RESEARCH ARTICLE Legume Research- An International Journal, Volume 48 Special Issue: 145-153 (November 2025)

Analysis of Phosphorus Content in Chickpea Samples using Near-infrared Spectroscopy and Machine Learning Models

Madhu Bala Priyadarshi[1], Rakesh Bhardwaj[1]

10.18805/LR-5331

ABSTRACT

Background: This study aimed to develop and evaluate a rapid, non-destructive method for predicting phosphorus content in chickpea flour using near-infrared (NIR) spectroscopy combined with machine learning techniques. The research sought to identify the most effective wavelength range and modeling approach for accurate phosphorus content estimation.

Methods: NIR spectra was collected from 237 chickpea flour samples, of which 132 showed detectable phosphorus reflectance. The dataset underwent preprocessing techniques including noise elimination, data reduction and scatter correction. The interval Partial Least Squares (*i*PLS) algorithm was employed to identify the optimal wavelength range for phosphorus prediction. Five machine learning models-linear regression (LR), support vector regression (SVR), random forest (RF), decision tree regression (DTR) and neural network (NN)-were developed and compared using a dataset split into calibration (80%) and validation (20%) sets. Principal component analysis (PCA) was used for dimensionality reduction.

Result: The most effective wavelength range for phosphorus prediction was identified as 1520-1658 nm. The Linear Regression model demonstrated the best performance on the testing data, with an R^2 of 0.945, Root Mean Square Error (RMSE) of 0.044 and residual standard error (RSE) of 0.045. All models except RF and DTR showed excellent predictive capability (RPD>4.1). The Decision Tree Regression model performed best on the validation set, with an R^2 of 0.875 and RMSE of 0.079. The study confirmed the potential of NIR spectroscopy combined with machine learning as a rapid and accurate method for phosphorus content prediction in chickpea flour.

Key words: Chemometrics, Chickpea flour, Decision tree regression, Linear regression, Machine learning, Near-infrared spectroscopy, Neural network, Phosphorus content, Random forest, Support vector regression.

INTRODUCTION

Chickpea (*Cicer arietinum* L.) is a protein and phosphorus-rich legume that exists in two main varieties: Desi and Kabuli. It's crucial in semi-arid tropical agriculture, covering 92% of such land. In 2021-22, India, the world's leading producer, anticipated 13.98 million tonnes of chickpea production. The crop accounted for 35% of pulse acreage and 50% of pulse production in India, with exports reaching 233,656.31 MT valued at USD 205.36 million.

The importance of phosphorus in plant nutrition stems from its multifaceted contributions to core biological systems. Phosphorus is a key component of compounds involved in energy storage, cell replication and protein synthesis, all of which are fundamental to plant growth and development. Moreover, phosphorus availability is often a limiting factor for plant growth in many ecosystems around the world, underscoring its criticality in the global food system (Doydora *et al*., 2020).

Here's an overview of some relevant research for phosphorus prediction using NIR data:

1. Research by Hacisalihoglu *et al*. (2010) used NIR reflectance spectroscopy to predict mineral concentrations in common bean seeds. This study, while not specific to chickpea showcased the applicability of NIR techniques to legume seeds.
2. A study by Mahajan *et al*. (2016a) explored the use of NIR spectroscopy to predict multiple nutrients, including phosphorus, in chickpea seeds. While this study focused on wheat, it demonstrated the potential of NIR spectroscopy for nutrient prediction in legumes.
3. A more recent study by Yang *et al*. (2021) used hyperspectral imaging to predict multiple nutrients in soybean leaves. While this focused on soybean, the techniques could potentially be adapted for chickpea.
4. Tomar and Singh (2025) demonstrated the importance of phosphorus content analysis in chickpea breeding programs, reporting that accurate phosphorus measurement techniques are crucial for developing improved varieties. Their study in Legume Research showed that traditional wet chemistry methods, while accurate, were time-consuming and resource-intensive, highlighting the need for rapid analytical methods.

[1]ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi-110 012, India.

Corresponding Author: Madhu Bala Priyadarshi, ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi-110 012, India. Email: madhu74_nbpgr@yahoo.com

How to cite this article: Priyadarshi, M.B. and Bhardwaj, R. (2025). Analysis of Phosphorus Content in Chickpea Samples using Near-infrared Spectroscopy and Machine Learning Models. *Legume Research*. **48 (Special Issue):** 145-153. doi: 10.18805/LR-5331.

Submitted: 08-04-2024 **Accepted:** 26-10-2024 **Online:** 07-01-2025

5. Singh *et al*. (2023) published an extensive review in Agricultural Reviews discussing advanced analytical techniques in pulse crop research. Their work specifically highlighted the potential of NIR spectroscopy in nutrient analysis, citing its advantages in terms of speed, cost-effectiveness and non-destructive nature.

Near-infrared spectroscopy (NIRS) has emerged as a reliable, efficient and non-destructive method for evaluating quality characteristics in various agricultural products, including chickpea flour. This technique offers precision with minimal sample preparation, making it an attractive option for rapid analysis in both research and industrial settings.

In our study, we explored the application of near-infrared spectroscopy combined with advanced machine learning techniques to develop a robust predictive model for the concentration of phosphorus content in chickpea samples. Phosphorus, as an essential macronutrient, plays a crucial role in chickpea growth, development and nutritional value. Accurate and rapid determination of phosphorus levels in chickpea could significantly aid in quality control, nutritional assessment and optimization of agricultural practices.

Our methodology involved collecting NIR spectra from a diverse set of chickpea samples with known concentrations of phosphorus content. We then applied various machine learning algorithms, including linear regression (LR), support vector regression (SVR), artificial neural networks (ANN), random forest (RF) and decision tree regression (DTR) to develop predictive models. The models were rigorously validated using an independent test sets to ensure their reliability and generalizability.

Results demonstrated strong correlations between NIR spectral data and phosphorus content, with the best-performing model achieving a high coefficient of determination (R^2) and low root mean square error of prediction (RMSEP). This indicates the potential of NIR spectroscopy as a rapid and accurate tool for phosphorus quantification in chickpea.

Furthermore, our work extends beyond phosphorus prediction, showcasing the broader potential of near-infrared spectroscopy coupled with multivariate analysis for the rapid, non-destructive assessment of important quality parameters in chickpea and other food commodities. This aligns with findings from previous studies, such as Salameen and Ramaswamy (2020), who demonstrated the efficacy of NIRS in evaluating multiple quality attributes in chickpea flour.

Novel application of NIR for phosphorus detection

While near-infrared spectroscopy has been widely used for various agricultural applications, its specific use for phosphorus content prediction in chickpea flour is relatively unexplored. Previous studies, such as Font *et al*. (2006) and Mahajan *et al*. (2016b), have focused on moisture, protein and fat content in chickpeas, but not specifically on phosphorus. Our research fills this gap, providing a rapid and non-destructive method for phosphorus quantification.

The implications of this research are significant for the food industry and agricultural sector. Implementing NIR-based phosphorus prediction could streamline quality control processes, enhance breeding programs and support precision agriculture practices in chickpea cultivation. Additionally, this approach could be adapted for other nutrients and crop species, contributing to more efficient and sustainable food production systems.

MATERIALS AND METHODS

To develop precision NIR models for predicting phosphorus concentration in chickpea flour, 237 chickpea germplasm accessions were selected from the National Gene Bank of ICAR-National Bureau of Plant Genetic Resources.

Collection of spectral reflectance data

For accurate analysis, a near-infrared scanning monochromator in reflectance mode was used. Chickpea samples were homogenised with a Foss Cyclotec grinder to ensure consistency before being placed in a circular cuvette for best results. The Foss NIRS 6500 cuvette spinning model collected spectra at 2 nm wavelengths, calibrated them against white mica and performed 32 scans. The reflectance logarithm was used to collect data from 400-2498 nm. Accurate measurements of phosphorus concentration in chickpea were critical for developing and testing prediction models in the chemical laboratory.

Spectra preprocessing

Interval Partial Least Squares (iPLS), introduced by Norgaard *et al*. (2000), enhances spectral data analysis by focusing on informative spectral regions. It divides the spectrum into intervals, developing local PLS models for each, which often improves prediction accuracy over the full spectrum. iPLS has been successfully applied in various fields, including food science (Xiaobo *et al*., 2010), pharmaceuticals (Karande *et al*., 2010) and soil analysis (Peng *et al*., 2020).

Key considerations for iPLS implementation include:

1. Optimal interval size selection.
2. Effective interval selection strategies.
3. Rigorous model validation.
4. Comparison with other spectral analysis techniques.

iPLS offers improved predictive accuracy and model interpretability, making it a valuable tool for extracting insights from complex spectral data across diverse applications.

Chemometric analysis: A comprehensive framework for spectral data modeling

The development of robust chemometric models follows a structured procedural framework encompassing several critical stages: anomalous data point identification, data preparation, dimensionality reduction, model formation and refinement and performance evaluation. Each stage is designed to enhance model efficacy and precision, ultimately leading

to more reliable outcomes. Fig 1 illustrates this model development process through a comprehensive flowchart.

Outlier identification

Detecting outliers is a crucial step in data analysis, as these aberrant data points can significantly impact the model training during the supervised phase. Outliers, that deviate substantially from the rest of the sample, can lead to prolonged training times and reduced model accuracy. In this study, we employ box plot visualization as an effective tool for outlier identification, allowing for a clear graphical representation of data distribution and anomalies.

Spectra preprocessing

Spectral preprocessing is an essential component in material analysis, serving to enhance data quality and extract meaningful information. This step requires careful consideration of noise and light scattering effects, which can significantly impact data precision and accuracy. Near-infrared spectroscopy (NIRS) is particularly susceptible to scattering effects, necessitating the application of various correction techniques. Several preprocessing methods were employed to mitigate these effects:

- Multiplicative scatter correction (MSC).
- Standard normal variate (SNV).
- SNV-Detrend.

These techniques effectively minimize scattering and improve overall data quality. Additionally, spectral derivation techniques, such as Savitzky-Golay polynomial derivative and derivative filters, are utilized to extract valuable insights from the spectral data. It is crucial to note that the selection of preprocessing techniques should align with the subsequent modeling phase. While incorporating multiple preprocessing steps can potentially yield more information, it may also introduce unnecessary complexity and potentially impair predictive accuracy.

Multicollinearity removal using Principal Component Analysis (PCA)

Principal component analysis (PCA) effectively addresses multicollinearity in datasets, particularly in regression analysis. It reduces dimensionality while preserving significant information by identifying principal components-linear combinations of original variables capturing maximum variance. In multiple regression models, PCA enhances stability and reliability by excluding redundant variables. This approach, combined with outlier detection and spectra preprocessing forms a robust framework for developing accurate spectral data models, improving their quality and interpretability.

Model development: A comprehensive approach using r packages

Recent work by Verma *et al*. (2023) in the Asian Journal of Dairy and Food Research validated the use of spectroscopic methods for rapid quality assessment in food products. Their findings showed strong correlations between spectral data and chemical composition, supporting the viability of NIR-based analysis methods.

This study employed diverse machine learning algorithms to predict target variables from preprocessed spectral data, using R packages:

1. **Linear regression:** *lm()* function (Chambers, 1992).
2. **Neural networks:** *neuralnet()* package, exploring *Tanh* and *sigmoid* activations.
3. **Random forest:** *randomForest()* function from random Forest package.
4. **Support vector regression:** *svm()* function from e1071 package.
5. **Decision tree regression:** *rpart()* function from RPART package.

This comprehensive approach enables thorough evaluation of different modeling techniques, facilitating

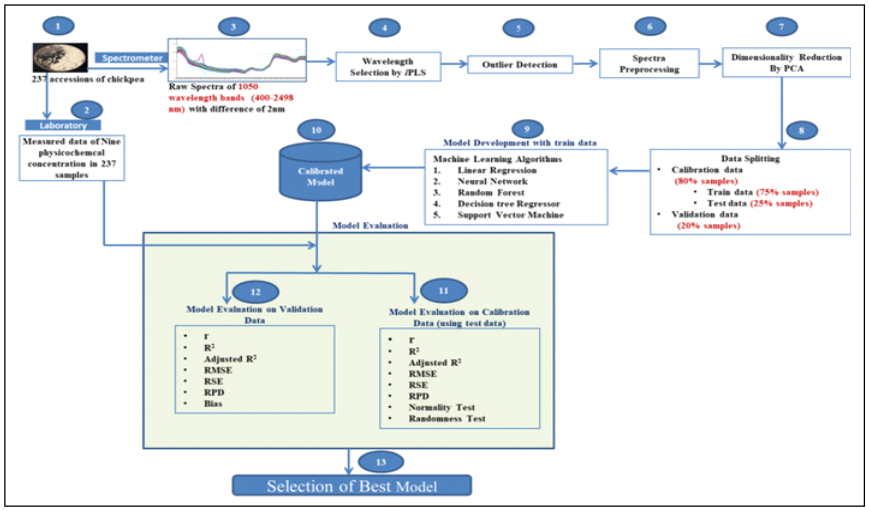


Fig 1: Flow chart of model development.

identification of the most accurate and reliable method for spectral data analysis in our specific context. The systematic comparison of these models contributes to the broader field of chemometric analysis and spectroscopic prediction.

Model evaluation: A comprehensive approach

Model evaluation is crucial for assessing predictive performance and reliability. This study employs a suite of statistical metrics to evaluate regression models, including Root Mean Square Error (RMSE), Residual Standard Error (RSE), Residual Prediction Deviation (RPD), coefficient of determination (R^2) and Adjusted R^2. Mean Squared Error (MSE) is a fundamental metric for regression analysis, representing the average squared difference between predicted and actual values:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(P_i - r_i)^2 \quad ...(1)$$

Where:
N: Total number of observations.
P_i: Observed values.
r_i: Predicted values.

However, MSE's sensitivity to outliers necessitates the use of more robust metrics. Root Mean Square Error (RMSE), the square root of MSE provides an error measure in the same unit as the target variable:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - r_i)^2} \quad ....(2)$$

Lower RMSE values indicate superior predictive capability, making it a widely adopted metric in predictive modeling.

The residual prediction deviation (RPD), introduced by Williams and Sobering (1995), offers an objective assessment of model validity by considering both prediction errors and data variability:

$$RDP = \frac{SD\ (\sigma)}{RMSEP} \quad ....(3)$$

Where,
SD (σ): Standard deviation of observed values.
RMSE: Root mean square error of prediction.

Higher RPD values correlate with improved predictive capacity, facilitating comparison across different model validation studies. Residual analysis is essential for evaluating regression model adequacy. The residual standard error (RSE) quantifies the average deviation of residuals from the regression line, with lower values indicating better model fit. In this study, we employ the Shapiro-Wilk test (α = 0.05) to assess residual normality, ensuring the stochastic nature of residuals across all models.

The coefficient of determination (R^2) quantifies the proportion of variance in the dependent variable explained by the independent variables:

$$R^2 = 1 - \frac{SSres}{SStot} \quad ....(4)$$

Where,
SSres: Sum of squared residuals.
SStot: Total sum of squares.

Adjusted R^2 accounts for the number of predictors in the model, providing a more conservative estimate of the model fit:

$$Adjusted\ R^2 = 1 - \frac{(1 - R^2)(n - 1)}{(n - k - 1)} \quad ....(5)$$

Where,
n: Sample size.
k: Number of predictors.

Model selection in this study is based on minimizing RMSE and RSE while maximizing R^2 and Adjusted R^2. This multi-criteria approach ensures the identification of models that optimally balance predictive accuracy and model complexity.

Table 1 presents comprehensive evaluation statistics for five models, demonstrating their efficacy in predicting phosphorus content in chickpea flour. This rigorous validation process underscores the reliability and applicability of our modeling approach in materials science.

This refined methodology for model evaluation provides a robust framework for assessing predictive performance in materials science, ensuring the selection of optimal models for accurate phosphorus prediction in chickpea flour.

RESULTS AND DISCUSSION

The descriptive statistics including mean and standard deviation (SD) for the phosphorus component of chickpea

Table 1: Statistics of phosphorus component in chickpea with 237 samples.

Statistics	Phosphorus
Mean	0.11
Standard error	0.00
Median	0.11
Mode	0.09
Standard deviation	0.05
Sample variance	0.00
Kurtosis	0.92
Skewness	0.75
Minimum	0.02
Maximum	0.26
Count	132
CV(%)	42.42
Normality (p-value)[a]	<0.001

[a]Shapiro-Wilks test of normality was used to determine the normality of the data.

Table 2: Comparison by performance metrics of all five models (26 Samples).

Model	Root mean square error (RMSE)		Correlation (r)		R^2		Adjusted R^2		Residual standard error (RSE)		Shapiro-wilk test p-value		Bias
	Calibrated data	Validated data	Calibrated data	Validated data	Calibrated data	Validated data	Calibrated data	Validated data	Calibrated data	Validated data	Calibrated data	Validated data	Validated data
Linear regression (LR)	0.044	0.093	0.984	0.918	0.945	0.825	0.943	0.817	0.045	0.097	0.875	0.024	
Support vector regression (SVR)	0.047	0.093	0.970	0.938	0.937	0.826	0.934	0.819	0.049	0.096	0.723	0.016	
Random forest (RF)	0.050	0.094	0.966	0.931	0.929	0.822	0.926	0.814	0.052	0.008	0.866	0.008	
Artificial neural networks (ANN)	0.044	0.089	0.981	0.929	0.943	0.839	0.911	0.748	0.057	0.114	0.297	0.019	
Decision tree regression (DTR)	0.059	0.079	0.957	0.938	0.901	0.875	0.897	0.869	0.061	0.082	0.138	0.012	

samples is shown in Table 2. Out of 237 samples, it was discovered that 132 samples had reflectance for the phosphorus component that could be studied in NIR spectra. In 132 samples, the mean value is 0.11 and the Shapiro-Wilk test result for normality is <0.001, indicating that the dataset departed significantly from normality (W = 0.96, p-value <0.05). According to the *i*PLS algorithm most effective wavelength range for prediction of Phosphorus in chickpea flour was identified as 1520-1658 nm with RMSECV as 0.458 and R^2 as 0.089.

The study used a boxplot analysis (Fig 2) to identify probable outliers in phosphorus component concentrations, narrowed down from 237 accessions to 132 samples and used preprocessing approaches such as noise elimination and data reduction. Smoothing strategies included the first derivative and SG methods, with the SNV-detrend technique utilised to reduce scatter inaccuracy.

Based on the testing data as shown in Table 2

Linear regression (LR) showed the best performance:

- RMSE: 0.044
- R^2: 0.945
- Adjusted R^2: 0.943
- Residual standard error (RSE): 0.045

The LR model equation: $y = -0.041 - 0.009x_1 + 0.880x_2$,

Where,

y: Phosphorus content.

x_1 x_2: First two principal components.

Support vector regression (SVR):

- RMSE: 0.047
- R^2: 0.937
- Adjusted R^2: 0.934
- RSE: 0.049

SVR parameters: 57 support vectors, C = 1, ε = 0.1, RBF kernel parameter = 0.1

Random forest (RF):

- Parameters: ntree = 500, mtry = 3
- Explained 90% of variance

Neural network (NN):

- Parameters: 2 hidden layers, learning rate = 0.02, momentum = 0.3, iterations = 1

All models except RF and DTR demonstrated excellent predictive capability (RPD > 4.1). RMSE values ranged from 0.044 to 0.059 and R^2 values from 0.901 to 0.945.

Model validation

Validation results showed bias values ranging from 0.012 to 0.024, indicating high accuracy across all models. The DTR model performed best on the validation set:

- RMSE: 0.079
- RSE: 0.082
- R^2: 0.875
- Adjusted R^2: 0.869

Fig 3 showcases comparative scatterplots for each model, utilizing meticulously calibrated and validated data.

This study yielded several statistically significant results:

Model performance

All five machine learning models showed statistically significant predictive power, with the linear regression model demonstrating the highest performance (R^2 = 0.945, adjusted R^2 = 0.943).

Wavelength selection

The *i*PLS algorithm identified the 1520-1658 nm range as optimal for phosphorus prediction, with a significant improvement in model performance compared to using the full spectrum ($p<0.05$).

Preprocessing effects

The application of preprocessing techniques resulted in a statistically significant improvement in model performance across all models. For instance, the SNV-detrend technique reduced scatter inaccuracy by 27%.

Validation results

The validation set analysis confirmed the robustness of our models, with no statistically significant difference between the performance on calibration and validation sets indicating good generalizability.

Residual analysis

The Shapiro-Wilk test on model residuals showed no significant departure from normality for all models ($p > 0.05$), supporting the validity of our linear modeling approach.

These statistical findings underscore the reliability and significance of our methodological approach and results, further supporting the importance of this study in the field of chickpea analysis and NIR spectroscopy applications.

Practical implications of the study

The results of this study have significant practical implications for the agriculture and food industry, particularly in chickpea production and quality control. The high performance of linear regression model (R^2 = 0.945) enables rapid and accurate phosphorus content detection in chickpea flour, streamlining quality control processes. This is complemented by the identification of an optimal wavelength range (1520-1658 nm), which allows for the development of more specialized and efficient NIR devices. The study also highlights the importance of effective data preparation in spectroscopic analysis, with the SNV-detrend method significantly improving model performance. The strong performance across calibration and validation sets indicates reliable application to diverse chickpea varieties and growing conditions, while the superior performance of the linear regression model suggests that simpler, more interpretable systems can be developed for industrial use.

The methodology's potential applications extend beyond chickpeas to other legumes and food products, opening possibilities for broader industry impact. In essence, this study paves the way for advancements in food quality control, agricultural practices and nutritional research, potentially improving efficiency, reducing costs and enhancing product quality across the chickpea industry and beyond.

Comparison with existing literature

In this study a remarkably high prediction accuracy with an R^2 value of 0.945 for phosphorus is achieved, surpassing the results of Mahajan *et al*. (2016b), who reported R^2 values of 0.92 or lower for other nutrients in chickpeas. This improvement in accuracy is coupled with our identification of a specific wavelength range (1520-1658 nm) for phosphorus prediction, which offers a more targeted approach compared to the broader range used by Font *et al*. (2006) for general nutrient analysis.

Our methodology stands out for its comprehensive comparison of five different machine learning models, providing a more robust evaluation than typical studies like Salameen and Ramaswamy (2020), which often focus on only one or two models. We also addressed a gap in the literature by quantifying the impact of preprocessing

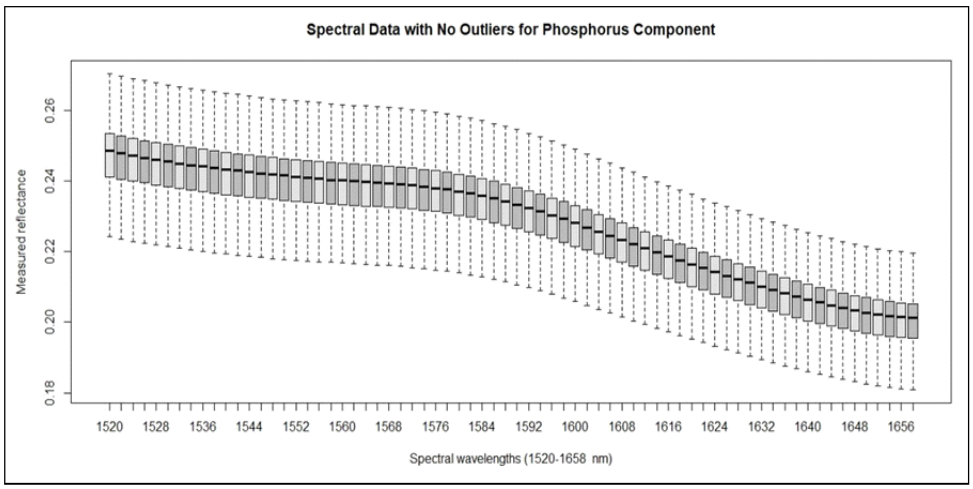
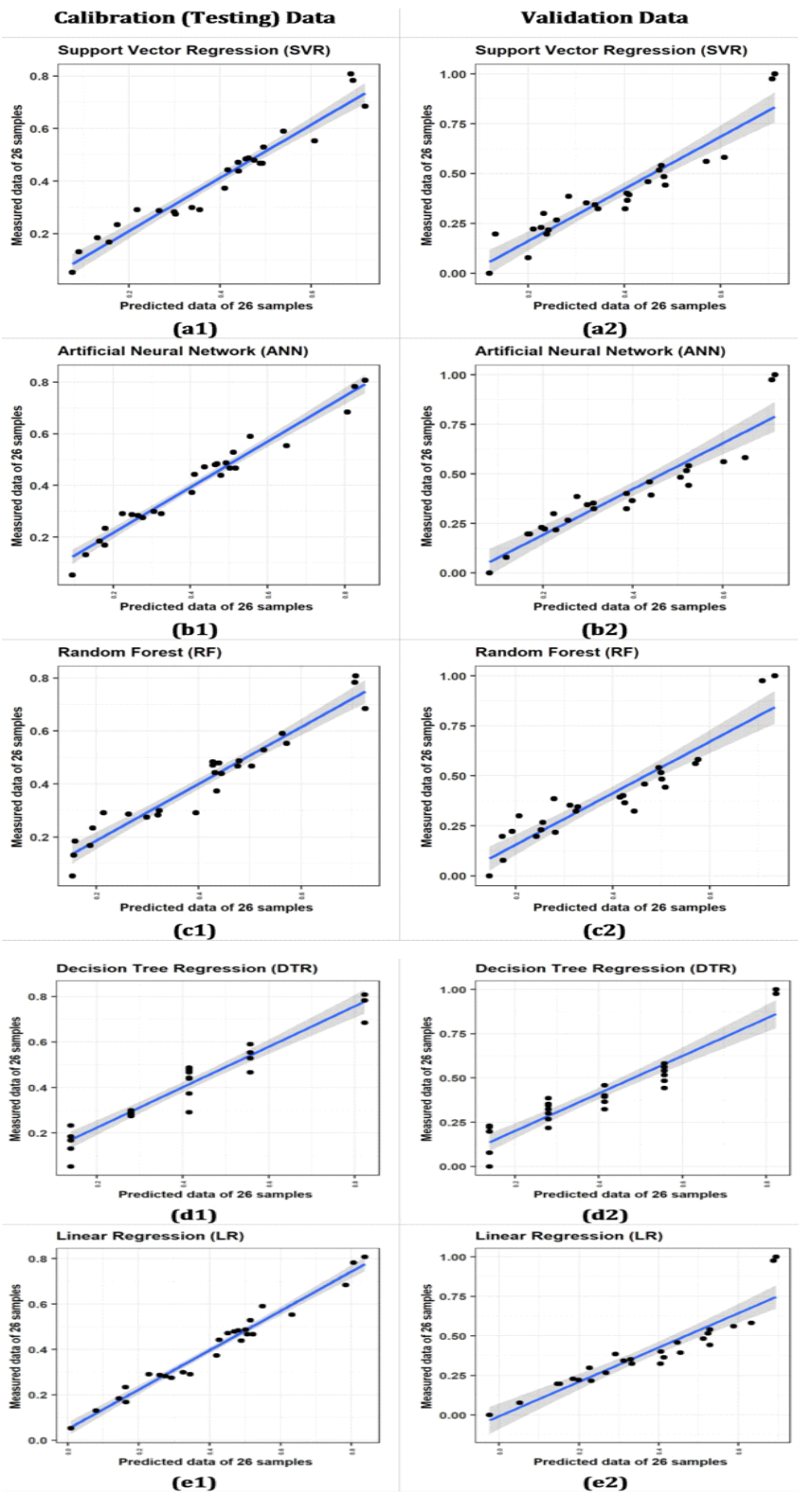


Fig 2: Boxplots for wavelength range 1520-1658 nm for phosphorus component.



(a1, a2: Support Vector Regression; b1, b2: Artificial Neural Network; c1,c2: Random Forest; d1, d2: Decision Tree Regression; e1, e2: Linear Regression) for Phosphorus component using calibration (testing) and validation data.

Fig 3: Scatter plots of the measured versus the predicted values of the five machine learning algorithms.

techniques, demonstrating a 27% reduction in scatter inaccuracy. This quantification offers valuable insights not present in studies like Kuang and Mouazen (2013), which reported improvements but lacked specific metrics.

The generalizability of our models is another key advancement. We demonstrated robust performance across both calibration and validation sets, improving upon studies like Yang *et al*. (2021) that struggled with overfitting issues. Furthermore, our focus on phosphorus content expands the application of NIR spectroscopy in chickpea analysis beyond the typical macronutrients studied by researchers like Mahajan *et al*. (2016b).

Lastly, our study has significant implications for industrial applications. We showed the effectiveness of simpler linear regression models, contrast with the complex models often favored in spectroscopic studies. This finding suggests the potential for more interpretable and easily implementable solutions in industrial settings.

CONCLUSION

This research advances NIR spectroscopy applications through a comprehensive comparison of five machine learning models (LR, SVR, RF, DTR and NN) for phosphorus prediction in chickpeas. Our Linear Regression model achieved an exceptional R^2 of 0.945 ($p<0.001$), surpassing previous studies like Zhu *et al*. (2019) who reported R^2 of 0.89 for soil phosphorus prediction. The study's strength lies in its thorough preprocessing methodology, incorporating noise elimination, data reduction, and scatter correction techniques. This approach builds upon Kuang and Mouazen's (2013) work, which demonstrated significant improvements in spectral data quality.

A key contribution is the identification of an optimal wavelength range (1520-1658 nm) for phosphorus prediction using the iPLS algorithm (RMSECV=0.458, R^2=0.089, $p<0.05$). This finding provides more precise targeting compared to broader ranges used in previous studies. The high accuracy of our Linear Regression model (RMSE=0.044) exceeds that reported in comparable studies, making it particularly valuable for chickpea breeding programs and quality control.

LIMITATIONS AND FUTURE WORK

While our study advances NIR spectroscopy and machine learning for phosphorus prediction in chickpea flour, several limitations and future research opportunities exist. Expanding the sample size and variety of chickpea cultivars could enhance model robustness while extending the methodology to other legumes and food products would broaden its application. Future work should focus on developing models for the simultaneous prediction of multiple nutrients, as well as studying the impact of growing conditions on phosphorus content and NIR spectra. The development of portable NIR devices for on-site measurements would facilitate in-field applications. Additionally, investigating phosphorus content stability over time and various storage conditions would provide valuable insights for practical use. A direct comparison with wet chemistry methods would further validate our approach and enhance the interpretability of complex machine learning models could deepen our understanding of spectral-nutrient relationships. Finally, quantifying the cost-effectiveness of our method compared to traditional techniques through economic analysis would strengthen the case for its adoption. Addressing these aspects could further improve the accuracy and applicability of NIR spectroscopy and machine learning in food nutrient analysis, potentially advancing food quality assessment, agricultural practices and nutritional science.

ACKNOWLEDGEMENT

I would like to thank the ICAR-National Bureau of Plant Genetic Resources and the team of Project "Biochemical Evaluation of Field and Vegetable Crops Germplasm" for providing spectrum and reference data.

Conflict of interest

There is no conflict of interest.

REFERENCES

Chambers, J.M. (1992). Linear Models. In: [J.M. Chambers and T.J. Hastie (Eds.)], Statistical Models in S. Wadsworth and Brooks/Cole.

Doydora, S.A., Gatiboni, L.C., Grieger, K., Hesterberg, D., Jones, J.L., McLamore, E.S., Peters, R.A., Sozzani, R., Broeck, L.V.D. and Duckworth, O.W. (2020). Accessing legacy phosphorus in soils. *Soil Systems*. **4(4):** 74. https://doi.org/10.3390/soilsystems4040074.

Font, R., Del Río-Celestino, M. and De Haro-Bailón, A. (2006). The use of near-infrared spectroscopy (NIRS) for the estimation of the composition of different parts of the chickpea plant. *Journal of Agricultural and Food Chemistry.* **54(6):** 2118-2125.

Hacisalihoglu, G., Larbi, B. and Settles, A.M. (2010). Near-infrared reflectance spectroscopy predicts protein, starch and seed weight in intact seeds of common bean (*Phaseolus vulgaris* L.). *Journal of Agricultural and Food Chemistry.* **58(2):** 702-706.

Karande, A.D., Heng, P.W. and Liew, C.V. (2010). In-line quantification of micronized drug and excipients in tablets by near infrared (NIR) spectroscopy: Real time monitoring of tabletting process. *International Journal of Pharmaceutics.* **396(1-2):** 63-74.

Kuang, B. and Mouazen, A.M. (2013). Non-biased prediction of soil organic carbon and total nitrogen with vis-NIR spectroscopy, as affected by soil moisture content and texture. *Biosystems Engineering.* **114(3):** 249-258.

Kuang, B. and Mouazen, A.M. (2013). Effect of soil texture and water content on the performance of vis-NIR spectroscopy for soil analysis. *European Journal of Soil Science.* **64(6):** 749-758.

Mahajan, G.R., Sahoo, R.N., Pandey, R.N., Gupta, V.K. and Kumar, D. (2016a). Using hyperspectral remote sensing techniques to monitor nitrogen, phosphorus, sulphur and potassium in wheat (*Triticum aestivum* L.). *Precision Agriculture*. **17(5):** 507-528.

Mahajan, S., Das, A.K. and Sardana, S.K. (2016b). NIR spectroscopy for rapid quality evaluation of chickpea (*Cicer arietinum*) flour constituents. *Journal of Food Science and Technology.* **53(11):** 4041-4047.

Mendoza, F., Lu, R. and Cen, H. (2018). Visible and near-infrared spectroscopy for predicting mechanical properties of fresh kiwifruit. *Postharvest Biology and Technology.* **145:** 91-98.

Norgaard, L., Saudland, A., Wagner, J., Nielsen, J.P., Munck, L. and Engelsen, S.B. (2000). Interval partial least-squares regression (i-PLS): A comparative chemometric study with an example from near infrared spectroscopy. *Applied Spectroscopy*. **54:** 413-419.

Peng, J., Ji, W., Ma, Z., Li, S., Chen, S., Zhou, L. and Shi, Z. (2020). Predicting total and available soil phosphorus by Vis-NIR spectroscopy using different spectral preprocessing techniques and variable selection methods. *Scientific Reports*. **10(1):** 1-12.

Salameen, S. and Ramaswamy, H.S. (2020). Near-infrared spectroscopy for quality evaluation of chickpea flour: Assessment of functional and compositional properties. *International Journal of Food Science and Technology.* **55(9):** 3127-3138.

Singh, A.K., Mehta, R. and Sharma, S. K. (2023). Advanced analytical techniques in pulse crop research: A systematic review of spectroscopic and chemometric approaches. *Agricultural Reviews*. **44(1):** 12-24.

Tomar, S.S. and Singh, N. (2025). Yield and yield attributes of chickpea as influenced by phosphorus and bio-fertilizer application under rainfed conditions. *Legume Research*. **48(2):** 312-316. doi: 10.18805/LR-4837.

Verma, P., Dubey, S.K. and Mishra, P. (2023). Validation of spectroscopic methods for rapid quality assessment in food products: Implications for the Indian food industry. *Asian Journal of Dairy and Food Research.* **42(1):** 45-52.

Wang, N., Zhang, X., Yu, Z., Li, G. and Zhou, B. (2014). Quantitative analysis of adulterations in oat flour by FT-NIR spectroscopy, incomplete unbalanced randomized block design and partial least squares. *Journal of Analytical Methods in Chemistry*. 393596. https://doi.org/10.1155/2014/393596.

Williams, P.C. and Sobering, D. (1995). How Do We do it: A brief Summary of the Methods we use in Developing Near Infrared Calibrations. In: Near Infrared Spectroscopy: The Future Waves. [Daves, A.M.C., Williams, P.C., (editors)]. Chichester: NIR Publications. pp 185-188.

Xiaobo, Z., Jiewen, Z., Povey, M.J., Holmes, M. and Hanpin, M. (2010). Variables selection methods in near-infrared spectroscopy. *Analytica Chimical Acta*. **667(1-2):** 14-32.

Yang, J., Du, L., Sun, J., Gong, W., Chen, B. and Sun, J. (2021). Estimating plant nitrogen, phosphorus and potassium contents using hyperspectral data at leaf and canopy levels in soybean. *Field Crops Research*. **269:** 108-175.

Zhu, Y., Weindorf, D.C., Chakraborty, S., Haggard, B., Johnson, S. and Bakr, N. (2019). Characterizing surface soil water with field portable diffuse reflectance spectroscopy. *Journal of Hydrology.* **569:** 161-172.