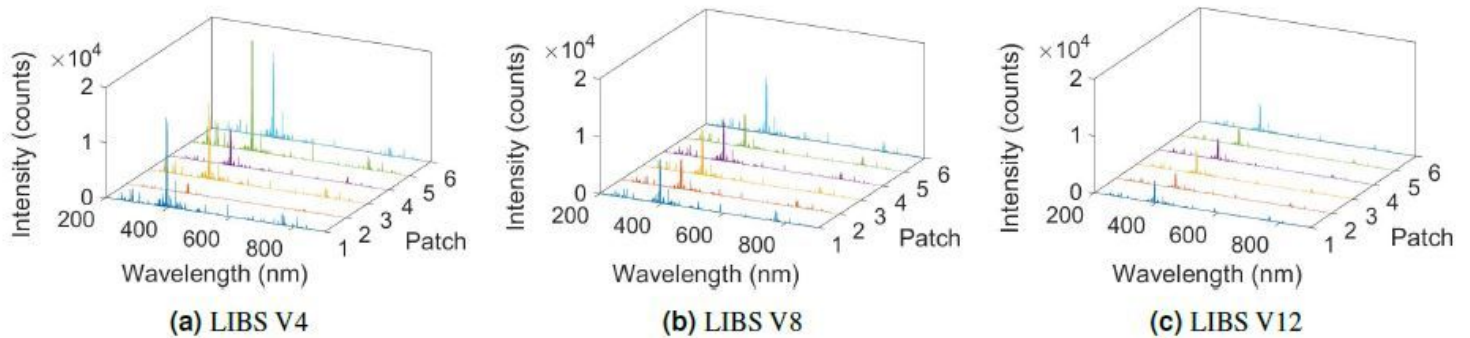
**(a)** Multispectral camera mounted on a Mavic 2 Pro UAV



**(b)** Captured multispectral images and soil sample locations

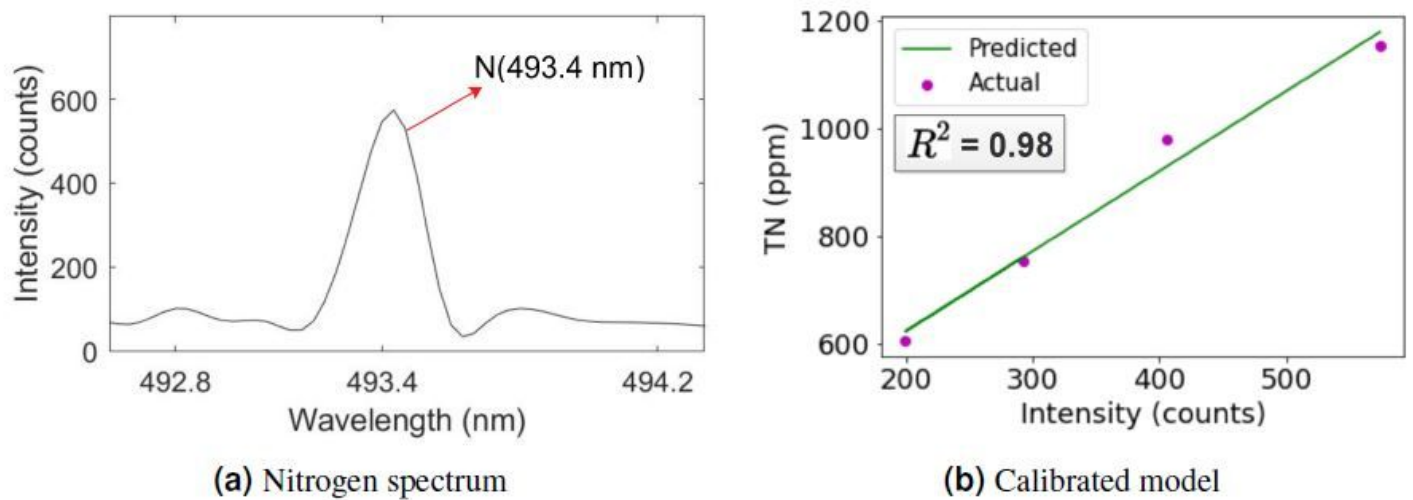
**Figure 2**

(a) Mavic 2 pro UAV with multispectral camera mounted. (b) The the flags shows the sample locations of the corresponding crops. The patches have crops including Peas, HRS Wheat, Millet, Soybean, Corn and HRW Wheat respectively.



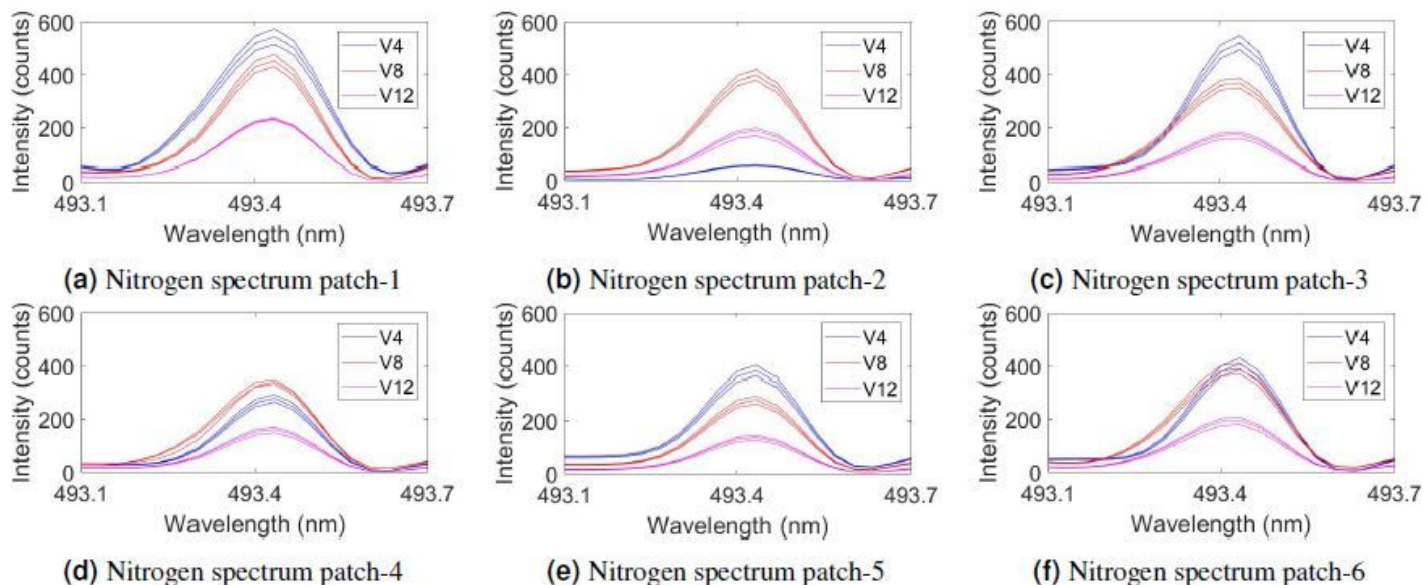
**Figure 3**

Emission lines of soil samples at the V4, V8, and V12 stages for six patches.



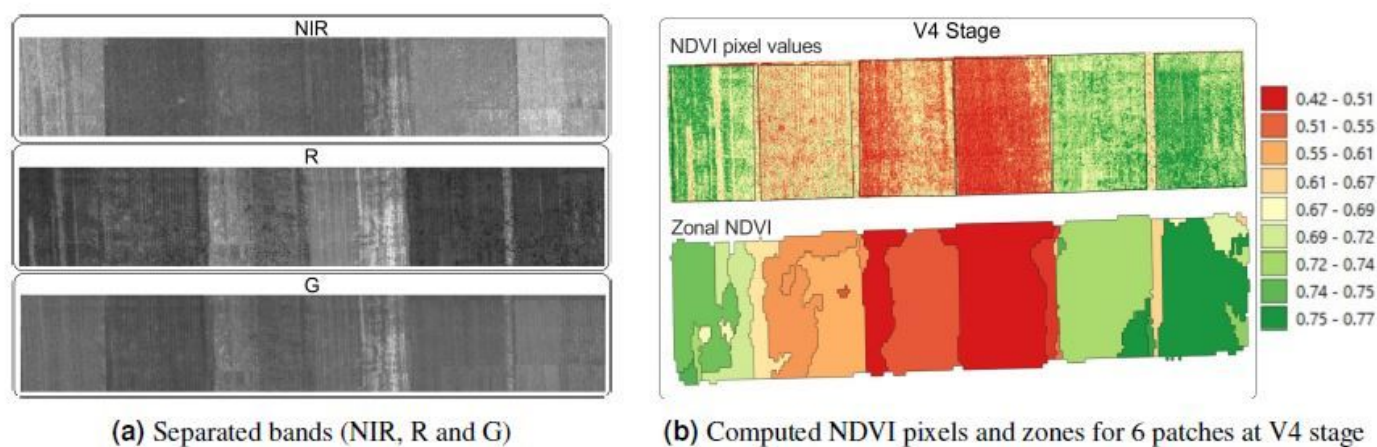
**Figure 4**

Calibration plot for computing soil TN using the peak intensity of the nitrogen spectrum at 493.4 nm



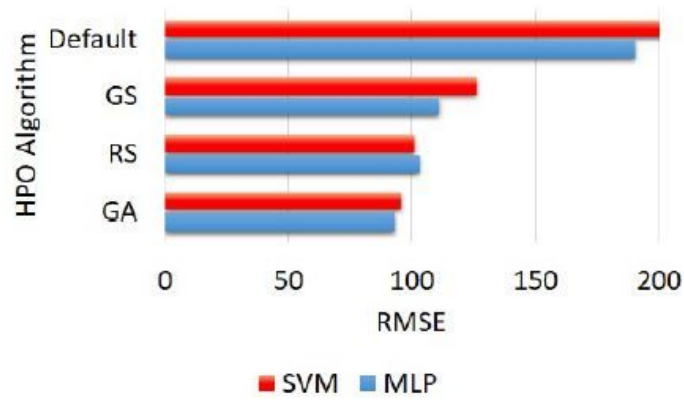
**Figure 5**

Nitrogen spectrum of the soil samples for six patches at the V4, V8 and V12 stages.

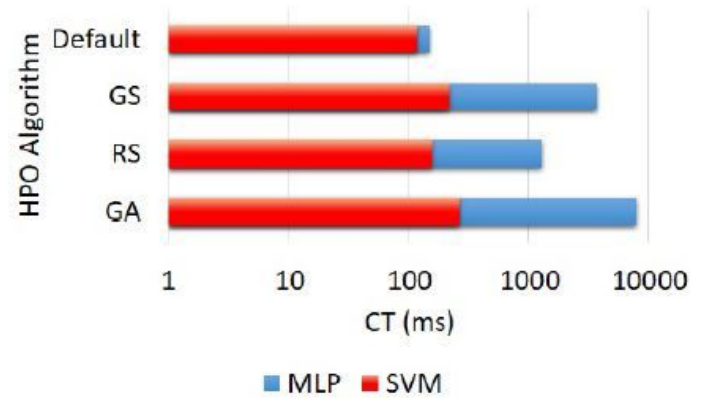


**Figure 6**

Band separation, and computed NDVI pixels and zonal NDVI



(a) RMSE for different HPO algorithms



(b) CT for different HPO algorithms

Figure 7

Performance comparison for different HPO algorithms