ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR404

Electro-Optics Properties of Intact Cocoa Beans based on Near Infrared Technology

Syehan Atilla Munawar*[1]
Madrasah Tsanawiyah Negeri 1/
MTsN Model Banda Aceh, Indonesia
Senior High School SMU Negeri 3 Banda Aceh, Indonesia

Nabilul Kamal[2]
SMA Modal Bangsa, Banda Aceh, Indonesia

Zalfa Maulidya Rihani[3]
Senior High School SMU Negeri 3 Banda Aceh, Indonesia

Nurmahni Harahap[4]
Madrasah Tsanawiyah Negeri 1/
MTsN Model Banda Aceh, Indonesia

T Muhammad Adzka Rahmatillah[5]
Madrasah Aliyah Negeri 1 /
MAN Model Banda Aceh, Indonesia

Junaidi IB[6]
Doctoral Program, Graduate School UIN Ar-Raniry
Banda Aceh, Indonesia

Corresponding Author:- Syehan Atilla Munawar*[1]

Abstract:- This study encapsulates the efficient prediction of moisture content in cocoa beans through Near Infrared Spectroscopy (NIRS) and Partial Least Squares (PLS) regression, showcasing a strong model fit with a high R square value of 0.92 and low Root Mean Square Error (RMSE) of 0.36% in calibration; these values underscore the model's accurate estimation of moisture levels. In the realm of electro-optics properties, this success highlights NIRS's capability in assessing key attributes like moisture content in cocoa beans based on their unique spectral signatures, emphasizing the technology's role in quality control for chocolate production. Furthermore, the precise predictions align with the broader objective of leveraging NIRS to evaluate and optimize the electro-optics properties of cocoa beans, fostering informed decision-making for enhanced processing and quality assurance in the cocoa industry.

Keywords:- *NIRS, Technology, Spectroscopy, Non-Destructive.*

I. INTRODUCTION

The electro-optic properties of cocoa beans based on near infrared spectroscopy (NIRS) marks a significant leap forward in agricultural product analysis, specifically for assessing the quality of cocoa beans, which are essential for chocolate production. This technique operates on the principle that various molecular bonds absorb near-infrared light differently, offering a unique spectral signature that allows for the determination of a substance's properties [1].

Applied to cocoa beans, NIRS enables the rapid and non-destructive analysis of key quality parameters such as moisture content, fat content, and polyphenols concentration [2,3]. This electro-optic method, which converts light into electrical signals for analysis, stands out for its speed, non-destructive nature, cost efficiency, and the capability for real-time decision-making. It significantly cuts down the time for analysis compared to traditional methods, without altering the beans, thereby preserving them for further use. Moreover, its cost-effectiveness and the ability for on-the-spot analysis promote immediate sorting and processing decisions.

Near infrared spectroscopy or abbreviated as NIRS, is a sophisticated analytical technique that has grown increasingly popular across a multitude of disciplines due to its non-destructive nature and ability to quickly assess the composition and characteristics of various materials. Fundamentally grounded in molecular spectroscopy, NIRS operates within the near infrared region of the electromagnetic spectrum, spanning approximately 780 nm to 2500 nm [4,5].

This method capitalizes on the selective absorption of near-infrared light by molecules with hydrogen bonds—such as O-H, N-H, and C-H—where the absorbed wavelengths correspond to overtone and combination vibrations of these bonds. When near-infrared light is projected onto a sample, it is either absorbed, transmitted, or reflected as illustrated in Figure 1 based on the chemical composition of the sample, creating a pattern unique to its constituents [6].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR404

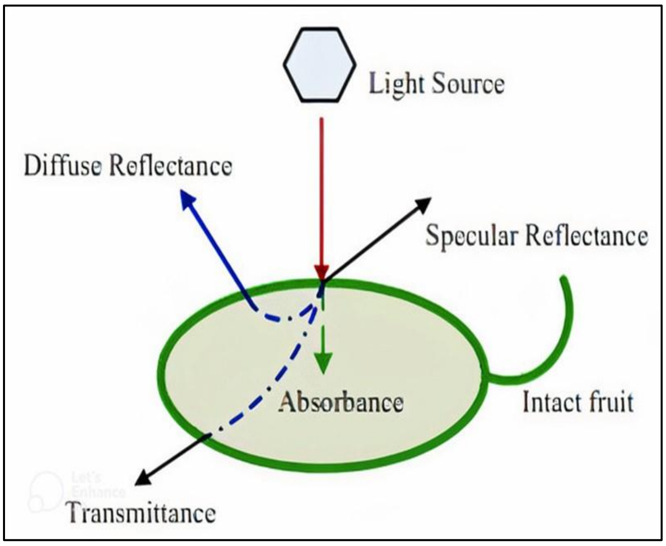

Fig 1 Interaction between NIR Radiation with Biological Object in Wavelength Range 780 – 2500 nm [7].

By analyzing the intensity of absorbed or reflected light, NIRS can quantify the concentration of specific compounds, adhering to the Beer-Lambert law, which relates absorbance directly to concentration, thus allowing for the precise determination of various components within a sample.

The NIRS's versatility also permits the identification of nuanced differences related to the beans' geographic origin, genotype, or post-harvest treatments, offering a comprehensive insight into their quality. A crucial step in adopting NIRS for cocoa bean analysis is calibration, where a model is developed to link spectral data from known samples to their measured properties, facilitating the property prediction of new samples based on their spectral information as described in spectral pattern in Figure 2. This integration of NIRS in the cocoa industry is a promising advancement that enhances the efficiency, accuracy, and depth of quality control and analysis processes in cocoa production.

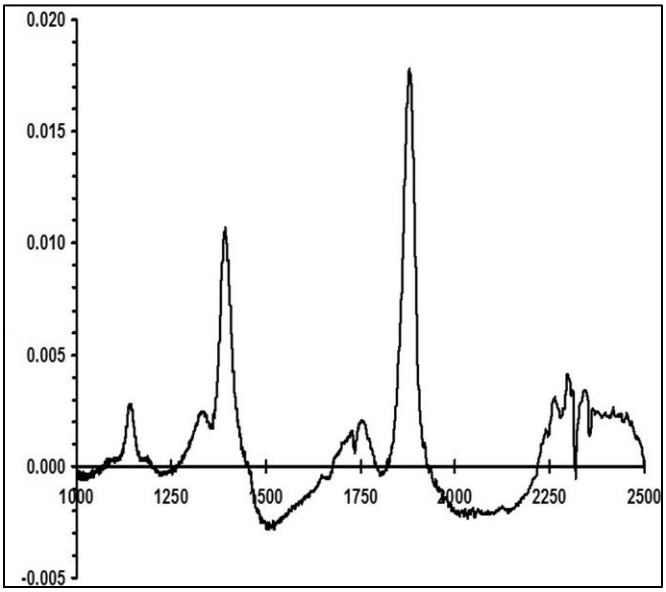

Fig 2 Spectra Pattern of Biological Object in the NIR Wavelength Region [8]

The versatility, rapidity, and non-destructive character of NIRS make it applicable in numerous fields. In agriculture, it serves to gauge moisture, protein, and fat levels in crops and grains, aiding in improved crop management and processing. The pharmaceutical sector employs NIRS for quality control measures and to ensure uniformity in medication. Within the food industry, it's instrumental in evaluating product quality, such as determining fat content in dairy, assessing the ripeness of fruits, or verifying the authenticity of oils. Medically, NIRS is valuable for monitoring oxygen saturation and blood hemoglobin levels, crucial for diagnosing various health conditions [3,9,10].

Industrial applications benefit from its material identification capabilities and quality control of products, including the measurement of coating thickness. Additionally, NIRS has a role in environmental monitoring, where it assists in identifying contaminants in water and soil. The advent of portable NIRS devices further broadens its application, enabling field-based testing and in-process control in manufacturing, underscoring the technology's capacity to revolutionize real-time monitoring and analysis where traditional methodologies may fall short due to time or cost constraints

II. MATERIALS AND METHODS

➤ *NIR Spectral Data Acquisition*

The acquisition of NIR spectra from cocoa bean samples involves several steps designed to ensure accurate, reproducible results that can provide valuable insights into the quality and characteristics of the beans. The process broadly encompasses sample preparation, spectral acquisition, and data analysis phases [11].

The prepared sample is placed in an appropriate container or holder that is compatible with the NIR instrument (*PSD NIRS iKakao USK*) being used. For powders, this might be a small cup or a rotating sample holder to present a uniform, flat surface to the instrument as illustrated in Figure 3.

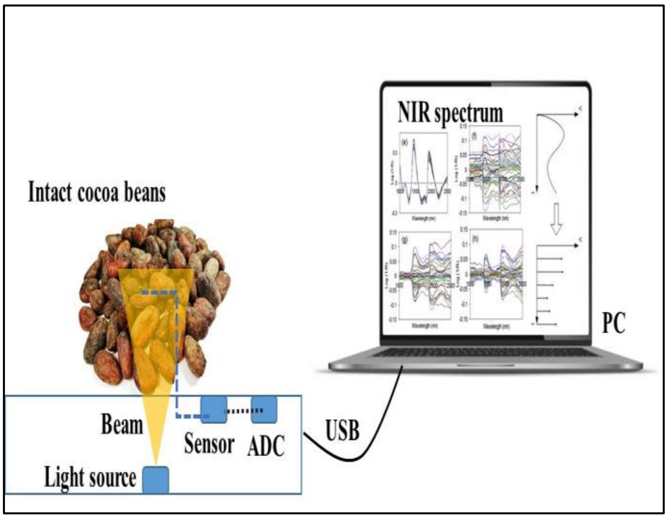

Fig 3 Spectral Data Acquisition of Intact Cocoa Bean Samples in NIR Region 1000 – 2500 nm [12].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR404

Before acquiring spectra from cocoa bean samples, the NIR spectrometer were calibrated. This involves running standard reference materials through the NIRS instrument to ensure accurate wavelength and absorbance readings.

- *Data Analysis*

The raw spectra obtained from previous phase, often require preprocessing to correct for baseline drift, remove noise, and normalize the data. Spectra smoothing was employed to pre=process the NIR spectral data of the cocoa beans samples.

The preprocessed spectra are analyzed to extract relevant information. This can involve comparing the spectra to calibrated models that relate specific spectral features to the concentrations of interest within the sample. Multivariate techniques, including multivariate calibration models like Partial Least Squares (PLS) was used to develop prediction models, as presented in Figure 4 used to determine quality parameters of intact cocoa beans, where in this study, we predict moisture contents [12,13].

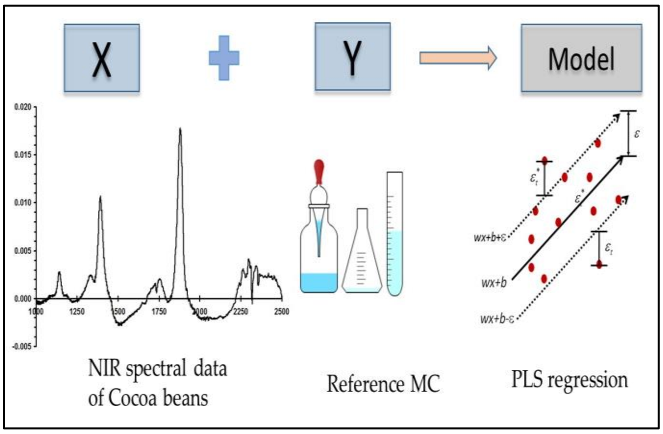

Fig 4 NIRS Based Model to Predict Moisture Content of Intact Cocoa Beans [11].

By applying these models to the NIR spectra of unknown samples, the content of moisture, fat, polyphenols, and other quality parameters can be quantified rapidly and non-destructively. NIRS with PLS handles the high dimensionality of NIR data effectively by extracting relevant information from complex spectral datasets, enabling the modeling of nonlinear relationships between spectra and properties.

PLS can address collinearity issues present in NIR spectral data, where multiple spectral variables are highly correlated. It helps in identifying the spectral features that contribute most to predicting the target properties.

- *NIRS Model's Performance Evaluation*

When evaluating the performance of a prediction model in Near-Infrared Spectroscopy (NIRS) applications for cocoa beans, several key metrics and validation methods can be employed to assess the accuracy, reliability, and robustness of the model. These evaluation techniques help ensure that the model can effectively predict the properties of cocoa beans based on their NIR spectral data [14,15].

The coefficient of determination or abbreviated as R^2 measures the proportion of variance in the predicted property that is explained by the model. A higher R^2 value close to 1 indicates a better fit between the predicted and actual values.

Beside R^2 coefficient, the root mean square error (RMSE) was also used to quantify the average difference between the predicted and actual values. Lower RMSE values suggest better predictive performance [16–18].

III. RESULTS AND DISCUSSION

The NIR region of the electromagnetic spectrum reveals spectral patterns that are indicative of various quality parameters crucial to the cocoa industry. These patterns emerge from the interaction of NIR light with the molecular constituents of the cocoa beans, particularly the vibrations of hydrogen-containing bonds such as O-H, N-H, and C-H as shown in Figure 5, which are abundant in the chemical composition of cocoa beans.

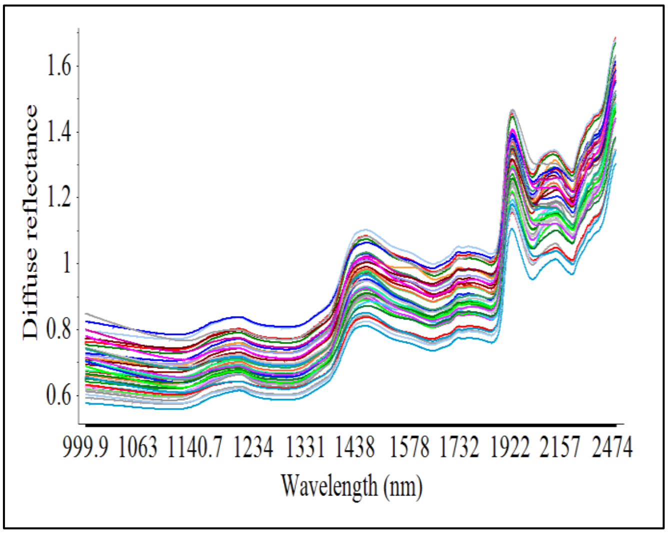

Fig 5 NIR Spectra Feature of Intact Cocoa Beans

One of the most critical factors in assessing cocoa bean quality is moisture content. Cocoa beans with high moisture levels are more susceptible to mold and fermentation during storage. The O-H bond in water molecules strongly absorbs NIR light, and the intensity of this absorption can be correlated with the amount of moisture in the cocoa beans.

Cocoa beans are prized for their fat (cocoa butter) content, which significantly impacts chocolate's flavor and texture characteristics. The C-H bonds present in the fats exhibit specific absorption patterns in the NIR spectrum. By analyzing these patterns, the fat content of cocoa beans can be quantitatively determined, providing valuable information for grading and pricing. The corrected spectral data using Smoothing approach is presented in Figure 6.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR404

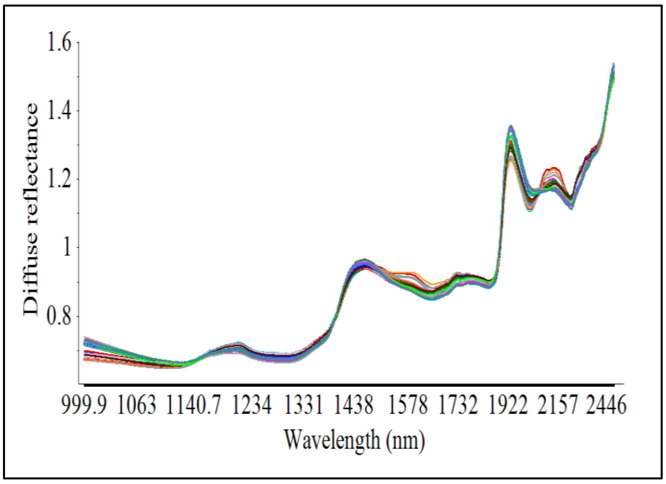

Fig 6 Corrected Spectral Data using Smoothing Approach

Polyphenols are important antioxidants in cocoa beans that contribute to the health benefits and bitterness of chocolate. The structure of polyphenols includes aromatic rings with attached O-H groups, which have distinct NIR absorption features. Monitoring these features allows for the assessment of polyphenol content, offering insights into the beans' flavor profile and nutritional value.

The fermentation and roasting processes significantly affect the flavor profile and quality of cocoa beans. Changes in the molecular structure of proteins, carbohydrates, and fats during these processes can be monitored using NIR spectroscopy. For instance, the Maillard reaction, a form of non-enzymatic browning during roasting, produces complex molecules that can be detected and measured through their unique NIR spectral signatures. Spectral data analysis using principal component analysis is presented in Figure 7.

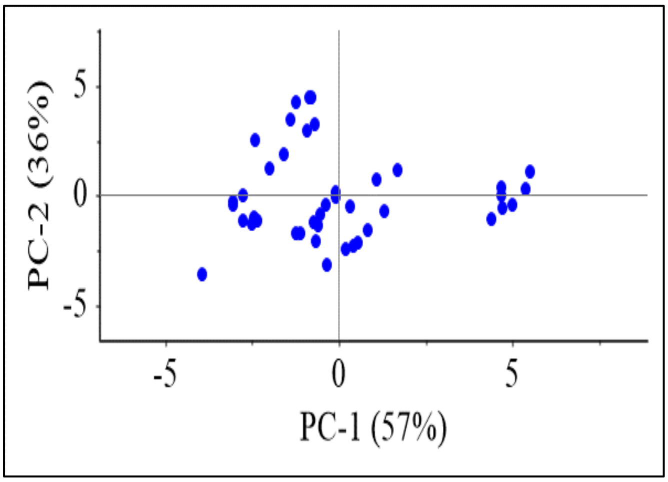

Fig 7 Data Analysis using PCA of NIR Spectrum

Principal Component Analysis (PCA) is an indispensable tool in NIR spectroscopy for simplifying complex spectral datasets from cocoa beans and extracting essential information. By reducing the dimensionality of the data, PCA efficiently condenses the multitude of spectral variables into a smaller set of principal components that retain the most significant variation. This process not only aids in noise reduction and filtering out irrelevant features but also enhances the signal-to-noise ratio, allowing for clearer spectral patterns to emerge.

With PCA, researchers can visualize similarities and differences among samples, identify key spectral regions contributing to variations, and cluster samples based on spectral similarities for pattern recognition and quality control. PCA also plays a crucial role in outlier detection, helping pinpoint unusual or erroneous samples that deviate from the norm.

Furthermore, when combined with multivariate techniques like Partial Least Squares (PLS), PCA serves to optimize calibration models for predicting properties such as moisture content, fat content, and polyphenol levels in cocoa beans. Overall, PCA facilitates data interpretation, model optimization, and quality assessment in NIR spectroscopy, offering valuable insights into the composition and quality attributes of cocoa beans for informed decision-making in cocoa processing and quality control processes.

Calibrating NIRS with (PLS to predict moisture content in cocoa beans involves a systematic process to develop a reliable model correlating the NIR spectral data of cocoa bean samples with their known moisture levels as presented in Figure 8. Initially, a diverse sample set with varying moisture content is selected and measured using both NIR spectroscopy and reference methods to create a dataset pairing spectral data with actual moisture values.

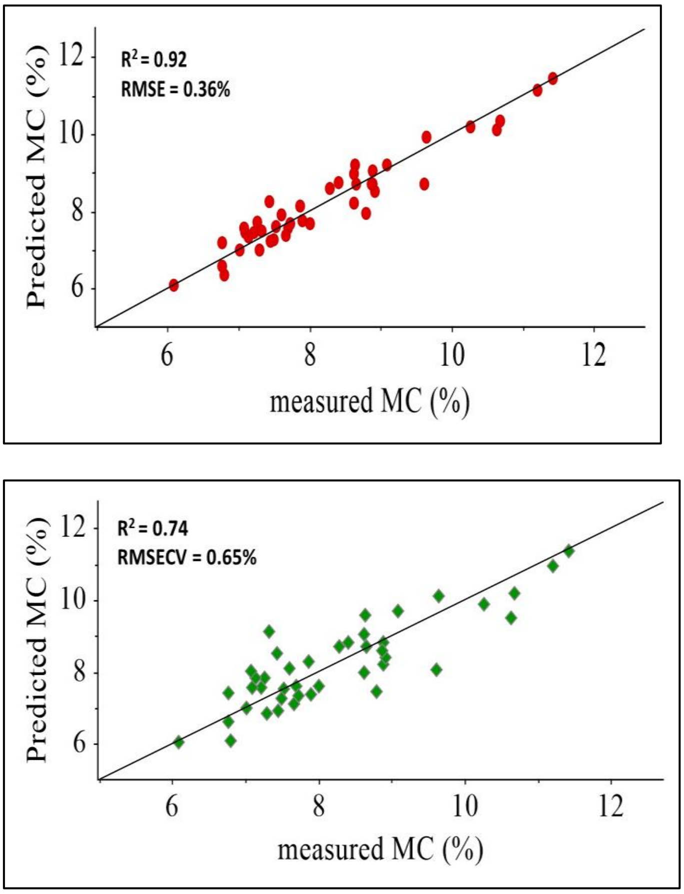

Fig 8 NIRS with PLS to Predict Moisture Content in Calibration and Cross Validation Performances

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR404

Preprocessing steps like baseline correction and noise filtering was applied to the NIR data to enhance its quality. The PLS regression algorithm is then utilized to establish a predictive model by extracting latent variables that best predict moisture content. Optimization of the model through parameter tuning and cross-validation validates its predictive performance. Evaluation metrics such as R^2 and RMSE assess the model's accuracy. Following successful validation, the calibrated PLS model can be applied to predict moisture content in new cocoa bean samples, aiding in quality assessment and process optimization in cocoa processing. Continuous monitoring and refinement ensure the model's ongoing accuracy and applicability. In essence, the calibration of NIRS using PLS for moisture content prediction in cocoa beans enables accurate, non-destructive moisture analysis for quality control and process improvement in the cocoa industry.

A scatter plot graph between measured moisture content (X-axis) and predicted moisture content (Y-axis) visually represents the relationship between the actual and predicted values. With an R square value of 0.92 for calibration and 0.74 for cross-validation, along with RMSE values of 0.36% and 0.65% respectively.

In the scatter plot for the calibration dataset, where the model was trained, we would expect to see the points aligning closely along a diagonal line. This alignment indicates a strong positive correlation between the measured and predicted moisture content values. With an R square value of 0.92, the points should cluster tightly around the best-fit line (slope of 1) with minimal scatter. The low RMSE of 0.36% further confirms the accuracy of the model in predicting moisture content during the calibration phase.

For the cross-validation dataset, where the model's performance was tested on unseen data, the scatter plot may show slightly more dispersion compared to the calibration data. An R square value of 0.74 suggests that the model explains 74% of the variability in the moisture content predictions, indicating reasonably good performance. The RMSE value of 0.65% represents the average deviation of the predicted values from the actual values, with lower values indicating better model fit. Despite some variability, the scatter plot should still demonstrate a positive relationship between measured and predicted moisture content, albeit with slightly more spread-out data points compared to the calibration data.

In the utilization of PLS regression for predicting moisture content in cocoa beans by means of NIRS, the loading plot plays a pivotal role in deciphering the crucial wavelengths influencing the prediction model as shown in Figure 9. This visualization tool graphically illustrates the weights or loadings of each wavelength in the NIR spectrum in relation to the latent variables extracted by the PLS algorithm.

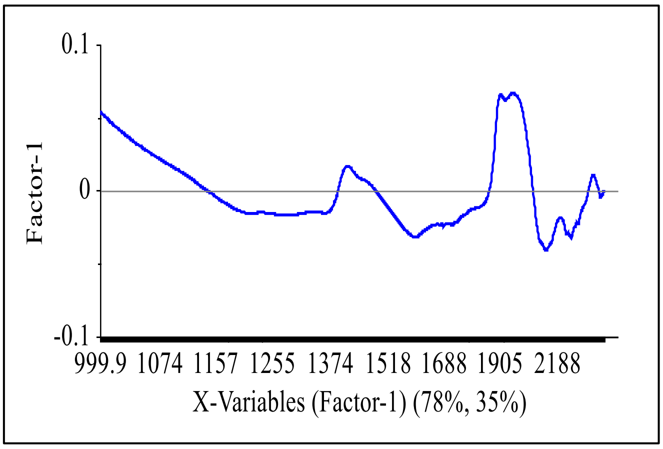

Fig 9 Loading Plot of PLS Regression to Determine Important Wavelength in the NIR Region for Moisture Prediction

By scrutinizing the loading plot, researchers can discern the wavelengths that hold the most significance in predicting moisture content accurately. Wavelengths exhibiting higher loadings are indicative of strong correlations with the predicted property, offering valuable insights into the spectral features that directly impact moisture prediction within cocoa beans. These key wavelengths, usually identified through their pronounced positive or negative loadings, serve as essential components for the interpretation, refinement, and optimization of the prediction model.

Furthermore, the loading plot aids in feature reduction by spotlighting the most influential spectral variables, streamlining the model for enhanced interpretability and performance. Researchers can utilize the insights gleaned from the loading plot to refine the PLS model, focusing on the critical wavelengths that align with variations in moisture levels. This refined model, enriched with valuable spectral information identified through the loading plot, forms the cornerstone for informed decision-making in cocoa processing and quality control endeavors. The strategic use of important wavelengths extracted from the loading plot not only optimizes the prediction model's accuracy but also guides researchers towards selecting key features for model optimization, ensuring robustness and efficacy in moisture content prediction for cocoa beans using NIRS.

The use of NIR spectroscopy to analyze cocoa beans provides a fast, non-destructive means of assessing key quality parameters. Advances in portable NIR technology also facilitate on-site quality control, enabling cocoa bean suppliers and chocolate manufacturers to make informed decisions regarding bean selection, processing conditions, and final product quality. The spectral data obtained from NIR analysis, when combined with sophisticated data analysis techniques, can predict bean quality attributes with high accuracy, thus enhancing the efficiency and sustainability of the cocoa production chain.

IV. CONCLUSION

Incorporating the performance metrics of calibration and cross-validation, highlighted by the R square and Root Mean Square Error (RMSE) values, further strengthens the conclusion. A high R square value of 0.92 for calibration signifies a strong model fit, indicating that 92% of the variability in moisture content predictions can be explained by the model, yielding accurate and reliable results. The low RMSE value of 0.36% in calibration confirms the model's precision in predicting moisture content in cocoa beans. Comparatively, the slightly lower R square value of 0.74 in cross-validation may indicate a slightly diminished but still acceptable model performance on unseen data, explaining 74% of the variability in moisture predictions. The RMSE value of 0.65% in cross-validation, while slightly higher than in calibration, still reflects a consistently good level of accuracy.

These performance metrics validate the effectiveness of the model in both calibration and cross-validation scenarios, reinforcing the reliability and robustness of the moisture prediction model for cocoa beans using NIRS. By combining the insights from the loading plot analysis with the exemplary performance metrics in both calibration and cross-validation, the conclusion underscores the model's accuracy, reliability, and practical relevance for moisture prediction in cocoa bean quality control and processing applications.

Leveraging the insights from the loading plot enables researchers to refine the model, optimize feature selection, and streamline the analysis process, leading to improved interpretability and model performance. In application, the refined PLS model, informed by key spectral features highlighted in the loading plot, not only enhances the quality of moisture predictions but also guides decision-making in cocoa processing and quality control activities. Overall, the strategic integration of loading plot analysis in PLS calibration for moisture prediction in cocoa beans using NIRS facilitates informed model optimization, robust prediction capabilities, and enhanced understanding of the spectral patterns associated with moisture content in cocoa beans

REFERENCES

[1]. A. Drees, J. Brockelt, L. Cvancar, M. Fischer, Rapid determination of the shell content in cocoa products using FT-NIR spectroscopy and chemometrics, Talanta. 256 (2023) 124310. https://doi.org/10.1016/j.talanta.2023.124310.

[2]. I.A. Santos, D.G. Conceição, M.B. Viana, G. de J. Silva, L.S. Santos, S.P.B. Ferrão, NIR and MIR spectroscopy for quick detection of the adulteration of cocoa content in chocolates, Food Chem. 349 (2021) 1–6. https://doi.org/10.1016/j.foodchem.2021.129095.

[3]. C. Hernández-Hernández, V.M. Fernández-Cabanás, G. Rodríguez-Gutiérrez, Á. Fernández-Prior, A. Morales-Sillero, Rapid screening of unground cocoa beans based on their content of bioactive compounds by NIR spectroscopy, Food Control. 131 (2022). https://doi.org/10.1016/j.foodcont.2021.108347.

[4]. C. Pasquini, Near infrared spectroscopy: A mature analytical technique with new perspectives – A review, Anal. Chim. Acta. 1026 (2018) 8–36. https://doi.org/10.1016/j.aca.2018.04.004.

[5]. A.A. Munawar, D. von Hörsten, J.K. Wegener, E. Pawelzik, D. Mörlein, Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics, Eng. Agric. Environ. Food. 9 (2016) 208–215. https://doi.org/10.1016/j.eaef.2015.12.004.

[6]. K.B. Walsh, V.A. McGlone, D.H. Han, The uses of near infra-red spectroscopy in postharvest decision support: A review, Postharvest Biol. Technol. 163 (2020) 111139. https://doi.org/10.1016/j.postharvbio.2020.111139.

[7]. A.A. Munawar, Kusumiyati, D. Wahyuni, Near infrared spectroscopic data for rapid and simultaneous prediction of quality attributes in intact mango fruits, Data Br. 27 (2019) 104789. https://doi.org/10.1016/j.dib.2019.104789.

[8]. A.A. Munawar, D.V. Hörsten, D. Mörlein, E. Pawelzik, J.K. Wegener, Rapid and non-destructive prediction of mango sweetness and acidity using near infrared spectroscopy, in: Lect. Notes Informatics (LNI), Proc. - Ser. Gesellschaft Fur Inform., 2013.

[9]. H.H. Farghal, S.T. Mansour, S. Khattab, C. Zhao, M.A. Farag, A comprehensive insight on modern green analyses for quality control determination and processing monitoring in coffee and cocoa seeds, Food Chem. 394 (2022) 133529. https://doi.org/10.1016/j.foodchem.2022.133529.

[10]. A. Biancolillo, S. Preys, B. Gaci, J.L. Le-Quere, H. Laboure, Z. Deuscher, V. Cheynier, N. Sommerer, N. Fayeulle, P. Costet, C. Hue, R. Boulanger, K. Alary, M. Lebrun, M. Christine Lahon, G. Morel, I. Maraval, F. Davrieux, J.M. Roger, Multi-block classification of chocolate and cocoa samples into sensory poles, Food Chem. 340 (2021). https://doi.org/10.1016/j.foodchem.2020.127904.

[11]. R. Hayati, Z. Zulfahrizal, A.A. Munawar, Robust prediction performance of inner quality attributes in intact cocoa beans using near infrared spectroscopy and multivariate analysis, Heliyon. 7 (2021) e06286. https://doi.org/10.1016/j.heliyon.2021.e06286.

[12]. Z. Zulfahrizal, A.A. Munawar, Near-Infrared Technology in Agriculture : Non-Destructive Determination of Inner Quality Parameters in Intact Cocoa Beans, Trends Sci. 18 (2021).

[13]. Yusmanizar, I.S. Setiasih, S. Nurjanah, M. Muhaeimin, B. Nurhadi, S. Rosniawaty, A.A. Munawar, Fast and Non-Destructive Prediction of Moisture Content and Chologenic Acid of Intact Coffee Beans Using Near Infrared Reflectance Spectroscopy, in: IOP Conf. Ser. Mater. Sci. Eng., Institute of Physics Publishing, 2019. https://doi.org/10.1088/1757-899X/506/1/012033.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR404

[14]. U. Siripatrawan, Y. Makino, Spectrochimica Acta Part A : Molecular and Biomolecular Spectroscopy Hyperspectral imaging coupled with machine learning for classification of anthracnose infection on mango fruit, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 309 (2024) 123825. https://doi.org/10.1016/j.saa.2023.123825.

[15]. P. Mishra, J. Verschoor, M.N. Vries, G. Polder, M.P. Boer, Portable near-infrared spectral imaging combining deep learning and chemometrics for dry matter and soluble solids prediction in intact kiwifruit, Infrared Phys. Technol. 131 (2023) 104677. https://doi.org/10.1016/j.infrared.2023.104677.

[16]. M. Wohlers, A. Mcglone, E. Frank, G. Holmes, Chemometrics and Intelligent Laboratory Systems Augmenting NIR Spectra in deep regression to improve calibration, Chemom. Intell. Lab. Syst. 240 (2023) 104924. https://doi.org/10.1016/j.chemolab.2023.104924.

[17]. S. Yu, K. Huan, X. Liu, L. Wang, X. Cao, Quantitative model of near infrared spectroscopy based on pretreatment combined with parallel convolution neural network, Infrared Phys. Technol. 132 (2023) 104730. https://doi.org/10.1016/j.infrared.2023.104730.

[18]. L. Bawitlung, D. Bhatt, A.K. Maurya, A. Pal, D.U. Bawankule, A. Mishra, P. Gupta, C.S. Chanotiya, Industrial Crops & Products Chemical investigation and biological activities of Patchouli (Pogostemon cablin (Blanco) Benth) essential oil, Ind. Crop. Prod. 188 (2022) 115504. https://doi.org/10.1016/j.indcrop.2022.115504.