

# Optimizing Nutrient Detection in Hydroponic Solutions through UV-Vis-NIR Spectroscopy and Machine Learning Modelling

Rozita Sulaiman<sup>1,\*</sup>, Nur Hidayah Azeman<sup>2</sup>, Athiyah Sakinah Masran<sup>1</sup>, Muhammad Qayyum Othman<sup>1</sup>, Nur Afifah Ahmad Nazri<sup>2</sup>, Azween Hadiera Hishamuddin<sup>3</sup>, Mohd Hafiz Abu Bakar<sup>4</sup> and Ahmad Ashrif A Bakar<sup>1,2</sup>

<sup>1</sup>Photonics Technology Laboratory, Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

<sup>2</sup>Institute of Microengineering and Nanoelectronics, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

<sup>3</sup>School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia.

<sup>4</sup>Institute of Power Engineering, Universiti Tenaga Nasional, 43000 Kajang, Selangor, UNITEN, Malaysia.

\*Corresponding author: rozita@mohr.gov.my\*; ashrif@ukm.edu.my

## Keywords

Spectroscopy  
Regression  
Macronutrient  
Hydroponics  
Efficiency

## Article History

Received: 5 May 2025

Accepted: 5 June 2025

Published: 30 October 2025

## Abstract

Accurate nutrient determination is crucial for optimizing hydroponic systems. Traditional methods like electrical conductivity (EC), ion-selective electrodes (ISE), and colourimetry are hindered by limited ion specificity, high maintenance, and prolonged analysis times. Spectroscopy-based AI methods often demand extensive sample sets, limiting their real-time applicability. This study presents a rapid UV-Vis-NIR spectroscopy technique integrated with machine learning (ML) for predicting nitrogen and phosphorus levels using minimal samples. Among five regression models tested Support Vector Machine Regression (SVR), Linear Regression (LR), Partial Least Squares Regression (PLSR), Random Forest Regression (RFR), and Gradient Boosting Trees Regression (GBTR), linear regression (LR) excelled for nitrogen prediction ( $R^2=0.9734$ ; RMSE=19.2345 ppm), although its performance slightly dropped during external validation ( $R^2=0.923$ ; RMSE=30.6666 ppm). Random forest regression (RFR) showed superior accuracy for phosphorus prediction ( $R^2=0.9578$ , 0.9468). Integration of Support Vector machine (SVM) classification with LR and RFR further improved accuracy ( $R^2=0.96$ ), even with just 26 samples. This ML-based spectroscopy model enables rapid, precise nutrient monitoring, enhancing crop yield and sustainability in hydroponics.

## 1. Introduction

Hydroponic cultivation is emerging as a sustainable solution to meet global food demands by 2050, offering efficient, high-yield production with reduced water and pesticide use. Nitrogen and phosphorus are essential macronutrients for plant growth, and maintaining the right nutrient balance is crucial for plant health. Conventional nutrient detection methods, such as electrical conductivity (EC), ion-selective electrodes (ISE), and colorimetric tests, are limited by poor ion specificity, high maintenance requirements, and time-consuming procedures. Moreover, current spectroscopy-based artificial intelligence (AI) models often require large datasets, rendering them impractical for small-scale or real-time applications. We developed a machine learning model to predict nitrogen and phosphorus concentrations with minimal sample concentrations. Five regression algorithms, LR, PLSR, SVR, RFR, and GBTR, were compared to identify the most effective regression model for each nutrient. Then, a single-rule algorithm was used to integrate the classification model with selected regression models. By enabling accurate predictions with fewer samples, this approach addresses the limitations of current methods, making it more practical, economical, and suitable for real-time hydroponic cultivation [1].



## 2. Method

This study developed machine learning models to predict nitrogen and phosphorus concentrations using UV-Vis-NIR spectroscopy. Twenty-six hydroponic nutrient samples with varying concentrations of sodium nitrate ( $\text{NaNO}_3$ ), potassium dihydrogen phosphate ( $\text{KH}_2\text{PO}_4$ ), and sodium nitrite ( $\text{NaNO}_2$ ) were prepared. Nitrogen levels ranged from 30–400 ppm (nitrate) and 150 ppm (nitrite), while phosphorus concentrations ranged from 5–105 ppm. Absorbance spectra were collected using an Ocean Optics USB4000 spectrometer controlled via SpectraSuite software. Each sample was scanned 250 times (200 for model training, 50 for external validation), resulting in 5000 total spectral readings. Spectral data underwent preprocessing (noise removal, baseline correction) and dimensionality reduction before ML model training. Support Vector Machine (SVM) was selected due to its robustness in classifying data with high dimensionality and small sample size, making it ideal for determining nutrient types before regression [1]. Five regression algorithms (LR, PLSR, SVR, RFR, GBTR) were evaluated. Additionally, an integrated workflow was designed: SVM first classifies the nutrient type (nitrogen or phosphorus), then the appropriate regression model is applied. This model architecture improves prediction by tailoring regression methods to each nutrient class.

## 3. Results and discussion

Figure 1(I) shows nitrogen absorption spectra (50–350 ppm) with a distinct peak at 305 nm, while Figure 1(II) displays phosphorus spectra (5–100 ppm), indicating phosphate ion behavior in a nitrate-nitrite solution.

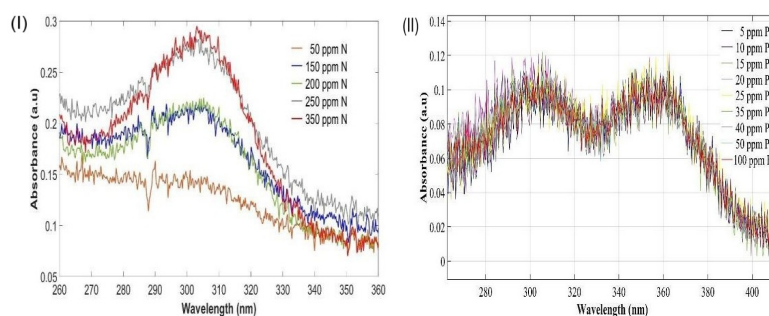


Figure 1. Absorption spectra: (I) nitrogen, (II) phosphorus at various concentrations.

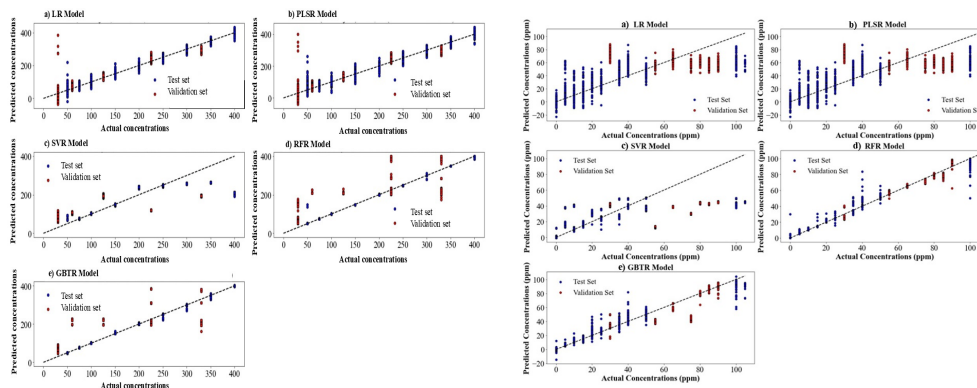


Figure 2. Scatter plot distribution of nitrogen (left) and phosphorus (right) for test and validation set using (a) LR (b) PLSR (c) SVR (d) RFR (e) GBTR Model.

Figure 2 shows the performance of different regression models in predicting nitrogen and phosphorus concentrations. For nitrogen, the LR model performed excellently with  $R^2 = 0.9734$  and  $\text{RMSE} = 19.2345$  ppm, indicating low error and high consistency with test data. However, when tested with external validation data, the model's performance slightly decreased, with  $R^2$  dropping to 0.923 and  $\text{RMSE}$  increasing to 30.6666 ppm, indicating a slight loss in generalisation capability.

As for phosphorus, the RFR model outperformed others, with low  $\text{RMSE}$  and high  $R^2$  values (0.9578 and 0.9468). Additionally, an integrated SVM classification model with LR and RFR regression achieved an  $R^2$  of 0.96, showing significant improvement in predicting nutrient concentrations compared to individual models, even



with fewer samples than those used in [2], [3]. This indicates that the pre-classification of nutrient types enhances the overall regression accuracy by improving nutrient-specific model targeting and reducing cross-variable interference.

#### 4. Conclusions

This study demonstrates a rapid and accurate approach for nutrient prediction in hydroponics using UV-Vis-NIR spectroscopy integrated with a single-rule machine learning model combining SVM, LR, and RFR. The LR model yielded high accuracy for nitrogen ( $R^2 = 0.9734$ ; RMSE = 19.2345 ppm), while RFR performed best for phosphorus ( $R^2 = 0.9578$ ; 0.9468). Despite a limited dataset, the integrated model achieved an overall  $R^2$  of 0.96, highlighting its potential for efficient and sustainable nutrient monitoring.

#### References

- [1] R. Sulaiman, N. H. Azeman, M. H. H. Mokhtar, N. N. Mobarak, M. H. Abu Bakar, and A. A. A. Bakar, "Hybrid ensemble-based machine learning model for predicting phosphorus concentrations in hydroponic solution," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 304, p. 123 327, Jan. 2024, ISSN: 1386-1425. DOI: [10.1016/j.saa.2023.123327](https://doi.org/10.1016/j.saa.2023.123327)
- [2] F. Monteiro-Silva, P. A. S. Jorge, and R. C. Martins, "Optical sensing of nitrogen, phosphorus and potassium: A spectrophotometrical approach toward smart nutrient deployment," *Chemosensors*, vol. 7, no. 4, p. 51, Oct. 2019, ISSN: 2227-9040. DOI: [10.3390/chemosensors7040051](https://doi.org/10.3390/chemosensors7040051)
- [3] F. Silva et al., "Reagent-less spectroscopy towards npk sensing for hydroponics nutrient solutions," *Sensors and Actuators B: Chemical*, vol. 395, p. 134 442, Nov. 2023, ISSN: 0925-4005. DOI: [10.1016/j.snb.2023.134442](https://doi.org/10.1016/j.snb.2023.134442)