Diagnosis of Grapevine Nitrogen Content Using Proximal Hyperspectral Imaging

Chenchen Kang[a], Nataliya Shcherbatyuk[b], Pierre Davadant[b], Manoj Karkee[a], Qin Zhang[a], Markus Keller[b]

a Center for Precision and Automated Agricultural Systems, Department of Biological Systems Engineering, Washington State University, Prosser, WA 99350, USA

b Department of Horticulture, Washington State University, Prosser, WA 99350, USA

A paper from the Proceedings of the

15th International Conference on Precision Agriculture

June 26-29, 2022

Minneapolis, Minnesota, United States

Abstract.

The nitrogen (N) content of grapevines could affect the grape yield and quality, which is a major concern in vineyards. Precise N fertilization plays an important role in the search for a balance between optimizing vigor and grape composition, controlling production costs and limiting pollution The tissue sampling in the laboratory is the way to get an accurate result, but it is destructive to plants, time and cost-intensive. The differences in leaf nutrient levels also alter spectral characteristics outside of the electromagnetic spectrum visible range, even before visible deficiency symptoms become apparent. The previous studies regarding remote sensing in vineyards are restricted to the UAVs, which capture an average reflectance observed across a vineyard block. These approaches do not reflect spatial variability within and between canopies. In this study, hyperspectral images acquired from a ground-based on-the-go utility vehicle were analyzed to evaluate the applicability of hyperspectral data in predicting the N content. One N trial was conducted in a Syrah vineyard with four treatment levels. During veraison, hyperspectral image data (400-1000nm) was collected in both sun-exposed and diffused light conditions. The leaf samples were collected the same day from the same vines, then N content was analyzed in the laboratory. Partial Least Squares (PLS) were applied to build the predictive regression models. Simulated Annealing (SA) was applied to select the key wavelengths related to N content. The quality of these models was hugely improved by using SA to select optimized bands for model establishment. When comparing the model performance in direct sunlight and diffused illumination conditions, both models have high accuracy and low RMSE. For both light conditions, the same spectrum regions are found in the red edge and NIR.

Keywords. Nitrogen content, hyperspectral imaging, salient wavelengths, grapevine

1. Introduction

Nitrogen (N) is a major and essential nutrient for plant growth, development and reproduction. Excess N could contribute to the contamination of environments, causing health risks for humans and other organisms. In viticulture, while the N fertilization alters the vine vigor and grape yield, excessive shoot growth and canopy density negatively impact berry pigments and sugar content, decreasing wine quality (Chaves et al., 2010). As a result, precise N fertilization plays an important role in the search for a balance between optimizing vigor and grape composition, controlling production costs and limiting pollution(Verdenal et al., 2021). The precision management of nitrogen in vineyards relies on monitoring of N along the growing season, which directs agronomic practices to adjust the fertilization to meet production objectives.

The current N assessment in vineyards is based on tissue sampling, in which the leaf samples are acquired and sent to the laboratory for analysis. These techniques are destructive and time-consuming. Waiting for results could lead to the delay of N fertilization when the N deficiency is found. Indirect remote and proximal sensing techniques such as spectral sensing provide opportunities for non-destructive, rapid and reliable assessment of crop nutrient status. The difference in leaf physiological characteristics would alter the spectral signatures when the plants interact with incident light. It is found that the reflectance in the visible light spectrum (400-700 nm) is determined mainly by the leaf pigments; while the near-infrared (NIR, 700-1000 nm) spectrum is mainly influenced by leaf structure (Guo, Yang, Guan, & He, 2017; Slaton, Raymond Hunt Jr, & Smith, 2001). Therefore, some studies focused on the visible light and NIR (VNIR) spectrum to retrieve leaf N content for different crops and achieved good results. Osco et al. (2020) used a spectroradiometer to measure the spectral reflectance (380 – 1020 nm) of the orange leaf; several machine learning algorithms were developed to predict the macro-and micronutrient content; the prediction for N could achieve high accuracy (R^2 = 0.912). While a spectroradiometer/spectrometer can only measure spectral signatures from individual points, the hyperspectral imaging (HSI) sensors can also capture the 2D distribution of the spectral signatures of objects. However, most of the studies using hyperspectral camera in the field condition obtained spectral data from airborn platforms. Wang et al. (2021) used visible to shortwave infrared (400 – 2400 nm) airborne hyperspectral imaging on a maize field; the built machine learning model could exhibit good accuracy (R^2 = 0.80) to predict nitrogen content. These approaches do not reflect spatial variability within and among canopies. The ground-based hyperspectral imaging applications are not robust and reliable enough in field conditions due to the changing solar illumination, clouds and shadows, and leaf inclination angles(Liu, Bruning, Garnett, & Berger, 2020). These factors can significantly affect the signals received by optical sensors and make interpreting plant parameters more complicated and unreliable. More robust techniques are needed to overcome these obstacles when using VNIR HSI for grapevine N and other nutrient content assessments.

The overall objective of this study is to assess the use of proximal HSI sensing approaches for retrieving grapevine leaf N content from a ground-based VNIR hyperspectral camera. This study was conducted at a vineyard with various N application rates. The specific objectives of this study are (1) to build predictive regression models using optimized combinations of VNIR wavelengths; (2) to compare the results of the datasets collected in different light conditions (direct sunlight and diffused light).

2. Materials and Methods

2.1 Study Area

The experimental study area was a Syrah grape vineyard block of the Ste. Michelle Wine Estates in Paterson, WA, USA. The field trial was built in the spring of 2021 with four different treatments of fertilizer rate: 0, 22.42, 44.83 and 89.67 kg/ha. The fertilizer was applied by fertigation (UAN32) in three split applications between the 6-leaf stage and fruit set.

Figure 1. Google Earth view of the experimental Syrah vineyard plot.

2.2 Sampling plan

Leaf-blades were collected at 50% of veraison from 30 randomly selected Syrah grapevines (Vitis vinifera). For leaf tissue collection, 50 youngest fully matured leaves on the shoot, approximately 5th – 6th leaf from the shoot tip were sampled. Tissue samples were washed and dried at 60 °C for 72 hours, ground, and sent to the commercial testing lab for macro-and micronutrient analysis. Total N (TN) was analyzed by dry combustion (LECO CN628, Dumas' method).

2.3 Imaging Campaign

A VNIR hyperspectral camera (274 spectral bands from 400 to 1000 nm)(Nano-Hyperspec®, Headwall Photonics, Bolton, MA, USA) mounted on a ground-based sensing platform was used to capture the image data. A reflectance panel was placed next to each object as the simultaneous white reference (shown in Figure 2). The images in sun-exposed conditions were taken first; then, a localized diffused illumination environment was set by an adjustable tarp roof, followed by capturing images in such conditions.

(a) (b)

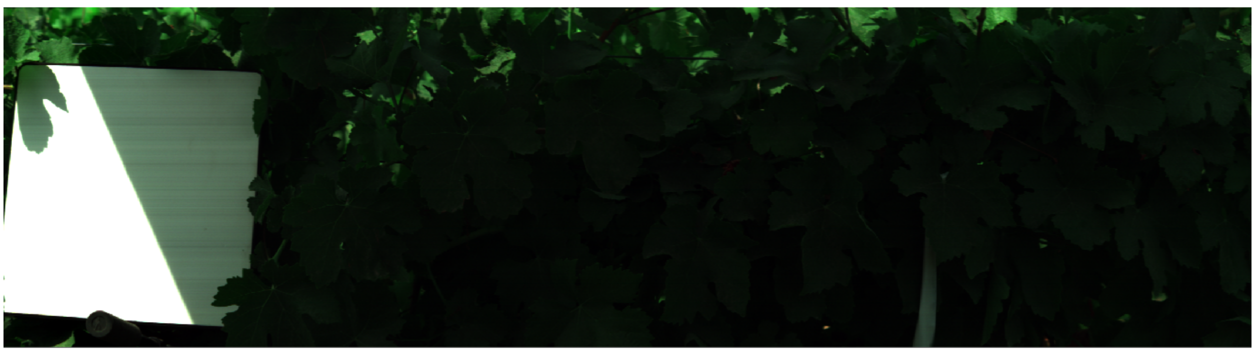
Figure 2. Hyperspectral image data acquisition in the vineyard. (a) Canopy in direct sunlight condition. (b) Canopy in a localized diffused illumination environment.

2.4 Data Processing

The data processing included reflectance correction, region-of-interest (ROI) creation and spectral extraction.

The raw image data was calibrated with dark and white references. The dark reference was captured when the lens was covered with its original covering. The unique white reference for each individual image was obtained from the reflectance panel in the view of each image. This process removes both the solar irradiance and atmospheric path radiance, which was done in SpectralView-HypersecIII version 3.1.5.1 (Headwall Photonics, Bolton, MA, USA).

(a)



(b)

Figure 3. Hyperspectral image data visualization in RGB. (a) Canopy in direct sunlight condition. (b) Canopy in a localized diffused illumination environment.

The ROI creation and spectral extraction were conducted in Python 3.9. Ten leaf blades were randomly selected using a 60 × 60 pixel bounding box as shown in Figure 4. The mean spectra of each leaf blade were extracted. The spectra of 10 ROIs were averaged to denote the mean spectral reflectance of the vine. Finally, the Multiplicative Scatter Correction and First Derivative Smoothing were used to smooth the spectra data.

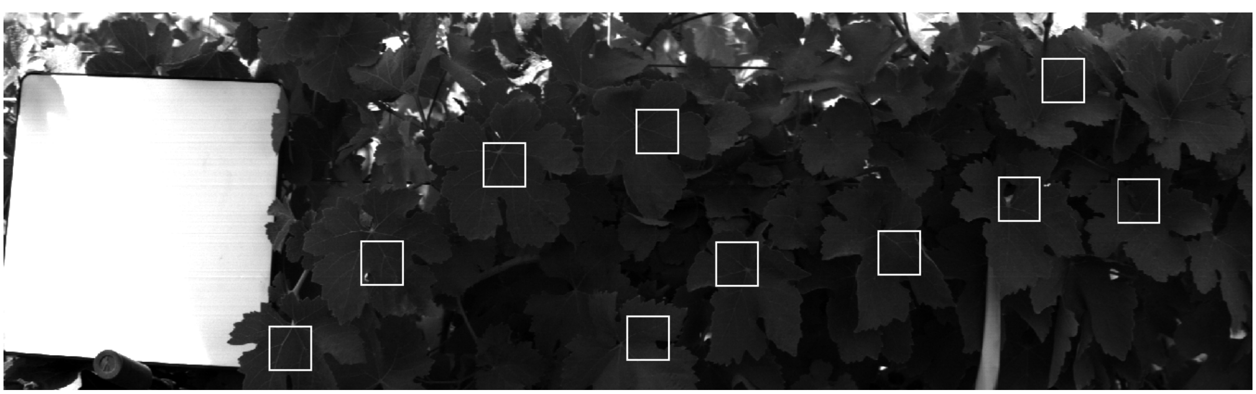


Figure 4. Ten leaf blades were randomly selected using a 60 × 60 pixel bounding box as ROIs

2.6 Data Analysis

2.6.1 Basic PLS models

The basic predictive regression models were firstly performed using Partial Least Squares (PLS). PLS could explore the linear relationship between the independent and dependent variables, especially when the number of independent variables is greater than the number of dependent variables. Cross-validation was performed to test the built models. Root mean square errors (RMSE) and the determination coefficients (R^2) were used to evaluate modeling performances.

2.6.2 Optimizing the bands via Simulated Annealing (SA)

Compared to the specific regression models, selecting the wavelength bands that better correlate with the grapevine N content is more practical for developing robust sensing solutions for near-real-time nutrient assessment in vineyards. The SA was used in this study to perform a random search in the space of possible band combinations.

A cost function was defined: the root mean square error in cross-validation (RMSECV). Unlike greedy algorithms which seek to monotonically decrease the cost function, SA is an optimization process that can occasionally allow an increase of the cost function. 20 bands were randomly picked to start the optimization process; a PLS model was fit, and RMSECV was calculated. In the iterative procedure, three bands were swapped at random at each step, and the RMSECV of the optimized PLS was extracted. If the new value of the RMSECV is lower than the older one, keep the swapped bands. Otherwise, keep it only in accord with some small probability, driven by the cooling parameter. After repeating a sufficient number of iterations, a PLS model was built with the optimized number of components.

3. Results and Discussion

3.1 Data Visualization

3.1.1 Response variables: TN

The value distribution of TN content is shown in Figure 5. TN values range from 2.3% to 3.1%. The majority of values are located between 2.6% to 2.9%.

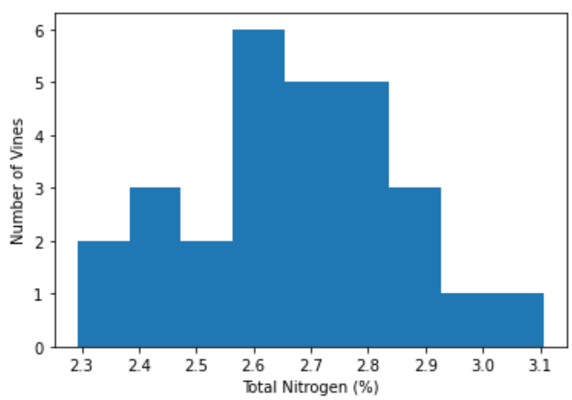
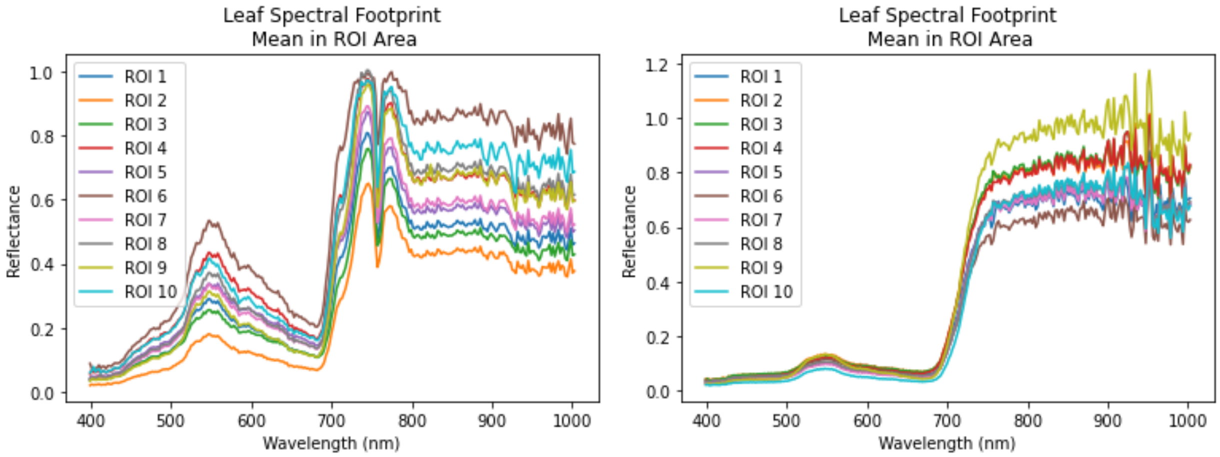


Figure 5. Histogram of response variable values of total N content (%).

3.1.2 Spectra Data

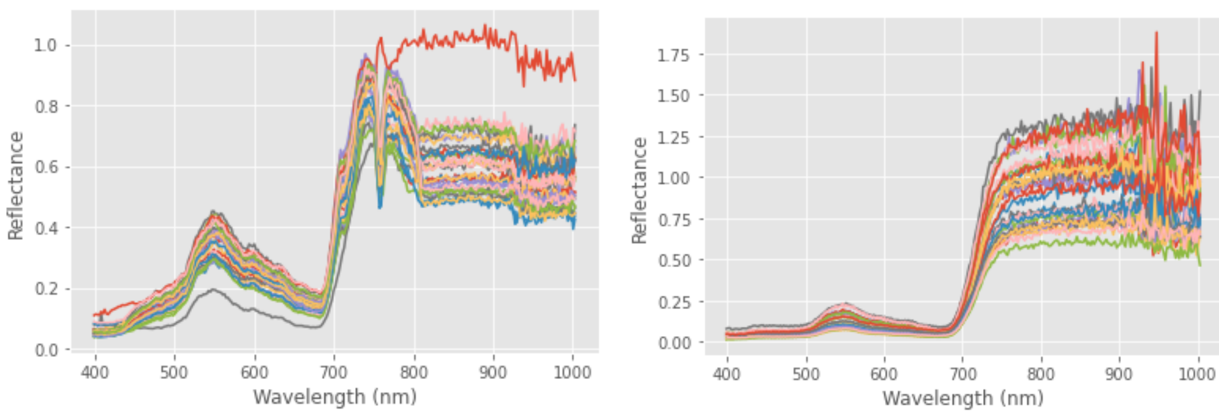
Figure 6 shows an example of the mean spectra by each ROI from one vine. In terms of scale, the reflectance of each leaf blade differs from one to another in direct sunlight conditions. However, it is very close to each other in the diffused light conditions. It is because each leaf gets different incident light due to varying leaf inclination angles in direct sunlight conditions, while the light conditions are more uniform in the shadow.



(a) **(b)**

Figure 6. Spectra data, mean spectra per leaf blade in (a) direct sunlight and (b) diffused light conditions.

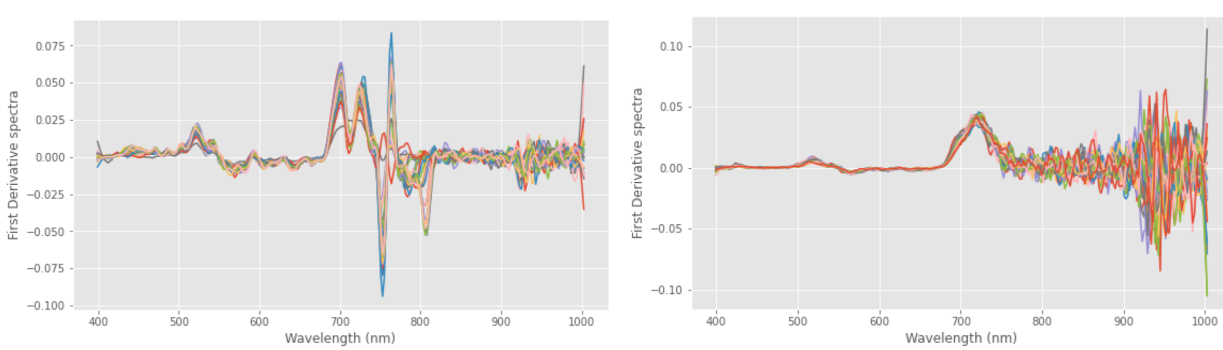
Figure 7 shows the average spectra obtained from 10 ROIs by each vine. In those direct sunlight conditions, spectral reflectance differs from vine to vine in regions of green (500nm), red edge (700-800nm) and NIR (800-1000nm). The spectra obtained in diffused light conditions only show differences when the wavelengths ≥ 750 nm and fewer signatures in the red edge area.



(a) **(b)**

Figure 7. Spectra data, mean spectra per vine in (a) direct sunlight and (b) diffused light conditions.

Figure 8 shows the first derivative spectra data after the smoothing process. The offset is gone, and the data look more bunched together.



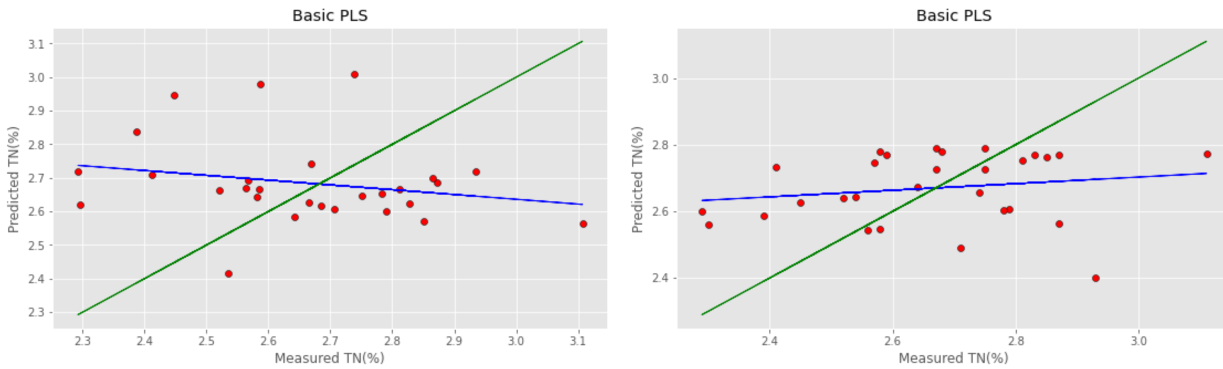
(a) **(b)**

Figure 8. First derivative spectra data, mean spectra per vine in (a) direct sunlight and (b) diffused light conditions.

3.2 Model Evaluation

3.2.1 Basic PLS Models

Figure 9 shows the cross-validation results of basic PLS regression models for TN prediction. Both models show a bad performance with low R^2 (-0.732 and -0.102, respectively) and high RMSE (0.251 and 0.197, respectively). These models built by the basic PLS method cannot distinguish the TN content with acceptable accuracy.

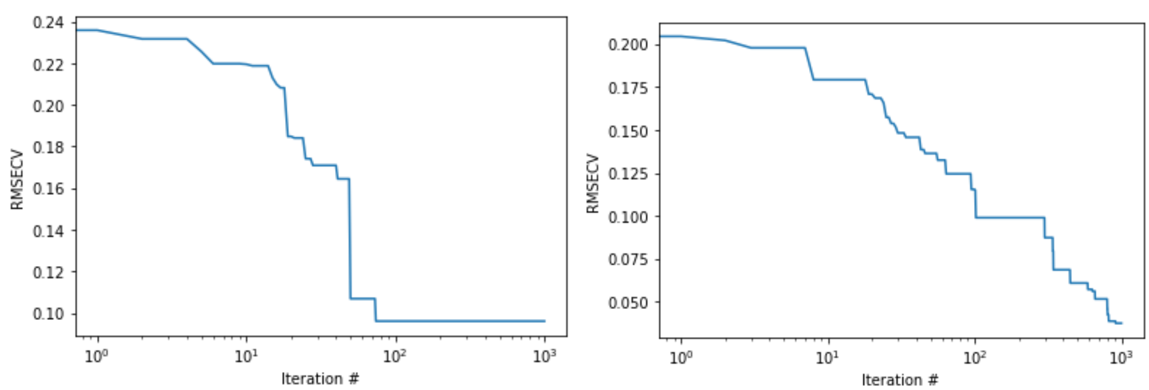


(a) (b)

Figure 9. Model evaluation via cross-validation for the prediction of TN (%) in (a) sun-exposed and (b) diffused light conditions.

3.2.2 Models with Optimized Bands Selected by SA

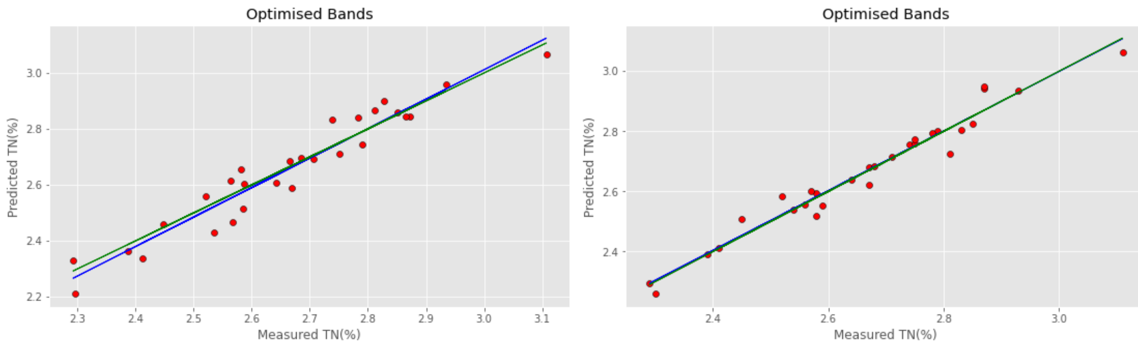
The RMSECV curves for the optimized bands' selection during the cross-validation progress are shown in Figure 10. In those direct sunlight conditions, the RMSE was down to 0.056 after 788 iterations. The RMSE for the model developed with the dataset in the diffused light was down to 0.038 after 1000 iterations.



(a) (b)

Figure 10. RMSECV curves during the cross-validation progress in (a) direct sunlight and (b) diffused light conditions.

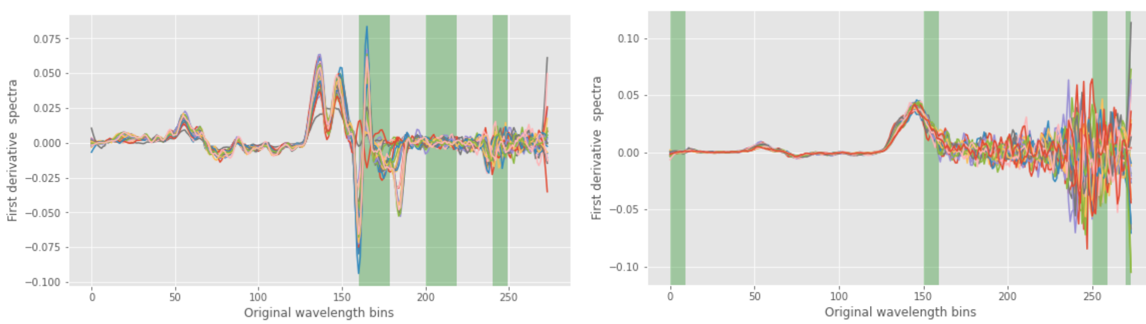
The cross-validation results of the optimized PLS regression models for TN prediction are shown in Figure 11. R^2 of both models is 0.914 and 0.960, respectively. This criterion is improved hugely compared to the models with the basic PLS method. This observation indicates the significant correlation between TN content and VNIR characteristics.



(a) **(b)**

Figure 11. Optimized model evaluation via cross-validation for the prediction of TN (%) in (a) direct sunlight and (b) diffused light conditions.

Figure 12 shows the selected wavelengths for both optimized models in relation to the first derivative spectra data. A number of observations can be made based on Figure 12. In those direct sunlight conditions, the salient bands are found in three regions, 700-780 nm, 840-880 nm and 920-950 nm, which correspond to the red edge, NIR and NIR, respectively. For the spectra data obtained in the diffused light, the salient bands are found not only in the red edge and NIR regions but also in one region in the visible spectrum (400 – 420 nm).



(a) **(b)**

Figure 12. Salient bands selection for the prediction of TN (%) in (a) direct sunlight and (b) diffused light conditions.

4. Conclusion

This article proposed a study employing proximal HSI sensing approaches to retrieve grapevine leaf N content from a ground-based VNIR hyperspectral camera. Two spectra datasets captured in direct sunlight and diffused light conditions were used to build PLS prediction models. The quality of these models was hugely improved by using simulated annealing to select optimized bands for model establishment.

When comparing the model performance in direct sunlight and diffused illumination conditions, both models have high accuracy and low RMSE. Different salient bands responding to the N levels were found in these two light conditions. However, the same regions are found in the red edge and NIR spectrum, which indicates that red edge and NIR spectrum may be more informative for quantifying grapevine N content during veraison.

Acknowledgments

This research was supported by the U.S. Department of Agriculture National Institute of Food and Agriculture's Specialty Crop Research Initiative Coordinated Agricultural Projects (CAP) grant (2020-51181-32159). We thank Patrick Scharf and Alan Kawakami for their skilled technical assistance. Chenchen Kang would like to thank China Scholarship Council (CSC) for sponsoring his study at Washington State University.

References

Chaves, M. M., Zarrouk, O., Francisco, R., Costa, J. M., Santos, T., Regalado, A. P., . . . Lopes, C. M. (2010). Grapevine under deficit irrigation: hints from physiological and molecular data. *Ann Bot, 105*(5), 661-676. doi:10.1093/aob/mcq030

Guo, J.-T., Yang, D.-C., Guan, Z., & He, Y.-H. (2017). Chlorophyll-catalyzed visible-light-mediated synthesis of tetrahydroquinolines from N, N-dimethylanilines and maleimides. *The Journal of organic chemistry, 82*(4), 1888-1894.

Liu, H., Bruning, B., Garnett, T., & Berger, B. (2020). Hyperspectral imaging and 3D technologies for plant phenotyping: From satellite to close-range sensing. *Computers and Electronics in Agriculture, 175*. doi:10.1016/j.compag.2020.105621

Osco, L. P., Ramos, A. P. M., Faita Pinheiro, M. M., Moriya, É. A. S., Imai, N. N., Estrabis, N., . . . Eduardo Creste, J. (2020). A Machine Learning Framework to Predict Nutrient Content in Valencia-Orange Leaf Hyperspectral Measurements. *Remote Sensing, 12*(6). doi:10.3390/rs12060906

Slaton, M. R., Raymond Hunt Jr, E., & Smith, W. K. (2001). Estimating near-infrared leaf reflectance from leaf structural characteristics. *American journal of botany, 88*(2), 278-284.

Verdenal, T., Dienes-Nagy, Á., Spangenberg, J. E., Zufferey, V., Spring, J.-L., Viret, O., . . . Van Leeuwen, C. (2021). Understanding and managing nitrogen nutrition in grapevine: a review. *OENO One, 55*(1), 1-43. doi:10.20870/oeno-one.2021.55.1.3866

Wang, S., Guan, K., Wang, Z., Ainsworth, E. A., Zheng, T., Townsend, P. A., . . . Jiang, C. (2021). Airborne hyperspectral imaging of nitrogen deficiency on crop traits and yield of maize by machine learning and radiative transfer modeling. *International Journal of Applied Earth Observation and Geoinformation, 105*. doi:10.1016/j.jag.2021.102617